THE STRUCTURE OF THE NITRILASE FROM *RHODOCOCCUS RHODOCHROUS* J1: HOMOLOGY MODELING AND THREE-DIMENSIONAL RECONSTRUCTION

Robert Ndoria Thuku

A minithesis submitted in partial fulfilment of the requirements for the degree of Magister Scientiae (Structural Biology) in the Department of Biotechnology, University of the Western cape

Supervisor: Associate Professor Bryan Trevor Sewell

March 2006

KEYWORDS

Nitrilase

Rhodococcus rhodochrous J1

Bioinformatics

Protein

Electron Microscopy

IHRSR

3D reconstruction

Docking

Oligomeric structure

Industrial enzyme

ABSTRACT

THE STRUCTURE OF THE NITRILASE FROM *RHODOCOCCUS RHODOCHROUS* J1: HOMOLOGY MODELING AND THREE-DIMENSIONAL RECONSTRUCTION

Robert Ndoria Thuku

M.Sc. minithesis, Faculty of Science, University of the Western Cape and Electron Microscope Unit, University of Cape Town

The nitrilases are an important class of industrial enzymes that are found in all phyla. These enzymes are expressed widely in prokaryotes and eukaryotes. Nitrilases convert nitriles to corresponding acids and ammonia. They are used in industry as biocatalysts because of their specificity and enantioselectivity. These enzymes belong to the nitrilase superfamily in which members share a common $\alpha\beta\beta\alpha$ structural fold and a unique cys, glu,lys catalytic triad with divergent N- and C-terminals.

There are four atomic structures of distant homologues in the superfamily, namely 1ems, 1erz, 1f89 and 1j31. All structures have two-fold symmetry which conserves the $\alpha\beta\beta\alpha$-$\alpha\beta\beta\alpha$ fold across the dimer interface known as the A surface. The construction of a 3D model based on the solved structures revealed the enzyme has two significant insertions in its sequence relative to the solved structures, which possibly correspond to the C surface. In addition there are intermolecular interactions in a region of a conserved helix, called the D surface. These surfaces contribute additional interactions responsible for spiral formation and are absent in the atomic resolution homologues.

The recombinant enzyme from *R.rhodochrous* J1 was expressed in *E. coli* BL21 cells and eluted by gel filtration chromatography as an active 480 kDa oligomer and an inactive 80 kDa dimer in the absence of benzonitrile. This contradicts previous observations, which reported the native enzyme exists as an inactive dimer and elutes as a decamer in the presence benzonitrile. Reducing SDS-PAGE showed a subunit atomic mass of ~40 kDa. EM and image analysis revealed single particles of various shapes and sizes, including c-shaped particles, which could not form spirals due to steric hindrances in its C terminal. Chromatographic re-elution of an active fraction of 1-month old J1 nitrilase enabled us to identify an active form with a mass greater than 1.5 MDa. Reducing SDS-PAGE, N-terminal sequencing and mass spectroscopy showed the molecular weight was ~36.5 kDa as result of specific proteolysis in its C terminal. EM revealed the enzyme forms regular long fibres. Micrographs (109) were recorded on film using a JEOL 1200EXII operating at 120 kV at 50K magnification. Two independent 3D reconstructions were generated using the IHRSR algorithm executed in SPIDER. These converged to the same structure and the resolution using the FSC 0.5 criterion was 1.7 nm.

The helix structure has a diameter of 13nm with ~5 dimers per turn in a pitch of 77.23 Å. Homology modeling and subsequent fitting into the EM map has revealed the helix is built primarily from dimers, which interact via the C and D surfaces. The residues, which potentially interact across the D surface, have been identified and these confer stability to the helix. The conservation of the insertions and the possibility of salt bridge formation on the D surface suggest that spiral formation is common among microbial nitrilases. Furthermore, the presence of the C terminal domain in J1 nitrilase creates a steric hindrance that prevents spiral formation. When this is lost – either by specific proteolysis or autolysis - an active helix is formed.

March 2006

DECLARATION

I declare that *The Structure of the Nitrilase from Rhodococcus rhodochrous J1: Homology Modeling and Three-Dimensional Reconstruction* is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Robert Ndoria Thuku March 2006

Signed: 

ACKNOWLEDGEMENTS

I would like to thank my supervisor Assoc. Prof B.T Sewell for his endless support, encouragement and guidance in the field of structural biology.

I am grateful to Dr. Brendon Price and the staff of the EMU for their assistance, technical skills transfer and advice. Many thanks also to Dr. Dean Brady of CSIR (Bio/Chemtek), my fellow students and friends for their support and helpful discussions.

I am indebted to my family for their love, patience, support and encouragement.

Finally, I gratefully acknowledge the financial assistance of the CSIR (Bio/Chemtek), the University of Cape Town and the Carnegie Corporation of New York.

LIST OF ABBREVIATIONS

3D PSSM	three-dimensional position specific scoring matrix
BSA	bovine serum albumin
CD	circular dichroism
CFE	cell free extract
D_1	one-fold dihedral point group
EM	electron microscopy
FSC	Fourier shell correlation
IEX	ion exchange
IHRSR	iterative helical real space reconstruction
IPTG	isopropyl-β-D-thiogalactopyranoside
J1 nitrilase	*Rhodococcus rhodochrous* J1 nitrilase
LB	Luria broth
MALDI-TOF	matrix-assisted laser desorption/ionization - time of flight
MRA	multi-reference alignment
MSA	multivariate statistical analysis
MW	molecular weight
NAD	nicotinamide adenine dinucleotide
OD	optical density
PDB	protein data bank
PROSA	protein structure analysis
PSIPRED	protein structure prediction
RMSD	root mean square deviation
S_σ	σ-fold screw symmetry
SDS-PAGE	sodium dodecyl suphate – polyacrylamide gel electrophoresis
SNR	signal-to-noise ratio
UV	Ultraviolet

TABLE OF CONTENTS

CHAPTER 1

NITRILASE REVIEW

1.1 Introduction

Nitrilases are important industrial enzymes that hydrolyze nitriles to the corresponding carboxylic acids and ammonia. Enzymes in the nitrilase superfamily are widely expressed in prokaryotes and eukaryotes. Plants synthesize nitriles such as indole-3-acetonitrile, cyanoglycosides, ricinine, cyanolipids and β-cyano-L-alanine, while nitrile-converting biocatalysts are used extensively to manufacture compounds such as acrylonitrile, adiponitrile and acetonitrile in the chemical industry (Kobayashi and Shimizu, 2000; Kobayashi et al., 1998). Diversa Corporation (US) has established a novel enzyme library of about 300 nitrilase sequences by screening large genomic libraries from various habitats (Burk MJ, 2002; DeSantis et al., 2002; Robertson et al., 2003). This corporation also attempted to crystallize about 200 nitrilases with no success (DeSantis et al., 2002). Despite the availability of a large number of nitrilase sequences, no crystal structure of a nitrilase exists. There is a need to understand the details of their biological functions, catalytic mechanisms and specificity in structural terms. The aim is to gain these insights using established techniques of structure determination and ultimately, to engineer improved enzymes with specific properties that maximize their use in industry.

1.2 Nitrilase superfamily

Initially, the carbon-nitrogen (C-N) hydrolyzing enzymes included nitrilases, cyanide hydratase, aliphatic amidase and β-ureidopropionase (Bork and Koonin, 1994). These enzymes have been classified into 13 distinct branches based on amino acid sequence analysis as the nitrilase superfamily (Pace and Brenner, 2001). Enzymes in the superfamily are used by prokaryotes and eukaryotes to hydrolyze and condense a variety of protein and non-protein C-N bonds (Brenner, 2002). Branch 1 enzymes consist of nitrilases, cyanide dihydratases and cyanide hydratases, which digest only nitriles. However, the *in vivo* substrates for nitrilases remain unknown (except for the cyanide dihydratases and cyanide hydratases). The majority of the remaining 12 branches exhibit amidase activity with varying specificity. Branches 2, 3, 4, 5 consist of aliphatic amidases, amino-terminal amidase, biotinidase, and β-ureidopropionase respectively. Branch 6 enzymes, namely the carbamylases, are specific for the decarbamylation of D-amino acids. Branch 7 and 8 enzymes consist of prokaryotic and eukaryotic NAD synthetases, which utilize either glutamine or free ammonia to catalyze the transfer of amide nitrogen to a wide range of substrates.

A recent study on the NAD^+ synthetase from *Mycobacterium tuberculosis* concluded glutamine-amidotransferase activity depends on an amino-terminal nitrilase domain, namely glutamine amide transfer (GAT) domain (Bellinzoni et al., 2005). Mutation of the active cysteine in the nitrilase domain inactivated the NAD^+ synthetase from *M.tuberculosis.* This suggests that the nitrilase domain is sometimes fused to another domain and may be part of a complex chain of reactions necessary for the production of metabolites in various biological systems (Pace and Brenner, 2001). Hence there is need to target these enzymes and associated domains as a potential source of novel antibacterial agents (Bellinzoni et al., 2005). Only branch 9 enzymes, namely the alipoprotein N-acyltransferases, catalyze an amide condensation reaction. The members in the superfamily share a common fold with a conserved catalytic cys-glu-lys site, which is surrounded by signature sequences conserved within each branch but differ between branches (Brenner, 2002). The functions of enzymes in branches 11,12 and 13 are unknown.

1.3 Nitrile-degrading enzymes

The nitrile-degrading enzymes consist of nitrilases and amidases. The microbial degradation of nitriles proceeds via two different enzymatic pathways (figure 1.1). Firstly, nitrilase, cyanide dihydratase and cyanide hydratase catalyze the direct hydrolysis of nitriles into the corresponding carboxylic acid and ammonia and secondly, nitrile hydratase convert nitriles to the corresponding amide which is then hydrolysed to the corresponding carboxylic acid and ammonia by an amidase (Kobayashi et al., 1992, 1998; Ramakrishna et al., 1999). While nitrilases hydrolyze larger nitriles (RCN), cyanide dihydratases and cyanide hydratases hydrolyze cyanide (HCN) only.

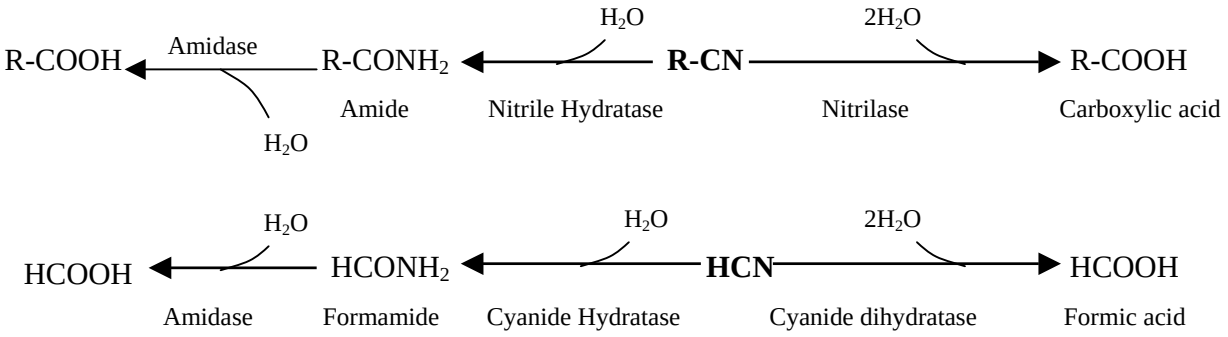


Figure 1.1. Enzymatic pathways for nitrile hydrolysis.

The first bacterial nitrilase was isolated from a soil bacterium, a pseudomonad, using ricinine a naturally occurring nitrile as the sole carbon source (Hook and Robinson 1964; Robinson and

Hook, 1964). This enzyme hydrolysed ricinine and 2-pyridones to the corresponding acid and ammonia. In plants, a nitrilase catalyzes the hydrolysis of indole-3-acetonitrile (IAN) to indole-3-aceticacid (IAA), a plant growth hormone (Kobayashi et al., 1995; Bartling et al., 1992). In the plant *Arabidopsis thaliana*, two metabolic pathways catalyzed by a nitrilase for the synthesis of IAA were identified (Bartling et al., 1994). *A. thaliana* has four nitrilase homologues; NIT1, NIT2, NIT3 and NIT4. NIT4 does not cleave IAN and is highly specific for β-cyano-L-alanine (Piotrowski et al., 2001).

Because of the potential uses of nitrilases in industry, most of the known enzymes have been isolated from a range of bacteria and fungi. The nitrilase of *Rhodococcus rhodochrous* J1 was expressed when isovaleronitrile is used as an inducer (Nagasawa et al., 1988). Interestingly, when these cells were cultured in a medium containing urea and cyclohexanecarboxamide in the presence of cobalt ions, high and low molecular weight nitrile hydratases are induced respectively (Nagasawa et al., 1991). Other organisms that express nitrilases when induced with isovaleronitrile include *Alcaligenes faecalis* JM3, *Rhodococcus* ATCC 39484 and *Rhodococcus rhodochrous* K22 (Nagasawa et al., 1990; Stevenson et al., 1992; Kobayashi et al., 1990). The nitrilase from *Rhodococcus rhodochrous* NCIMB 11216 is induced with propionitrile as the sole carbon and nitrogen source (Hoyle et al., 1998).

Cyanide dihydratases are classified under branch 1 enzymes of the nitrilase superfamily, which hydrolyze cyanide (free form or metal cyanide complexes) to formate and ammonia. These enzymes have been identified in *Alcaligenes xylosoxidans* subsp. *denitrificans* (Ingvorsen et al., 1991), *Bacillus pumillus* C1 (Meyers et al., 1993; Jandhyala et al., 2003), *Pseudomonas fluorescens* NCIMB 11764 (Kunz et al., 1994), and *Pseudomonas stutzeri* AK61 (Watanabe et al., 1998). The enzyme from *Alcaligenes xylosoxidans* subsp. *denitrificans* (Ingvorsen et al., 1991) was induced by cyanide. None of these enzymes was reported to form formamide in the reactions they catalyze.

Another group of branch 1 nitrilases, namely the cyanide hydratases, hydrolyze cyanide to formamide. The enzyme, cyanide hydratase, was first identified in the fungus *Stemphylium loti*, a pathogen of cyanogenic plants (Fry and Millar, 1972). This enzyme is not related to the metal containing nitrile hydratases, which have no relation to enzymes in the nitrilase superfamily. The cyanide hydratases, which have been studied so far, include those from *Gloeoecercospora sorghi* (Fry and Munch, 1975; Wang et al., 1992; Wang and VanEtten, 1992), *Fusarium lateritium* (Cluness et al., 1993), *Fusarium solani* (Barclay et al., 1998), *Fusarium solani* IMI 196840 (Nolan et al., 2003) and *Leptosphaeria maculans* (Sexton and Howlett, 2000).

Amidases are enzymes involved in nitrogen metabolism in prokaryotic and eukaryotic cells (Banerjee et al., 2000). While nitrilases cleave the C-N bond of a nitrile, amidases cleave the C-N bond of an amide to produce the corresponding carboxylic acid and ammonia. These enzymes comprise branch 2 of the nitrilase superfamily and examples include amidases from *Pseudomonas aeruginosa*, *Rhodococcus erythropolis*, *Helicobacter pylori* and *Bacillus stearothermophilus* (Pace and Brenner, 2001). Mutational and comparative modeling studies on the amidase from *Pseudomonas aeruginosa* concluded it has a cys-glu-lys catalytic triad that aligned with that of the NitFhit structure (PDB ID, 1EMS) from the worm *Caernorhabditis elegans* (Novo et al., 1995, 2002; Farnaud et al., 1999; Pace et al., 2000). There exists a class of amidases that are structurally or sequentially similar to nitrilases. The bacteria *Rhodococcus rhodochrous* J1 is reported to have both nitrilase and amidase enzymes, which hydrolyze benzonitrile and benzamide respectively using similar catalytic sites (Kobayashi et al. 1998a). There is no sequence similarity between the amidase from *R. rhodochrous* J1 with the amidase from *P. aeruginosa*. However, the nitrilase from *R. rhodochrous* J1 has significant sequence similarity with the aliphatic amidase from *P. aeruginosa* and both have cysteine as the active nucleophile (Kobayashi et al., 1998b). Notable differences between nitrilases are evident in their substrate specificity, enzyme structure, pH, and temperature optima (O'Reilly and Turner, 2003).

1.4 Substrate specificity and activity

Studies using a range of substrates have shown that some nitrilases are specific for aromatic and aromatic-substituted nitriles, arylacetonitriles, aliphatic nitriles or bromoxynil. On the basis of their physical properties and substrate specificities, various authors have adopted a broad classification as shown in Table 1.

1.4.1 *Aromatic nitrilases*

Aromatic nitrilases have been purified from the bacteria *Pseudomonas* sp. (Hook and Robinson, 1964; Robinson and Hook, 1964), *Norcadia sp*. NCIB 11215 and NCIB 11216 (Harper 1985; Harper 1977), *Arthrobacter sp*. Strain J1 (Bandyopathay et al., 1986*)*, *R. rhodochrous* J1 (Kobayashi et al., 1989), *Klebsiella pneumonae* subsp. *ozaenae* (Stalker et al., 1988) and the fungus *Fusarium solani* (Harper, 1977). This class consists of nitrilases that hydrolyze benzonitrile and related aromatic substrates. The best characterized nitrilase is that from *R. rhodochrous* J1, which was found to be specific for aromatic or heterocylic nitriles such as benzonitrile, 3-chlorobenzonitrile and 2-thiophenecarbonitrile. This enzyme was purified from isovaleronitrile-induced cells of *R. rhodochrous* J1 and exhibits regiospecificity

by hydrolyzing 1,3-dicyanobenzoate and 1,4-dicyanobenzene to 3-cyanobenzoate and 4-cyanobenzoate respectively (Kobayashi et al., 1988). Most of these enzymes are inactive on aliphatic nitriles except the nitrilase of *R. rhodochrous* (formerly Norcadia) NCIMB 11216, which hydrolyzes benzonitrile and propionitrile, an aliphatic nitrile (Hoyle et al., 1998). The nitrilase from *R. rhodochrous* J1 could hydrolyze acrylonitrile upon substrate-induced activation of the native organism with benzonitrile (Nagasawa et al., 2000). The nitrilase from isovaleronitrile-induced cells of *Rhodococcus* sp. ATCC 39484 was reported to prefer hydrophobic aromatic substrates (Stevenson et al., 1992). Upon binding the substrate, the activity in the enzyme was enhanced. This would suggest that the binding of an aromatic moiety in the active site triggers a conformational change that leads to enzyme association and increased activity. All these enzymes prefer aromatic substrates substituted with methyl or halogen groups at the *meta*- and *para*-position, with poor or no activity with *ortho*-substitutions (Harper 1977b; Stevenson et al., 1992). This could possibly be due to steric conflicts. The only atom tolerated at this position is fluorine, as reported for the nitrilase from the bacterium *Norcadia* sp. (*Rhodochrous*) NCIMB 11215 (Harper, 1985).

1.4.2 *Aliphatic nitrilases*

These have been purified and characterized from their native organisms with different substrate specifity. The nitrilase from crotonitrile-induced cells of *Rhodococcus rhodochrous* K22 has high activity with acrylonitrile and glutaronitrile (Kobayashi et al., 1992). The nitrilase from *Acinetobacter* sp. AK226 has high activity with benzonitrile and highest with acrylonitrile (Yamamoto and Komatsu, 1991). The rate of hydrolysis of 3-phenylpropionitrile versus that of benzonitrile is increased by a factor of 270 in the recombinant AtNIT1 nitrilase of *Arabidopsis thaliana* (Osswald et al., 2002). The AtNIT1 nitrilase prefers aliphatic nitriles and showed increased hydrolysis with increasing size of substrate. This enzyme also produced 95% amide product upon hydrolysis of 3-nitroacrylonitrile. Further substrate analysis showed that indoleacetonitrile (IAN) was not the preferred substrate for AtNIT1, 2, 3 and 4 nitrilases (Vorwerk et al., 2001). The nitrilases from *Acidovorax facilis* 72W (Gagavan et al., 1999) and *Synechocystis* sp. Strain PCC6803 (Heinemann et al., 2003) are both reported to hydrolyse fumarodinitrile (*trans*-1,2-dicyanoethene). Particularly, the enzyme from the cyanobacterium *Synechocystis* sp. Strain PCC6803 had 120 times higher relative activity with fumarodinitrile compared to that observed with benzonitrile. It also had higher activity with some aromatic, heterocylic and aliphatic nitriles. Addition of various concentrations of methanol (0-20%), an organic solvent, resulted in increased conversion rates for long chain hydrophobic aliphatic nitriles. Other enzymes in this class include those from *Comamonas testosteroni* (Levy-Schil et

al., 1995) and *Pseudomonas* sp. S1 (Dhillon et al., 1999). The nitrilase from *Pseudomonas* sp. S1 was reported to hydrolyze cyanide (KCN) to formic acid besides aliphatic and some aromatic nitriles (Dhillon and Shivaraman, 1999a; Dhillon et al., 1999b).

1.4.3 *Bromoxynil-specific nitrilase*

The nitrilase of *Klebsiella pneumoniae* subsp. *ozaenae* is highly specific for the herbicide bromoxynil. The native and recombinant enzymes could completely convert 3,5-dibromo-4-hydroxybenzonitrile (bromoxynil) to 3,5-dibromo-4-hydroxybenzoic acid and use the liberated ammonia as the sole nitrogen source (McBride et al., 1986; Stalker et al., 1988). This enzyme shows no activity at all with benzonitrile, but good activity with bromoxynil, chloroxynil and ioxynil (Stalker et al., 1988). The nitrilase of *Norcadia* sp. NCIMB 11215 was reported to have little but detectable activity with bromoxynil and ioxynil (Harper 1985).

1.4.4 *Arylacetonitrilases*

These enzymes constitute the fourth class and display no activity with benzonitrile. The nitrilase from *Alcaligenes faecalis* ATCC 8750, which was induced with *n*-butyronitrile showed highest activity with *p*-aminobenzylcyanide (Yamamoto et al., 1992). This enzyme is enantioselective and produces *R*-(-)-mandelic acid from mandelonitrile or benzaldehyde and HCN (Yamamoto et al., 1991). The nitrilase from *Alcaligenes faecalis* JM3 is reported to hydrolyze arylcetonitriles such as 2-thiophenacetonitrile, *p*-tolyacetonitrile, *p*-fluorobenzylcyanide, 3-pyridylacetonitrile with high activity when *p*-chlorobenzylcyanide is used as substrate (Nagasawa et al., 1990). The nitrilase from *Pseudomonas fluorescens* DSM 7155 was induced when phenylacetonitrile was used as the sole carbon source (Layh et al., 1998). It exhibited high activity with phenylacetonitrile and no activity with benzonitrile. Whereas the majority of nitrilases produce no amide by-product, the arylacetonitrilase from *P. fluorescens* DSM 7155 produced 3-5% of the equivalent amide from nitrile hydrolysis but was not able to utilize it as a substrate. Recently, Kiziak et al (2005) purified a recombinant enantioselective nitrilase from *Pseudomonas fluorescens* EBC191. This enzyme exhibited high relative activity with para-substituted phenylyacetonitriles and also hydrolysed heterocyclic and bicyclic arylacetonitriles. The enzyme converted (R,S)-mandelonitrile to R-mandelic acid with an enantiomeric excess (ee) of 31%, whereas 2-phenylpropionitrile was converted to the corresponding S-acids and amides with ee of 65%. The nitrilase also produced 43% and 28% amide by-products when racemic *O*-acetylmandelonitrile and (R)- *O*-acetylmandelonitrile were used as substrates respectively. This nitrilase shares 47% amino acid sequence identity with the

arylacetonitrilase from *A.faecalis* JM3 and both have very little activity with benzonitrile and aliphatic nitriles.

1.4.5 *Cyanide dihydratases and cyanide hydratases*

Cyanide dihydratases and cyanide hydratases are specific for cyanide (HCN) and their products upon hydrolysis are formate and formamide respectively. The cyanide dihydratase from *Alcaligenes xylosoxidans* subsp. *denitrificans* (Ingvorsen et al., 1991) was tested with acetonitrile, benzonitrile, propionitrile among others, but was seen to hydrolyze cyanide only. Other cyanide dihydratases have also been identified in *Bacillus pumilus* C1 (Meyers et al., 1993), *Pseudomonas fluoroscens* NCIMB 11764 (Kunz et a., 1994) and *Pseudomonas stutzeri* AK61 (Watanabe et al., 1998). The filamentous fungus *Fusarium lateritium* produces cyanide hydratase when induced with cyanide. This enzyme showed 3000-fold activity towards KCN as compared to benzonitrile (Nolan et al., 2003). The cyanide hydratases isolated from *Fusarium oxysporum* N-10 (Yanase et al., 2000) and *Fusarium solani* (Barclay et al., 1998) are reported to hydrolyse tetracyanonickelate (II) {$K_2[Ni(CN)_4]$; TCN}, a metal-cyano complex. The cyanide hydratase from *F.oxysporum* N-10 had 2330 times more relative activity with KCN than with TCN. No activity was detected with benzonitrile.

1.5 Enzyme subunit size and quaternary structure

The subunit and complex molecular mass for many nitrilases has been determined as indicated in Table 1. SDS-PAGE or sequencing techniques have shown that nitrilases consist of a single subunit of about 35-40 kDa that aggregates to form an active oligomer with a mass greater than 300 kDa. Exceptions include the nitrilase from *Fusarium solani* with a 76 kDa subunit size (Harper, 1977a), 32 kDa nitrilase from *Alcaligenes faecalis* ATCC 8750 (Yamamoto et al., 1992) and a 30 kDa nitrilase from *Arthrobacter* sp. Strain J1 (Bandyopadhyay et al., 1986). The results from techniques of gel filtration chromatography, light scattering and electron microscopy indicate that majority of the nitrilases have between 10 to 18 subunits. Exceptions to this rule include the nitrilases from *Athrobacter* sp. Strain J1 and *Rhodococcus rhodochrous* PA-34 (Bhalla et al., 1992), which exist as monomers, and the nitrilases of *Klebsiella pneumoniae* subsp. *ozaenae* (Stalker et al., 1988) and *Pseudomonas fluorescens* DSM 7155 (Layh et al., 1998) that exist as dimers or trimers. The native form of the fungal cyanide hydratases from *G.sorghi*, *F.solani* and *F.lateritium* is reported to be between 300 – 10000 kDa (Cluness et al., 1993; Nolan et al., 2003; Fry and Munch, 1975; Wang and VanEtten, 1992b), while that from *L.maculans* is 160 kDa. It is important to note that the sizes for some nitrilases were determined in the absence of the substrate. Also, majority of the nitrilases that have been

Table 1: Characteristics of purified nitrilases, substrate specifity and relative activity.

Substrate Class	Organism	Substrate(s)	% Relative activity	Subunit (SDS/PAGE) (kDa)	Complex (kDa)	Amino Acids	Subunit	pH	Temp Opt ^{o} C	% Amide by-product	Reference
Aromatic	*Arthrobacter* sp. Strain J1	Benzonitrile *p*-Tolunitrile	100 125	30	30	----	1	8.5	40	----	Bandyopadhyay et al., 1986
	Bacillus pallidus DAC 521	Benzonitrile Crotonitrile	100 80.3	41 (+72 chaperone)	600		14	7.6	65	----	Alamatawah et al., 1999
	Fusarium solani IMI196840	Benzonitrile Acrylonitrile	100 6.6	76	620		8	7.8-9.1		----	Harper 1977a
	Fusarium oxysporum f.sp. melonis	Benzonitrile Acrylonitrile	100 35	37	550	----	14	6-11	40	4-6 for benzonitrile	Goldlust and Bohak, 1989
	Norcadia (rhodoccocus) NCIMB 11215	Benzonitrile *m*-nitrobenzonitrile	100 841.7	45	560	----	12	7-9.5	30	----	Harper 1985
	Norcadia (Rhodococcus) NCIMB 11216	Benzonitrile *m*-nitrobenzonitrile	100 174.8	45.8	560	----	12	8	30	----	Harper 1977b Hoyle et al., 1998
	Rhodococcus rhodochrous ATCC39484	Benzonitrile 2-furancarbonnitrile	100 171	40, 40.3 (mass spectrometry)	560	----	14	7.5	40	2 (with phenyl-acetonitrile	Stevenson et al., 1992
	Rhodococcus rhodochrous J1	Benzonitrile Acrylonitrile	100 128	41.5	80 & 410 480 & >1.5MDa	366	2, 10	7.6	45	0.00022 for benzamide	Kobayashi et al., 1989, 1992a; Nagasawa et al., 2000 Thuku and Sewell 2005; this work.
	Rhodococcus rhodochrous PA-34	Benzonitrile acrylonitrile	100 22.4	45	45	----	1	7.5	35	----	Bhalla et al., 1992
Aliphatic	*Acidovorax facilis* 72W	Fumaronitrile Benzonitrile	100 4.6	40	570	369	14	8-9	65	----	Gavagan et al., 1999; Chauhan et al., 2003
	Acinetobacter sp. AK226	Acrylonitrile Racemic Ibu-CN Benzonitrile	144 100 94	41,43	580	----	14	8	50	----	Yamamoto and Komatsu 1991
	Arabidopsis thaliana AtNIT1	3-phenylpropionitrile Benzonitrile	729 2.7	38	450	346	12	9	35	95 with 3-nitroacrylonitrile	Osswald et al., 2002
	Comamonas testosteroni sp.	Adiponitrile Benzonitrile	100 4	38,38.7	----	354	----	7	30		Levy-Schil et al., 1998
	Pseudomonas sp. S1	Acrylonitrile acetonitrile Benzonitrile	100 31.5 3.8	41	41	----	----	----	----	----	Dhillon et al., 1999
	Rhodococcus rhodochrous K22	Acrylonitrile Benzonitrile	348 27.1	41	650	383	14-16	5.5	50	----	Kobayashi et al., 1990,1992a
	Synechocystis sp. Strain PCC6803	Fumaronitrile Benzonitrile	12000 100	ca. 40	390	346	10	7	50	----	Heinemann et al., 2003

Bromoxynil-specific nitrilase	*Klebsiella pneumoniae* ssp. Ozaenae	Bromoxynil Benzonitrile	100	37	74	349	2	9.2	35	----	Stalker et al., 1988
Arylaceto-nitrilase	*Alcaligenes faecalis* ATCC 8750	*p*-Aminobenzylcyanide Benzonitrile	1670 1.1	32	460	----	14	7.5	45	----	Yamamoto et al., 1992
	Alcaligenes faecalis JM3	2-thiopheneacetonitrile *p*-Chlorobenzylcyanide	100 188.7	38.9, 44	260, 275	356	6	7.5	45	----	Nagasawa et al., 1990; Kobayashi et al., 1993
	Pseudomonas fluorescens DSM 7155	Phenylacetonitrile 2-(Methoxy)-mandelonitrile	100 10	38, 40 (+ 57 CPN60 chaperonin)	130	----	2 (or 3)	9	55	3-5% for 2-(Methoxy)-mandelonitrile	Layh et al., 1998
	Pseudomonas Fluorescens EBC191	2-phenylvaleronitrile Benzonitrile	5600 0.25	37.7 (calculated) from sequence)	----	350	----	6.5	50	43% 28%	Kiziak et al, 2005
Cyanide dihydratase	*Alcaligenes xylosoxidans* ssp. *denitrificans* strain DF3	Cyanide (HCN)	81 (μmol/min/mg)	39,40	>300	----	----	7.5-8.2	26	----	Ingvorsen et al., 1991
	Bacillus pumilus C1	HCN	97 (μmol/min/mg)	37 (Mass spectrometry)	672	330	18 (EM)	7.8-8.0	37	----	Jandhyala et al., 2003
	Pseudomonas stutzeri AK61	HCN	54.6 (μmol/min/mg)	38	532	334	14 (EM)	7.5	30	----	Watanabe et al., 1998a, 1998b; Sewell et al., 2003
Cyanide hydratase	*Fusarium lateritium*	KCN Benzonitrile	100 0.033	43	1217	356	----	7.5	----	100	Cluness et al., 1993; Nolan et al., 2003
	Fusarium solani	KCN $K_2Ni(CN)_4$,$K_4Fe(CN)_6$	1.7(μmol/min/mg)	45	>300	363	----	7.5	25	100	Barclay et al., 1998; Barclay et al., 2002.
	Fusarium Oxysporum N-10	KCN $K_2[Ni(CN)_4]$	233000 100	40	160	----	4	7.5	30	100	Yanase et al., 2000
	Gloeoecercospora sorghi	HCN	555 μmol/min/mg)	45, 40.9	>300, 2000-10000	368	----			100	Fry and Munch 1975; Wang and Van Etten 1992b

Note: The activity with benzontrile was often found to be 100% except in those nitrilases that do not hydrolyze aromatic substrates at all. The % relative activity is quoted directly from the referenced publications. In most cases, this was determined for various substrates through enzyme kinetic studies or use of activity assays.

studied so far were isolated from the native organism. Harper (1997a) reported that the nitrilase of *Norcadia* (*Rhodococcus*) NCIMB 11216 aggregates into a dodecamer with a molecular mass of 560 kDa in the presence benzonitrile. This phenomenon was also observed in the nitrilases of *Rhodococcus rhodochrous* ATCC 39484, *Alcaligenes faecalis* ATCC 8750, and *Rhodococcus rhodochrous* J1 (Stevenson et al., 1992; Yamamoto et al., 1992; Nagawasa et al.; 2000).

In Rhodococci, nitrilase activation following assembly has been shown to be dependent on enzyme concentration, pH and temperature. The homodimeric nitrilase of *Rhodococcus rhodochrous* J1 is reported to associate into a decamer in the presence of benzonitrile (Nagasawa et al., 2000). Upon assembly, this enzyme hydrolyzed acrylonitrile among other aliphatic unsaturated nitriles. The rate of hydrolysis of acrylonitrile was half that benzonitrile. The dimer form was inactive on acrylonitrile, which did not induce assembly. The benzonitrile-induced association was inhibited by mercury chloride. However, this was not the case when benzonitrile was absent at high enzyme concentrations. Enzyme association was also observed when purified enzyme was dialysed against ammonium sulphate, glycerol and at increased temperatures or enzyme concentrations. It has been postulated that in dehydrating conditions such as salts and organic solvents, the enzyme undergoes a conformational change that exposes more hydrophobic sites that bring about subunit association (Nagasawa et al., 2000). The nitrilase of *Rhodococcus* ATCC 39484 fully assembles into a complex of 14 subunits in the presence of the substrate, benzonitrile and benzaldehyde, a strong inhibitor (Stevenson et al., 1992). Partial assembly was observed in the presence of benzoic acid, the reaction product that induced dimerisation, while benzyl alcohol, a weak inhibitor of the enzyme induced the formation of hexamers.

The nitrilases of *Bacillus pallidus* DAC521 (Alamatawah et al., 1999) and *Pseudomonas fluorescens* DSM 7155 (Layh et al., 1998) co-purify with chaperone-like proteins. The purified enzyme from *P. fluorescens* DSM 7155 showed three subunit bands on SDS-PAGE of molecular weight 57 kDa, 40 kDa and 38 kDa. N-terminal sequencing indicated that the 57 kDa protein shared 97% homology with GroEL, while the other two proteins are related to known nitrilases but are different from each other. The nitrilase of *Fusarium oxysporum* f.sp. *melonis* (Goldlust and Bohak, 1989) is reported to show many bands on non-denaturing PAGE. Activity assays showed these bands were active. However, only a single peak corresponding to the 14-mer was observed upon gel filtration. Other 14-mers include the nitrilases from *Bacillus pallidus* DAC521 (Almatawah et al., 1999), *Rhodococcus rhodochrous* K22 (Kobayashi et al.,

1990, 1992), *Acidovorax facilis* 72W (Gagavan et al., 1999; Chauhan at al., 2003), and *Acinetobacter* sp. AK226 (Yamamoto and Komatsu, 1991). Heinemann et al., (2003) reported a nitrilase from *Synechocystis* sp. Strain PCC6803, which showed a single band of about 40 kDa on SDS-PAGE and single peak of about 390 kDa upon gel filtration. This suggests that the enzyme complex is made up of 10 subunits.

The cyanide dihydratase isolated from *Alcaligenes xylosoxidans* ssp. *denitrificans* strain DF3 showed two bands of 39 and 40 kDa with identical amino terminal sequences on reducing SDS-PAGE (Ingvorsen et al., 2003). Recently, Jandhyala et al. (2003) reported a molecular weight of 37 kDa for the cyanide dihydratase of *Bacillus pumilus* C1. This enzyme forms a terminating 18-subunit spiral at pH 8 and a left-handed helix at pH 5.4. The rates of cyanide hydrolysis were similar to the K_m and V_{max} values reported earlier by Meyers et al. (1993). The only other purified cyanide dihydratase, namely $CynD_{stu}$ or cyanidase from *Pseudomonas stutzeri* AK61 forms a 14-subunit spiral with two-fold symmetry (Sewell et al., 2003). The amino-terminal has significant sequence homology to the enzymes from *A. xylosoxidans* ssp. *denitrificans* strain DF3 and *B. pumilus* C1 (Watanabe et al., 1998). One major band was observed on SDS-PAGE corresponding to a molecular weight of 38kDa.

The molecular mass of the active cyanide hydratase from *Fusarium oxysporum* N-10 was determined by gel permeation chromatography on a HiPrep Sephacryl S-300 column to be 160kDa (Yanase et al., 2000). The complex molecular mass is much lower compared to that of other fungal cyanide hydratases (Table 1). However, a single band on SDS-PAGE had a molecular mass of 40kDa, which is similar to the subunit mass of other fungal cyanide hydratases, suggesting that the enzyme is a homotetramer.

1.6 Enzyme Structure and Structural Homology

There are four atomic structures of distant homologues in the nitrilase superfamily. These include the NitFhit protein (PDB ID, 1EMS) from *Caenorhabditis elegans* (Pace et al., 2000), N-carbamyl-D-aminoacid amidohydrolase (DCase or D-NCAase; PDB ID, 1ERZ) from *Agrobacterium* (Nakai et al., 2000; Chen et al., 2003; Wang et al., 2001), Nit3 CN hydrolase (PDB ID, 1F89) from yeast (Kumaran et al., 2003), and the hypothetical protein Ph0642 (PDB ID, 1J31) from *Pyrococcus horikoshii* (Sakai et al., 2004). Structures of 1EMS and 1ERZ reveal that these enzymes form tetramers with 222 symmetry. There are two contact surfaces between the tetramers that are almost perpendicular to each other. A larger contact surface, referred to as the A surface and a smaller contact surface, B. The A surface of 1EMS is made

up of two α-helices and a length of β-sheet, while that of 1ERZ consists of three α-helices. In both structures, surface B is made up of two anti-parallel β-sheets which are externally located next to the central core of β sheets in each monomer. The worm NitFhit structure is a fusion protein of a nitrilase homologue (Nit) to Fhit, a protein that plays a role in apoptosis.

The structures of 1F89 and 1J31 share a conserved catalytic cys-glu-lys triad with 1EMS, 1ERZ and indeed, with other members in the nitrilase superfamily. These structures resulted from efforts of structural genomics and their functions are unknown. The nitrilase atomic homologues form dimers across a conserved two-fold symmetric interface, which has been called the A surface (Sewell et al., 2003). The A surface in 1F89 is held together by salt bridges and hydrogen bonds comparable to those observed in 1EMS and 1ERZ. Details of their structural alignment using GenTHREADER (Jones, 1999) are discussed in the next chapter.

Nitrilases and amidases are reported to have a wide substrate range which may be due to the shape and volume of the active site cavity. The volume in 1F89 is larger compared to that of 1ERZ suggesting that this enzyme has a broader substrate range (Kumaran et al., 2003). This was not determined in the NitFhit structure because of a mercury atom, which was bound to the active cysteine and might have altered the geometry of the active site (Pace et al., 2000). The structure of the amidase, namely N-carbamoyl-D-amino acid amidohydrolase (D-NCAase) from *Agrobacterium radiobacter*, which was solved with bound substrate, revealed that there were limited conformational changes in structure upon binding the substrate (Chen et al., 2003). This enzyme belongs to the nitrilase superfamily and shares a similar fold with the NitFhit structure. Structural analysis of its free and substrate-bound forms revealed residues that were directly bound to the substrate within a large pocket, while modeling studies concluded that this enzyme has preference for bulky substrates with long hydrophobic chains.

Sequence alignment of the nitrilase protein sequences from *P. stutzeri* AK61, *B. pumilus* C1 and *G. sorghi* against solved structures using GenTHREADER (Jones, 1999) showed that they share a similar fold (Sewell et al., 2003). A dimer model of the cyanide dihydratase from *P.stutzeri* AK61 based on the solved structures was used to interpret a low-resolution map. This revealed a left-handed 14-subunit spiral and highlighted residues involved in stabilizing its quaternary structure. Two major insertions in its sequence relative to those of the Nit domain of 1EMS and 1ERZ are also proposed to contribute to spiral formation. The region of these interactions is referred to as the C surface and is absent in the crystal structures. The nitrilase atomic homologues exist as dimers or tetramers whereas microbial nitrilases exist as

higher oligomers. Sewell et al (2003) found that the number of subunits required to form one turn of the spiral is 10 and suggested this would be the smallest microbial nitrilase.

There is growing evidence that formation of the spiral leads to activation of the enzyme. A pH-dependent transition from a terminating spiral at pH 8 to a long helix at pH 5.4 was observed in the cyanide dihydratase of *B. pumilus* C1 (Jandhyala et al., 2003). Interestingly, an increase in enzymatic activity of about 10% is also observed at pH 6, while no increase in activity was observed in the enzymes from *G. sorghi* and *P. stutzeri* AK61 (Jandhyala et al., 2005). The enzymes from *G.sorghi* and *P.stutzeri* AK61 do not undergo a structural transition either. In the case of Rhodococcal nitrilases, particularly the nitrilase from *Rhodococcus rhodochrous* J1, which is the focus of this study, it has been shown that the dimers are inactive, but that the higher oligomers have activity (Harper, 1985; Nagasawa et al., 2000; Stevenson et al., 1992). Evidence suggests interplay between activation and oligomerization and lays the foundation for a detailed investigation into this phenomenon.

1.7 Mechanism of catalysis

Most nitrilases exhibit maximum activity at an optimal pH range of 7.5-8.5. The pH values for some nitrilases are indicated in Table 1. The nitrilases of *Klebsiella pneumoniae* subsp. *ozaenae* and *Acinetobacter* sp. AK226 were reported to have sharp maximum pH values at 9.2 (Stalker et al., 1988; Yamamoto and Komatsu, 1991). Jandhyala et al. (2003) reported the pH-dependent transition of the cyanide dihydratase from *Bacillus pumilus* C1 between a short spiral form and a long regular helix. The activity for this enzyme was lost at a pH above 8.4, but the presence of Cr^{3+} and Tb^{3+} at a concentration of 200 μM restored activity identical to that at its optimum pH (Meyers et al., 1993).

Site-directed mutagenesis on the catalytic residues in the aliphatic amidase from *Pseudomonas aeruginosa* led to inactivity and instability of the enzyme (Novo et al., 2002). This amidase shares a similar structural and catalytic framework to the solved structures of the nitrilase homologues. Crystal structures of the unbound modified biocatalyst from *Agrobacterium radiobacter*, namely N-carbamoyl-D-amino acid amidohydrolase (DNCase or DCase) and also in complex with the substrate N-carbamoyl-D- *p*-hydroxyphenylglycine (HPG) were solved at a resolution range of 2.0-2.40 Å by molecular replacement (Chen et al., 2003). Circular dichroism studies showed that mutation of the active Cys172 did not affect the conformation of the active site, secondary structure or the stability of the enzyme. The substrate-bound mutants revealed that the carbamoyl group of the substrate is close to the thiol group of cysteine.

Further evidence on the role of the active cysteine residue in catalysis is provided by effect of inhibitors on enzyme activity. Thiol-complexing agents inactivated the nitrilase of *Pseudomonas fluoroescens* DSM 7155, while thiol reducing agents enhanced its activity (Layh et a., 1998). Thiol reducing agents were also observed to enhance activity in the enzymes of *Alcaligenes faecalis* JM3 (Yamamoto et al., 1992) and *Rhodococcus rhodochrous* J1 (Kobayashi et al., 1989).

Crystal structures of 1EMS (Pace et al., 2000), 1ERZ (Nakai et al., 2000) and 1F89 (Kumaran et al., 2003), suggest that catalytic residues are located at the bottom of a pocket in which these residues are observed to be pointing towards it center (Brenner, 2002). In these structures, the active site is located at the bottom of a pocket in which the three active residues point towards its center. It has also been suggested that the residues located in the loops above this pocket play a role in determining specificity (Chen et al., 2003; Novo et al., 2002). Other studies on the nitrilases of *Alcaligenes faecalis* JM3 (Kobayashi et al., 1993), *Pseudomonas stutzeri* AK61 (Watanabe et al., 1998c) and *Rhodococcus rhodochrous* J1 (Kobayashi et al., 1992b) using far-UV circular dichroism, kinetic and molecular mass analyses concluded there was no significant structural conformational change upon substitution of the active cysteine with alanine or serine. However, a study by Novo and others (2002) concluded that modification of the catalytic glutamate or lysine residues induced conformational changes as well as enzyme instability.

As early as 1964, a reaction mechanism for hydrolysis of ricinine by a nitrilase of *Pseudomonas* species had been proposed (Hook and Robinson, 1964). Mechanistic studies on the nitrilase of *Rhodococcus* ATCC 39484 indicated that the reaction proceeds through covalent intermediates (Stevenson et al., 1992). Based on their results and known chemical reaction between thiols and nitriles, Stevenson et al (1992) suggested that the thiol group of the cysteine initiates a nucleophilic attack on the nitrile carbon followed by protonation of the nitrogen to form a thioimidate. Subsequent two-step hydrolysis of the thioimidate produces the corresponding acid and ammonia. Using ion-spray mass spectroscopy, a covalent intermediate whose mass was consistent with either the thioimidate or acylenzyme was detected (Stevenson et al., 1990). A modification of the earliest nitrilase reaction mechanism has been proposed (Kobayashi et al., 1998a,b), in which the tetrahedral intermediate is shown to break down reversibly to produce an amide instead of an acid product via the usual route. This correlates with the observation that some nitrilases (highlighted in Table 1) produced small amounts of amide products with selected substrates.

On the basis of the evidence presented above and insights from the atomic structures of the nitrilase homologues, a refined nitrilase reaction mechanism is proposed (figure 1.2). Glutamate acts as a general base through its carboxyl group, which abstracts a proton from the thiol group of cysteine. This activates cysteine for nucleophilic attack on the nitrile carbon followed by protonation of the nitrogen. Lysine acting as a general base activates a water molecule for further attack on the nitrile carbon to produce a tetrahedral intermediate. The oxyanion of the tetrahedral intermediate may be stabilized by the positive charge of lysine. This intermediate was reported to break down uncharacteristically to produce the corresponding amide (2%), instead of the usual acid product upon hydrolysis of phenylacetonitrile by the nitrilase from *Rhodococcus* ATCC 39484 (Stevenson et al., 1992). This suggests that breakdown of this intermediate is determined by relative strength of the C-S and C-N bonds (indicated in red as A and B in Figure 1.2). In most nitrilases, C-N cleavage is preferred leading to formation of a thioester or the acylenzyme with the release of ammonia. Activation of a second water molecule by glutamate initiates attack on the thioester leading to the formation of a second tetrahedral intermediate. This tetrahedral intermediate eventually breaks down to produce the carboxylic acid with release of the enzyme.

Based on the hydrolysis of different substrates to produce amide or acid products by AtNIT1 nitrilase from *Arabidopsis thaliana*, Osswald et al. (2002) postulated a structural relationship between the type of substrate and amide formation. In another study, it was concluded that the ratio of acid to amide produced is variable and depends on the microbial species, substrate structure, incubation time and the age of the culture batch (Brady et al., 2004). An amidase model built according to the NitFhit structure showed that the catalytic residues were not only conserved but also aligned well with those of 1EMS (Novo et al., 2002). Furthermore, equivalent distances between the catalytic residues from both structures were also observed. As already discussed, modification of these residues in the amidase from *P.aeruginosa* affected its activity. Looking at figure 1.2, it follows that the cleavage of the C-S bond will lead to formation of an amide. Thus, a catalytic mechanism for the amidase reaction is proposed (Figure 1.3). While nitrilases and amidases act on different substrates, the end products upon hydrolysis are similar. Furthermore, both reactions have common tetrahedral intermediate and an acylenzyme as shown in figures 1.2 and 1.3. There is need to investigate the structural implications of these mutations for a better understanding of how these enzymes differ with respect to their catalytic mechanism and substrate specificity.

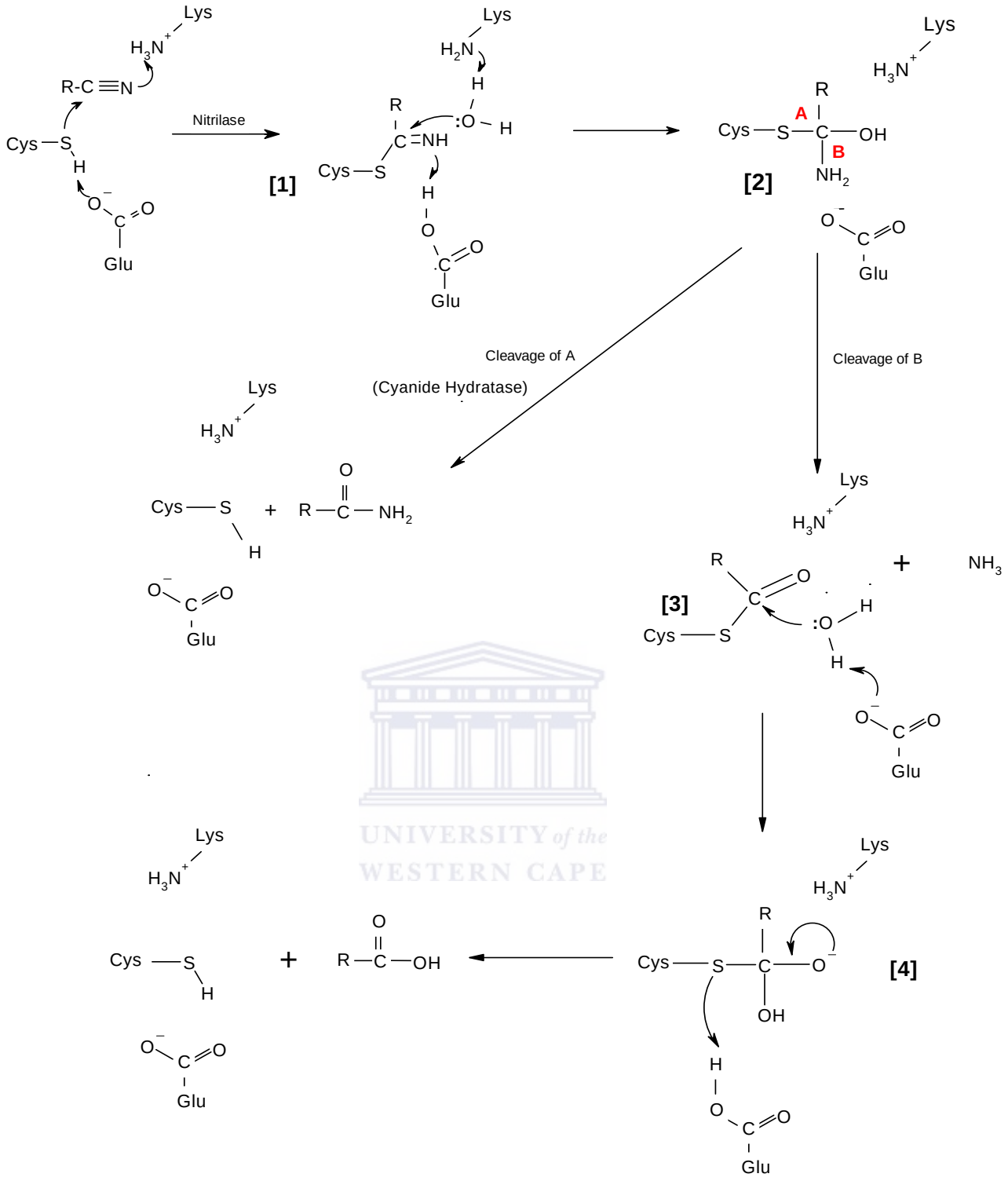


Figure 1.2: Proposed nitrilase reaction mechanism. Glutamate, the general base, activates nucleophilic attack on the nitrile by the sulfhydryl moiety of cysteine. This leads to the formation of a thioimidate [1], the first covalent intermediate. Lysine acting as a general base activates a water molecule by increasing the nucleophility of the lone pair. This water attacks the nitrile carbon leading to the formation of a tetrahedral intermediate [2], which breaks down spontaneously via cleavage of C-S or C-N bonds (indicated in red as A and B). The usual route in the nitrilases involves cleavage of the C-N and not C-S bond to produce the acylenzyme [3] and ammonia. Glutamate activates the second hydrolysis, which leads to another tetrahedral intermediate [4]. This spontaneously decomposes to produce a carboxylic acid and the enzyme.

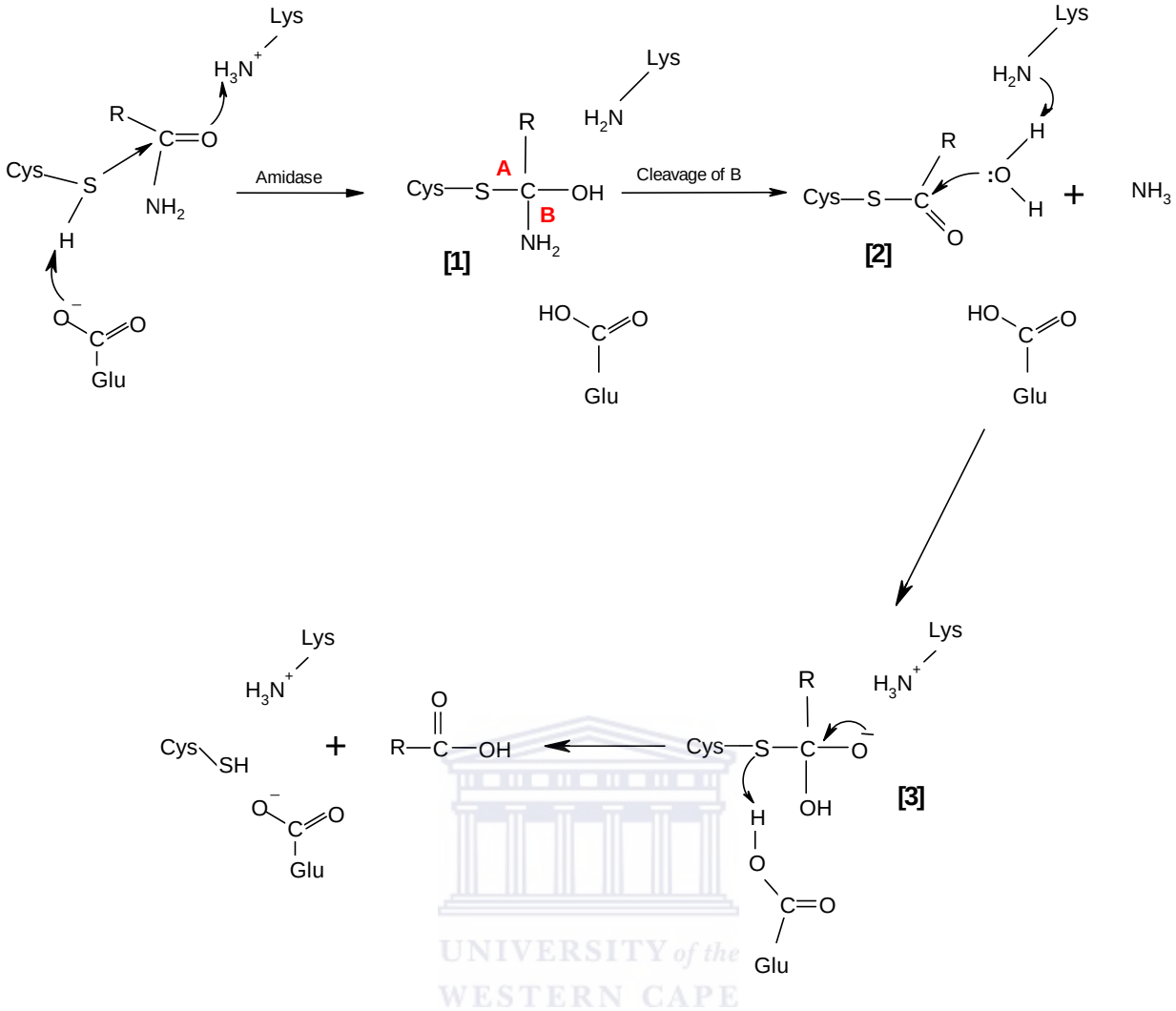


Figure 1.3: Proposed amidase reaction mechanism. Glutamate, the general base activates cysteine for nucleophilic attack on the amide carbon. This results in a tetrahedral intermediate [1], which is either cleaved at bonds A or B (indicated in red above). Cleavage of bond A regenerates the substrate and the enzyme. However, bond B is often cleaved to give an acylenzyme [2] and ammonia. Lysine activates a water molecule, which attacks the acylenzyme resulting in a second tetrahedral intermediate [3]. This eventually breaks down to give a carboxylic acid and regenerates the enzyme.

1.8 Industrial Applications

Conventional chemical methods are still used in industry for the production of carboxylic acids and amides from corresponding nitriles. These reactions occur in harsh conditions often at high temperatures that damage catalysts and produce undesirable by-products such as salts and cyanide (Banerjee et al., 2002). This has forced industries to use nitrile biocatalysts because of their stereoselective properties and their ability to catalyse a whole range of substrates. The scope of products synthesized includes a whole range of aliphatic, alicylic, aromatic and heterocylic carboxylic acids and their amides (Mylerova and Martinkova, 2003). Nitrile-

converting biocatalysts are active within a narrow pH and temperature range in neutral or slightly alkaline conditions, and produce the desired product with high selectivity at high conversion (Hann et al., 2004). Microorganisms that produce the desired enzymes constitutively are the biocatalysts of choice (Martinkova and Kren 2002). Furthermore, the differences in substrate specificity among nitrilases provides for flexibility in the choice of biocatalyst.

The nitrilase of *Rhodococcus rhodochous* J1 catalyzes the direct hydrolysis of 3-cyanopyridine to nicotinic acid, which has wide applications in pharmaceutical industry (Mathew et al., 1998). The nitrilase of *Comamonas testosteroni sp*. hydrolyses the dinitrile adiponitrile into cyanovaleric acid and adipic acid, which are intermediates in synthesis of Nylon (Levy-Schill et al., 1995). The thermophilic nitrilase of *Bacillus pallidus* DAC 521 is reported to directly hydrolyse 3-cyanopyridine into nicotinic acid at temperatures close to 60 ^{o}C (Almatawah et al., 1999). The nitrilase from *Acinetobacter* sp. AK226 has been shown to hydrolyse (R,S)-(±)-ibuprofen nitrile to (S)-(+)-Ibuprofen, a non-steroidal non-inflammatory drug (Yamamoto et al., 1990). R-(-)-Mandelic acid was produced from racemic mandelonitrile by *Alcaligenes faecalis* ATCC 8750 (Yamamoto et al., 1992). These enzymes are among other nitrilases from *R. rhodochrous* NCIMB 11216 (Gradley and Knowles, 1994) and the arylacetonitrilase of *P. fluorescens* DSM 7155(Layh et al., 1998) that perform enantioselective transformations to produce optically pure compounds for the pharmaceutical industry. Regioselectivity for hydrolysis of the (E)-isomer of several (E,Z)-3-substituted-acrylonitriles has been demonstrated using the recombinant nitrilase AtNIT1 from *Arabidopsis thaliana* (Effenberger and Osswald, 2001). The aliphatic nitrilase from *Acidovorax facillis* 72W (Hann et al., 2004), catalyses the regioselective hydrolysis of (E,Z)-2-methyl-2-butenenitrile, producing only (E)-2-methyl-2-butenoic acid.

Nitrile-metabolizing enzymes are also used in the degradation of herbicides such as bromoxynil, commonly used in agricultural industry. A gene encoding the nitrilase of *Klebsiella pneumoniae* ssp. *Ozaenae* has been cloned into tobacco and tomato to confer resistance to the bromoxynil (Stalker et al., 1988). Nitriles used in the chemical and agricultural industry are pollutants with different degrees of toxicity. However, cyanide dihydratases and cyanide hydratases readily convert cyanide into non-toxic products and do not require any co-factors. Nitrile-converting whole-cell biocatalysts are utilized in degradation of acrylonitrile in polymers, and in waste effluents from acrylonitrile synthesizing industries (Wyatt and Knowles, 1995).

The cyanide hydratase of *Fusarium solani* was shown to utilize the metal cyanides $K_4Fe(CN)_6$ at pH 4 and $K_2Ni(CN)4$ at both pH 4 and 7 as a source of nitrogen (Barclay et al., 1998). This enzyme may be used for bioremediation of metal-cyanide complexes in industrial effluents. The fungi *Gloeoecercospora sorghi* detoxifies cyanide in the host plant sorghum through direct hydrolysis by cyanide hydratase (Wang et al., 1992a, b). The enzyme of *Pseuodomonas paucimobilis* was used to remove weak-acid-dissociable cyanide from the effluent of a mine in South Dakota, USA (Dubey and Holmes, 1995). The enzyme from the fungus *Fusarium oxysporum* N-10 is capable of degrading free cyanide (KCN), aliphatic mono- or dintriles and some aromatic nitriles. It posses broad substrate activity, and could be potentially useful in the bioremediation of contaminated soil and waste water (Yanase et al., 2000).

1.9 Motivation

The importance of nitrilases as industrial enzymes makes it essential to understand factors, which influence their stability and efficacy. The evidence presented above has implicated the oligomeric state of these enzymes in their activity. In particular, the nitrilase of *Rhodococcus rhodochrous* J1 has been reported to be inactive as a dimer but active as a decamer (Nagasawa et al, 2000). Although this behaviour is typical of the Rhodococcal nitrilases, it is certainly the best characterized. The work of Sewell et al (2003) has shown that the structural insights obtained from three-dimensional electron microscopy of the nitrilase from *Pseudomanas stutzeri* AK61 can be interpreted at an atomic level and can give insight into the process of oligomerization. For this reason it is proposed to determine the oligomeric structure of the recombinant nitrilase from *R.rhodochrous* J1. The aim of this study is firstly to purify the recombinant protein and secondly, to determine and interpret the quaternary structure using a combination of single particle electron microscopy methods and homology modeling.

CHAPTER 2
BIOINFORMATICS

2.1 Introduction

The structure of a protein provides insights into the location of probable active sites, identification of distantly related proteins and regions of the protein involved in maintaining its structure. However, there is a significant difference between the number of known protein sequences (~1,000,000) and experimentally determined structures (~27,000). Comparative or homology modeling provides a means to narrow this gap. This is a computational method, which involves the construction of structurally conserved regions of a target protein, the addition of loops and side chain optimization from a known homologous structure (Bower et al., 1997). The method aims to build a useful three-dimensional (3D) model of an unknown protein based on one or more atomic structures (Sanchez and Sali, 2000). Approximately 40% of all protein sequences can have at least one domain modeled on a related known protein structure. Some proteins may have low sequence identities yet they share a similar fold and a closely related function. This is because the 3D structure of proteins from the same family is more conserved than their amino acid sequences. Members in the nitrilase family have low sequence similarity but high structural fold conservation. Even though no crystal structure of a microbial nitrilase exists to date, there are four structures of distant homologues that provide an opportunity to pursue structural insights via homology modeling. While an atomic resolution structure provides for its functional characterization, a 3D model of the nitrilase from *Rhodococcus rhodochrous* J1 will reveal the inherent structural fold, residues that determine substrate specificity and lead to the design of mutants. The underlying objective is to use this model to interpret a 3D map of its quaternary structure as determined by single particle electron microscopy techniques.

2.2 Review of protein structure modeling

Sequences are homologous if they are related by divergence from a common evolutionary origin. These evolutionary relationships make it possible for researchers to apply the principles of comparative modeling, which require some detectable sequence similarity between the target and template and the construction of a correct alignment between them. There are a variety of servers and programs that automate the modeling process. Few examples include SWISS-MODEL (www.expasy.ch/swissmod/SWISS-MODEL.html), WHATIF (www.cmbi.kun.nl/whatif), 3D-PSSM (www.sbg.bio.ic.ac.uk/~ 3dpssm), and PDB-Blast

(www.bioinformatics.burnham-inst.org/pdb_blast). However, a useful model is often obtained by manually controlling this process due to decisions such as choice of appropriate templates, constructing a correct alignment, specifying external restraints and loop modeling that a researcher has to make during the process. There are four sequential steps in the modeling process as shown in figure 2.1.

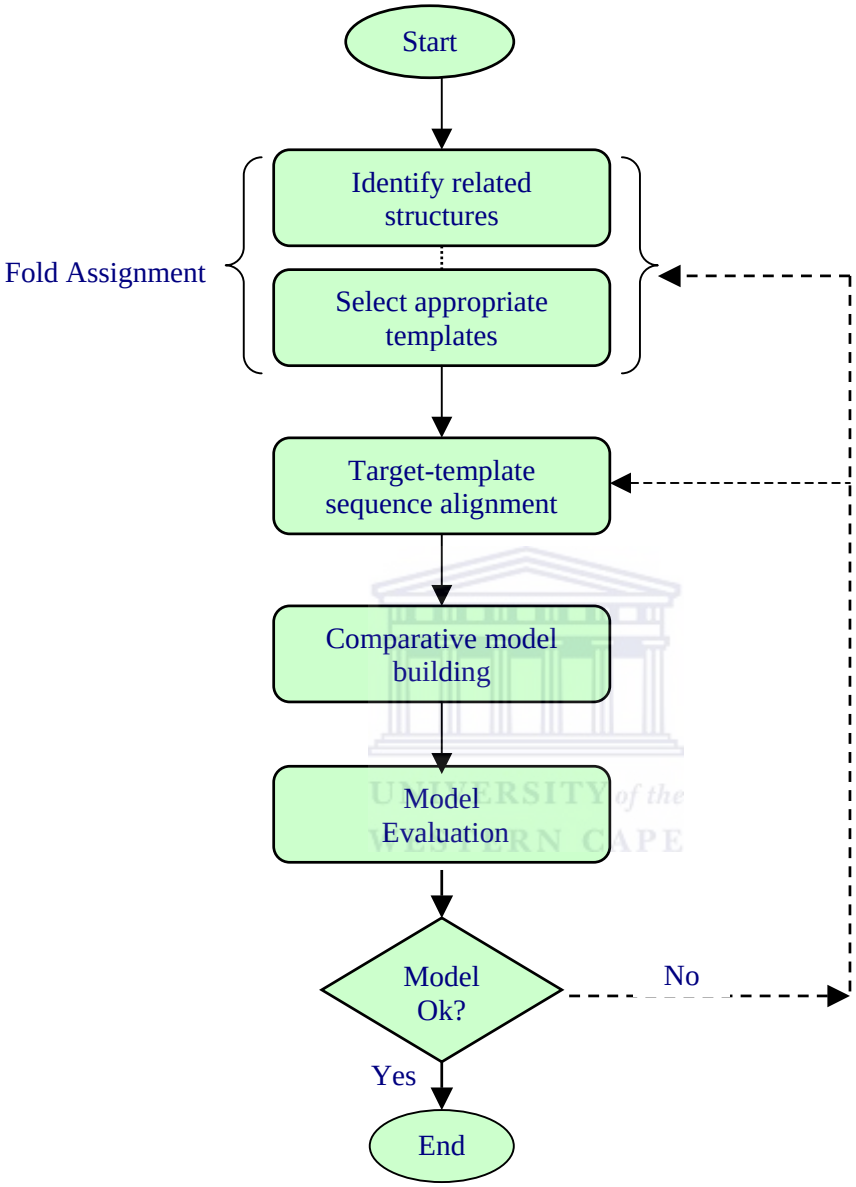


Figure 2.1. Steps in comparative protein structure modeling

2.2.1 Fold assignment

This involves the search for related known protein structures in the Protein Data Bank (http://www.rcsb.org/pdb/) by querying with a target sequence and the selection of appropriate templates (Sanchez et al., 2000). This search can be carried out in three ways. Firstly, using pair-wise sequence comparison methods, which compare the target sequence with each of the sequences in the database independently. Examples include FASTA (Pearson, 2000) and

BLAST (Altschul et al., 1997). Secondly, multiple sequence comparison methods such as PSI-BLAST (Altschul et al., 1997), find all sequences in the databases that are related to and easily align with the target. These methods produce a multiple alignment of hits versus target sequence, improve the sensitivity of the search and also provide information about the location and pattern of evolutionarily conserved regions of the protein (Fiser, 2004). Thirdly, threading methods evaluate the compatibility between the target sequence and solved structures in the PDB. These methods are useful when a sequence profile is not easy to construct because only a few number of known sequences clearly related to the target are available. Popular programs include GenTHREADER (Jones, 1999) and FUGUE (Shi et al., 2001). Subsequently, appropriate templates that address a particular modeling problem are selected based on factors such as sequence similarity, quality of solved structures, protein family and similarity between the environment of template and that of the target, including solvent, ligand, pH and quaternary interactions (Sanchez and Sali, 2000).

2.2.2 Target-template sequence alignment

This is the most important step in comparative modeling. The aim is to establish the best possible structural equivalences between the target and template by constructing a correct alignment between them. Automatic alignment programs such as CLUSTALX (Thompson et al., 1997; Jeanmougin et al., 1998) and MULTIALIGN (http://cbrg.inf.ethz.ch/Server/MultAlign.html) produce local alignments according to the best score that correlates both target and template sequences. These programs are tuned to find remote relationships with homologues but not the optimal alignment. The alignment is usually correct for sequences sharing more than 40% identity. Below 40%, the alignment has gaps and will require user intervention to minimize the number of misaligned residues. The accuracy of the alignment is the most important factor affecting the quality of the model. The alignment can be improved by avoiding gaps in secondary structure elements, in buried regions or between residues that are far apart in space. The programs GenTHREADER (Jones, 1999) and FUGUE (Shi et al., 2001) are designed to take such structural information into consideration. Sanchez and Sali (2000), report that a misalignment by only one residue position results in an error of approximately 4Å in the model. This is because the current comparative modeling programs cannot recover from errors in the alignment.

2.2.3 Model building

This relies on the sequence alignment between target and template to produce a 3D model of the target protein. This model will include all main-chain, side-chains and their attached non-

hydrogen atoms including the insertions and deletions relative to the template structures. There are three automated methods for model building. Firstly, modeling by rigid-body assembly constructs the model from a few core or structurally conserved regions, loops and side-chains obtained from an analysis of related atomic structures. Secondly, modeling by segment matching determines approximate positions of conserved atoms between the templates to calculate positions of the atoms in the target sequence. Third, modeling by satisfaction of spatial restraints uses distance geometry or optimization techniques to satisfy spatial restraints obtained from the alignment between the target and template. This approach is implemented in the program MODELLER (Sali and Blundell, 1993). The spatial restraints are obtained from homology-derived restraints on dihedral angles and bond distances in the target sequence extracted from the alignment, from stereochemical restraints (bond lengths and bond angles) obtained from a molecular mechanics force field as well as from the statistical preferences for dihedral angles and non-bonded atomic distances derived from a representative set of known protein structures (Fiser, 2004; Fiser and Sali, 2003; Eswar et al., 2003). This program is automated and will allow for easy recalculation of a model when a change is made in the alignment. It also allows the use of several templates and provides tools for incorporating prior knowledge about the target sequence such as predicted secondary structure profiles. It also allows for *ab initio* modeling of loops or insertions, some of which may play a role in the functions of the protein. The accuracies of the above three methods are similar if used optimally (Fiser and Sali, 2003; Marti-Renom et al., 2002).

2.2.4 Model evaluation

This process helps to assess what information can be extracted from the model. This evaluation considers not only the sequence similarity, but also structural and energetic criteria. It is critical for testing and identifying the most accurate model. Errors in a model will increase with decreasing sequence similarity between the target and the template. The most common errors include errors in side-chain packing, shifts of correctly aligned residues, shifts of a region without an equivalent template, errors due to misalignments and the use of incorrect templates (Sanchez and Sali, 2000). These errors are infrequent for sequences with more than 40% identity because the alignment is easy to construct, has few gaps and structural differences would occur in the loops and the side chains. Below 40%, frequent misalignments and insertions are observed and these give rise to extensive gaps within the alignment. The problem of finding related sequences and construction of a correct alignment arises where sequence identity is below 30% (Fiser, 2004). One way of finding unreliable regions in a model is to evaluate its stereochemistry using programs such as RAMPAGE (Lovell et al., 2003),

PROCHECK (Laskowski et al., 1993) or WHATCHECK (Hooft et al., 1996). These programs perform an internal evaluation, which ascertains whether or not a model satisfies the restraints used to calculate it. The model may alsobe externally evaluated using programs such as PROSAII (Sippl, 1993) and VERIFY3D (Eisenberg et al., 1997). These programs will predict unreliable regions of the model by generating graphs of energy distributions, which pinpoint problematic sections in the model. Errors in the model will appear as peaks of positive energy in the profile (Sanchez and Sali, 2000). PROSAII also provides a Z-score, which reflects the quality of the structure. Bad Z-scores indicate there are erroneous sections in the model or structure (Sippl, 1993). Ultimately, a good protein model depends on some prior experimental knowledge about the structure and its function (Sanchez and Sali, 2000).

If the model obtained is unsatisfactory, the cycle of the four steps discussed above is iterated until a reliable model is obtained. If no suitable model is obtained, then the *ab initio or de novo* protein structure prediction methods may be used. These methods predict the native structure of a protein from sequence alone by exploring the conceivable protein conformations in an attempt to find the global free energy minimum accessible during the lifespan of that protein (Fiser, 2004). These methods are complex, not easily applicable and are limited to a number of protein sequences with about 100 residues.

2.3 Current scope and homology

Several nitrilases have previously been studied for the preparation of carboxylic acid derivatives with wide applications in the pharmaceutical industry, but little progress has been made in harnessing nitrilases as viable synthetic tools. In the past five years, there has been increasing interest about the potential uses of nitrilases as biocatalysts and also in the treatment of toxic waste in industry. Most of the microbial nitrilases have been defined at the sequence level and were shown to possess a characteristic α-β-β-α fold, similar topologies, sequence motifs and a conserved catalytic triad (Pace and Brenner, 2001; Brenner, 2002). Even though there are four atomic structures of distant homologues in the nitrilase superfamily, no crystal structure of a microbial nitrile exists. About 20 nitrilase enzymes have been characterized in detail at a biochemical level. While most of the enzymes were found to have the ability to catalyze stereoselective hydrolysis of various nitrile substrates, their mechanism of catalysis is not well understood. Furthermore, none of these substrates are known to be their natural substrates. The crystal structure of an identical structure to DCase, namely the N-carbamoyl-D-amino acid amidohydrolase (D-NCAase) from *Agrobacterium radiobacter* with bound substrate provided clues that a cys-glu-lys triad is involved in hydrolysis of the carbamoyl

moiety, while residues in the substrate-binding pocket are crucial in determining substrate specificity (Chen et al., 2003). Therefore, there is need for more structural studies on members of the nitrilase superfamily, which would shed light on their specificity and their ability to hydrolyze structurally distinct molecules.

Until recently, there have not been any comparative modeling studies in nitrilases. Homology modeling based on two atomic structures of 1EMS and 1ERZ enabled the construction of 3D dimeric models of the cyanide dihydratases of *Bacillus pumilus* C1 and *Pseudomonas stutzeri* AK61, which were used to interpret low-resolution maps of their quaternary structures (Jandhayala et al., 2003; Sewell et al., 2003). The models revealed a similar fold to the solved structures, two significant insertions in their sequences relative to the crystal structures, the number of subunits that form the complex, and that the dyad axis at the A surface of the dimer coincided with that of the complex. Residues present in these insertions together with those mating across the A surface were suggested to be involved in subunit interactions that lead to the formation of a left-handed 14-subunit spiral of the cyanide dihydratase of *P.stutzeri* AK61 (Sewell et al., 2003). Hence comparative modeling provides an alternative way to try and uncover the mystery, which is inherent within the nitrilase superfamily. This is only possible if there is some detectable homology with at least one solved structure.

2.4 Methods

2.4.1 Search for templates

Protein sequences of 20 biochemically characterized branch 1 nitrilases were threaded through the PDB using GenTHREADER (Jones, 1999), which produced four candidates, namely NitFhit from *C.elegans* (PDB ID, 1EMS), DCase from *Agrobacterium sp*. KNK712 (PDB ID, 1ERZ), Yeast Nit3 protein (PDB ID, 1F89) and hypothetical protein Ph0642 from *Pyrococcus horikoshii* (PDB ID, 1J31), with an average confidence score of 0.97 despite a low sequence identity of 20%.

2.4.2 Multiple sequence alignment

Structural alignment of the four structures was done using ALIGN (Cohen, 1997). This program superimposes the α-carbon coordinates of two structures while taking account of the insertions and deletions present in their sequences. Only the Nit domain of the NitFhit structure was used in the alignment, which showed high structural conservation of the central core of β sheets, catalytic triad, positions of the external helices and those located at the A surface.

Structural variability is observed in the externally located loops and in the C-terminal tail. Using these principles and guided by predictions from GenTHREADER (Jones, 1999) and PSIPRED (McGuffin et al., 2000; 2003), an optimal alignment of 24 sequences from the nitrilase superfamily is shown in figure 2.2.

2.4.3 Model building and side-chain optimization

The alignment file from the previous step was converted into PIR format for subsequent model building using MODELLER (Sali and Blundell, 1993; Sali et al., 2000). Predictions of secondary structure and symmetry impositions were added as spatial restraints in the modeling top file under the "special_restraints" routine. Chains A and B of the target and templates were kept separate using chain breakers ("/") during the modeling cycle. Twenty cycles of model building were executed with continuous scanning of the log file for errors. This process was iterated until a satisfactory model was obtained. The best model was chosen as that having the lowest score of the Modeller objective function. This value does not necessarily indicate the quality of the model; rather it is a measure of how well the model satisfies the restraints used to calculate it (Fiser and Sali, 2003). The models were visualized with PyMOL (DeLano, 2002). The orientations of the side-chains were optimized using SCWRL 3.0 (Bower et al., 1997; Catunescu et al., 2000).

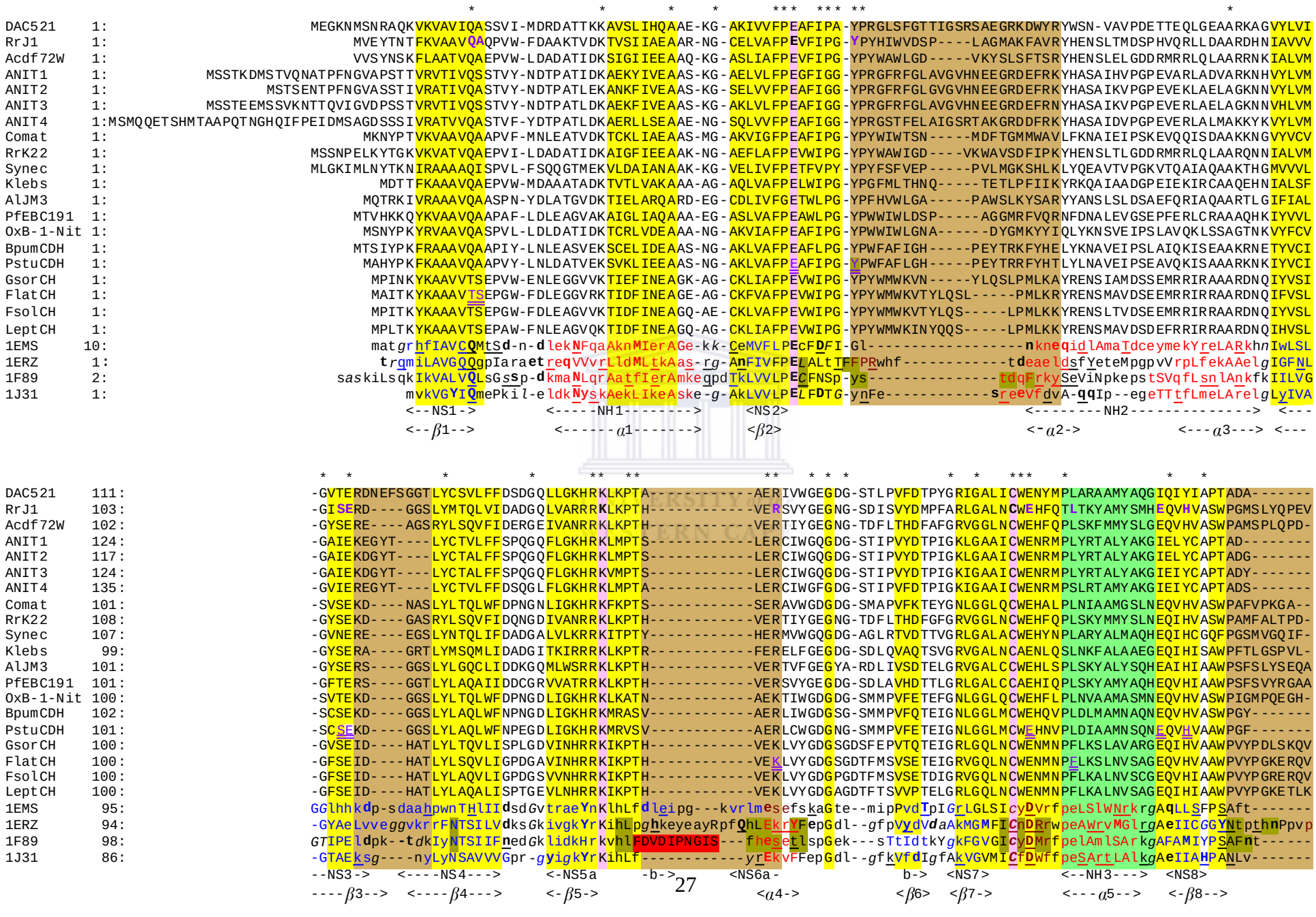
2.4.4 Model evaluation

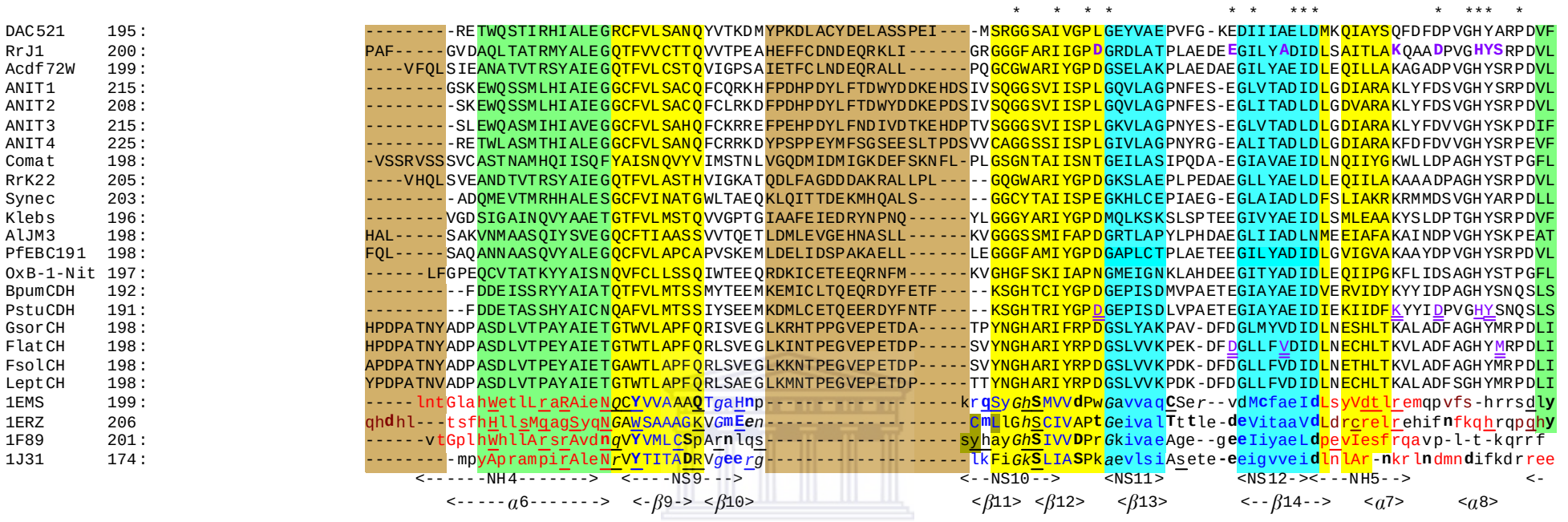
Accuracy of the model was assessed using RAMPAGE (Lovell et al., 2003), PROCHECK (Laskowski et al., 1993) and also using PROSAII (Sippl, 1993). The coordinates of the final model were superimposed with those of the crystal structures using ALIGN (Cohen, 1997), to assess the root mean deviation of main chain atoms from the templates.

2.5 Reliability of comparative modeling methods

Comparative modeling methods are designed to arrange the backbone of the target sequence to that of the template by using the sequence alignment to determine the position of each corresponding residue. This is not trivial because there are many residue substitutions, deletions and insertions between the aligned sequences (Fiser and Sali, 2003). This suggests that these methods are unreliable in predicting side chain conformations or portions of the target sequence that do not align well with the template.

Multiple sequence alignment of 20 characterized nitrilases vs. crystal structures of 4 distant homologues in the nitrilase superfamily





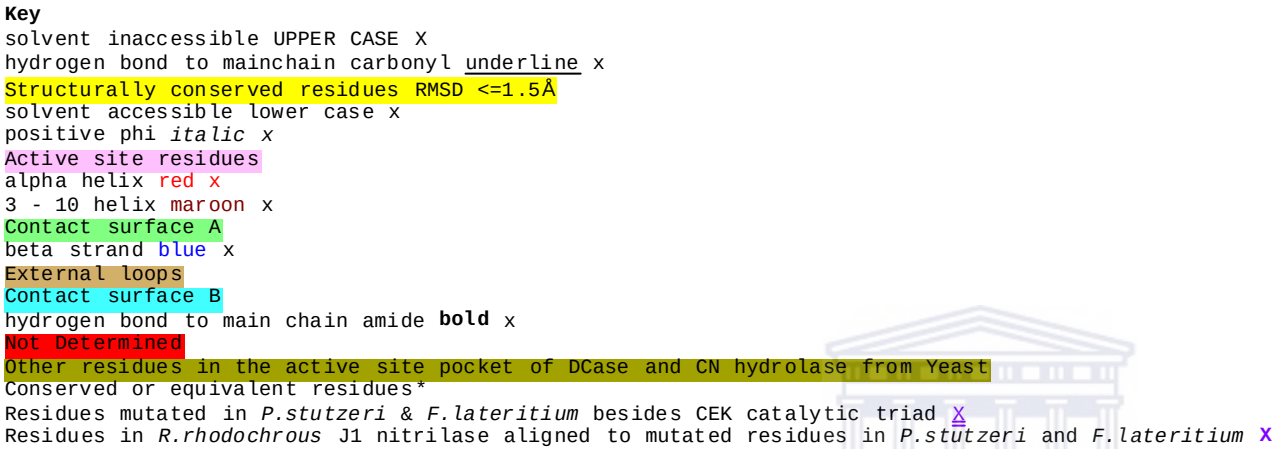


Figure 2.2: Multiple sequence alignment of 24 sequences of members in the nitrilase superfamily.
Nitrilases: *Bacillus pallidus* DAC521 (DAC521), *Rhodococcus rhodochrous* J1 (RrJ1), *Acidovorax facilis* 72W (Acdf72W), *Arabidopsis thaliana* Nit1, Nit2, Nit3, Nit4 (ANIT1/2/3/4), *Comamonas testosteroni* (Comat), *Rhodococcus rhodochrous* K22 (RrK22), *Synechocystis sp*. strain PCC6803 (Synec), *Klebsiella pneumoniae* subsp. *ozaenae* (Klebs), *Alcaligenes faecalis* JM3 (AlJM3), *Pseudomonas fluorescens* EBC191 (PfEBC191) and *Bacillus* sp. Strain OxB-1 (OxB-1-Nit).
Cyanide dihydratases: *Bacillus pumilus* C1 (BpumCDH) and *Pseudomonas stutzeri* AK61 (PstuCDH).
Cyanide hydratases: *Gloeocercospora sorghi* (GsorCH), *Fusarium lateritium* (FlatCH), *Fusarium solani* (FsolCH), and *Lephtospaheria maculans* (LeptCH) and
Crystal structures of distant homologues: Nit domain of NitFhit structure from *Caenorhabditis elegans* (1EMS), N-carbamyl-D-amino acid-amidohydrolase from *Agrobacterium* sp. strain KNK712 (1ERZ), Yeast Nit3 CN hydrolase (1F89), and *Pyrococcus horikoshii* hypothetical protein Ph0642 (1J31).
The alignments of the four crystal structures are presented in JOY notation (Mizuguchi et al., 1998). Secondary structures of the Nit Domain of NitFhit from *C.elegans* (Pace et al., 2000) and N-carbamyl-D-amino acid-amidohydrolase from *Agrobacterium* sp. strain KNK712 (Nakai et al., 2000) are annotated below the alignment respectively. The regions highlighted in yellow are conserved to within 1.5Å. The catalytic residues numbered E54, K127 and C169 according to Nit numbering are highlighted in pink. The external loops where nitrilases have significant insertions and deletions relative to the solved structures along their sequences are highlighted in brown. Contact surfaces A (highlighted in green) and B (highlighted in blue) corresponding to those of the Nit domain of 1EMS and 1ERZ structures are also shown. The alignment was constructed manually using results from GenTHREADER (Jones, 1999) and ALIGN (Cohen, 1997).

Two criteria determine the quality of a model. First, correctness of a model is determined by the quality of sequence alignment used to guide the modeling process. Errors in the sequence alignment will result in errors in spatial arrangement of these residues in the model. Secondly, accuracy of a model is related to the percentage sequence identity and structural equivalences between the target and template. This is limited by the structural deviation between the template structures used in the alignment. A model is considered "accurate enough" when the root mean square deviation (RMSD) is within the spread of deviations observed for structures used in the alignment. The RMSD gives the overall differences in the protein backbone structure by assessing the root mean standard deviations of the positions of the main-chain atoms. Baker and Sali (2001), report that models based on more than 50% sequence identity have about a 1Å root mean square (RMS) error for the main-chain atoms, while those based on sequence identity between 30-50% tend to have an RMS error of 1.5 Å. Below 30%, the model will have increasing alignment errors. Accuracy of models will also be limited by the modeling routine applied. Structural differences between models and solved structures may arise from errors inherent within the modeling routines and also from variations due to the molecular environment and the data collection methods used to determine the protein crystal structures. For instance, X-ray crystal structures show differences in data collection, data refinement, experimental conditions, crystal packing and the presence or absence of ligands.

All protein structure prediction methods, comparative modeling routines and molecular dynamics force field derive information from solved structures, which are known to have some errors. Therefore, verification of the modeling procedure is a prerequisite for good results. In this case, the method outlined above (section 2.4) was tested and verified by constructing a model of one of the solved structures and comparing it with its original.

2.6. Verification of the modeling procedure

2.6.1 A 3D model of the hypothetical protein Ph0642 of Pyrococcus horikoshii

The crystal structure of the hypothetical protein Ph0642 from *Pyrococcus horikoshii* (PDB ID, 1J31) is a product of structural genomics and its biological function remains unknown. This structure was solved at a resolution of 1.6Å. The protein is dimeric and shares a similar fold to 1EMS, 1ERZ and 1F89. Its sequence was aligned with those of the three crystal structures and used for building a 3D model using the procedure discussed in section 2.4. No structural information from the crystal structure was included in the modeling routine. While imposing symmetry, a dimeric model of 1J31 was constructed with 262 residues each monomer.

2.6.2 Results:

Structural alignment and modeling

A search for structural alignment with the sequence from *P.horikoshii* using GenTHREADER (Jones, 1999) produced 3 crystal structures, namely 1ERZ, 1EMS and 1F89 with an average score of 0.97. The average sequence identity is 27% with an average RMSD of 1.55Å when their coordinates were superimposed using ALIGN (cohen, 1997). The overall alignment is shown in figure 2.2, while figure 2.3 summarizes the structural information between them.

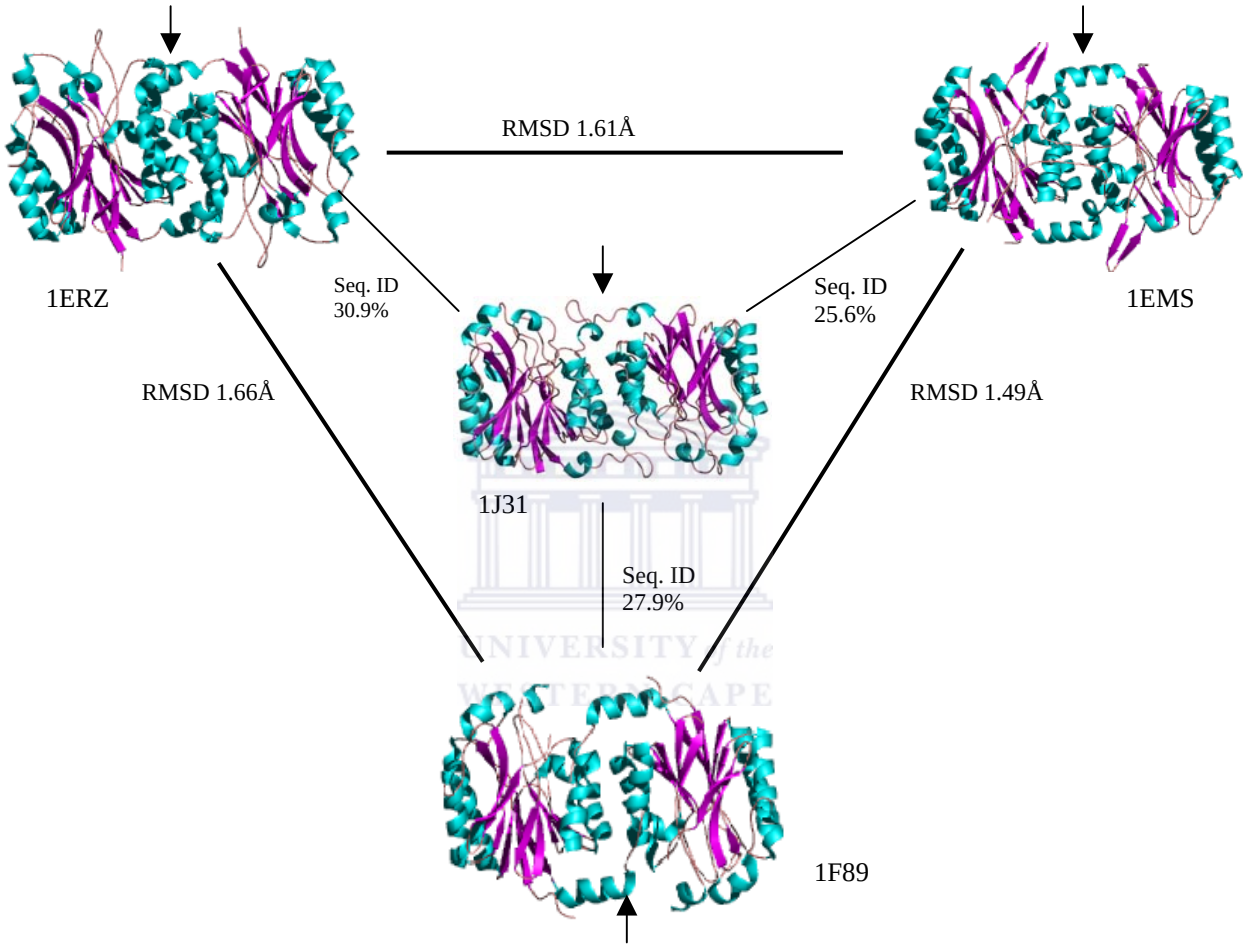


Figure 2.3. Cartoon representation of dimeric structures of N-carbamyl-D-amino acid amidohydrolase from *Agrobacterium sp*. KNK712 (PDB ID, 1ERZ), Nit domain of NitFhit from *C.elegans* (PDB ID, 1EMS), and Yeast Nit3 protein (PDB ID, 1F89) (Nakai et al., 2000; Pace et al., 2000; Kumaran et al., 2003) and the structure of the hypothetical protein Ph0642 from *P.horikoshii* (PDB ID, 1J31), which is shown at the center. The structures are depicted in colours that highlight the $\alpha\beta\beta\alpha$ fold and the two-fold non-crystallographic axis at the dimer interface, surface A (black arrow). Sequence identity (Seq. ID) and RMSD is also indicated. These figures have been generated using PyMOL v0.97 (DeLano, 2002).

The coordinates of the final model were superimposed with those of the crystal structure as shown in figure 2.4. The two structures fit well particularly in the central core of β-sheets, in the externally located α-helices and also at the A surface.

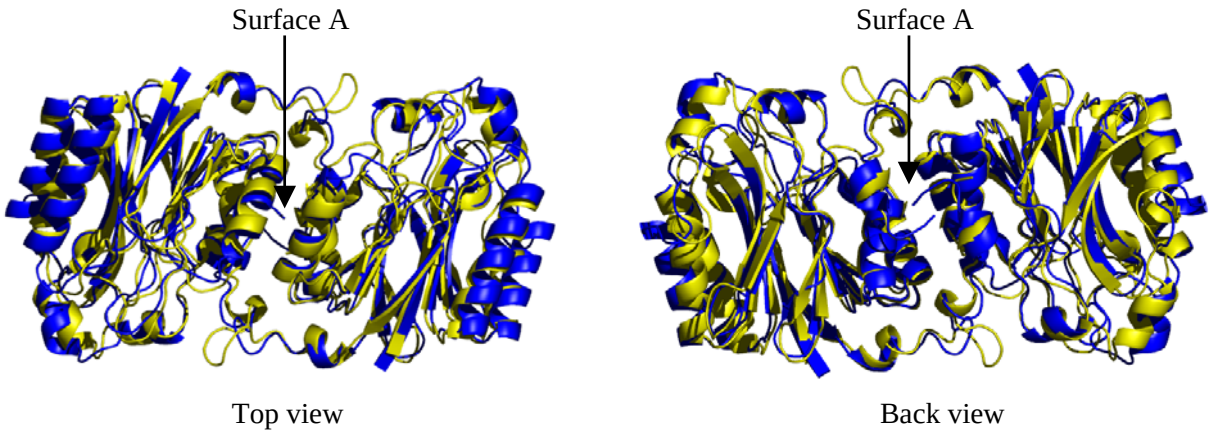


Figure 2.4. Cartoon representation of the superimposed coordinates of the model (blue) generated with spatial restraints and crystal structure (yellow) of hypothetical protein Ph0642 from *Pyrococcus Horikoshii* (PBB ID, 1J31). The average RMSD is 1.32Å. Structurally variable regions occur in the loops, coils and the C-terminal tail. The monomers associate via the conserved two-fold symmetric A surface (black arrow). This figure has been generated using PyMOL 0.97 (Delano, 2002).

There were visible changes in the orientation of the side chains after optimisation using SCWRL (Bower et al., 1997; Catunescu et al., 2003). Furthermore, an inspection of the catalytic triad showed conservation of residue positions, particularly the catalytic triad in both the model and crystal structure of 1J31 (figure 2.5).

With such a low sequence identity it is expected that the reliability of this model would decrease because of structural deviations between family members, which often occur in the loop regions and sometimes as shifts in the positions of secondary structure. For this reason, the structure was validated using RAMPAGE (Lovell et al., 2003). The percentage of the residues of the model in the favoured, allowed and outlier regions was 91.2, 5.4 and 3.5, while the crystal structure had 96.7% of the residues in the favoured region, 3.3% in the allowed region, and none in the outlier region. A similar check using PROCHECK (Laskowski et al., 1993), showed that majority of the residues were in favourable regions similar to those observed for the crystal structure. The Ramanchandran plots for both structures are shown in figure 2.6.

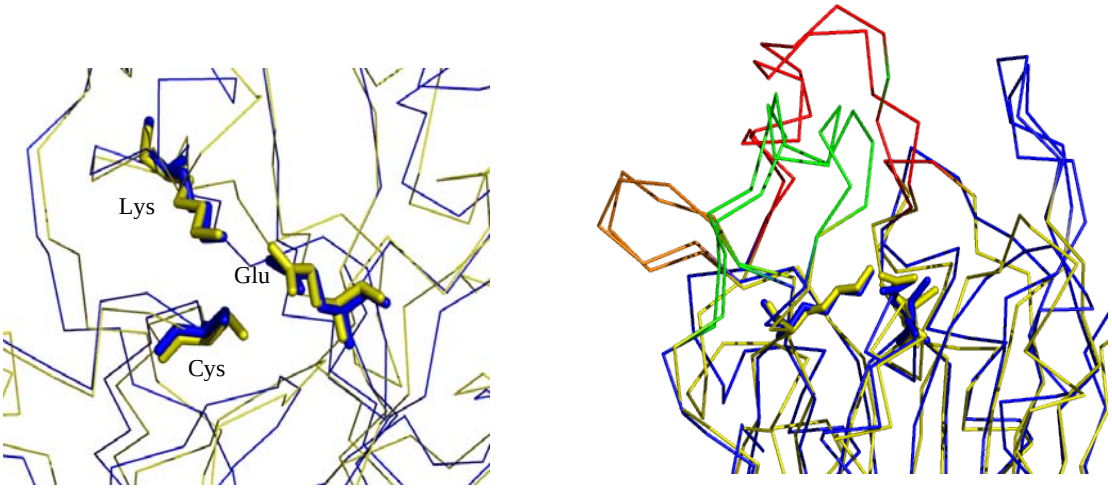


Figure 2.5. Left: Superimposed catalytic triad of the model (blue) and the crystal structure (yellow) of hypothetical protein Ph0642 from *P.horikoshii*. The catalytic residues, namely Glu42, Lys113 and Cys146 are shown as sticks and the model is a close fit in this region. Right: catalytic residues are seen to be pointing upwards into a cavity that is enclosed by 3 loops shown in red, green and blue. No attempt was made to model these loops. These figures have been generated using PyMOL v0.97 (DeLano, 2002).

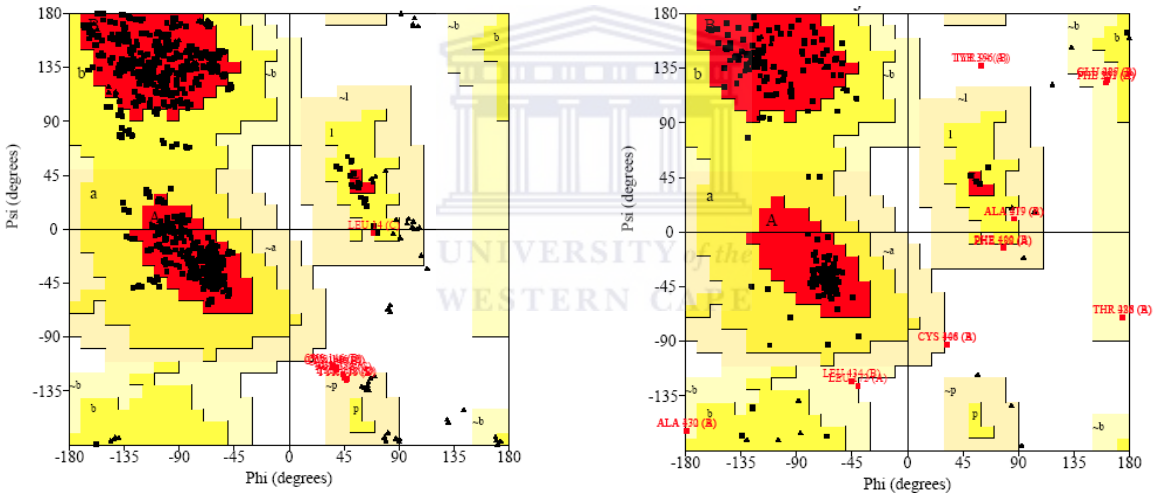


Figure 2.6. Ramanchandran plots for the crystal structure (left) and model (right) of the hypothetical protein Ph0642 from *Pyrococcus horikoshii* (PDB ID, 1J31). 87.4%, 8.7%, 3.0% and 0.9% of model residues compared to 89.4%, 9.5%, 0.6% and 0.5% of the residues in the crystal structure are found in the favoured, allowed, generously allowed and disallowed regions respectively.

The difference in residue distribution (figure 6) arises due to probable errors in the alignment, improper orientation of side chains and variability in the loop regions, which leads to a poor fit and consequently, poor stereochemistry in the model. However, the Ramachandran plot gives a good assessment of model accuracy. Further evaluation was carried out using PROSAII (Sippl, 1993). The energy profiles for the model and the solved structure of 1J31 are shown in figure 2.7. The Z scores for the model and the atomic structure of 1J31 are –5.36 and –6.64

respectively. The comparability of these scores shows that correct templates were used for model building.

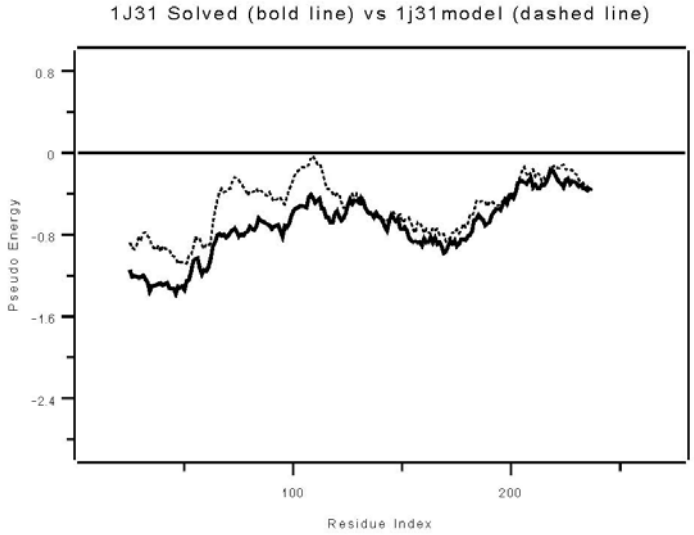


Figure 2.7. A combined energy graph of the model and atomic structure of 1J31. The graphs are smoothed by a window size of 50 residues. Both structures have similar pseudo energy profiles, which are below zero. Some variations in the energy profile of the model (dashed) correspond to the variable loop regions, as highlighted in brown colour in figure 2.2.

In conclusion, the model shows good stereochemistry and conserves the overall native protein fold common to this family of proteins. The Z scores also indicate that the correct choice of templates was made resulting in a model, which has been evaluated to be close to the solved structure. Such a model is good enough to provide a valuable interpretation and verifies the reliability of the modeling procedure discussed in section 2.4.

2.7 A 3D model of the structure of the nitrilase from *Rhodococcus rhodochrous* J1

The nitrilase of *Rhodococcus rhodochrous* J1 (from here on referred to as J1 nitrilase) is a member of branch 1 of the nitrilase superfamily (Pace and Brenner, 2001). As already discussed, this enzyme exists as an inactive dimer, which becomes active upon oligomerization in the presence of benzonitrile (Nagasawa et al., 2000). This phenomenon has previously been studied in other Rhodococci nitrilases (Harper, 1997; Stevenson et al., 1992). Three-dimensional electron microscopy of the cyanide dihydratases from *P.stutzeri* AK61 and *B.pumilus* C1 show that 3D models can be used to interpret their low-resolution maps and provide insights about interactions that stabilize their quaternary structures (Jandhyala et al., 2003; Sewell et al., 2003). Despite these efforts, the structural gap, that is, the difference between the number of known sequences and experimentally determined 3D structures, is wide within this family. Nevertheless, availability of atomic structures of distant homologues within the nitrilase superfamily provides an opportunity to build a 3D dimeric model of J1 nitrilase. It

is hoped that this model will provide clues of residues located in the substrate-binding pocket and provide an interpretation of its 3D map.

2.7.1 Structural alignment and homology

A search for structural alignment using the protein sequence of J1 nitrilase against known structures in the PDB using GenTHREADER (Jones, 1999) produced four candidates with an average score of 0.97 and sequence identity of 19.4%. The results are summarized in Table 2.1 and figure 2.8.

	1F89	**1J31**	**1EMS**	**1ERZ**
1F89 (2.4Å)	282	83	89	60
1J31 (1.6Å)	28.2%	262	67	79
1EMS (2.8Å)	31.4%	25.6%	296	62
1ERZ (1.7Å)	21.8%	30.2%	20.9%	303

Table 2.1. Comparison statistics of the four nitrilase atomic homologues. The number of amino acids in each structure is highlighted in red, sequence identity (blue), identical amino acids residues (green) and the resolution of each structure is highlighted in brackets in the left hand column.

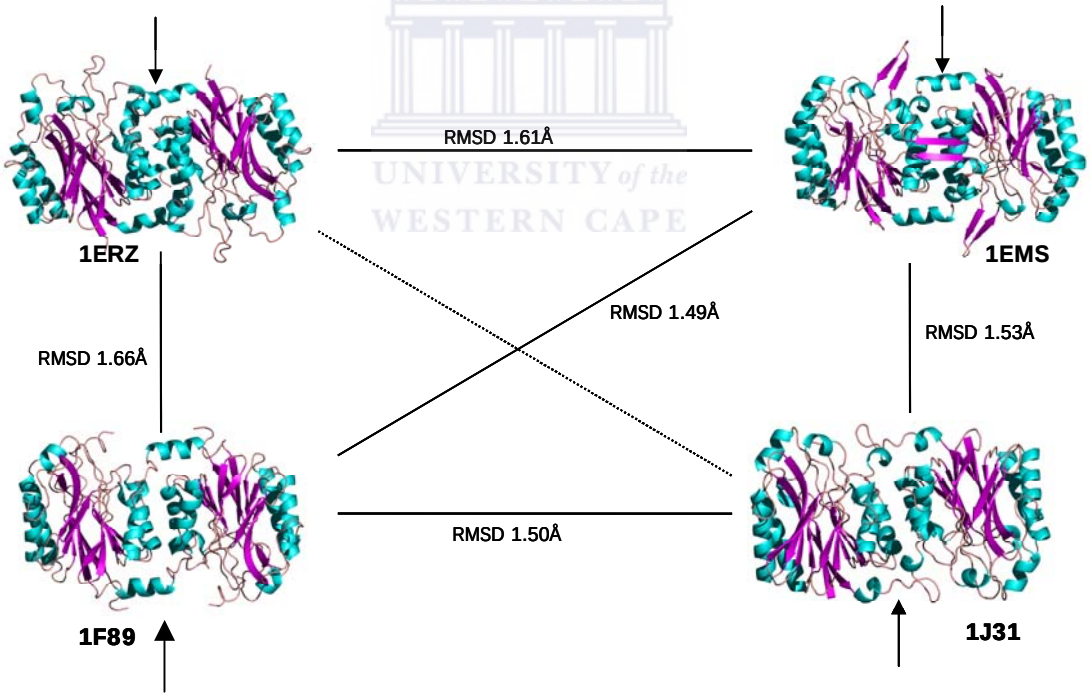


Figure 2.8. Cartoon representation of the structures of 1EMS (Pace et al., 2000), 1ERZ(Nakai et al, 2000), 1F89 (Kumaran et al., 2003) and 1J31 (Sakai et al., 2004). The $\alpha\beta\beta\alpha$ fold between monomers is conserved across the two-fold dimeric interface referred to as the A surface (black arrows). The figures were generated using PyMOL v0.97 (DeLano, 2002).

Pair-wise alignment of the atomic structures of 1EMS, 1ERZ, 1F89 and 1J31 using ALIGN (Cohen, 1997) showed high fold conservation with structural variability in the loops and the C-terminal tail (figure 2.9). The average RMSD is 1.55Å. Guided by predictions from GenTHREADER (Jones, 1999) and ALIGN (Cohen, 1997), the resulting alignment is shown in figure 2.10.

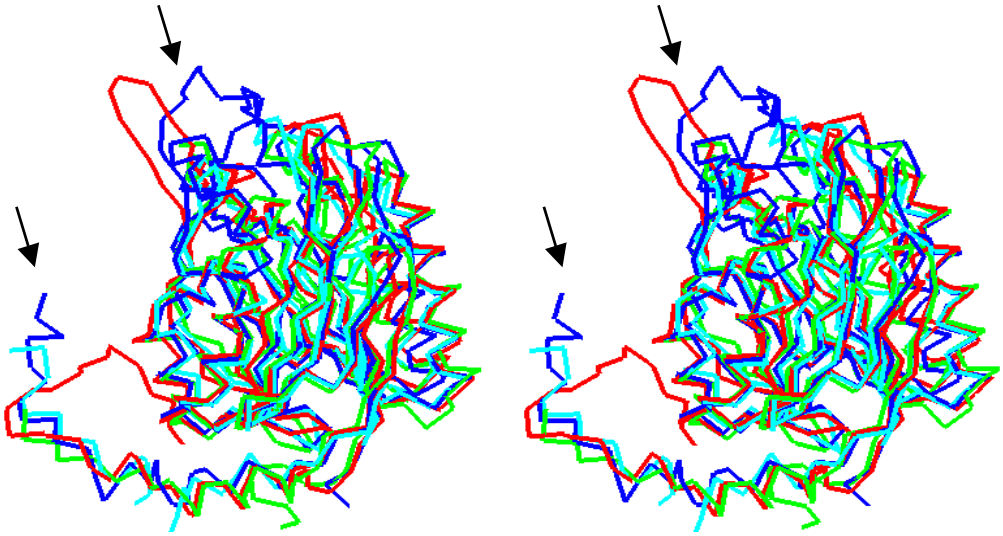
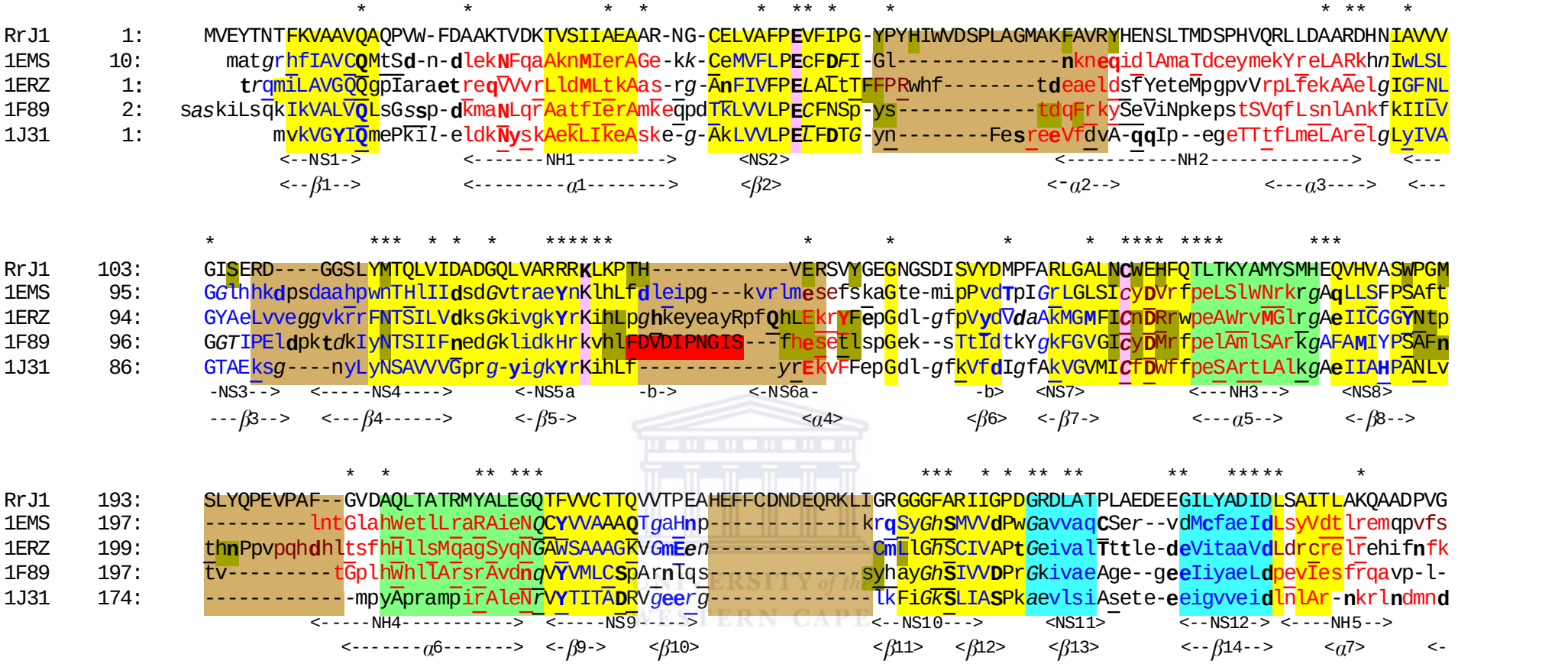


Figure 2.9. Stereo view of a Cα ribbon representation of the superimposed coordinates of chain A of the four atomic structures of 1EMS, 1ERZ, 1F89 and 1J31, are shown in red, blue, green and cyan respectively. The central region of β sheets and external α helices is highly conserved. Poor fits occur in the loop regions and the C-terminal tail (black arrows). The figures were generated using PyMOL v.0.97 (DeLano, 2002).

The J1 nitrilase has a total of 366 amino acid residues in its sequence, while 1ERZ is the longest template with 303 amino acids. This allowed for 313 residues of J1 nitrilase to be modeled because of lack of a template to build the extra 53 residues in its C-terminal tail. Although it is common practice to use those templates that were solved at high resolution and have significantly better sequence identity, it has been shown in section 2.6 that the use of several templates leads to a model which is significantly closer to the target structure than the templates. A dimer model of J1 nitrilase was constructed using MODELLER (Sali and Blundell, 1993), according to the method described in section 2.4 and 2.6. The final model is shown in figure 2.11.



Key

solvent inaccessible UPPER CASE X
hydrogen bond to mainchain carbonyl underline x
Structurally conserved residues (RMSD <=1.5Å)
solvent accessible lower case x
positive phi *italic x*
Hydrogen bond to main chain amide **bold** x

Beta strand blue x
Contact surface A
External loops
Contact surface B
Not Determined
Residues predicted in the active site pocket of *R.rhodochrous J1* nitrilase

Active site residues
alpha helix red x
3 - 10 helix maroon x
Conserved or equivalent residues(*)

Figure 2.10. Sequence alignment of *R. rhodochrous* J1 nitrilase (RrJ1) against sequences of the four nitrilase atomic homologues, namely 1EMS, 1ERZ, 1F89 and 1J31. Secondary structural elements of 1EMS and 1ERZ are presented below the alignment respectively. The protein sequences of the four crystal structures are presented in JOY notation (Mizuguchi et al., 1998). The three active site residues E54, K127 and C169 according to 1EMS numbering are shown in pink. This alignment is based on predictions using GenTHREADER (Jones, 1999) and pairwise alignment of the atomic structures using ALIGN (Cohen, 1997).

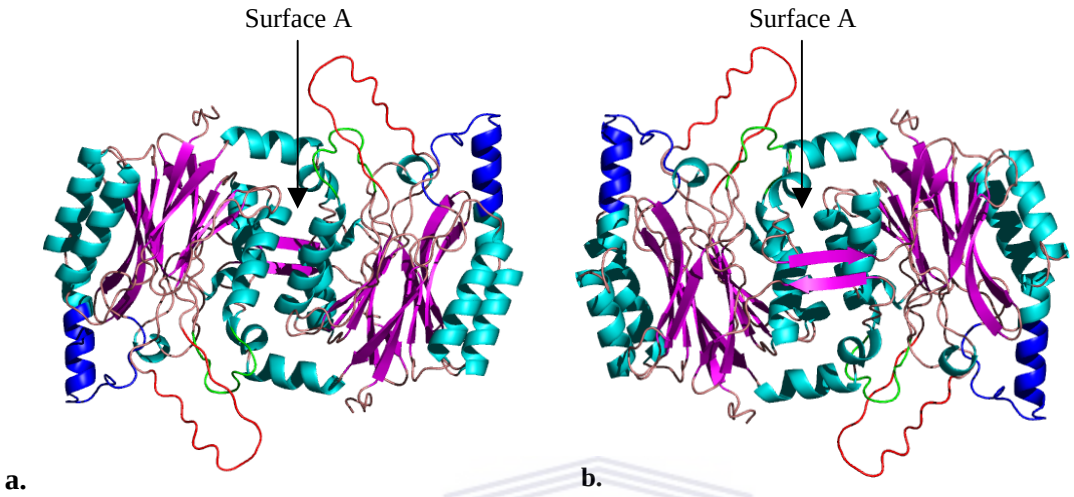


Figure 2.11. Cartoon representations of a dimer model of *Rhodococcus rhodochrous* J1 nitrilase. The $\alpha\beta\beta\alpha$-$\alpha\beta\beta\alpha$ structural fold is visible from the front (a) and back (b) of the model. Three long insertions, namely residues 54-73, 193-202, and 234-249 are highlighted in blue, green, and red respectively. Two helices from each monomer and a beta strand interact across a conserved two-fold axis located at the A surface (black arrows). The figures were generated using PyMOL v0.97 (DeLano, 2002).

2.7.2 Evaluation of the model

Superposition of the coordinates of chain A of the model with each of the corresponding chains of the templates using ALIGN (Cohen, 1997), revealed extensible conservation of the core β-sheets, α-helices and the catalytic triad (figure 2.12). Major differences are seen to occur in the region of insertions (figure 10) relative to the solved structures. Stereochemical analysis of the final model using PROCHECK (Laskowski et al., 1993) revealed that 82.4% of the residues were in most favoured regions, 13.6% and 1.1% in the additional and generously allowed regions, while 2.9% were in the disallowed regions. No attempt was made to model the long insertions having about 12-14 amino acids. Subsequent protein structure evaluation using PROSAII (Sippl, 1993) shows that the model has a similar energy profile to those of the atomic structures of distant homologues (figure 2.13).

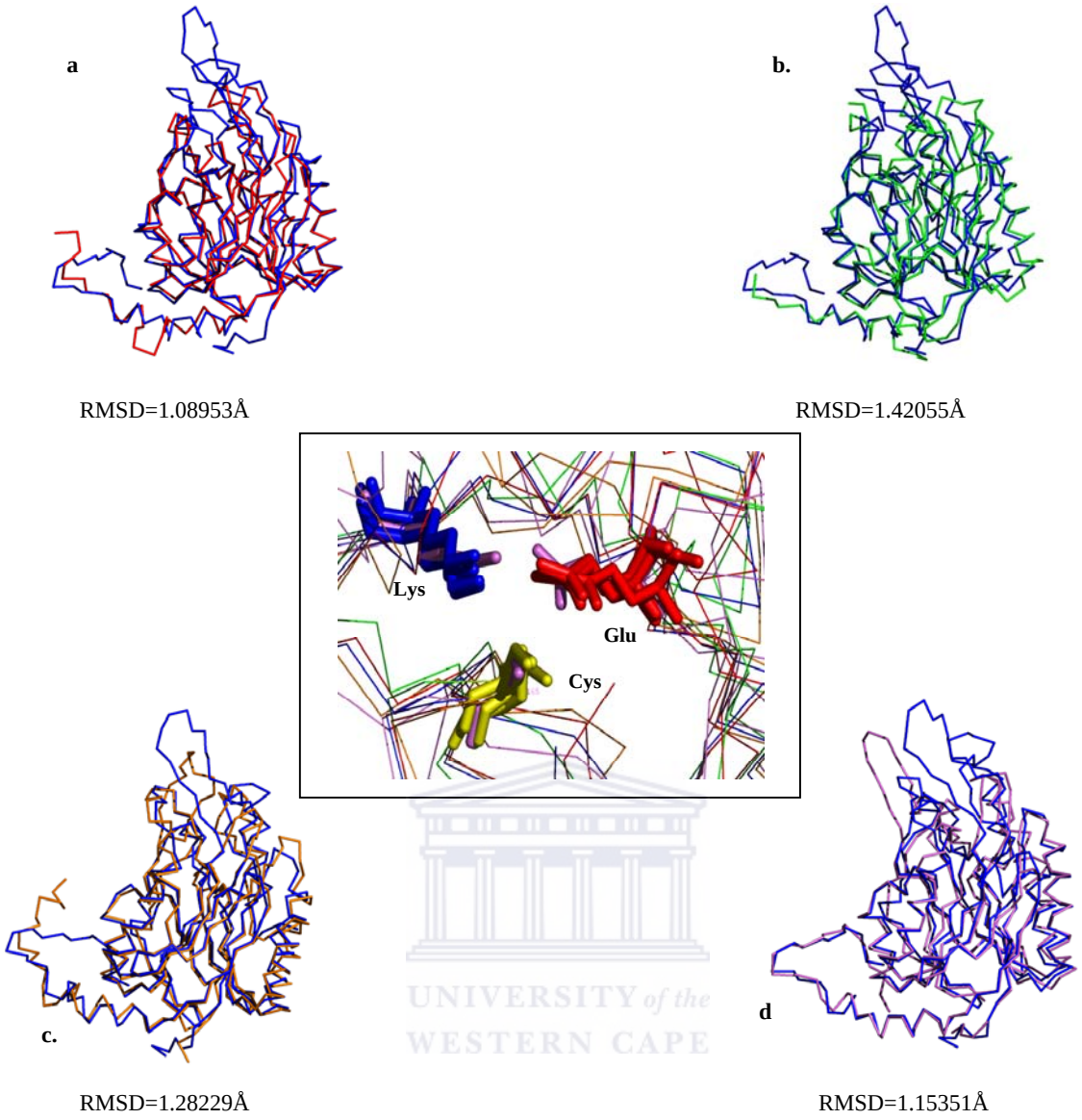


Figure 2.12. Ribbon representation of the superimposed coordinates of a monomer of J1 nitrilase model (blue) with monomers of the crystal structures of **a**. 1J31 (red); **b**. 1F89 (green); **c**. 1ERZ (orange) and **d**. 1EMS (violet). The two long insertions in J1 nitrilase sequence form extensive loops that are difficult to model. In the center, a close-up stick representation of the structurally conserved catalytic triad between these structures is shown. Residues Glu48, Lys131 and Cys165 in the sequence of J1 nitrilase are highlighted in violet. The figures were generated using PyMOL v0.97 (DeLano, 2002).

A few positive peaks are present in the energy profiles of the atomic structures and absent for J1 nitrilase model. This may indicate errors within the model of J1 nitrilase, which may arise due to errors in the alignment or from improper modeling of the actual interacting environment between residues in the atomic homologue structures.

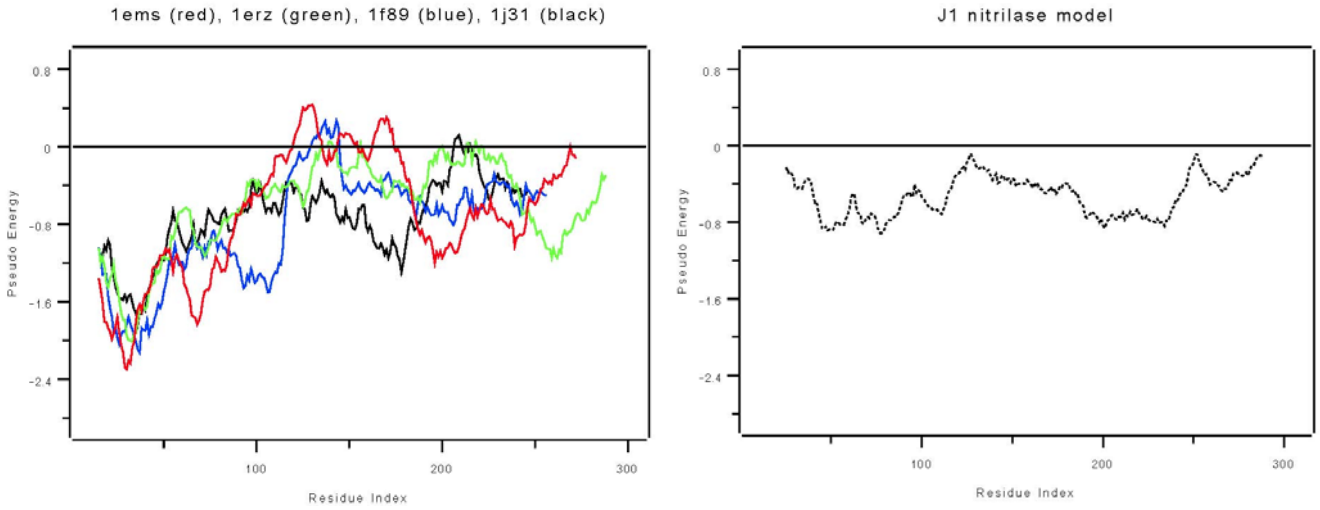


Figure 2.13. Energy profiles of the atomic homologues in the nitrilase superfamily are shown (right) and that of J1 nitrilase model (left). The graphs are smoothed by a window size of 50 residues. All the profiles appear to be similar. This indicates conservation of the protein fold within the family.

It may also be possible that the atomic structures have errors in these regions or their side chains adopt unusual conformations. The Z scores for J1 nitrilase model is –7.91 whereas for 1EMS, 1ERZ and 1F89 is –6.42, -6.36 and –7.49 respectively. These scores demonstrate the model inherits useful structural information from the templates, the correct choice of templates and assures quality of the final model.

2.8 Discussion

2.8.1 Insights from the model

The dimer model of J1 nitrilase provides interesting insights about this enzyme. One of the long insertions in its sequence, residues 54-73 extends the amino terminal of the alpha helix (labeled NH2 in NitFhit or 1EMS structure). The second insertion between residues 234-249 forms an extended loop of 14 residues between two beta sheets labeled NS9 and NS10 in the structure of 1EMS (figure 2.2). The residues in these regions are predicted to be involved in symmetric interactions between dimer subunits across the groove in the spiral structures of $CynD_{pum}$ and $CynD_{stu}$, which has been called the C surface (Jandhyala et al., 2003; Sewell et al., 2003; Sewell et al., 2005). Furthermore, it is predicted that residues located at these positions will contribute additional stabilizing interactions similar to those observed in the spiral structures of $CynD_{pum}$ and $CynD_{stu}$ (Sewell et al., 2005).

Residues 234-249 and residues 193-202 form two loops positioned above the active site cavity. It is predicted that these loops play a role in the process of oligomerization and may restrict access of the substrate to the active site (figure 2.14).

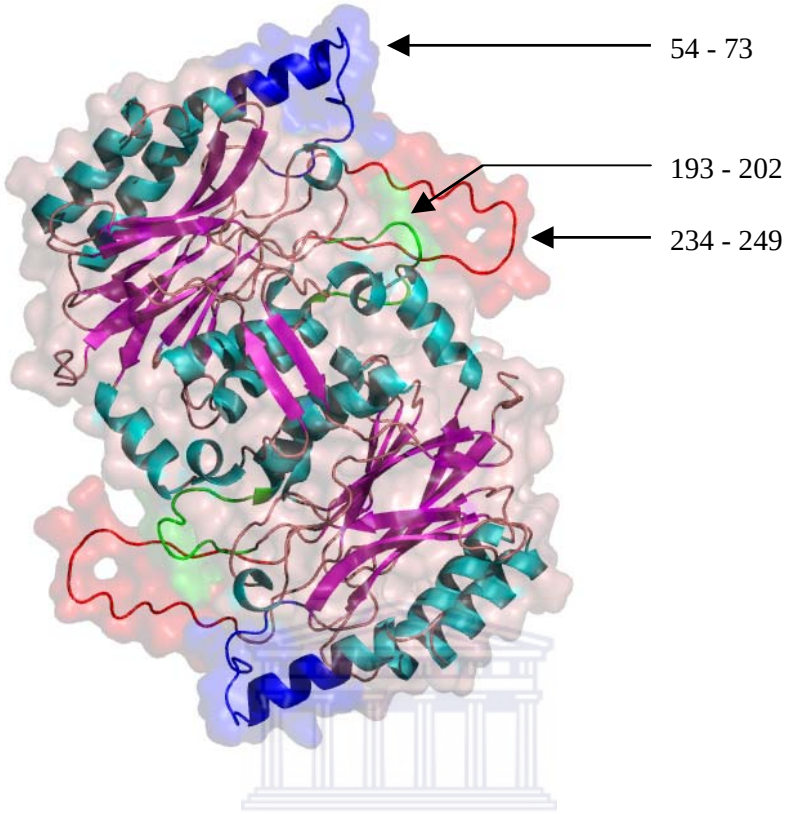


Figure 2.14. Molecular surface representation of the model of the nitrilase of *R.rhodochrous* J1. Residues 54-73, 193-202, and 234-249 are shown in blue, green and red respectively. The model has 626 residues. The surface is rendered transparent to highlight the $\alpha\beta\beta\alpha$ structural fold in each monomer. This figure was generated using PyMOL v0.97 (DeLano, 2002).

Hence, the substrate would access the active site pocket from the side while the tunnel formed by the two loops (highlighted above) is predicted to form an "escape route" for the products formed following hydrolysis of the substrate (figure 15). Each monomer has one active site.

2.8.2 Active site environment

Nitrilases are known to hydrolyze various nitriles but their mechanism of action as well as their specificity is not well understood. The volume of the active site pocket would determine the size of the substrate, whereas the residues lining the active site pocket determine the specificity. The J1 nitrilase is known to hydrolyze benzonitrile, acrylonitrile among other aliphatic nitriles (Nagasawa et al., 2000).

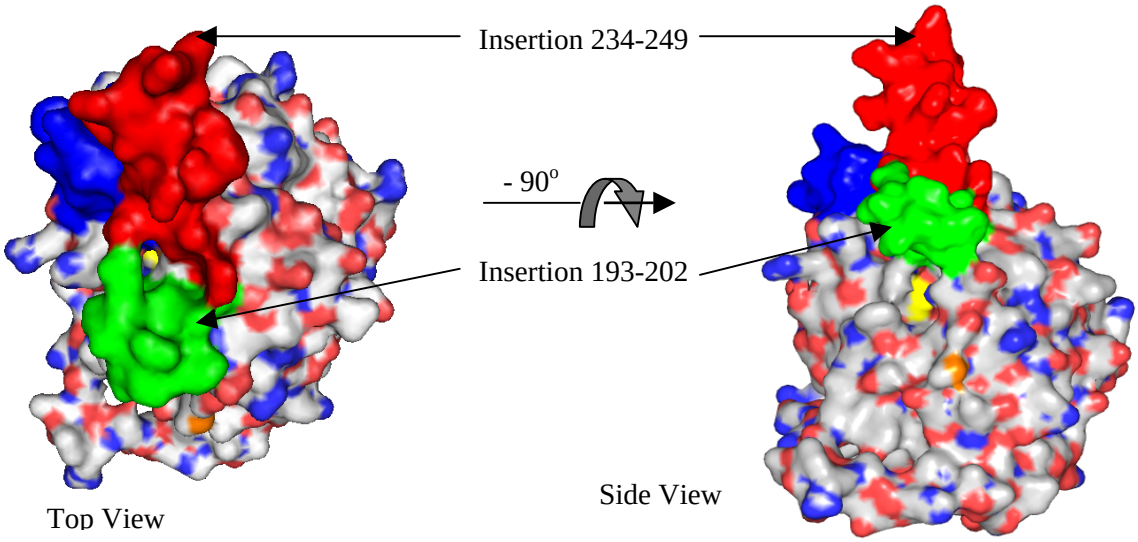
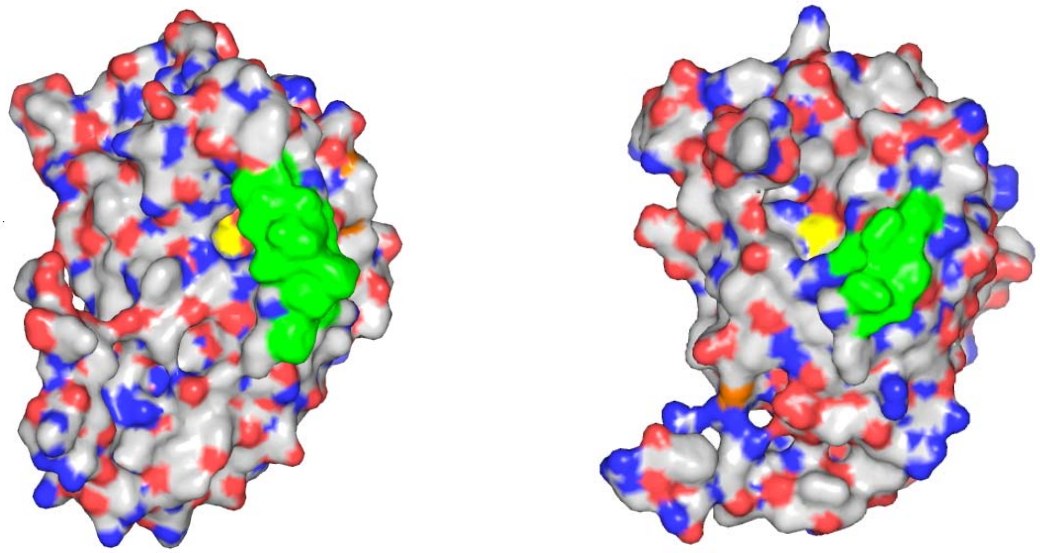


Figure 2.15. Molecular surface representation of the model of *R.rhodochrous* J1 nitrilase. The active Cys165 residue is highlighted in yellow and is located deep inside the cleft formed by residues in the loops and residues in the substrate-binding pocket. It is predicted that the substrate would access this cleft through one of the openings and upon hydrolysis the products would exit through the other. The externally extended helix shown in blue (residues 54-73) is located away from the active site. This figure was generated using PyMOL v 0.97 (Delano, 2002).

The size of these substrates suggests that the volume of the cavity surrounding the active site of J1nitrilase is smaller compared to that calculated in the crystal structure of Yeast (1F89) CN hydrolase (Kumaran et al., 2003). Furthermore, an insertion of residues 193-202 in the sequence of J1 nitrilase reduces the size of this cavity since these residues are observed to form a loop that occludes the opening of this cavity (figure 2.15). This corresponds to an insertion (residues 201-213) in structure of 1ERZ in which a similar effect is observed. Therefore, it follows that the size of the active site cavity in J1 nitrilase and 1ERZ is smaller than observed for 1F89 (figure 2.16). This also suggests that their substrate specificity is not broad as that of 1F89.

2.8.3 Binding Pocket

The catalytic residues, namely Glu48, Lys131 and Cys165 are located deep inside a cleft, which is predicted to be lined with residues Tyr54, His57, Phe70, Tyr74, Ser105, Met114, Thr135, His136, Glu138, Tyr142, Asn164, Trp166, His168, Trp189, and Met192 (as highlighted in figure 2.10 and 2.17). This prediction is based on structural alignment to the residues reported to be lining the active site pocket in the atomic structures of 1F89 (Kumaran et al., 2003), 1ERZ (Nakai et al., 2002) and 1FO6 (Chen et al., 2003).



1ERZ 1F89

Figure 2.16. Molecular surface representation of the structures of *Agrobacterium sp*. KNK712 (1ERZ) (left) and Yeast Nit3 protein (1F89; right) with many charged residues located on their surface The active cysteine residue (highlighted in yellow) is located deep in the active site cavity. The insertion between residues 201-213 in 1ERZ and the corresponding deletion in 1F89 are highlighted in green. 1F89 has a larger cavity because it does not have the loop overhang above the active site cleft unlike 1ERZ. This suggests a larger substrate for 1F89 compared to the J1 nitrilase and 1ERZ. The figures were generated using PyMOL v0.97 (DeLano, 2002).

Most of these residues point towards the center of the pocket, are hydrophobic and buried and may determine the type of substrate. The nitrilase of *R.rhodochrous* J1 is known to have preference for aromatic substrates.

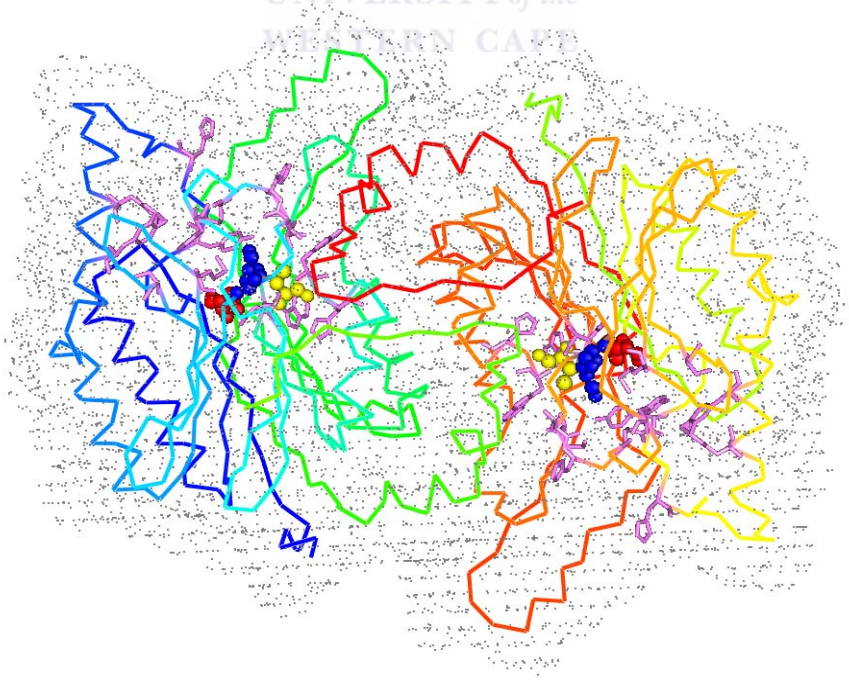


Figure 2.17. Alpha carbon representation of the model of the nitrilase from *Rhodococcus rhodochrous* J1. The catalytic residues, namely E48, K131 and C165 are highlighted in ball representation in red, blue and yellow respectively. Residues predicted to be in the substrate-binding pocket are shown in stick representation in violet colour. This figure was generated using PyMOL v0.97 (DeLano, 2002)

2.8.4 Effect of previous nitrilase mutations on structure

Earlier mutational studies have shown that a single change to any one of the three catalytic residues results in instability, reduced activity or inactivation of the enzyme (Novo et al., 2002). Further mutations on residues located away from the active site in the cyanide dihydratase from *P.stutzeri* AK61 (Watanabe et al., 1998), and the cyanide hydratase from *F.lateritium* (Nolan et al., 2003), which include a change in residue length, addition or removal of a charged residue or introduction of a kink are seen to disrupt the structurally conserved core of beta sheets (O'Reilly and Turner, 2003). Interestingly, two mutations in the cyanide dihydratase of *P.stutzeri* AK61, namely Glu180 to Gln180 and Asp247 to Asn247 showed a 1.2 and 1.7-fold increase in enzymatic activity respectively. These residues align with the Glu182 and Asp260 in the sequence of J1 nitrilase (Figure 2.2 and Table 2.2). It is not clear why these mutations led to increased activity in the cyanide dihydratase of *P.stutzeri* AK61. However, Glu182 is located close to the A surface and aligns with Ala 186, 188, 186 and 163 in the atomic structures of 1EMS, 1ERZ, 1F89 and 1J31 respectively, which do not form spiral structures (figure 2.10).

Reduced activity	Reduced stability	Increased activity
Q14, A15, Y54	E48, S105	E182
E106, K131, C165	E167, K289	D260
L172, H185, E273	D293, H297	
A278, S299	Y298	

Table 2.2. Corresponding residues in the nitrilase from *R.rhodochrous* J1 based on mutational studies on the enzymes from *P.stutzeri* AK61 (Watanabe et al., 1998) and *F.lateritium* (Nolan et al., 2003). The residue numbers are as highlighted in figure 10.

On the other hand, Asp260 is situated at the A surface and points towards the interface between the two interacting monomers. Additional interactions designated C, D and E surfaces would be required to stabilize spiral. It remains to be seen which residues contribute to these surfaces in J1 nitrilase.

2.8.5 Interactions at the A surface

The structure of the CN hydrolase from Yeast (1F89) is held together by salt bridges and hydrogen bonds between residues located on two helices (NH3 and NH4 in Nit numbering, figure 2.10) from each monomer at the A surface (Kumaran et al., 2003). These interactions are also present in the other three atomic resolution homologues. In the J1 nitrilase model, residues

Y214 and E217 are predicted to form a salt bridge, while there is a possibility of a strong hydrogen bond contact between symmetry related Met213. However, it is not possible to predict other residues in the model that align with those reported to form salt bridges in 1F89 due to poor modeling of side chains.

A component of a beta sheet from each monomer at the C-terminal of 1EMS interacts across the A surface (Pace et al., 2000). Furthermore, a series of mutants at the C surface in both $CynD_{pum}$ and $CynD_{stu}$ showed that these enzymes lost activity when part or the entire C terminal was truncated (Sewell et al., 2005). These results correlate with a region of density on the inner surface of their low-resolution maps, which was observed to be sufficient to accommodate the extra 30-35 residues in the C-terminal tail (Jandhyala et al., 2003; Sewell et al., 2003, 2005). Although 53 residues in the C-terminal of J1 nitrilase were omitted from model building, this line of evidence suggests that part these residues are also involved in interactions across the A surface. Indeed, the J1 nitrilase model is observed to have anti-parallel beta sheets in its C-terminal, which interact at the A surface (figure 2.14).

2.8.6 Spiral forming surfaces

In combination with interactions at the A surface, Sewell et al (2003, 2005), found further interactions arising from two significant insertions in the sequence of $CynD_{stu}$ which interact across a pseudo dyad symmetry axis extending the spiral. This surface is symmetric between subunits and is referred to as the C surface (Sewell et al., 2003). The insertions correspond to residues 54-73 (alpha helix NH2 in Nit numbering, figure 2.10) and 234-249 (bend between beta sheets NS9 and NS10, figure 2.10) in the sequence of J1 nitrilase. In the structures of $CynD_{stu}$ and $CynD_{pum}$, further interactions across the groove were observed, which confer stability to the spiral via formation of symmetric salt bridges (Sewell et al., 2003; Jandhyala et al., 2003). These interactions are located near the carboxy terminal end of the NH2 helix (figure 2.10), comprising residues 82EAVQK86 and 90EAAKRNE96 in $CynD_{stu}$ and $CynD_{pum}$ respectively, and are referred to as the D surface (Sewell et al., 2005). Sequence modification at this surface in $CynD_{pum}$ had no effect on cyanide dihydratase activity (Sewell et al., 2005). The corresponding residues in J1 nitrilase are 91-97 having the sequence DAARDHN (figure 2.10). This sequence suggests the possibility of two salt bridges across the groove between dimers, which would confer stability to the spiral. Whether this is true, ultimately depends on using the homology model to interpret a 3D map of its quaternary structure depending on the resolution achieved. Another asymmetric interaction across the groove involving residues 92RKNK95 (NH2 helix on one side) and the highly conserved residues 266EID268 (NS12 beta

sheet, see figure 2.10) was modeled to be the cause for terminating the spiral of the $CynD_{stu}$ (Sewell et al., 2005). This interaction is referred to as the E surface and it provides the first acceptable explanation into the mystery surrounding self-terminating homo-oligomeric spirals. This interaction is yet to be observed in the octadecameric structure of $CynD_{pum}$, but it is not expected to be present in J1 nitrilase.

2.8.7 Design of mutants

An interpretation of the structures of the cyanide dihydratases from *B.pumilus* C1 and *P.stutzeri* AK61 showed that the active enzymes possess approximately 5-fold symmetry along the spiral axis, which suggests that the native enzyme cannot be crystallized. Professor BT Sewell of the University of Cape Town, has proposed that breaking the active dimers at the A surface and forming inactive dimers across the C surface, may result in enzymes possessing the required symmetry (2, 3, 4, or 6-fold) for crystallization. Using insights from the model and structural alignment (figure 2.2), modifications to the A and D surfaces were proposed. Dr. Arvind Varsani, a molecular biologist at the University of Cape Town, designed each construct in *E.coli* cells strain BL21 bearing plasmid pET26b into which the modified gene of the wild type enzyme is incorporated. The mutations were confirmed by DNA sequencing and activity assays. These are listed in Table 2.3.

Table 2.3. Modifications to the nitrilase of *R.rhodochrous* J1

Mutant	Surface	Activity	Change and location
D91A	D	Active	Mutation of D91 → A91
D91A/H96N/N97Q	D	Not determined	Mutation of D91, H96, N97 to A91, N96 and Q97
D91E/R94K/D95R /H96N/N97E	D	Not determined	Mutation of D91, R94, D95, H96, N97 to E91, K94, N96 and Q97
Y178S/R179K	A	Inactive	Mutation of Y178, S179 to R178 and K179
Y214E	A	Inactive	Mutation of Y214 → E214
Delta 317	C-terminal	Active	Stop codon at N317
Delta 340	C-terminal	Active	Stop codon at R340

The results above show that J1 nitrilase does not require most of its C-terminal for activity. Furthermore, any modification at the A surface results in loss of activity as observed in the cyanide dihydratases from *B.pumilus* C1 and *P.stutzeri* AK61 (Sewell et al., 2005). More detailed biochemical studies of these mutants will be undertaken in future.

2.9 Conclusion

This chapter demonstrates that we can confidently construct a comparative model despite the low sequence identity among members in the nitrilase superfamily. J1 nitrilase conserves a similar fold across a two-fold symmetric surface and the catalytic residues are observed to be at equivalent distances similar to those in the atomic homologues. The model was used to predict residues that are likely to contribute to spiral formation and stabilization. Modification of residues at the A surface inactivates the enzyme and it is expected that the modifications at the C or D surfaces may also have a similar effect. This nitrilase is also functional without part of its C-terminal whose components are predicted to interact at the A surface. Even though these methods require some experimental knowledge about the target protein, they are theoretical and cannot replace known experimental methods of structure determination. However, the underlying objective is to use the model to interpret a low-resolution map of the quaternary structure of J1 nitrilase.

CHAPTER 3

PURIFICATION OF RECOMBINANT *RHODOCOCCUS RHODOCHROUS* J1 NITRILASE

3.1 Introduction

This chapter describes the expression and purification protocols of the recombinant nitrilase from *R.rhodochrous* J1 (J1 nitrilase). These are based on methods previously described in Nagasawa et al., 2000. The objective was to obtain purified homogeneous protein for structure determination using techniques of electron microscopy and 3D image processing.

3.2 Materials and Methods

3.2.1 Origin of bacterial plasmid

The host cell is *E.coli* strain BL21(DE3)pLysS, carrying plasmid pET30a, into which the gene of the wild type enzyme is incorporated. This recombinant clone was a generous gift from Dr Charles Brenner, Departments of Genetics and Biochemistry, Dartmouth Medical School, New Hampshire, US.

3.2.2 Growth conditions

The host cell, *E.coli*, was grown in Luria Bertoni broth (LB) (10 g/L Tryptone, 5 g/L Yeast extract, 5 g/L NaCl and 1 mL/L of 1M NaOH in distilled water), which was autoclaved and cooled to room temperature. A small amount of a glycerol stock of *E.coli* cells stored at –70^{o}C was used to inoculate 5 ml of liquid preculture containing 25 μg/mL Kanamycin and 200μg/mL Chloramphenicol and grown overnight at 30 ^{o}C while shaking. 1 mL of overnight culture was then diluted into 5-Litre Erlenmeyer flasks containing 1L of LB, 25 μg/mL Kanamycin with no chloramphenicol. The liquid culture (occupying 20 % of flask capacity) was grown at 37 ^{o}C until OD_{600} ~ 1. Cells were induced with the addition of IPTG to a final concentration of 1mM and the experiment was allowed to occur overnight at 30 ^{o}C while shaking.

3.2.3 Harvesting and lysis

Cells were collected by centrifugation for 10 minutes at a temperature of 4 ^{o}C at 4000 × g and resuspended in 1 L of fresh LB containing 200 μg/ml of Chloramphenicol. This is a pre-cautionary step to wash of residual IPTG and in the process solubilize any protein that was

expressed in inclusion bodies. This was grown for ~3 hours at 37 ^{o}C or until OD_{600} was ~1. Cells were again harvested by centrifugation and resuspended in 10 % of the original volume in 100 mM KH_2PO_4 pH 7.8 (buffer A), followed by the addition of 50 mM 2-mercaptoethanol and 2 ml of lysozyme (10 mg/ml) dissolved in buffer A. Lysis was allowed to occur at 37 ^{o}C for 30 minutes while shaking and then the lysate was cooled on ice for 30 minutes. The cells were disrupted by sonication for 4 minutes (15 seconds each cycle with a cooling period of 15 seconds between bursts) using a Misonix Sonicator® 3000. An ethanol/ice slurry was used to cool the sample during sonication. One tablet of protease cocktail inhibitors (Roche Diagnostics GmbH, Mannheim, Germany) was added and then the cell homogenate was centrifuged at 20000 × g for 30 minutes. The cell free extract (100 ml) was diluted with buffer A (260 ml) and 10 %(v/v) ethanol (40 ml) pre-cooled at -20^{o}C. This was stored overnight at 4^{o}C.

3.2.4 Anion exchange chromatography

All chromatography was done at either at room temperature or 4^{o}C. The active fractions were sometimes subjected through a second round of chromatography in order to try and obtain a homogeneous protein preparation. In addition to the Bradford assay (Bradford, 1976), the protein concentration was sometimes determined photometrically at 280nm using the known extinction coefficient of J1 nitrilase of 0.93 mg^{-1} $cm^{-1}ml^{-1}$.

3.2.4.1 Pilot experiment by step elution

In order to determine the appropriate salt concentration for elution of J1 nitrilase, 1 ml of freshly prepared cell extract was adsorbed on a miniature Q-Sepharose column with a volume of 2 ml. This column was packed in a syringe barrel and stoppered with glass wool. The column was equilibrated with buffer A prior to loading the sample and then exhaustively flushed with the same buffer again. A step gradient of buffer A was then passed through the column with successive KCl concentrations of 100 mM, 200 mM, 600 mM and 800 mM. Two successive eluate fractions were collected at each step and assayed for activity (see section 3.2.7).

3.2.4.2 Q-Sepharose column chromatography

A 5×30cm column was packed with 300 ml of Q-Sepharose (Amersham Biosciences) and equilibrated with two column volumes of 100 mM KH_2PO_4 containing 200 mM KCl, 10%(v/v) ethanol, pH 7.8 (buffer B) at a flow rate of 3 ml/min. 80 ml of cell free extract were then loaded onto the column and unbound protein was washed off using buffer B. J1 nitrilase was

then eluted off the column using 100 mM KH_2PO_4, 400 mM KCl, 10%(v/v) ethanol, pH 7.8, at a flow rate of 2 ml/min. The protein concentration of the eluate was measured at an absorbance of 280 nm and recorded by an in line UV/Vis detector (Gilson Inc, USA), coupled to a PC. Fractions (8.75 ml/tube) were collected using an in line Gilson FC203B fraction collector and selected fractions corresponding to the elution peak were assayed for protein concentration and activity. The Bradford assay (Bradford, 1976) with BSA as standard, was used to determine the protein concentration of selected fractions, while the proteins were separated on a 12.5 % (w/v) denaturing SDS-PAGE using the discontinous buffer system of (Laemmli, 1970) and stained with silver.

3.2.5 Ultrafiltration

The active fractions were pooled and concentrated by ultrafiltration using a 50 ml Amicon stirred cell with a Millipore PM10 10 kDa exclusion membrane under nitrogen gas pressure and stored at 4 ^{o}C. The ultrafiltrate was then filter sterilised using a 0.22 μm filter before gel filtration chromatography.

3.2.6 Gel filtration chromatography

All chromatography was done using columns that were previously calibrated with the usual standards (Bio-Rad Laboratories). The absorbance of the effluent was recorded at 280 nm.

3.2.6.1 First Elution using the Sephacryl S400 and S300 HR columns

A sample of J1 nitrilase (2ml) from the ultrafiltrate was injected into a packed 2.6×100 cm S400 HR column (Amersham Biosciences), equilibrated and eluted using 100 mM KH_2PO_4, 200 mM NaCl, pH 7.8 (buffer C) at a flow rate of 0.8 ml/min. This column has a fractionation range of $2\times10^4 - 8\times10^6$ and the bed volume was ~445 ml. The fractions collected (9 ml/tube) were immediately assayed for protein concentration and activity as discussed in section 3.2.4. The active fractions were separated on reducing 12.5 %(w/v) SDS polyacrylamide gel and bands were visualized by silver staining. This experiment was repeated twice and the results were reproducible.

The active fractions were pooled and subjected to Amicon ultrafiltration. Subsequently, 2 ml of protein was injected into a pre-packed 1.6×60 cm S300 HR column (AP Biotech), previously equilibrated with two column volumes of buffer C at a flow rate of 0.5 ml/min. The column has a fractionation range of 1×10^4-1.4×10^6. Fractions collected (2 ml/tube) were immediately assayed as discussed above. In all cases, fractions with high nitrilase activity were investigated

by negative stain electron microscopy. The purified protein was stored at 4 ^{o}C for CD and substrate characterization studies.

A second preparation of J1 nitrilase was set up and purified as discussed above. However, the purified enzyme was passed twice through the Sephacryl S400 HR column instead of the S300 HR column. Details of this purification are listed in Table 3.2 (see section 3.3.1). The purified enzyme was then investigated by negative staining microscopy by diluting the sample with a specific buffer (see section 3.2.8). Another sample (0.08 mg/ml) was also subjected to CD studies (see section 3.2.10). The purified enzyme was finally stored at 4^{o}C.

3.2.6.2 Elution of one-month old purified J1 nitrilase

An active sample of J1 nitrilase (from the second preparation) kept refrigerated for about 4 weeks was concentrated by ultrafiltration (~3 ml). Subsequently, 500 µl of a filter-sterilized sample (2.18 mg/ml) was injected into a 7.8×30 cm TSK G5000PWXL column (Tosoh Bioscience), pre-equilibrated and eluted with buffer C at a flow rate of 1ml/min. This column has a fractionation range $<1.0\times10^{7}$ Da for globular proteins. Fractions (0.25 ml/tube) corresponding to the elution peak were collected and analysed for protein content and nitrilase activity. The purified protein in the active fractions was then negatively stained and visualized using an electron microscope.

3.2.7 Measurement of enzyme activity

Nitrilase activity was determined by analyzing the release of ammonia using the Bertholet-reaction (Hiller and Van Slyke, 1933), which produces blue-coloured indophenol following the reaction between chloroamine and phenol (Bolleter et al., 1961). Reactions were carried out in 1ml volumes containing 2 µl of enzyme solution, 988 µl of buffer A and 10 µl of 100 mM benzonitrile (dissolved in 1 ml ethanol to increase its solubility), similar to a protocol described by Piotrowski et al., 2001. The reaction was allowed to occur for 1 hour at room temperature, after which 40 µl of phenol-alcohol, 40 µl of sodium nitroprusside and 100 µl of freshly prepared oxidizing solution (1 part NaOCl to 4 parts alkaline complexing agent [10 g sodium citrate, 0.5 g sodium hydroxide made up to 50 ml with distilled water]) were added to the mixture in an eppendorf tube. The reaction mixtures were incubated for 1 hour at room temperature and the absorbance was recorded at 620nm. A standard curve was prepared in the linear range 0.01-0.06 mM using NH_4Cl. One unit of the enzyme is defined as the amount that converts 1 µmole of benzonitrile in one minute to produce an equivalent amount of benzoate and ammonia.

3.2.8 Negative stain electron microscopy

All samples were negatively stained at room temperature except where the purified protein was first incubated at 35 ^{o}C for 30 minutes before staining as reported by Nagasawa et al (2000).

3.2.8.1 Sample preparation

Two drops (10 µl) of distilled water and two drops (10 µl) of 2 % uranyl acetate were prepared on parafilm^{TM}. 4 µl of purified J1 nitrilase solution was then pippetted onto a fresh glow-discharged grid previously coated with a thin carbon support film for 20 seconds under vacuum. Adsorption of sample onto the grid held between metal tweezers was allowed to occur for 20-30 seconds. In order to reduce precipitation between phospate buffer and uranyl acetate, grids were subjected to successive water washes followed by staining with acidic uranyl acetate. At each step, excess sample, wash and stain were blotted off using Whatman filter paper. The grids were air-dried before electron microscopy. The concentration of salt in the buffer (200 mM NaCl) was reduced by dilution with distilled water or specific buffers, including 50 mM TEA, 1 % glutaraldehyde pH 7.9 and 10 mM tris containing 50 mM NaCl, pH 8. Some of the samples were stained in the presence of 10 mM benzonitrile at either room temperature or at 35 ^{o}C for 30 minutes similar to conditions reported by Nagasawa et al (2000).

3.2.8.2 Electron microscopy and scanning

Electron micrographs were taken using the JEOL 1200EX II microscope operating at 120 kV and previously aligned according to standard practice at a nominal magnification of 50000 times. The instrument was corrected for astigmatism at high magnification (150000) and the defocus was zero. Images were recorded under low dose conditions to reduce exposure of specimen to damaging radiation. Micrographs were then digitized at 10 µm using a Leafscan^{TM}45 scanner, giving 2 Å/pixel.

3.2.9 N-terminal sequencing and Mass spectroscopy

Following results from gel filtration chromatography and subsequent gel electrophoresis of one-month old J1 nitrilase (see section 3.3.3 and figure 3.9), the amount of proteolysis at either end in the sequence of J1 nitrilase was investigated. A sample of J1 nitrilase (0.75 mg/ml) was sent to Commonwealth Biotechnologies, Inc (Richmond, Virginia) for N-terminal sequencing and MALDI mass spectral analysis. 100 µl of sample was subjected to 10 cycles of Edman degradation to determine the amino acid sequence. For mass spectrometry, 1 µl of undiluted sample was mixed with 1 µl of matrix (ferulic acid) and then spotted onto a sample plate for mass spectrometry. This sample was desalted to improve the signal.

3.2.10 Circular Dichroism spectroscopy

Circular Dichroism (CD) spectroscopy was used to determine whether J1 nitrilase undergoes structural changes upon binding the substrate benzonitrile. CD measures the differences in absorption of the left- and right-circularly polarized light as it passes through an optically active or chiral sample (Johnson, 1990). This technique is good for determining the folding of a protein's secondary and tertiary structure, conformational stability of a protein under stress (e.g. thermal stability, pH stability and stability to denaturants), comparing the structures of a protein from different sources (e.g. different species, different mutants of same protein) and also determining whether protein-protein interactions alter the conformation of the protein (Johnson, 1990).

CD spectra of a sample of purified J1 nitrilase (~8.5 mg/ml) diluted with buffer A were recorded on a JASCO J-810 spectropolarimeter. Measurements were made using a 0.02cm path length cell over a wavelength range of 190-320 nm. Sample concentrations were approximately 0.8 mg/ml. For each sample, the final spectrum was automatically averaged over 10 accumulations. In addition, spectra for buffer A, 10 mM benzonitrile and a mix of the two were also determined and subtracted from the final spectrum.

3.2.11 Substrate characterization

This part of the experiment was designed similar to studies previously conducted on the native enzyme purified from an extract of isovaleronitrile-induced cells of *Rhodococcus rhodochrous* J1 (Kobayashi et al., 1989). A reaction mixture containing buffer A and 10 mM substrate was prepared in a glass vial followed by the addition of 2 µg of an 8 mg/ml solution of J1 nitrilase. The substrates were dissolved in 1 ml of methanol to increase their solubility. The reactions were stopped after 1 hour (first batch) and after 2 hours (second batch) by the addition of 60 % of 0.1 % Trifluoroacetic acid mixed with 40 % acetonitrile adjusted with concentrated HCl to pH 2 for subsequent HPLC separations using an LC-MS system. The samples were automatically injected into the Walters Integrity HPLC System and eluted through a Chromolith speedrod column at a flow rate of 0.5 ml/min. The eluate was the same acidied solution used to stop the reactions. An in line Waters 486 UV detector monitored the absorbance of aromatic compounds at 215 nm. This experiment was carried out in the CSIR Bio/Chemtek labs in Modderfontein, Johannesburg.

3.3 Results

The purification of J1 nitrilase was based on the method of Nagasawa et al (2000). However, preliminary results showed that some steps in the purification such as ammonium sulphate fractionation and hydrophobic interaction chromatography could be omitted.

3.3.1 Expression and purification

In the early experiment, J1 nitrilase was expressed from 4 litres of culture, but this was later reduced to 1 litre because the host cells expressed abundant recombinant protein, which was more than adequate for structural analysis by single particle methods. The expressed enzyme showed high activity with benzonitrile. The results of the purification are summarized in Table 3.1. No quantitative activity values for the cell free extract and one-month old protein are given. However, these were active in the presence of benzonitrile.

Table 3.1 Purification Table for recombinant J1 nitrilase (from second preparation). Activity for cell free extract was not determined.

Recombinant J1 nitrilase	Volume (ml)	Concentration (mg/ml)	Total Protein (mg)	Total activity (units)	Specific Activity (units/mg)
Cell free extract	80.00	31.72	2537.63		
Q-Sepharose	8.75	13.44	117.61	1757.00	14.94
Sephacryl S400 (1^{st})	5.53	7.58	41.88	785.55	18.76

3.3.2 Anion exchange chromatography

A series of step gradients ranging from 100-800mM KCl applied to a miniature Q-Sepharose column completely eluted the enzyme in the range between 200-600mM (figure 3.1). The bound J1 nitrilase was eluted as a sharp peak using 0.1M KH_2PO_4 containing 400mM KCl and 10%(v/v) ethanol, pH 7.8 (figure 3.2). Reducing SDS PAGE corresponding to the active fractions revealed a characteristic band of about 40kDa and a higher band greater than 55kDa. A preparation from the fraction with the highest activity was negatively stained (as described in section 3.2.8) and investigated by electron microscopy (figure 3.3). This revealed a heterogeneous preparation of protein aggregates. The presence of particles, which resemble GroEL in the electron micrographs correlates with the higher band observed on reducing SDS-PAGE.

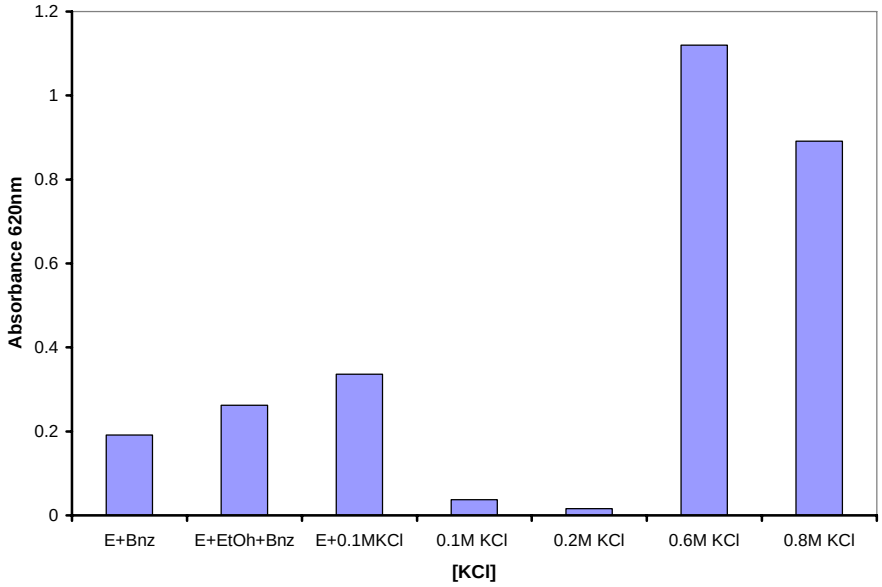


Figure 3.1. Activities of fractions from a miniature Q-Sepharose column loaded with cell free extract (E) of J1 nitrilase. KCl step gradients were applied at the concentrations indicated above. The effect of 10% (v/v) Ethanol (EtOH) on activity in presence of benzonitrile (Bnz) is also shown. Measurements were in duplicate

3.3.3 Gel Filtration

Unlike the method of Nagasawa et al (2000), no benzonitrile was present in the eluant in all the steps of gel filtration chromatography. Initially, the results of the first gel filtration revealed partially separated peaks (figure 3.4). Activity was localized in the smaller peak with a mass of approximately 480 kDa, whereas the higher peak of ~80 kDa had no activity. The molecular weight of the protein was calculated from a previous calibration of the column using only three standards. Reducing SDS-PAGE of active fractions revealed similar bands to those observed in the previous step, whereas negative stain electron microscopy showed heterogeneous protein of varying size (figure 3.5). Subsequent re-elution of a sample from the active peak two days later using the Sephacryl S300 HR column resulted in a single peak. Reducing SDS polyacrylamide gel showed protein common bands of about 40 kDa (figure 3.6). No repartitioning between the 480 and 80 kDa species was observed suggesting the 480 kDa oligomer is stable. The active fraction was investigated by negative staining, which revealed aggregates of various sizes (figure 3.7).

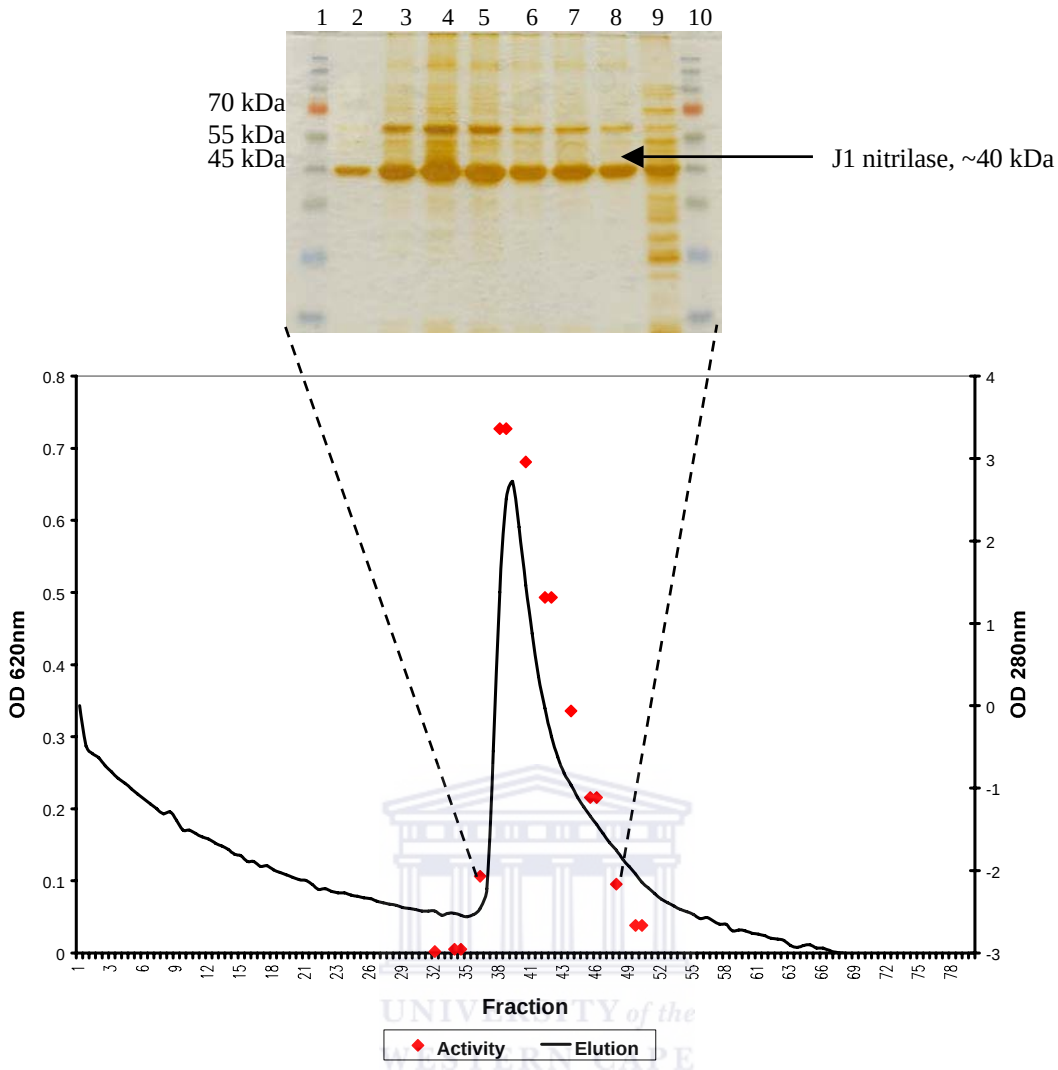


Figure 3.2.Elution of recombinant *R.rhodochrous* J1 nitrilase from the Q-Sepharose ion exchange column using 400mM NaCl. Corresponding reducing SDS polycrylamide gel of the partially purified recombinant *R.rhodochrous* J1 nitrilase is also shown. Lane 1 and 10, molecular weight markers; Lane 2-8, partially purified fractions 36-42 and Lane 9, Cell free extract.

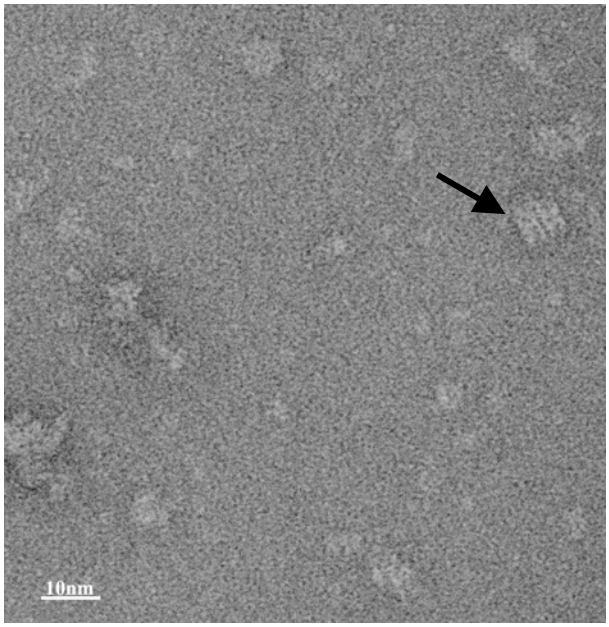


Figure 3.3. Electron micrograph of fraction 38 (buffer B) diluted 10-fold with distilled water. Black arrow points at a side view of a particle with four stripes characteristic of GroEL.

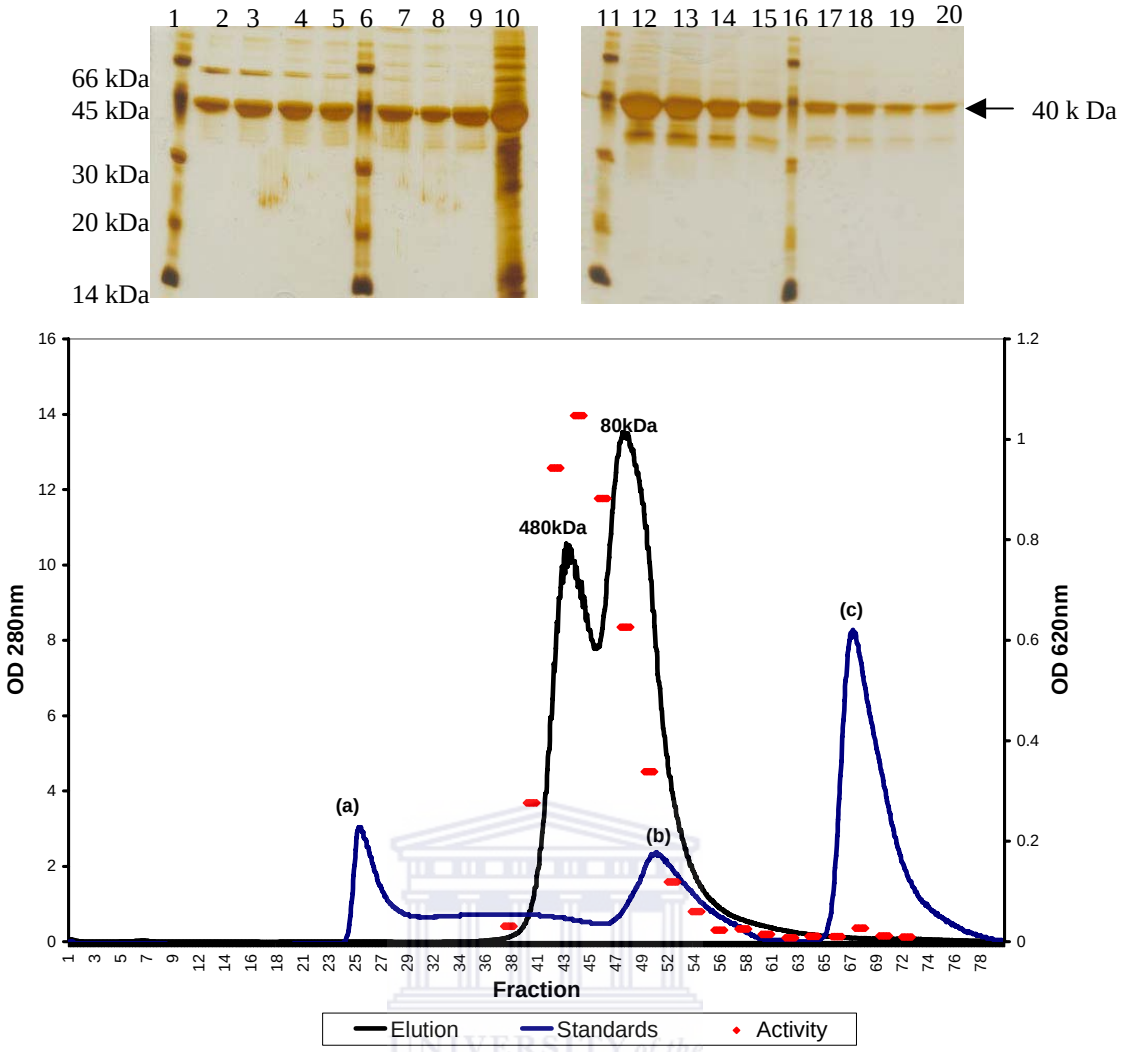


Figure 3.4. First gel filtration profile and corresponding 12.5 % reducing SDS PAGE. High activity is localized in the first peak. OD_{280} of calibration standards [Blue dextran (a), BSA (b), Vitamin B_{12} (c)] is multiplied by a factor of 4. Lanes 1, 6, 11 and 16 are standards; lane 10, Cell free extract; other lanes, fractions 40–66 of J1 nitrilase. Bands below 40 kDa suggest proteolyis.

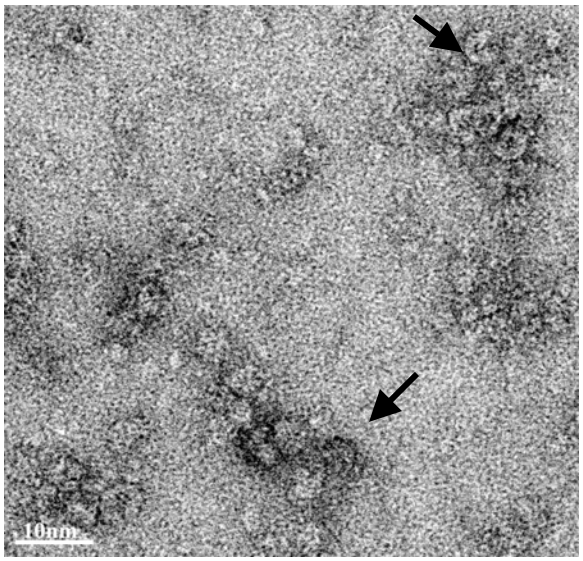


Figure 3.5. Electron micrograph of fraction 46 (see figure 3.6). Black arrow points at protein aggregates of varying size, some are c-shaped.

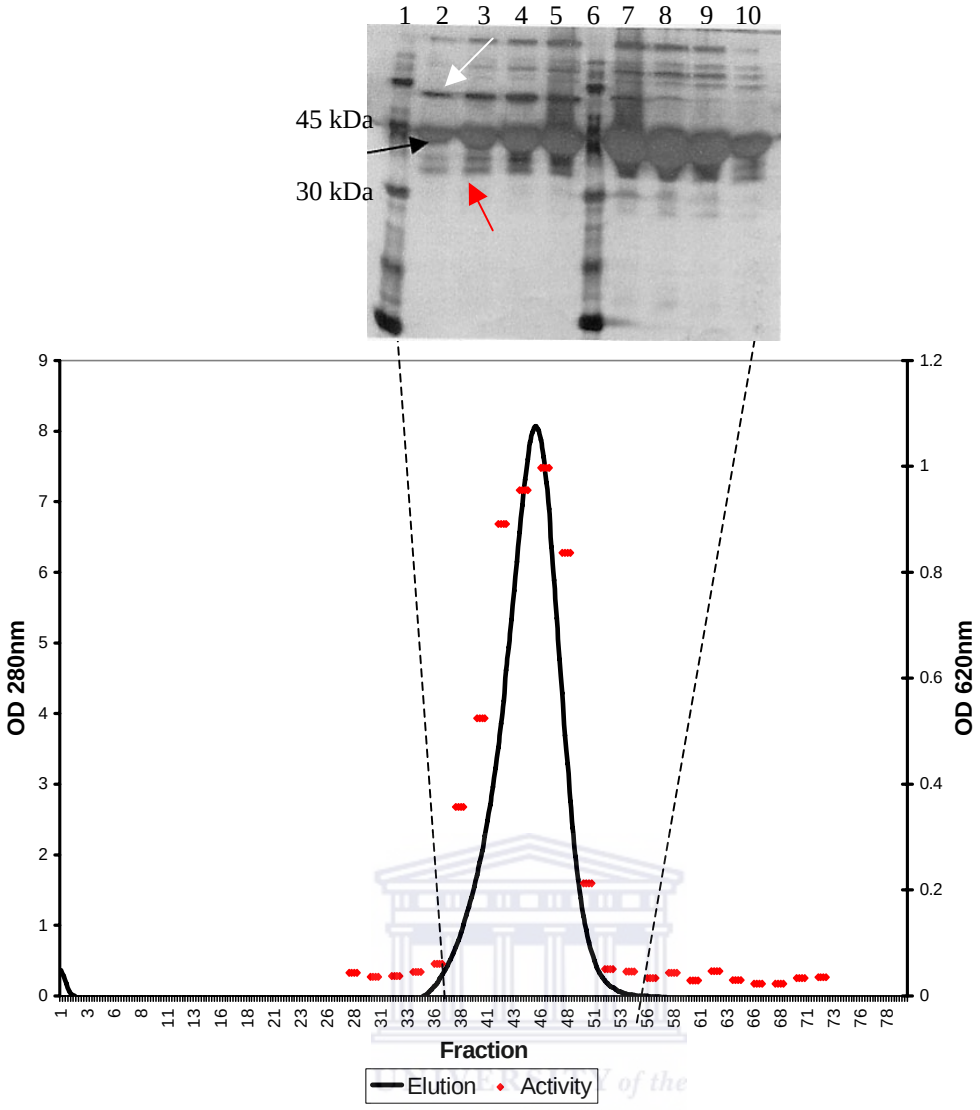


Figure 3.6. The re-elution profile of pooled fractions 41-45 (480 kDa peak) from the elution shown in figure 3.5. Reducing SDS PAGE: black arrow, 40 kDa J1 nitrilase; white arrow, putative GroEL band and red arrow, degraded protein. Lane 1and 6 molecular standards; other lanes, fractions 36-50.

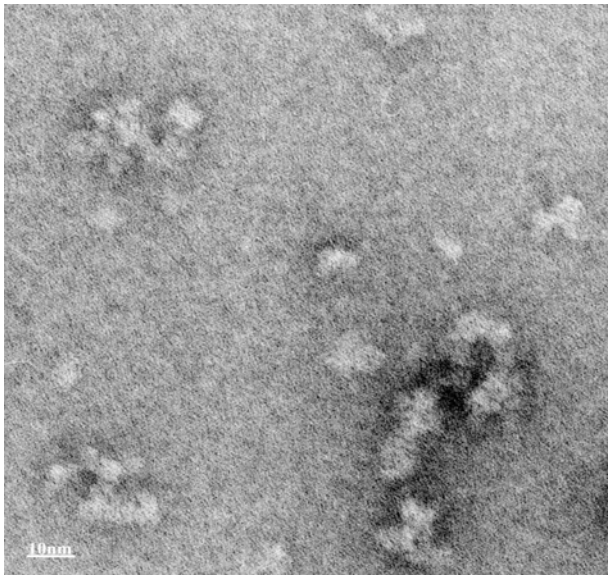


Figure 3.7. Electron micrograph of fraction 46 (see figure 3.6) in 0.01M KH_2PO_4, 0.02 M NaCl pH 7.8. Protein aggregates of varying size can be seen.

The results of gel filtration of the one-month old J1 nitrilase produced a higher leading peak in which activity was localized and smaller trailing peaks with no activity (figure 3.8). Reducing SDS-PAGE showed a single band, which was calculated from a plot based on mobility relative to standards to be approximately 36.5 (±0.6) kDa. This suggests a protein lower in mass compared to that from earlier purification steps because of proteolysis. Negative staining microscopy of the active fraction revealed a homogeneous preparation of enzymatically active long fibres of J1 nitrilase (figure 3.9). A 10-fold dilution with distilled water (to reduce salt concetration) was found to increase the length of the fibres. This may indicate the fibres are maintained by ionic interactions between protein subunits. The active peak eluted close to the void volume and was determined to have a molecular weight greater than 1.5 MDa in comparison to globular standards (figure 3.10).

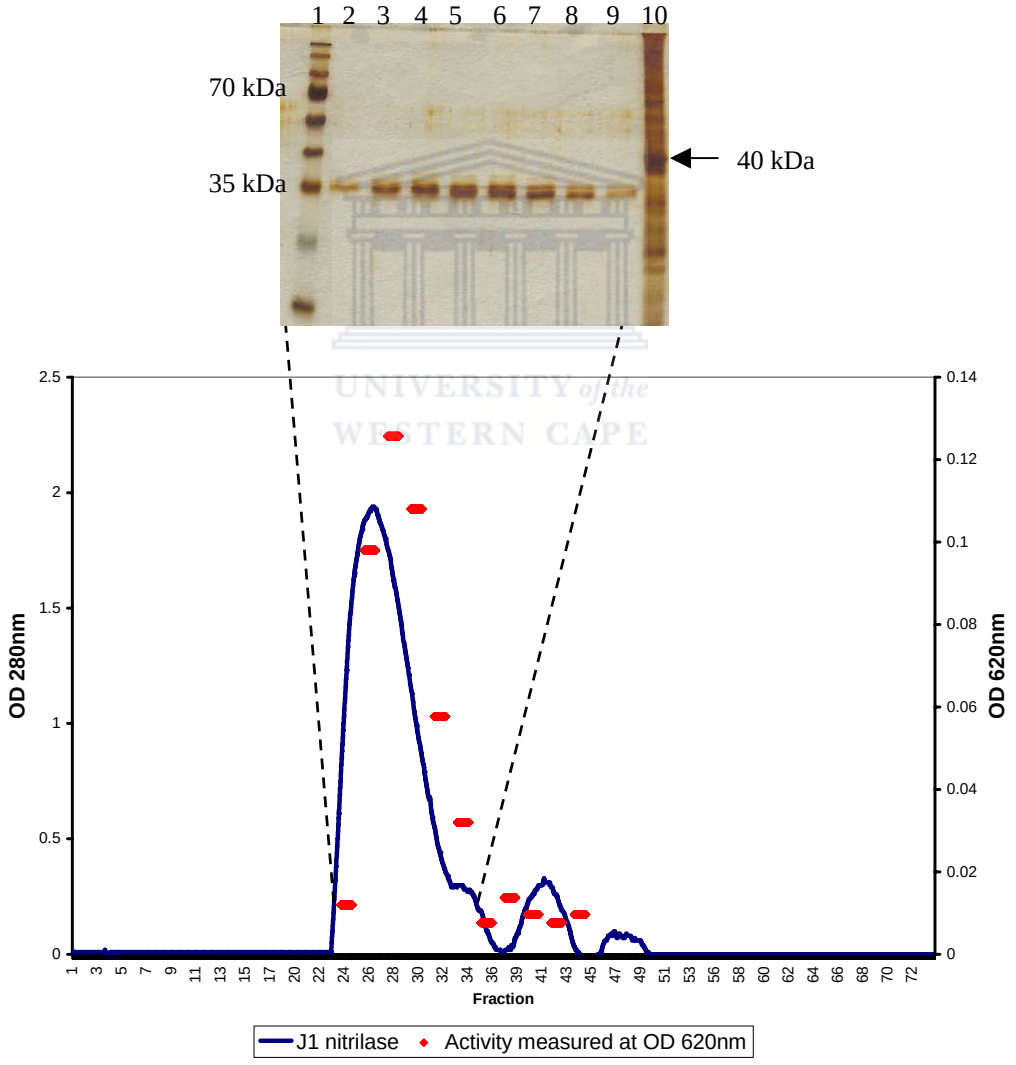


Figure 3.8. Elution of 1-month old J1 nitrilase and corresponding reducing SDS PAGE. Only one band of ~36.6 kDa visible. Lane 1, molecular standards; lane 2-9, fraction 25-24 and lane 10, Cell free extract.

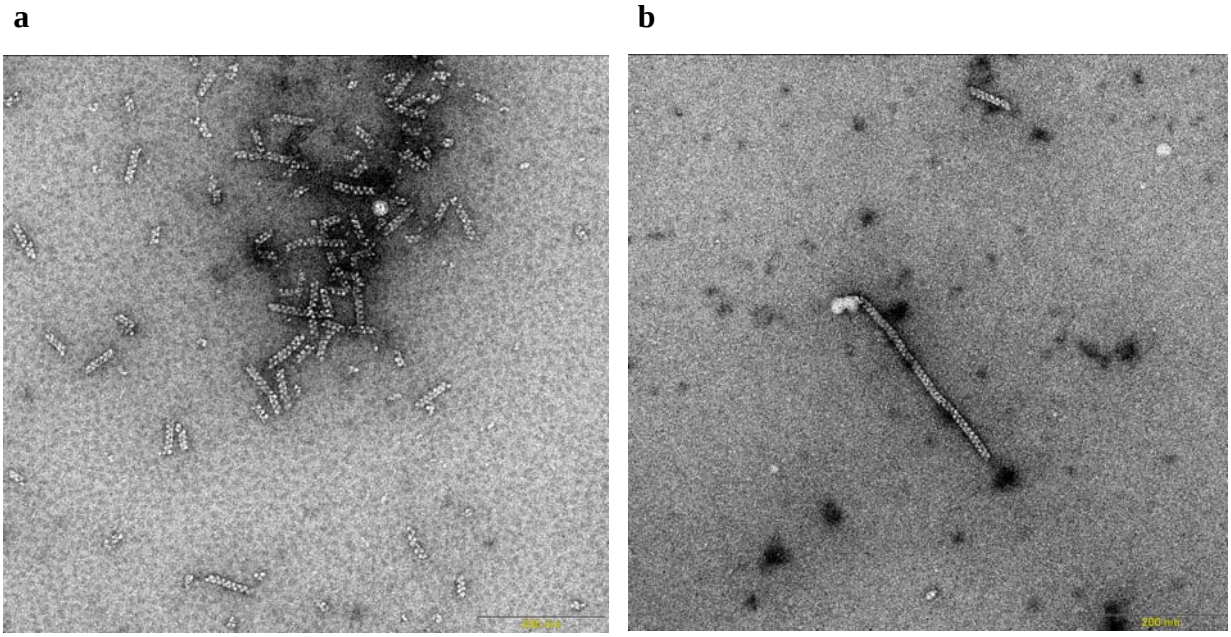


Figure 3.9. Negatively stained fibres of J1 nitrilase from fraction 30 (0.45 mg/ml) buffered in 100 mM KH_2PO_4, 200 mM NaCl pH 7.8. (a) Undiluted sample, short fibres with salt crystals in the background; (b) Same sample diluted 10-fold with distilled water, long fibre. Micrographs were recorded using Leo912 TEM at a magnification of 50000×.

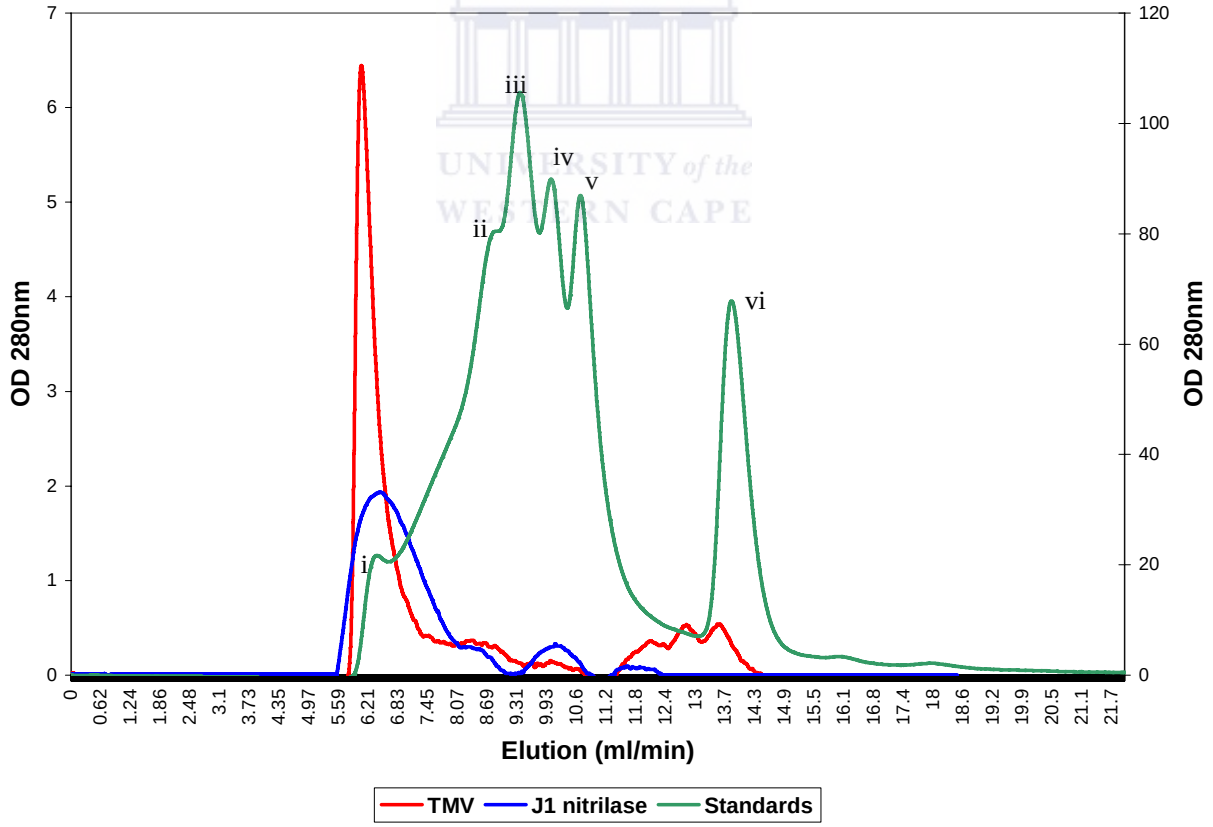


Figure 3.10. Calibration of the elution of J1 nitrilase (figure 3.8). The molecular weight of J1 nitrilase (>1.5 MDa) was determined under similar elution conditions against standards: Tobacco Mosaic Virus (TMV, 34 MDa), (i) Blue Dextran (2 MDa), (ii) Thyroglobulin (670 kDa), (iii) Bovine gamma albumin (158 kDa), (iv) Chicken Ovalbumin (44 kDa), (v) Equine Myoglobin (17 kDa) and (vi) VitaminB_{12} (1.35 kDa).

3.3.4 N-terminal sequencing and Mass spectroscopy

The results of mass spectral analysis showed a single peak of very low intensity with an apparent mass m/z = 36082.2 (figure 3.11). This agrees well within 0.5 kDa with the mass determined from SDS PAGE, but different from that of the cell free extract. The molecular weight of J1 nitrilase calculated from its sequence (GenBank accession no. D11425) is 40.189 kDa.

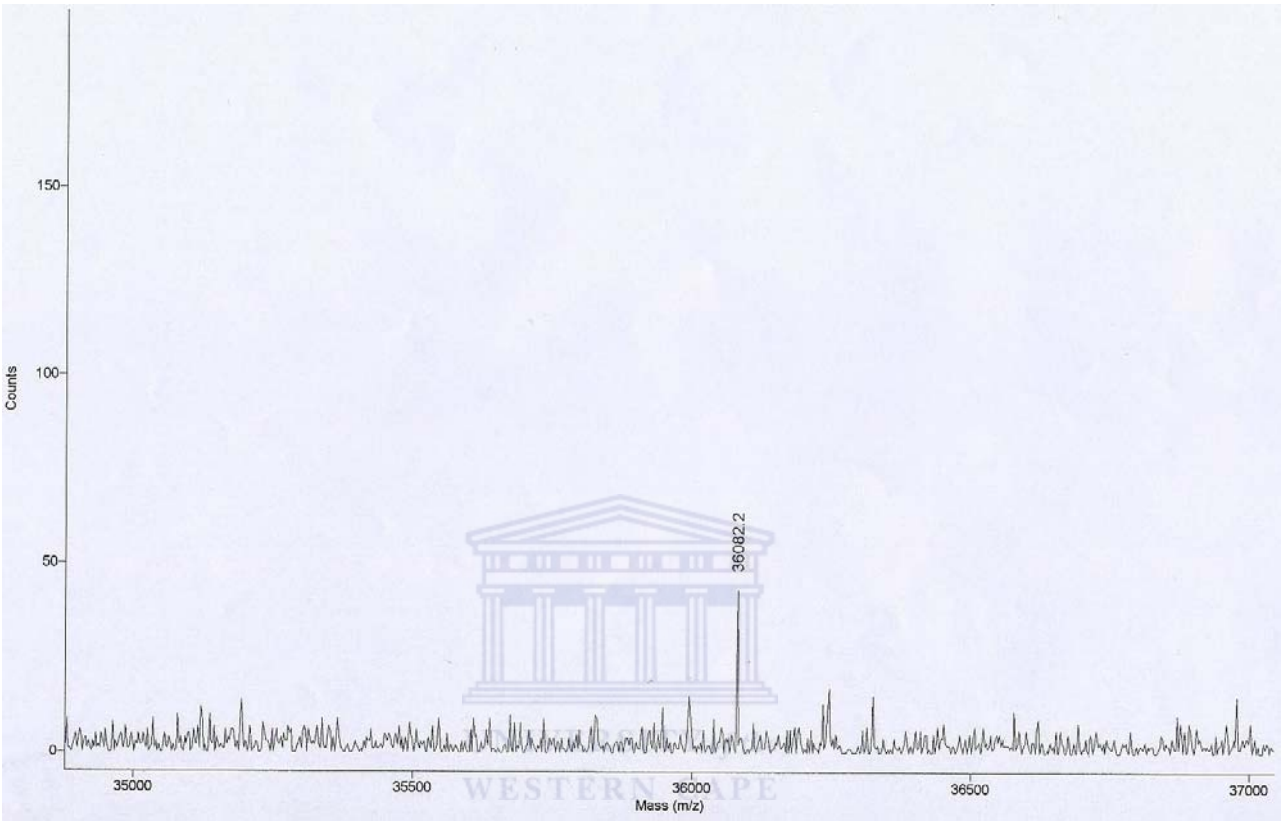


Figure 3.11. MALDI-TOF mass spectrum of a desalted sample of the purified nitrilase from *R.rhodochrous* J1. The matrix used for mass spectrometry is ferulic acid.

N-terminal sequencing was difficult but it was perfomed by two different sequencing sources. Sequencing of the first 10 amino acids, namely MVEYTNTFKV resulted in the assignment of the following residues: M 0 [F,E] Y 0 N 0 K K 0. The residues indicated by a zero were not detected while multiple residues (indicated in square brackets) were detected during the process. Overall 6 of the 10 residues at the N-terminal were assigned thus confirming that the long fibres (see figure 3.9) are definitely J1 nitrilase. Proteolysis must therefore occur in the C terminal. The sharp band on both SDS-PAGE and MALDI spectroscopy indicate that there is specific cleavage of about 39 amino acids in its C terminal to equal a mass of ~36 kDa.

3.3.5 CD spectroscopy

The purpose of the CD experiment was to see whether the putative conformational changes leading to activation upon oligomerization could be detected spectroscopically. The CD spectra of J1 nitrilase in the absence of benzonitrile are shown in figure 3.12. In the far-UV spectral region (190-250nm), the chromophore is the peptide bond and this spectrum is typical of a well-folded protein.

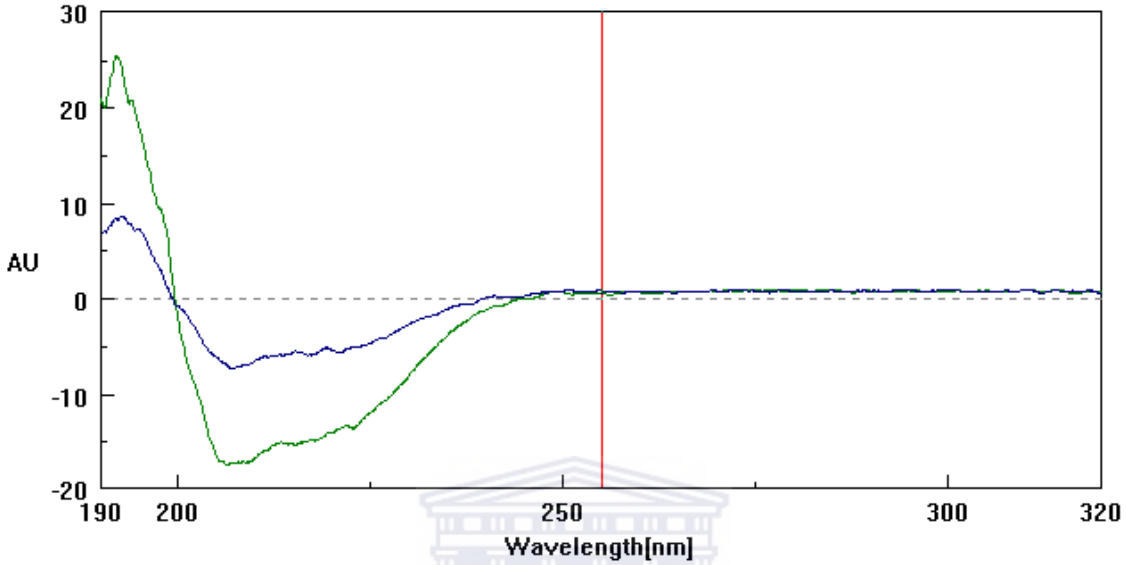


Figure 3.12. CD spectra of purified nitrilase from *R.rhodochrous* J1 using a path length cell of 0.20 cm. The protein concentration was 8.5 mg/ml, which was diluted 10 times (blue curve) and 5 times (green curve). AU = Absorbance Units.

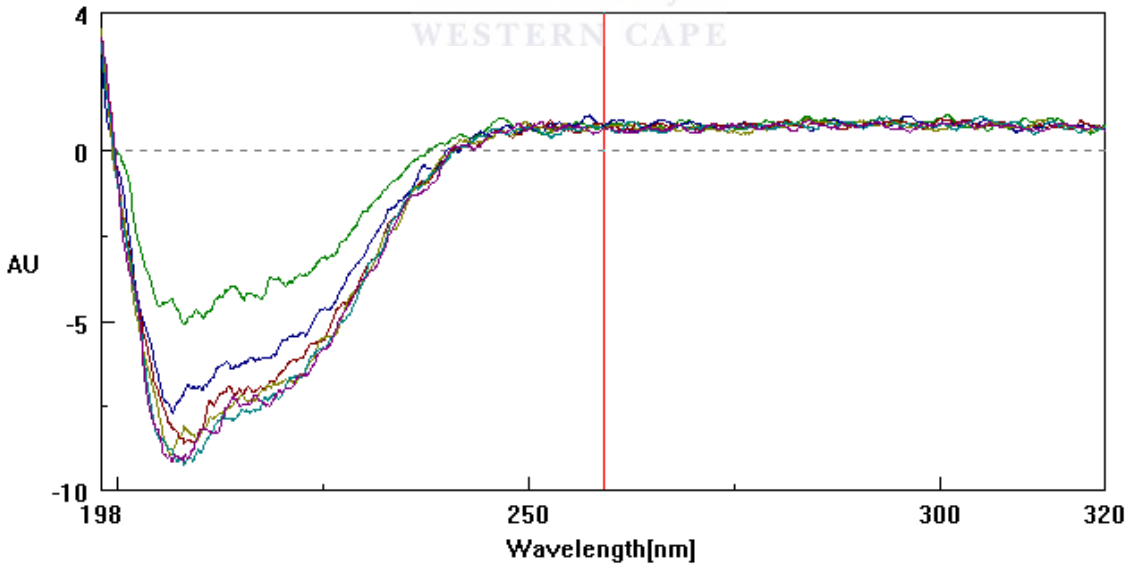


Figure 3.14. CD spectra of *R.rhodochrous* J1 nitrilase in the absence of substrate benzonitrile (teal line) and after addition of benzonitrile (after 5 mins (green), 15-20 mins(blue), >25 mins (brown), >30 mins (pale green), >45 mins (Cyan) and 55 mins (Purple)). There is a structural change on addition of benzonitrile and this change reverses as substrate is consumed.

A shift in spectra was observed upon addition of benzonitrile into the protein solution (figure3.13). This shift arises when benzonitrile is bound causing a change in overall structure due to protein-substrate and/or protein-protein interactions. However, this change reverses as the substrate is consumed until the spectrum resembles that obtained in the absence of benzonitrile.

3.3.6 Substrate characterization

The ability of the recombinant enzyme to catalyze the hydrolysis of various substrates was examined (Table 3.2).

Table 3.2. Conversion of aromatic and aliphatic nitriles by the nitrilase from *Rhodococcus rhodochrous* J1. Degree of conversion of 10mM substrate was assessed after two hours by HPLC. – No conversion, asterisk (*) denotes substrates whose conversion may need further analysis since these have previously not been reported.

Substrate	**% remaining after 2 hours**
Benzonitrile	0
4-chlorobenzonitrile	0
4-cyanopyridine	0
2-phenylyglycinonitrile	50
3,5-dibromo-4-hydroxybenzonitrile (Bromoxynil)	100
Benzilidenemalonitrile*	33
Fumaronitrile	60
α-methylbenzylcyanide	100
(R)- α-acetoxyphenylacetonitrile*	9
3-hydroxy-4-phenylvaleronitrile	100
3-phenylpropionitrile (hydrocinnamonitrile)	88
Propionitrile	-
Acrylonitrile	100
Benzylcyanide (phenylacetonitrile)	100
Mandelonitrile	50
Adiponitrile	-
3-hydroxy-3-phenylpropionitrile	100

After 1 hour, 87% of the benzonitrile and 70% of the 4-chorobenzonitrile was converted, while no conversion for the other substrates was observed possibly due to short reaction times (data not shown). The nitriles converted after 2 hours include benzonitrile, 4-chlorobenzonitrile, 4-cyanopyridine (all three to completion), mandelonitrile, 3-phenylpropionitrile

(hydrocinnamonitrile), (±)-2-phenylglycinonitrile and fumaronitrile. Aliphatic nitriles (e.g. acrylonitrile, propionitrile) and 3,5-di-bromo-4-hydroxybenzonitrile were not converted at all.

3.3.7 Negative staining microscopy

The re-elution of active samples of purified protein revealed single particles of J1 nitrilase with a size range between 10-20 nm (see chapter 4). SDS PAGE showed a single band of about 40 kDa, which indicates purity at the protein level. However, negative stain EM revealed heterogeneity at the structural level. The 1-month old J1 nitrilase was observed to homogeneous and reconstruction of its 3D map is discussed in chapter 5.

3.4 Discussion

The recombinant nitrilase from *R.rhodochrous* J1 has been purified to homogeneity from abundantly expressing *E.coli* cells. The activity of the purified enzyme is comparable to that obtained from previous work (Kobayashi et al., 1989). The activity of the filamentous form was not determined due to low enzyme concentration. However, this form is expected to exhibit higher activity because more active sites are available to the substrate than those present in a short homo-oligomer. The purified enzyme showed only one band on SDS PAGE with a subunit atomic mass of about 40 kDa. This mass corresponds to molecular weight previously reported in other Rhodococcal nitrilases (Harper 1985; Kobayashi et al., 1989, 1992b; Nagasawa et al., 2000; Stevenson et al., 1992) and certainly, is common among other nitrilases in the range 35-41 kDa. The enzyme was eluted in the absence of benzonitrile as two distinct and apparently stable forms of 480 kDa and 80 kDa respectively. The 480 kDa form was active whereas the 80 kDa form was inactive. This observation agrees with previous work where the native enzyme was purified as an inactive dimer, which assembles into a higher form and subsequently acquires activity (Harper, 1985; Nagasawa et al., 2000; Stevenson et al., 1992). However, previous authors have not reported a higher oligomeric structure in the absence of substrate. One possible explanation is that the 480 kDa complex of the recombinant enzyme was not isolatable from the native organism. Although reducing SDS PAGE showed one characteristic nitrilase band, negative stain electron microscopy revealed the recombinant enzyme forms previously unobserved protein aggregates of varying shape and size, whereas intermediate oligomeric nitrilase forms were not detected in the native enzyme using techniques of gel filtration, ultracentrifugation and light scattering (Nagasawa et al, 2000).

A common c-shaped form of J1 nitrilase is observed in almost all electron micrographs. This enzyme has significant insertions in its sequence relative to solved structures, which have

previously been shown to contribute intersubunit contacts at the C surface responsible for spiral formation in the cyanide dihydratases from *B.pumilus* C1 and *P.stutzeri* AK61 (Sewell et al., 2003, 2005). The association of dimers across the C surface is prerequisite towards the formation of higher oligomers. However, the c-shaped form cannot form a helix (or at least one turn of the spiral) because of possible steric hindrances arising from 53 extra residues in its C terminal relative to the nitrilase atomic homologues.

CD spectroscopy showed there is a change in spectra upon addition of benzonitrile into the active protein solution, and that this change is reversed as substrate is consumed. This suggests the enzyme undergoes a conformational change in the presence of substrate.

Subsequent chromatographic re-elution of one-month-old purified enzyme led to the identification of a novel form - long enzymatically active fibres of J1 nitrilase. Reducing SDS PAGE and mass spectroscopy showed it contained a single band of about 36 kDa (±0.6 kDa) as a result of proteolyisis. N-terminal sequencing confirmed it is J1 nitrilase and that no residues are cleaved from this end. This suggests there is specific cleavage in its C terminal as a result of proteolysis. The 480 kDa state of J1 nitrilase has a long C terminal in which steric hindrances prevent the formation of a regular helix.

Enzyme association among previously studied Rhodococcal nitrilases is attributed to increased hydrophobic interactions due to high salt concentration and organic solvents, presence of aromatic substrates (benzonitrile, 4-chlorobenzonitrile, phenylacetonitrile), increased temperatures and enzyme concentration (Harper, 1985, Nagasawa et al., 2000; Stevenson et al., 1992). Similarly, the recombinant enzyme was observed to form varying oligomers in high salt concentration (200mM NaCl) by electron microscopy. However, long fibre formation is observed in low salt concentrations due to increased ionic interactions, which may be required to stabilize the observed helical arrangement. The regular long fibres are more stable, robust and would exhibit high specific activity on a wide range of substrates than the shorter spiral form. This will be determined in the immediate future.

The recombinant nitrilase of *R.rhodochrous* J1 exhibits a substrate selectivity profile similar to the purified enzyme from the native organism. The enzyme prefers aromatic substrates and to a lesser extent long chain acylaliphatic nitriles. However, short chain aliphatics (propionitrile, acrylonitrile) are not converted which agrees with previously published work (Kobayashi et al., 1999). Arylaliphatic nitriles substituted at the α-position (mandelonitrile, 2-

phenylpropionitrile, 2-phenylglycinnonitrile) were generally not hydrolysed, while 3-phenylpropionitrile (a long chain aliphatic) was hydrolysed to a less extent possibly due to inhibition by the bulky substituent. No activity was detected with 3-hydroxy-3-phenylpropionitrile, 3-hydroxy-4-phenylvaleronitrile, phenylacetonitrile and with 3,5-dibromo-4-hydroxybenzonitrile (Bromoxynil, a herbicide with two halogen atoms at the meta positions). The subsituents on the ring are sterically hindered and thus inhibitory to susbstrate hydrolysis by the enzyme. The results suggests that the presence and position of the aromatic ring including the attached alkyl side chain play a role on the extent of substrate hydrolysis. It would be interesting to study the substrate profile of the long enzymatic fibres with a whole range of substrates in future.

Previous studies have shown that the purified enzyme exhibits a narrower substrate range than resting cells of Rhodococcus strains (Kobayashi et al., 1989). Furthermore, cell free nitrilases are known to be less stable than whole cell biocatalysts, are prone to substrate or product inhibition and have low conversion rates due to insolubility of the substrate in water. These drawbacks limit the commercial application of nitrile-converting enzymes. Nevertheless, there are many great opportunities for use of J1 nitrilase in the manufacture of stereochemically pure compounds such as (R)-mandelic acid and nicotinic acid (Mathew et al., 1998).

3.5 Conclusion

The purification protocol presented in this chapter is reproducible. It has been established that the recombinant enzyme occurs as two distinct and apparently stable oligomeric states. The 480 kDa state is active whereas the 80 kDa state is inactive. This observation differs from previously published work (Nagasawa et al., 2000). Negative stain electron microscopy revealed the 480 kDa state occurs as c-shaped oligomers, which are unable to form longer spirals due to steric hindrances in its C terminal. Loss of part of the C terminal (about 39 amino acids) leads to the formation of an active helix, which is stabilized by salt bridges. This helix has potential biotechnological applications. More detailed analysis of these oligomeric states will be conducted in the immediate future.

CHAPTER 4

ANALYSIS OF IMAGES OF THE 480kDA PARTICLES OF *RHODOCOCCUS RHODOCHROUS* J1 NITRILASE

4.1 Introduction

Nitrilases are found in all phyla. These enzymes are expressed widely in prokaryotes and eukaryotes. Despite the number of known sequences (>200), very little is known of their 3D structure. Previous 3D EM studies (Jandhyala et al., 2003; Sewell et al., 2002, 2003, 2005) revealed microbial nitrilases form higher oligomers as a result of interactions not present in the nitrilase homologues (1ems, 1erz, 1f89 and 1j31), which have been determined to atomic resolution. Furthermore, microbial nitrilases often have more residues in their C terminal compared to the crystal structures. The frequent presence of the insertions responsible for interactions in the spirals suggests the formation of higher oligomers is a common phenomenon. In spite of these differences, enzymes in the nitrilase superfamily share a common fold, putative catalytic site and a structurally conserved core of beta sheets with divergent N- and C-terminal domains (Brenner, 2002). The Rhodococcal nitrilases are widely studied among members in the superfamily. The native enzyme was reported to exist as an inactive dimer, which assembles to form an active higher oligomer in the presence of substrate (Harper, 1985; Nagasawa et al., 2000; Stevenson et al., 1992). Even though techniques of gel filtration, light scattering and sedimentation revealed the enzyme complex has 10 – 12 subunits, their quaternary structure is not known. A combination of negative stain EM and 3D image processing revealed the shape, symmetry and the basis of quaternary spiral formation in the cyanide dihydratase from *P.stutzeri* AK61 (Sewell et al., 2003). The spiral was left-handed with ~5-fold symmetry along its vertical axis. At present, 3D EM is the appropriate technique for studying the structure of the enzyme complex among nitrilases. Size exclusion chromatography revealed recombinant J1 nitrilase elutes as an active 480 kDa state suggesting a dodecamer. While SDS PAGE showed a single band of ~40 kDa, negative stain EM revealed heterogeneous population of distinct particles. This is the first time a Rhodococal nitrilase has been visualized by EM. Analysis of these particles using single particle methods is described in this chapter.

4.2 Review of single particle methods

Over the past three decades, EM and image reconstruction has proven to be a versatile tool to study the structure of proteins and macromolecular complexes. Most functions in the cell are not carried out by single molecules but by 'molecular machines' containing multiple subunits with specific functions (Nogales and Grigorieff, 2001). The size of these complexes makes them inaccessible to structure determination by X-ray crystallography because purification yields insufficient amounts for crystallization and phasing large complexes is difficult. Current research efforts are focused towards determining the structures of individual subunits and domains that define a complex. However, information in the context of a fully functional assembly may be lost. 3D EM provides information on the overall shape, symmetry, assembly process, distribution of subunits and can reveal changes in conformation within the assembly that define function. Recent developments in EM techniques and computer technology allow visualization at a resolution that enables secondary structure to be seen. This is more true for symmetric particles such as viruses than for asymmetric particles such as the ribosome, because the effective signal-to-noise ratio (SNR) is enhanced by a factor of its symmetry (Jiang and Ludtke, 2005). The structure of GroEL (with C7 symmetry) at 6 Å resolution represents the highest achieved resolution to date using single particle methods (Ludtke et al., 2004). At such resolution, secondary structural elements such as beta sheets and alpha helices can be located in the 3D map. Briefly, single particle methods generally involve the steps described below.

The preparation of the specimen under investigation: This is done either by negative staining or cryo-electron microscopy (cryo-EM), that is, the vitrification of samples in liquid ethane at liquid nitrogen temperatures. This process is necessary in order to prevent structural collapse upon dehydration or adsorption to the carbon-coated grid, improve the specimen contrast and also reduce radiation damage. Cryo-EM is desirable because of the high resolution achieved. Vitrification induces particles to adopt random orientations on the grid, which can be exploited using the angular reconstitution method (van Heel, 1987) to generate a 3D reconstruction, particularly for a homogeneous sample. The specimen is preserved close to its native environment without structural distortion. Unfortunately, images usually have poor SNR (arising from low dose exposures and background noise) making this technique unsuitable for the study of small particles (<250 kDa), whereas random particle orientations pose a drawback for heterogeneous samples (Ohi et al., 2004). The technique of negative stain EM is easy and is best suited for the study of heterogeneous particles because of its good SNR or high contrast in the images. Upon adsorption to the grid, particles may adopt one or a limited number of preferred

orientations. A major drawback of negative stain EM is structural distortion that is often observed in the resulting 3D maps due to flattening (caused by heavy metal stain) and incomplete stain embedding.

Visualization and collection of EM images: Areas with an even distribution of stain or thin layer of ice and with uniformly spaced particles (not too close to each other) are located on the grid and subsequently photographed on film or recorded by a charged-coupled device (CCD) detector (Tao and Zhang, 2000). Grids are viewed under a transmission electron microscope operating at a minimum accelerating voltage of 80 kV. The sample is normally imaged under conditions of low-dose to minimize beam damage to the specimen. Depending on the desired resolution of the image, parameters such as accelerating voltage, magnification, electron dose and defocus can be varied (Baker and Johnson, 1997).

Iterative alignment and classification methods: This is the most CPU intensive part of single-particle methods. The raw images need to be approximately aligned against references before classification. During alignment, the orientation parameters (single in-plane rotation angle, x and y shifts) between the different views of the particle are determined. The particles are then centered using a reference-free strategy in which particles are averaged together in order to improve the SNR for subsequent classification. Particles with similar views are grouped together as belonging to the same class using either multivariate statistical analysis (MSA) approaches (van Heel and Frank, 1981) or rotationally invariant K-means clustering (Penczek et al., 1996). Alignment and classification need only be iterated a few times to bring the particles into register in order to generate consistent groups and corresponding averages with improved SNR. The class averages are then used to determine the relationships between different views of the particles (orientation parameters) using various approaches (Radermacher, 1987, 1988; van Heel, 1987; Penczek et al., 1994), which are then used to generate a 3D structure by back projection.

For a comprehensive review of single particle reconstruction, the reader is referred to van Heel et al (2000), Ruprecht and Nield (2001) and Frank (2002).

4.3 Methods

4.3.1 Electron microscopy

A sample of purified J1 nitrilase was prepared for negative stain EM as discussed in chapter 3. All images were recorded on Kodak SO-163 film in low-dose mode. Good quality micrographs

were digitized with a pixel size of 10 μm using a Leafscan45, giving 2 Å per pixel at the specimen level.

4.3.2 Image preprocessing

Particles were manually picked and extracted in 160 × 160 pixel boxes and later binned by a factor of two using SPIDER routines (Frank et al., 1996). All processing was executed using the program SPIDER, running on a Linux Box with 2.8 GHz Xeon processors and 4 GB of RAM. Particles were band-pass filtered between $(15\ Å)^{-1}$ and $(240\ Å)^{-1}$ in order to ensure the success of subsequent alignment steps. The high-pass truncation filter suppresses disturbing low frequencies representing variations in the average densities arising from the amount and uniformity of staining, whereas the low-pass Gaussian filter removes noise at high spatial frequencies (Ruprecht and Nield, 2001). While the low-frequency components are irrelevant during structural interpretation, the high spatial frequencies contain fine details about structure and may be introduced at a later stage of processing (van Heel et al., 2000). After filtering, particles were normalized to an average density of zero and a standard deviation of one. A soft-edge circular mask with a diameter greater than the maximum particle dimension but less than the image box was imposed in order to minimize the influence of neighbouring particles or background noise, which may interfere with the alignment using cross-correlation procedures. The images were then stacked together and served as a starting point for subsequent alignment and classification steps.

4.3.3 Iterative alignment and classification of particles

The image stack was translationally and rotationally centered using an iterative reference-free approach (Penczek et al., 1992) implemented in SPIDER. Initially, particle parameters were determined by aligning them against an average calculated by summing over the entire data set and in subsequent alignment cycles, according to the center of gravity of their global average. The centered images were then classified into 100 classes using rotationally invariant K-means clustering (http://www.wadsworth.org/spider_doc/spider/docs/man/apca.html). This process generated class averages from which a representative set was selected to serve as references for multi-reference alignment (MRA). During MRA, the raw images are cross-correlated with the references and assigned to the reference that yielded the highest correlation coefficient. The images were then re-aligned and re-classified and the process iterated four times. At the end of each cycle, classes and corresponding averages were visualized using WEB (Frank et al., 1996). Particles that did not match class members or those with poor coefficients of correlation were deleted. Reference-free alignment was performed on a few unique classes to check for self-

consistency and also improve the SNR of the class average. Classes with a mixed population of members were re-classified and the process iterated. A summary of the iterative alignment and classification scheme is shown in figure 4.1.

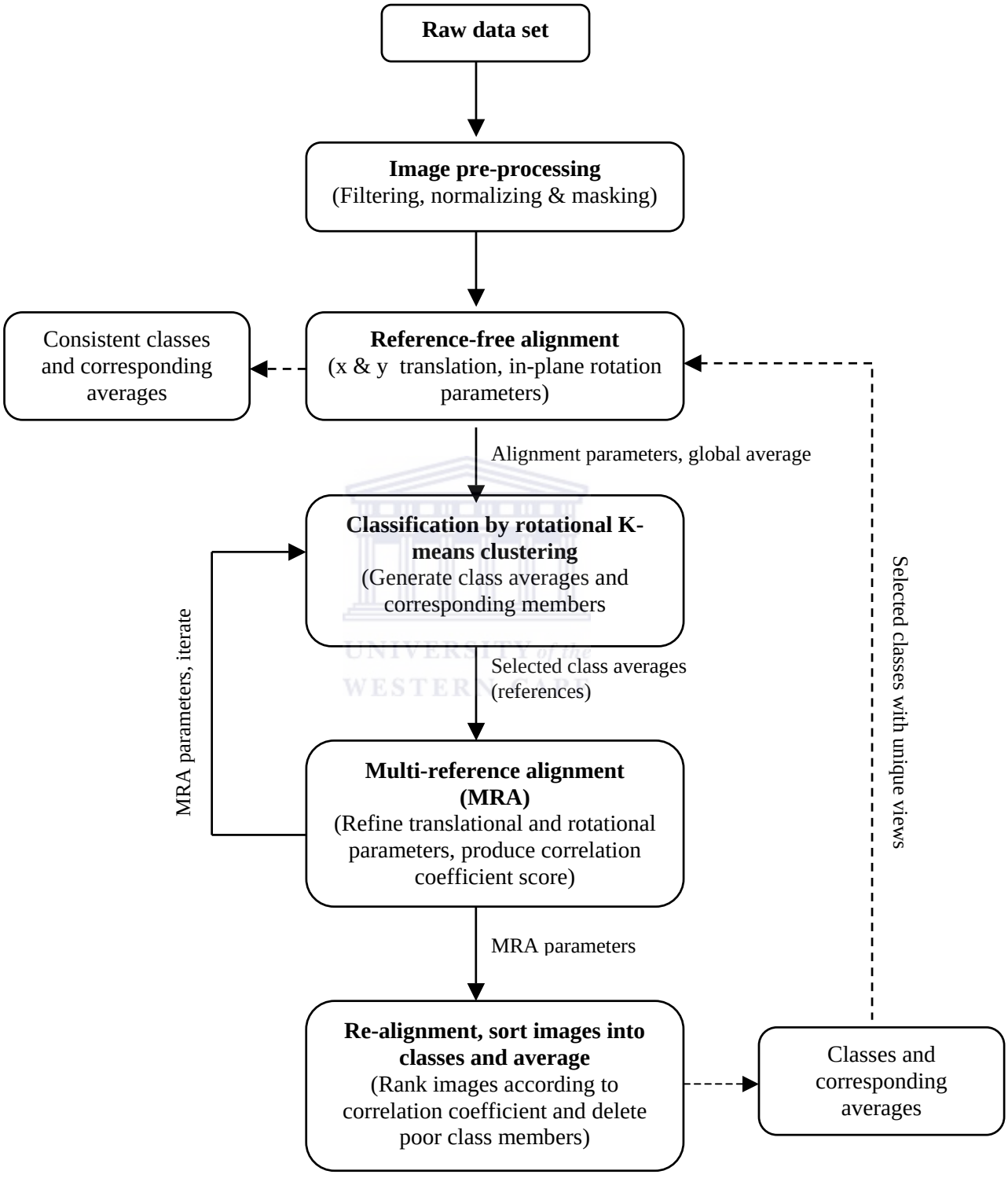


Figure 4.1. The data processing scheme used to align and classify the images.

A 3D reconstruction of the 480 kDa state of J1 nitrilase was not pursued because of firstly, sample heterogeneity and secondly, purification of one-month old J1 nitrilase revealed a homogeneous preparation of long regular fibres whose structure was reconstructed using single particle methods (see chapter 5).

4.4 Results

4.4.1 Electron microscopy

Negative stain EM of the active 480 kDa state of J1 nitrilase revealed a heterogeneous distribution of distinct particles, which could clearly be identified in the low dose micrographs. The size of the particles was observed to vary between 140-260 Å.

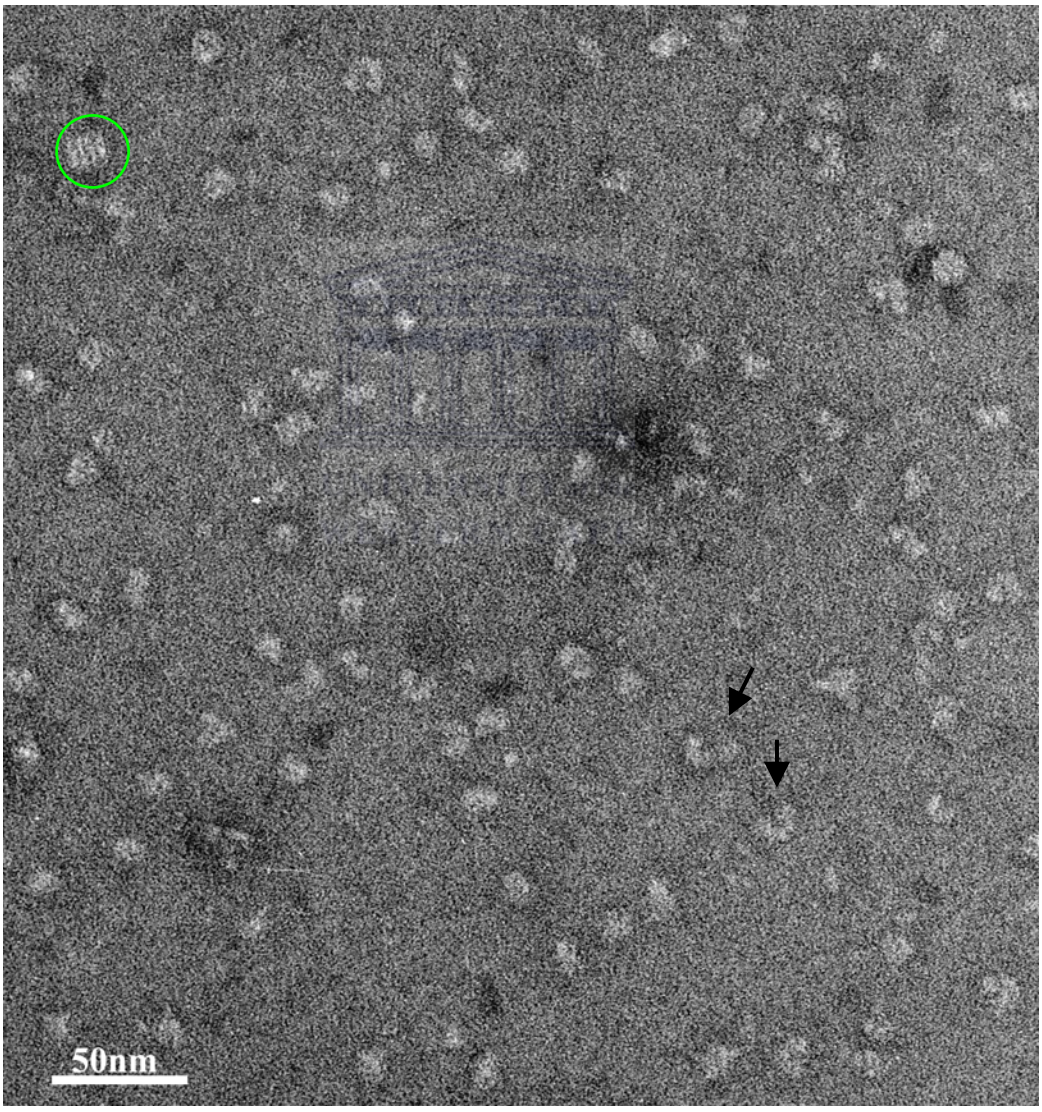


Figure 4.2. A low dose negatively stained electron micrograph of recombinant *R.rhodochrous* J1 nitrilase. Various forms of J1 nitrilase can be seen with possible slight contamination from GroEL (green circle). The c-shaped particles are also highlighted (black arrows). Micrograph was recorded at a magnification of 50000× and sampling is 2 Å/pixel.

4.4.2 Image pre-processing

Well-separated particles (11759) were manually selected from 74 micrographs using the program Ximdisp (Smith, 1999). The images were windowed using SPIDER routines, visualized using WEB (Frank et al., 1996) and averaged in 2 × 2 pixels giving a box size of 80 × 80 pixels and a sampling of 4 Å/pixel. An example of preprocessed raw images of varying shape and size is shown in figure 4.3 (a) and (b).

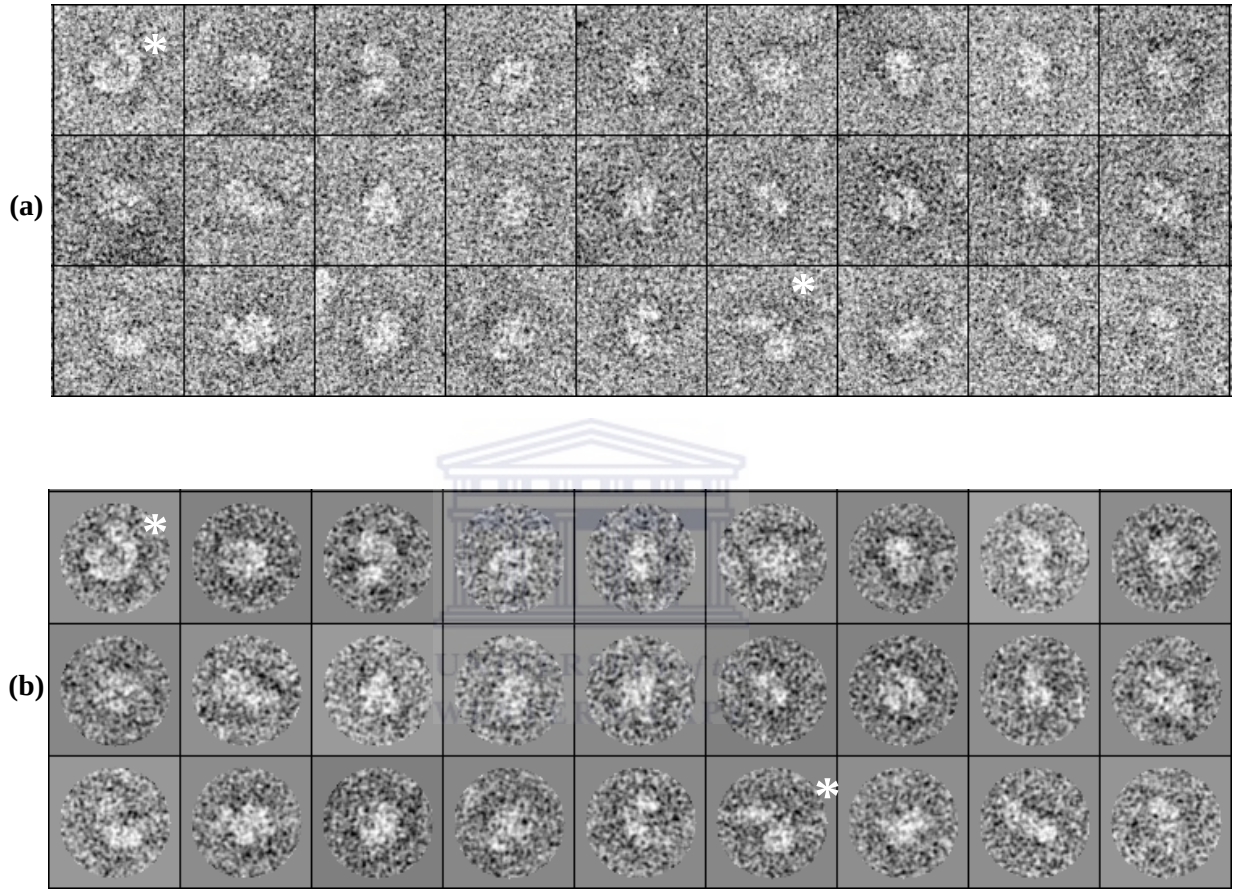


Figure 4.3. Distinct particles of the negatively stained 480 kDa (from the peak in figure 3.5) *R.rhodochrous* J1 nitrilase. (**a**). Representative images from a gallery of selected particles. **(b**). Corresponding set of filtered, normalized and circularly masked particles. Interesting c-shaped particles of varying size (discussed later in the text) are highlighted with asterisks. Box size is 80 × 80 pixels and the scale is 4 Å/pixel.

4.4.3 2D alignment and classification

The preprocessed particles were subjected to four rounds of alignment and classification specifying 100 output classes. The raw images were iteratively centered against a changing global average followed by rotational alignment and classification using SPIDER routines. A few class averages (25) were selected and used to refine particle parameters by MRA. Subsequent re-alignment, sorting and classification revealed class averages representing unique particle

corresponding to orientations of distinct particles on the carbon-coated grid. As can be seen in Figure 4.1(a), there is great variation in size and shape. Poorly correlated members within a selected class were deleted and the remainder were subjected to iterative reference-free alignment to improve the SNR of the class average.

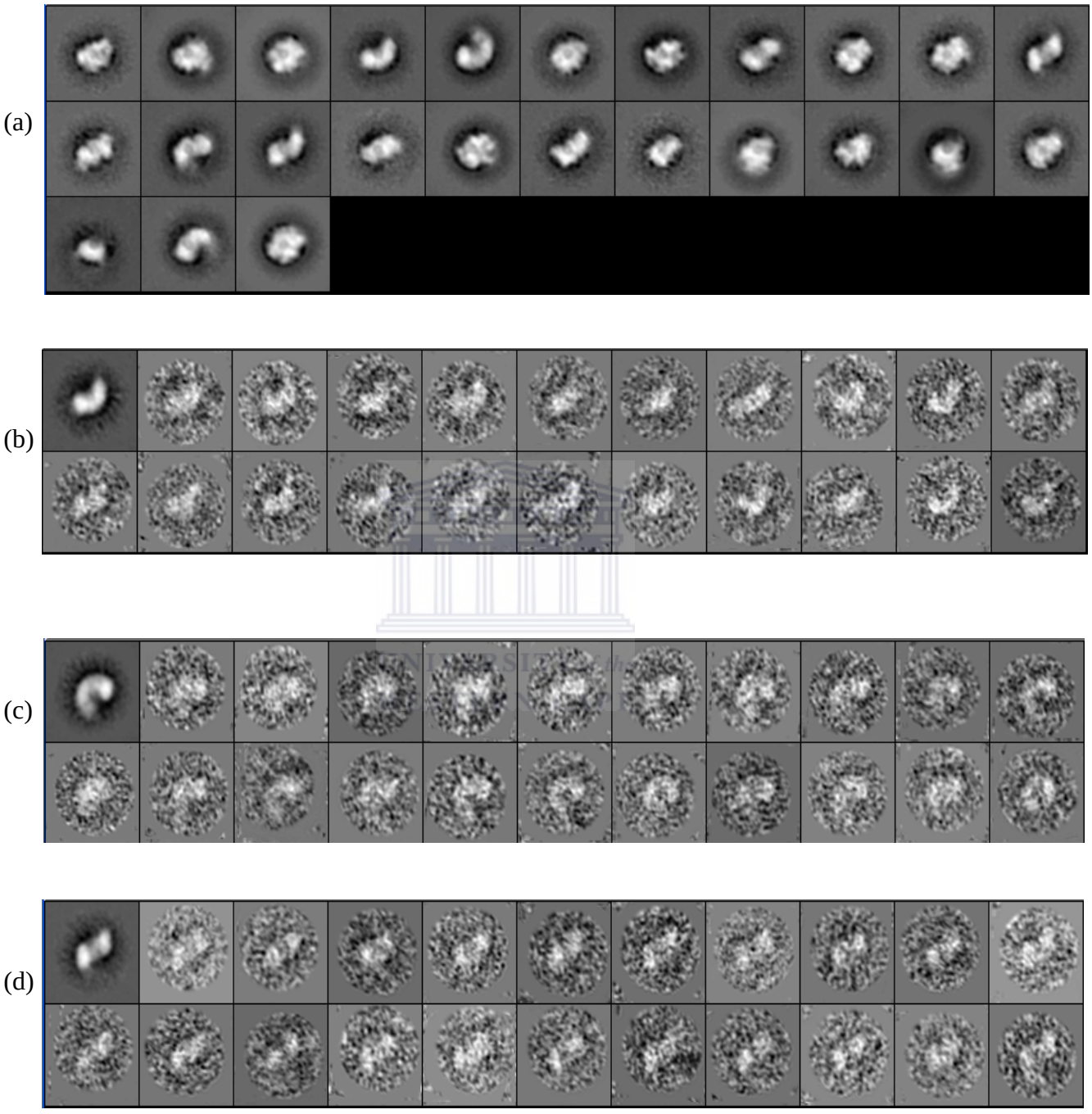


Figure 4.4. Results obtained after iterative alignment, sorting and classification. (a) 25 class averages representing common particle views of varying shape and size from the raw data set. Representative side-views of c-shaped class members (class averages 4 and 5 as seen in (a)) and their corresponding class average with improved SNR are shown in (b) and (c) respectively. The diameter of these particles varies between 9.6 – 13 nm. (d) Top view of the particles in (c) and corresponding average. The gallery of particles in (b) to (d) show incomplete spiral formation (open rings)

This generated consistent class members similar to their global average with improved SNR; otherwise the class was re-classified into various groups according to the perceived heterogeneity. Representative classes and their corresponding class averages representing the different views of the particles are shown in figure 4.4 (b) to (d). Some c-shaped particles that resemble open rings can be seen in figure 4.4 (c) and these represent an incomplete process towards spiral formation.

4.5 Discussion

Negative stain EM combined with single particle image analysis revealed quaternary polymorphism in the 480 kDa oligomer of J1 nitrilase. Even though a single band was observed in SDS PAGE, particles of various sizes and shapes were visible in the low dose micrographs. None of these particles clearly resemble those previously reported in the cyanide dihydratases of *B.pumilus* C1 and *P.stutzeri* AK61 (Jandhyala et al., 2003; Sewell et al., 2003). Nevertheless, an attempt was made to identify common shapes using techniques of alignment and classification in single particle methods.

The enzyme complexes among the Rhodococcal nitrilases are built primarily from dimers (Harper, 1985; Stevenson et al., 1992; Nagasawa et al., 2000) and therefore, variation in size would correspond to the addition or removal of these dimers (~80 kDa) in a dynamic oligomerization process. The variation in shape cannot be clearly explained, however, some shapes can be reconciled if the principles reported by Sewell et al (2005) are applied. As can be seen in figure 4.4 (c) and (d), some particles resemble open rings. The corresponding class averages revealed these vary in size between 90-130 Å. The diameter of the larger open rings of J1 nitrilase (~13 nm) corresponds to that of its helical structure (see chapter 5). This suggests they are the same species and that no significant shrinkage of the protein occurred after staining and drying of the stain. Therefore, the c-shaped particles show J1 nitrilase has a tendency to form spiral structures.

Hence, it is necessary to pursue further experiments in future whose objective would be to obtain a substantial data set of the c-shaped particles in order to pursue a reliable 3D reconstruction. A 3D reconstruction of the helical form of J1 nitrilase is described in the next chapter.

CHAPTER 5

THREE-DIMENSIONAL RECONSTRUCTION OF THE HELICAL FORM OF *RHODOCOCCUS RHODOCHROUS* J1 NITRILASE

5.1 Introduction

Many biological macromolecules such as F-actin, microtubules, myosin thick filaments, bacterial pili and flagella exist as helical assemblies (Egelman, 2000). These helical polymers may not crystallize because in a crystal, the allowed helical symmetries involve 2, 3, 4 or 6 subunits per turn, which pack into a space group that maintains this exact symmetry over long distances. Outside of a crystal, there is no constraint on a helix to have an integral number of subunits per turn. Helical assemblies of proteins are moderately ubiquitous and therefore it is not surprising that electron microscopy and 3D reconstruction techniques were first used to determine the structure of T4 bacteriophage tail (DeRosier and Klug, 1968). Over the last three decades, these methods have continued to improve both technically, through the refinement of existing algorithms and in achievable resolution. Early methods involved Fourier-Bessel inversion but single particle methods developed by Egelman (2000) have many advantages.

5.2 Review of helical reconstruction

A helix is constructed by repeating a subunit using a screw operation involving a coupled rotation and axial translation. The transformation between two adjacent subunits is defined by two parameters, which then define the helix (figure 5.1). These include the axial rise between subunits, h or delta Z (Δz), and the azimuthal rotation between subunits, delta phi ($\Delta\varphi$). Thus, helical symmetry is defined as. The translation along the helix axis (dotted line in figure 1(a)), needed to bring one subunit into exact superposition with another subunit is known as the helical repeat (c). For an integral number of subunits/turn (u/t), c is given by $u \times \Delta z$. The pitch (P) is the distance the helix takes to make one complete turn through 360^o, that is, $P = (360^o/\Delta\varphi) \times \Delta z$. It follows that $360^o/u = \Delta\varphi$, the azimuthal rotation between subunits. Of course there is no reason that a helix should have an integral number of subunits per turn.

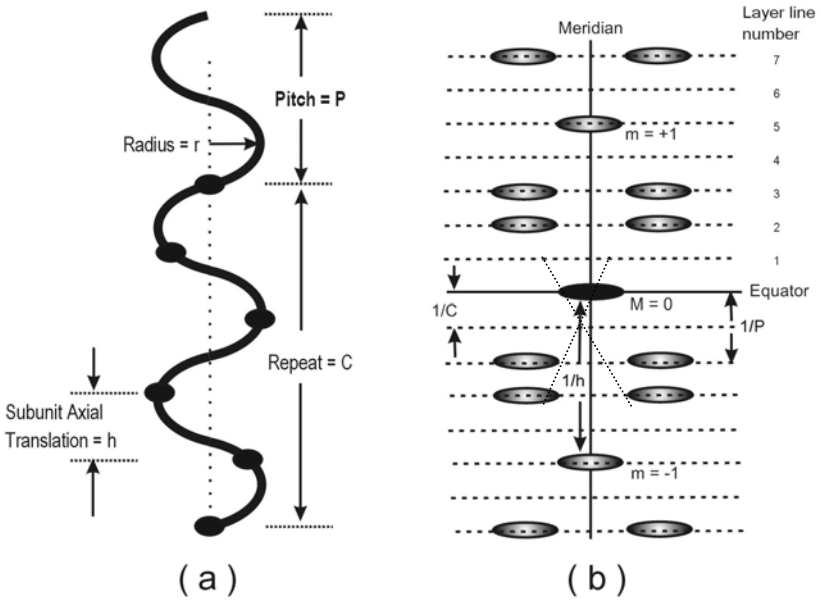


Figure 5.1 Parameters of a helical structure. (a) A helix with 5 subunits per two turns (5/2 or 5_2), and (b) distribution of diffracted intensities that would be obtained from the helix in (a). The first meriodional peak (m=1) is on the 5th layer line (number of subunits, u = 5) and the strongest peaks close to the meridian are on the 2nd layer line (number of turns t=2). Related layer line positions are at m±2. Figure was adapted from http://www.ccp13.ac.uk/software/program/Helix/HELIX_PAR.htm.

Helical structures can be studied using two methods, namely X-ray fibre diffraction and electron microscopy. Helical analysis using X-ray fibre diffraction is not easy, but it was used to elucidate the structure of the tobacco mosaic virus to 2.9 Å using the technique of isomorphous replacement (Namba et al., 1989). Hence for many years, electron microscopy has been the method of choice for 3D helical analysis using either Fourier-Bessel or single particle methods.

5.2.1 Fourier-Bessel method

In single particle methods, many views of the particle under investigation are averaged together in order to compute the 3D density. However, in an ideal helix, there exists many identical subunits and each subunit is in a different and predictable orientation to its neighbours. In theory, one should be able to obtain the helical structure from just one image of a filament (http://www.vir.gla.ac.uk/staff/bhellad/imageprocessing.htm). Particles for 3D processing are selected on the basis of the quality of the optical diffraction pattern and the helical parameters are determined from indexing the pattern. The Fourier transform of a 2D projection of a helical structure produces a diffraction pattern consisting of lines of intensity perpendicular to the helix axis called layer lines (figure 5.1(b)). This projection contains all the information necessary to reconstruct the helical structure using Fourier-Bessel methods. The scattering along each layer line is made up of Bessel functions (written as $J_n(x)$, where n is called the

order and x is the argument). A Bessel function is the form a wave adopts in situations of cylindrical symmetry and this is similar to the ripples observed when a pebble is thrown into a pool of water. Bessel functions characteristically begin with a strong peak and then oscillate like a damped sine wave as x increases. The position of the maximum amplitude depends on order n of the Bessel function. Crick showed that for a continuous helix, the order n of Bessel function occurring on a certain layer line is the same as the layer line number l, counted from the middle of the diffraction pattern (http://www.mpimf-heidelberg.mpg.de/~holmes/fibre/branden.html). Therefore, because the order of Bessel function n increases with the layer line number, so does the maximum amplitude which decreases towards higher resolution, resulting in a characteristic "X" shaped helical diffraction pattern (figure 5.1 (b)). In fact, DNA was known to be a helix long before its detailed model was produced due to its characteristic X-shaped diffraction pattern. The position of the first strong peak is inversely proportional to the radius of the helix, whereas the spacing of the lines is reciprocal to the pitch (P) of the helix. Therefore, a reciprocal relationship exists between the layer line separation and the pitch whereby a small separation, large P whereas a large separation, small P. Real helices are not continuous; rather they consist of repeating groups of subunits defined by $\Delta\varphi$ and Δz. The Bessel order n on the l^{th} layer line is given by the selection rule $l = um + tn$. This is an integer equation, which alternatively defines helical symmetry as u subunits in t turns of a repeat (m can take all positive and negative integer values). The helical selection rule is used to index a diffraction pattern by determining a particular set of Bessel orders for each layer line, which provides clues about the helical symmetry. Once Bessel orders have been assigned to each layer line, the real space helical structure can be obtained by back transforming the 3D Fourier transform.

Although the procedure is relatively straight forward, in theory the application of Fourier-Bessel methods is not easy due to the complexity of biological helical assemblies. [**1**] The pitch of the helix may vary along its length or between filaments. For instance, RecA protein has a varying pitch between 80-105 Å (Egelman and Stasiak, 1988). This can be overcome by selecting shorter segments. However, the widths of the layer lines will broaden as the length of the helix decreases resulting in Bessel overlap. This is a common problem for many real structures at low-resolution that makes it difficult to interpret the diffraction pattern. [**2**] The helix may not have precisely defined helical symmetry such as F-actin that has a random variable twist (Egelman et al., 1982). Furthermore, there is no reason for simple symmetry to be expressed as a ratio of small integers and therefore the use of methods that impose an average pitch and twist will fail. [**3**] Real filaments are flexible and disordered, that is, they are

never straight over long distances. However, one can computationally straighten filaments for small curvature using a "straightening algorithm" which has the potential to introduce artifacts (Egelman, 1986). [**4**] The absence of a space group to impose long-range order will result in deviations from average helical symmetry because the subunits are displaced away from positions that would be dictated by helical symmetry. This would force the subunits to interact only with their neighbours and thus "liquid-like" order will dominate. The use of conventional Fourier-Bessel methods to average over disordered filaments will limit the achievable resolution (Trachtenberg et al., 2005). In spite of these bottlenecks, improvements in existing methods designed to overcome disorder have led to the determination of the structures of the acetylcholine receptor (Miyazawa et al., 1999; 2003) and the bacterial flagellum filament (Yonekura et al., 2003) at 4Å resolution. [**5**] Structural distortion or uneven staining of the filaments may cause asymmetry in the diffraction pattern between the left and right halves. This is due to differences in scattering between the "far-side" (part furthest from the source) and the "near-side" (part closest to the source). This makes it more difficult to interpret each layer line. [**6**] Indexing a helical diffraction pattern can be difficult and ambiguous because the diffraction spots tend to be weak and diffuse in comparison with the intense sharp spots characteristic of extended, well-ordered crystalline arrays. Images of thin filaments in ice often have poor SNR, which may be too weak to enable indexing and further averaging (Trachtenberg et al., 2005). In practical terms, very few images give perfect diffraction. [**7**] Small changes in symmetry can result in large changes in the "selection rule". For instance, actin has u/t = 13/6, c = 355 Å and $\Delta\varphi = 166.1538^{o}$, but if we change $\Delta\varphi$ by a value of 0.128^{o}, then $\Delta\varphi = 166.2818^{o}$, u/t = 1299/600 and c = 35643 Å.

There are few examples where real space methods have been preferred over Fourier-Bessel methods to overcome the problem of disorder and flexibility. Bluemke and co-workers (1986) used a real space method to study the structure of a sickle-cell haemoglobin (HbS) filament. In their method, HbS particles with variable pitch were cross-correlated with test models having a constant pitch in order to determine the local pitch of the particle. The new estimate of the pitch was incorporated into the next cycle and the process iterated until the cross-correlation coefficient between the model and the particle was maximized. However, the advantages of using these methods are less obvious for highly ordered specimens such as tobacco mosaic virus. A reliable alternative is to reconstruct filaments as single particles using the Iterative Helical Real Space Reconstruction (IHRSR) method (Egelman, 2000), which borrows ideas from single particle reconstruction techniques.

5.2.2 IHRSR method

This approach is based upon refinement and imposition of the local helical geometry of the object under investigation during each cycle (figure 5.2). The IHRSR method overcomes common problems encountered using conventional Fourier-Bessel methods such as indexing, unbending and the need to find straight filaments over long distances (Egelman, 2000).

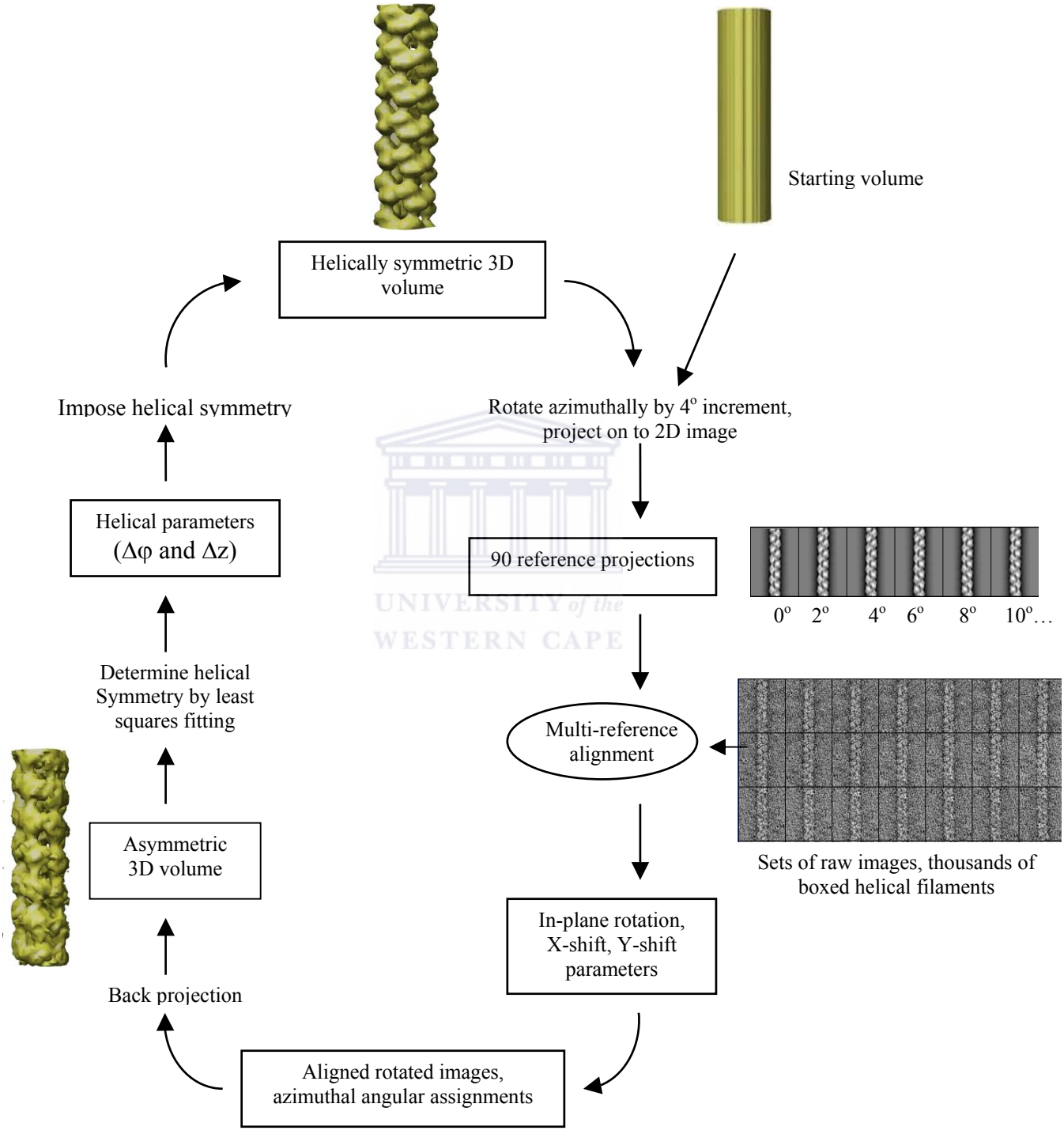


Figure 5.2. The IHRSR cycle (Egelman, 2000).

The procedure begins with a reference structure (either a blank cylinder or any symmetric helix), which is then projected into 2D images (same size as the raw images) by rotating about the filament axis by 4^{o} to generate 90 reference projections. The projections are cross-correlated with the raw images using a multi-reference alignment approach implemented in the program SPIDER (Frank et al., 1996). This procedure determines five particle parameters (azimuthal angle, in-plane rotation angle, x and y shifts, cross-correlation coefficient) and will assign each image to the reference that yielded the highest cross-correlation value. The raw images are then aligned and back projected using the known Euler angles to generate an asymmetric 3D reconstruction (figure 5.2). Helical symmetry is then determined *ab initio* using a least squares minimization search, which uses $\Delta\varphi$ to determine the value of Δz and vice versa. Once determined, these parameters are imposed to generate a helically symmetric volume. This volume becomes the reference for the next cycle and the procedure is iterated until convergence is reached.

Initially, the IHRSR method ignores the polarity of the filaments (uni- or bi-directional), and thereby speeds up the selection of filaments. Irrespective of the starting model, the radius of convergence is large. However, this cannot be too large for objects, which do not have helical symmetry. The algorithm can be applied to structures with a small number of subunits per turn or where Bessel overlap would complicate helical analysis. Several authors have shown that the IHRSR approach is quite powerful over conventional methods for more detailed analysis of disordered and heterogeneous helical filaments (Yang et al., 2003; Schoehn et al., 2004; Trachtenberg et al., 2005). The IHRSR method is relatively straight forward, easy to implement, can work with weakly scattering specimens and allows for separation of different classes of short helical segments.

5.3 Methods

5.3.1 Sample preparation, electron microscopy and digitization

The sample was prepared for negative stain electron microscopy as discussed in chapter 3. Micrographs were recorded on film using a JEOL 1200EX II microscope operating at 120 kV under conditions of low-dose at a nominal magnification of 50000×. These were subsequently digitized on a Leafscan45 as previously discussed (Chapter 3) at a scale of 2 Å per pixel.

5.3.2 Selection of filaments

Filament segments were cut out from 109 micrographs using Boxer, a program implemented within the EMAN software package (Ludtke et al., 1999). This program allows for rotational

alignment of selected filaments, which are then saved into a stack. Particles were windowed in 256×256 boxes at a scale of 2 Å per pixel in order to minimize interpolation errors during subsequent alignment steps. An overlap of 96% between boxed filaments was used because filaments related by an axial displacement can provide important information about the helical symmetry. All boxed filaments were saved in a stack.

5.3.3 Pre-processing

The stacked raw images (13,506) were reduced in size by a factor of two by first applying a low-pass filter followed by decimation using the DC operation in SPIDER (Frank et al., 1996). This procedure avoids interpolation errors, which would result in loss of signal-to-noiseratio (SNR) in the images. All subsequent processing steps were performed using routines executed in SPIDER (Frank et al., 1996). Boxed filament segments (128 × 128) were high-pass filtered to remove noise at high spatial frequencies greater than diameter of the filament. Subsequently, filament segments were approximately centered to their global average by applying x-shifts only (integer values to reduce interpolation errors) using a reference-free approach. This procedure was iterated until a well-centered image stack was obtained.

5.3.4 Image processing and 3D reconstruction

Two independent 3D reconstructions were generated using the IHRSR algorithm (Egelman, 2000), implemented on a Linux Intel server with 2.8 GHz Xeon processors and 4 Gb of physical memory. Different volumes were used as references at the start of the IHRSR cycle. First, a blank cylinder (radius of 40 pixels) was generated using the MO 3 operation in SPIDER (Frank et al., 1996) and second, a helix was generated partially in FORTRAN using 19 dimer models of J1 nitrilase (see chapter 2) and subsequently filtered to a resolution of 15 Å. Both references were boxed in 128 × 128 pixels at a scale of 4Å per pixel. The pre-processed images were aligned using a multi-reference alignment procedure by cross-correlation with 2D projections of initially, the starting models and subsequently, the resulting 3D symmetric reconstruction at the end of each IHRSR cycle (see section 5.2.2). The number of reference projections (N) required obtaining a desired theoretical resolution d (in this case 15 Å), while rotating about a single axis for an object of diameter D is derived from a single-axis tilt series reconstruction, where N is $\pi D/d$ (Crowther et al., 1970). Therefore, the minimum increment (in degrees) to obtain equally spaced projections is expressed as $(180 \times d)/\pi D$, giving 6.6^{o} for J1 nitrilase helix with a diameter of 130 Å. The increment was over sampled by a factor of ~two and the reference projection rotated about the z-axis by 4^{o} to generate 90 projections. During each cycle of iterative alignment, poorly correlated filaments

were excluded from subsequent steps based on deviations from set translational (x and y shifts $\leq \pm 2$ pixels) and rotational values ($\leq 10^{o}$ from either 0^{o} or 180^{o}).

5.3.5 Convergence

The convergence of the algorithm was determined by monitoring two statistics continuously during the iterative procedure. These are the average coefficient of cross-correlation between all raw images and the reference projections and secondly, parameters defining helical symmetry (delta phi, delta z and the pitch).

5.3.6 Resolution criteria

The common approach for determining resolution is to split the image data set in half, align these against a common reference and then generate two 3D reconstructions, which are then compared using the Fourier Shell Correlation criterion 'FSC' (Harauz and van Heel, 1986). The FSC criterion measures correlation as a function of resolution shells between the two half reconstructions. Unfortunately, this approach is susceptible to model bias and noise can be correlated during the alignment process because image sets are aligned against the same reference. Therefore, resolution was determined by measuring the correlation between truly independent reconstructions, each from 13,506 images, and generated from different starting models. This strategy avoids a possible extension of the true resolution due to the high correlation of neighboring shells in the Fourier Transform (FT) when the thickness of the shell is less than 1/D, where D is maximum particle dimension (Yang et al., 2003). Because few biological complexes are spherical, D is often greater than the extent of the particle in many directions. Therefore, the FT will be oversampled in some few directions even when sampled at 1/D resulting in high correlation of neighbouring shells and causing an extension of the true resolution. The 'conservative' 0.5 FSC measure, which has been estimated to have a SNR of about 1 (Penczek, 1998), was taken as the limit of resolution. We consider this as arbitrary, but standard criterion for estimation of the reconstruction quality. The models were then low-pass filtered to correspond to the determined resolution.

5.3.7 Map visualization and docking

Results from the reconstruction process were visualized using WEB (Frank et al., 1996) and UCSF Chimera (Peterson et al., 2004; Goddard et al., 2005). The threshold for contouring the 3D density of the final reconstruction was determined using the calculated molecular weight (M) of J1 nitrilase subunit (40,189Da), total number of dimers in the map, height (h) of the

map and an average value for protein density of 0.73cm^3/g. Thus, each dimer would enclose a volume V, expressed as:

$$V\ (nm^3) = [(0.73\ cm^3/g)\ (10^{21} nm^3/cm^3)/6.023 \times 10^{23} Da/g] \times M\ (Da) \times 2$$

It follows that total volume (nm^3) of the 3D map is equal to = $(h/\Delta z) \times V\ (nm^3)$.

Automatic fitting of a helix made up of 9 dimers of J1 nitrilase was done using the Contour-based Low-Resolution CoLoRes program implemented in the SITUS software package (Chacon and Wriggers, 2002). The 9 dimer model was generated by applying the helical symmetry operators to a dimer model whose two-fold axis was located on the x axis. Once the helical parameters were determined, the $D_1\ S_{\sim 4.9}$ symmetry of the helix allows only two additional degrees of freedom for fitting such a model, namely the azimuthal rotation about the x-axis and the translation about the same axis. The clarity of the map with the fitted model allowed for subsequent structural interpretation.

5.4. Results

5.4.1 Negative stain EM and image pre-processing

The purification of one-month old J1 nitrilase and subsequent negative staining revealed a homogeneous, open filamentous form with a diameter of approximately 13 nm and undefined length (figure 5.3). The filaments were observed to be generally straight over long distances with a few broken sections and curvature, which are thought to arise from adsorption of the sample to the grid and drying during the staining process (Ohi et al., 2004). The helical structure of J1 nitrilase was predicted to be left handed based on previous metal shadowing experiments on the filamentous forms of the cyanide dihydratase from *B.pumilius* C1 (Jandhyala et al., 2003) and the cyanide hydratase of *G.sorghi* (Margot Scheffer, personal communication). This prediction will be confirmed in the future through metal shadowing experiments.

Short filament segments (13,506) were cut out of the micrographs, pre-centered by applying only the x-shifts and stacked together. The quality of the micrographs could be directly gauged from the power spectrum, which was computed by summing over these short segments. As seen in figure 5.4, the computed transform has a characteristic "X" diffraction pattern typical of a helix and there is no splitting due to layer line broadening as a result of selecting short helical segments.

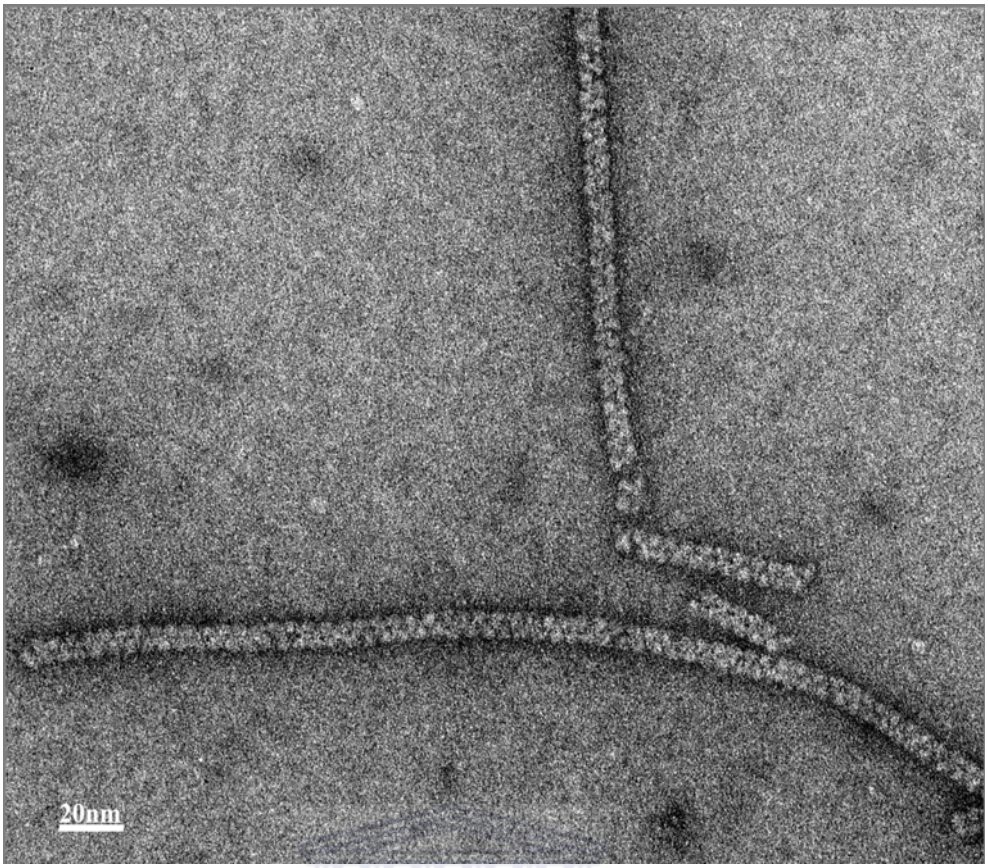


Figure 5.3. Electron micrograph of negatively stained helical fibres of the nitrilase of *R.rhodochrous* J1.The samples were prepared and recorded as described in chapter 3. Magnification = 50000×.

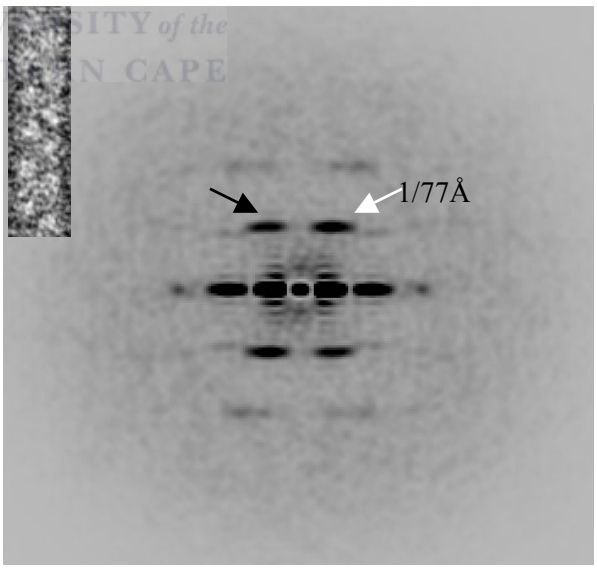


Figure 5.4. The transform computed from a set of 13,506 approximately centered short segments (example shown on top left) measuring 128 pixels in length. Broad layer lines (black arrow) can be seen on either side of the equator. The pitch (1/77 Å), measured from the equator to the first layer line, was used to determine the starting value for $\Delta\phi$ by first executing the algorithm using a subset of 1000 images till convergence was reached.

5.4.2 2D projections and iterative alignment

The stack of centered images was then filtered and subjected to 40 cycles of iterative alignment and refinement against the projections of starting model initially, and the 3D model generated

at the end of each cycle. This process generated a refined set of raw images (figure 5.5) and the procedure was terminated when convergence was reached.

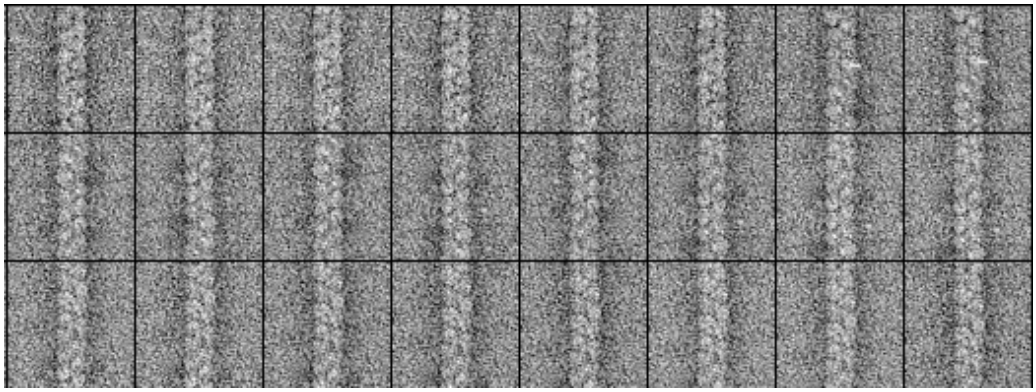


Figure 5.5. A representative set of aligned filament segments taken from a dataset of 13,506 that were used for 3D reconstruction. Each image is 128 × 128 pixels at a scale = 4 Å/pixel.

During multi-reference alignment, the search for translational (x and y) parameters was confined to within ±2 pixels and filaments with shifts greater than 3.5 pixels (as defined in the procedure) would be rejected. As can be seen in figure 5.6 ((a) and (b)), majority of the segments (>7000) had shifts approximately centered at zero in the x direction whereas their y shifts were within ±2 pixels. Furthermore, filaments with deviations greater than 10^{o} from either 0^{o} or 180^{o} would also be rejected. Fortunately, no single image in the data set was rejected because of high correlation to the corresponding references within the range of defined rotational and translational parameters (figure 5.6 (c) to (e)). At the end of 40 cycles, the distribution of raw images among the 90 reference classes was observed to be statistically flat (figure 5.6 (f)) because the search range specified during iterative alignment was able to separate images with 96 % overlap.

5.4.3 3D reconstruction

The aligned images (figure 5.5) were used for 3D reconstruction as outlined in the scheme in figure 5.2. The starting azimuthal angle was negative based on the prediction that J1 nitrilase also forms a left-handed helix (see section 5.4.1). The IHRSR cycle (Egelman, 2000) was iterated 40 times and terminated when the statistics showed convergence. Starting from different reference points, the two independent reconstructions converged to the same solution (figure 5.7 (a) and (b)). Visualization of the 3D map and corresponding projections showed the helix had striking two-fold symmetry on a global dyad axis perpendicular to the helix axis. Two-fold averaging was performed by estimating the x and y coordinates of the dyad axis of one volume and superimposing the density with its two-fold related image to generate a two-fold symmetric 3D reconstruction.

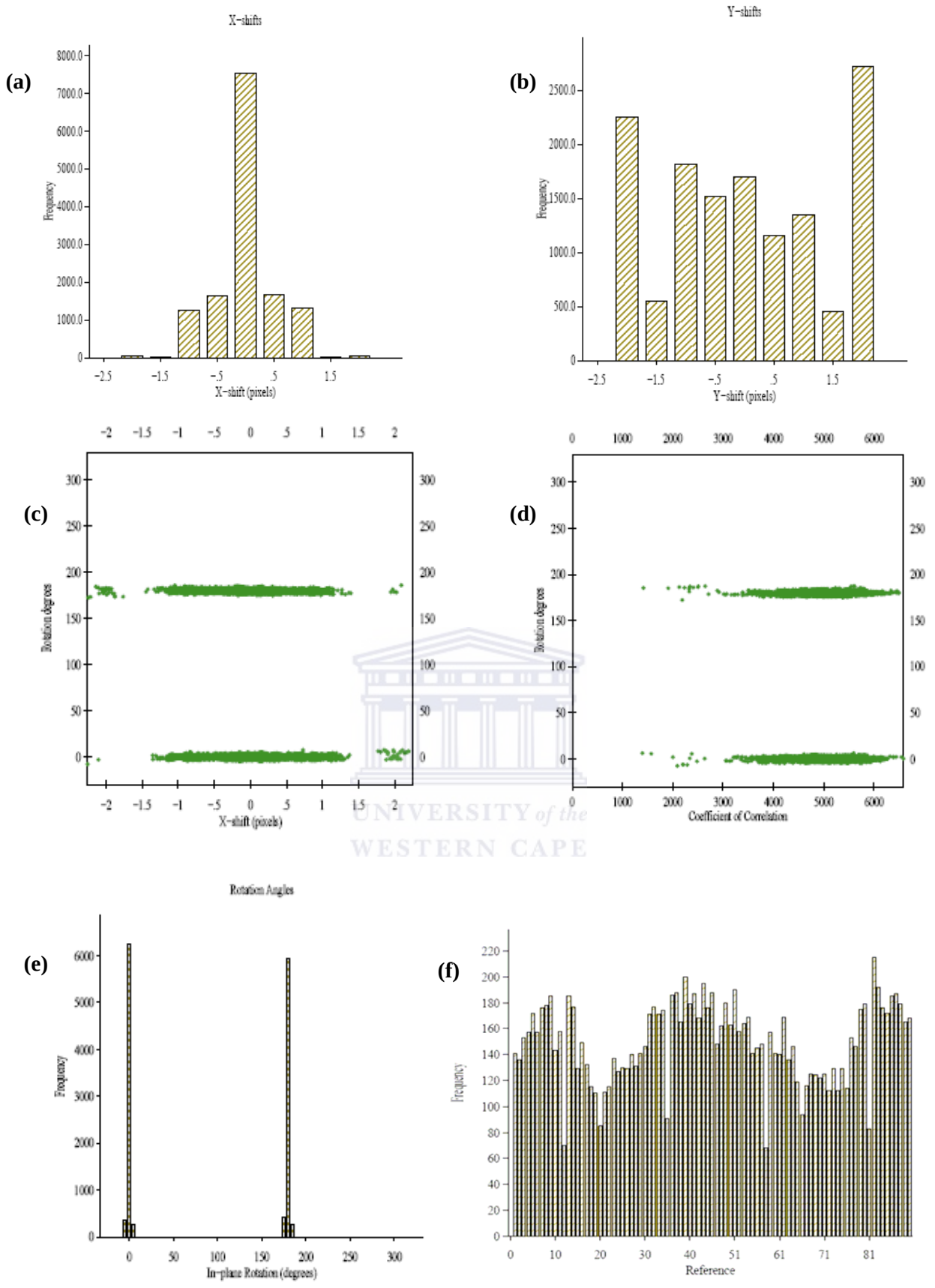


Figure 5.6. Graphical representation of multi-reference alignment statistics. Translational shifts for the raw data set are shown in (**a**) and (**b**). (**c**). Images with in-plane rotation angles of ~0^o or 180^o have high correlation scores and can be seen in (**d**) to be approximately centered between ±1.5 in the x-direction with a few outliers. **(e).** Distribution of in-plane rotation angles. **(f)**. Statistically flat distribution of 13,506 images among 90 reference classes.

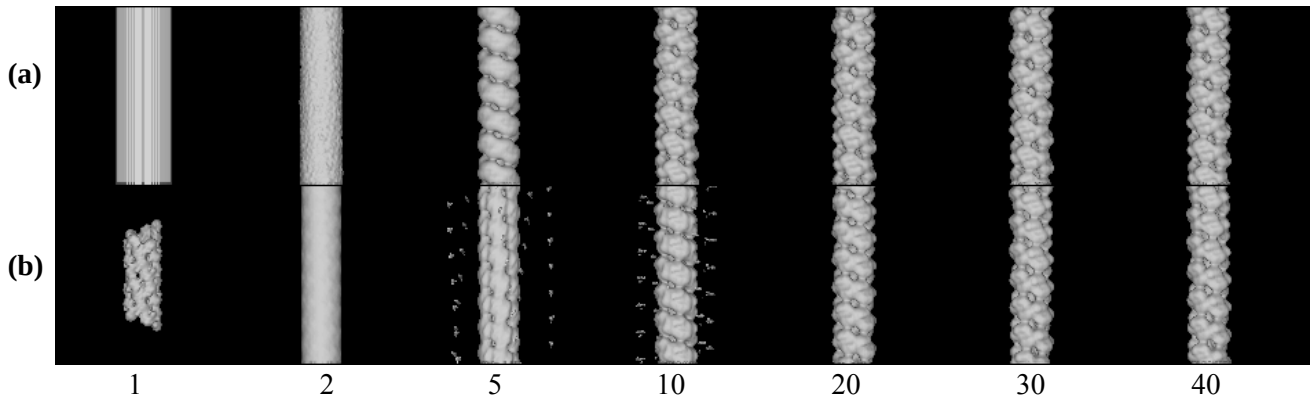


Figure5.7. Surfaces of reconstructed volumes starting from different models. (**a**). Blank cylinder, 128 pixels long and initial $\Delta\varphi$ = -75.0^o and Δz = 17 Å (**b**). A helix made up of 19 dimer models of J1 nitrilase, 82 pixels long and initial $\Delta\varphi$ = -73.0^o and Δz = 13 Å. The numbers correspond to the model generated at the end of that cycle. Models between cycles 30 and 40 are the same and have a two-fold axis perpendicular to the filament axis. Two-fold averaging was performed on the final model.

The quality of the two independent reconstructions using the FSC 0.5 criterion was determined to be 17 Å.

5.4.4 Convergence statistics

Continuous monitoring of the helical parameters (delta phi, delta Z, pitch) and the average cross-correlation value between the raw images and the projections of the reference showed the algorithm converged independently to generate identical maps after 40 cycles. This revealed a helix having 4.88 dimers per turn in a pitch equal to 77.23 Å. The final model (after 40 cycles) was significantly different from the starting models.

A least-squares search of the asymmetric volume generated by back projection in order to determine helical symmetry parameters is shown in figure 5.8 (a). Subsequently, imposition of helical symmetry during each cycle generated a final volume (after 40 cycles) that was significantly different from the starting models (figure 5.8 (b)). As can been seen in figure 5.9 (a) and (b), each dimer in the final model has an azimuthal rotation of 73.65^o coupled to an axial rise of 15.8 Å. The average cross-correlation value ("goodness-of-fit") reaches a nearly asymptotic value after 25 cycles (figure 5.9 (c)). This provides an internal measure of assessing the result based upon the comparison of raw images with reference projections during the alignment process. There is no change in structure between 25 and 40 cycles, consistent with lack of change in the helical parameters over these cycles.

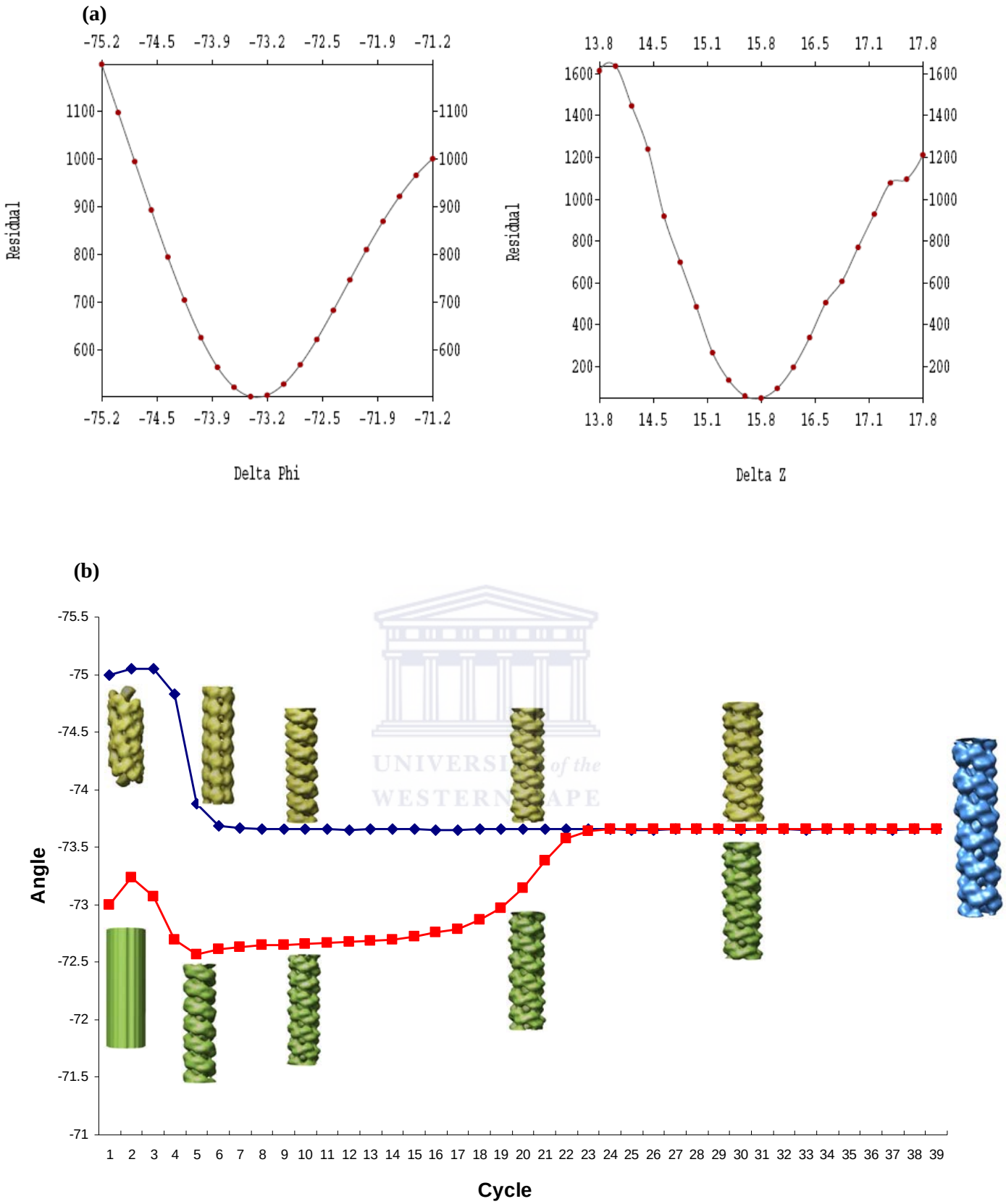


Figure 5.8. Convergence of the IHRSR procedure after 40 cycles. (**a**) Plots of the residual mean-square deviation in density as a function of changes in delta phi and delta z (after 40 cycles). (**b**) Starting from two different structures with different values of Δφ, the procedure converges to same structure after 40 cycles. The models shown correspond to cycles 1, 5, 10, 20, 30 and 40 respectively. Two-fold symmetry was imposed on the final volume (shown in blue).

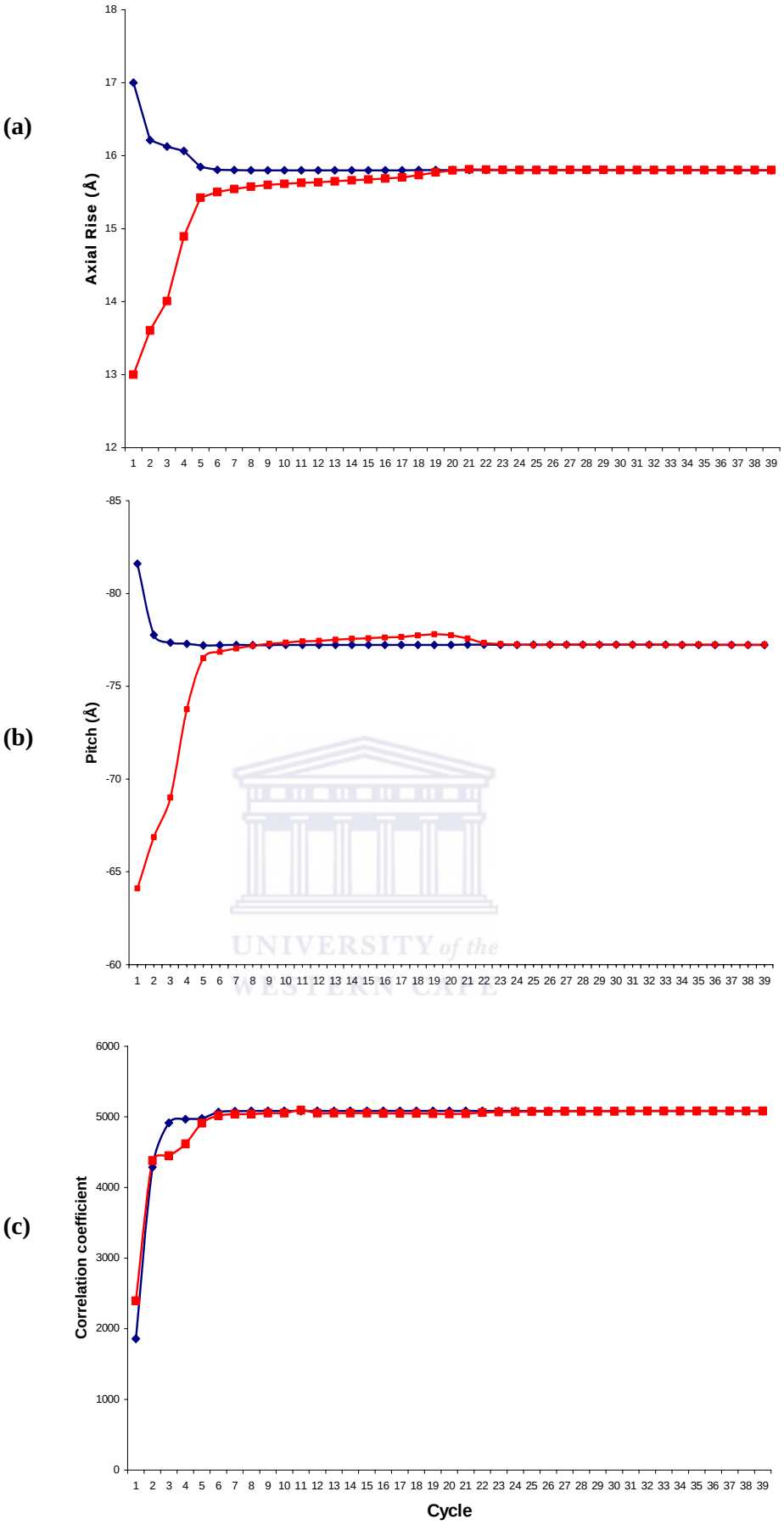


Figure 5.9. Convergence of the helical parameters. (**a**) The rise per subunit Δz. (**b**) The helical pitch, a function of Δφ and Δz, can also be seen to stabilize after 22 cycles. **(c).** The average cross-correlation value between the raw images and reference projection. Δφ = -73° (red squares) and –75° (blue diamonds) for all plots.

5.4.5 Resolution and quality of the reconstruction

The resolution of the models from two independent reconstructions was determined using the 'conservative' 0.5 cutoff of the FSC curve (Bottcher et al., 1997; Conway et al., 1997) to be 17 Å (figure 5.10). As expected, there was an improvement in resolution after imposing two-fold symmetry on the final reconstruction (data not shown) due to good correlation of high spatial frequencies between the two 3D volumes due to "edge effects" arising from use of the same mask. The models were low-pass filtered to 17 Å in order to smooth the surface.

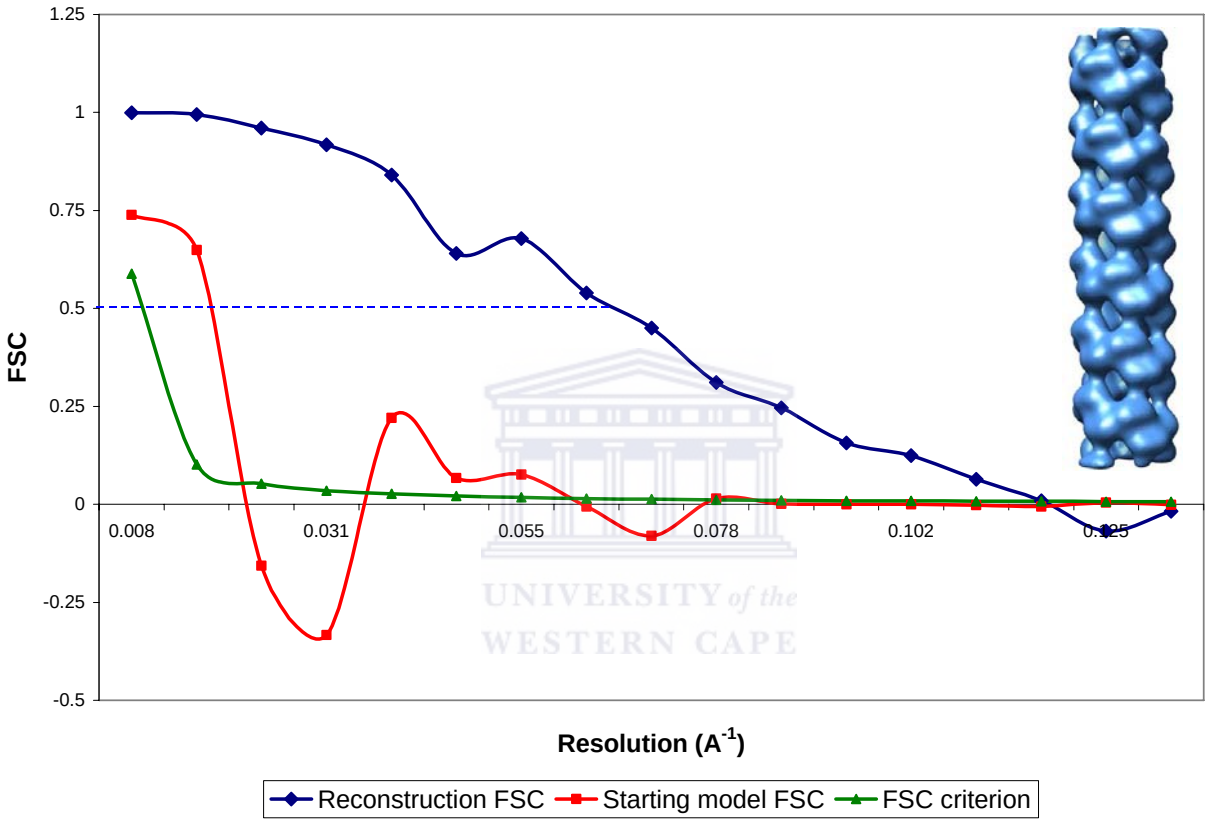


Figure 5.10. Resolution of two independent reconstructions, each from 13,506 images. Initial references (red curve) have a resolution (using FSC=0.5) of 54Å, whereas final reconstructions (navy blue curve) have a resolution of 17Å. The final filtered volume with imposed two-fold symmtery is shown on the right.

Regions of evenly distributed density corresponding to ridges and grooves that define each dimer and the connectivity between them, can be seen in model presented in figure 5.10. The reliability of the iterative reconstruction process can also be gauged from the similarity between the projections of the model with the raw images. All the views were identical because of tilting about the z axis (vertical axis of the helix) with uniform sampling of all the angles and hence there is no missing information. As can be seen in raw images (figure 5.5) and in the reconstruction (figure 5.10), there exists a two-fold axis, which is located at the center of each dimer, between the ridges.

5.4.6 Comparison of power spectra and the n, l plot

Another measure for correctness of the solution is to compare the power spectrum of raw images with the computed transform of the reconstruction (figure 5.11 (a) and (b)).

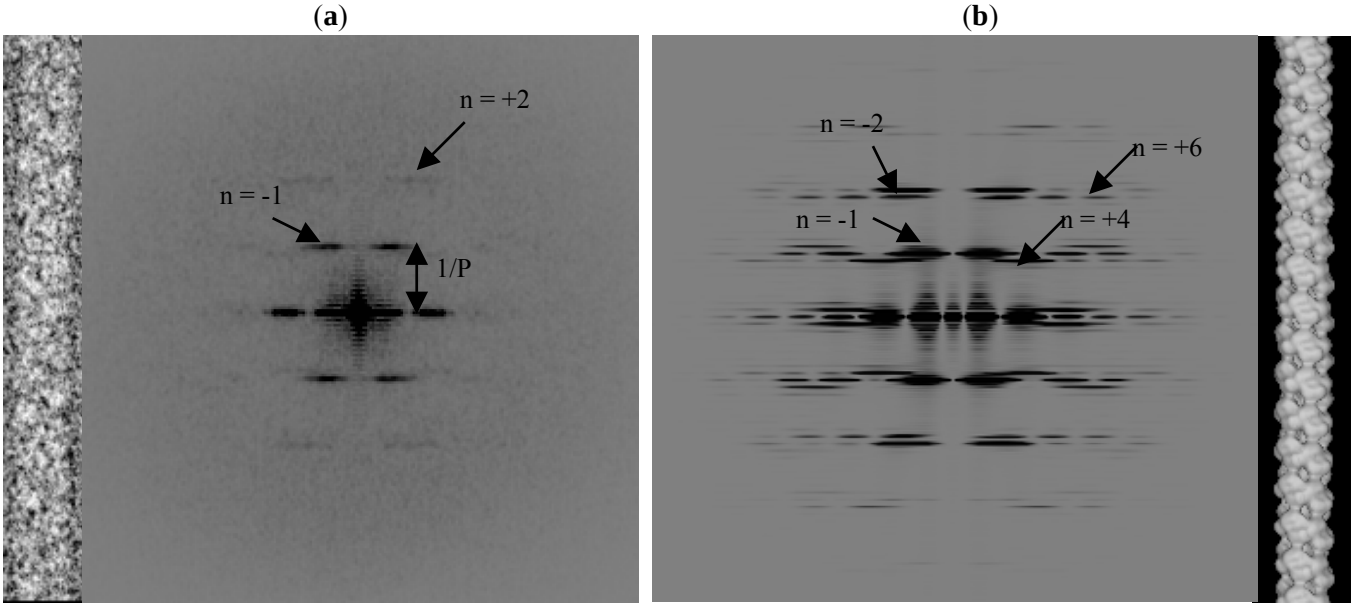


Figure 5.11. Calculated power spectra from: (**a**) 953 aligned segments, 512 pixels long (shown on the left) and (**b**) from the final 3D Model, 1000 pixels long (shown on the right) with imposed two-fold symmetry. The distance from the equatorial to the first layer line (1/P, double headed black arrow) corresponds to the pitch (~1/77Å).

All the transforms shown in 5.11 have a characteristic "X" diffraction pattern typical of a helix. The computed transform from the short filament segments (figure 5.4) does not have any splitting due to layer line broadening. Indexing the pattern requires determining the order of the Bessel functions for only two layer lines. This is ambiguous for short segments and would complicate helical analysis using Fourier-Bessel methods. In order to try and separate the layer lines, a Fourier Transform was computed using longer segments (figure 5.11 (a)). The resulting pattern is similar to that obtained using short segments, however, it is difficult to see beyond the second layer line because of noise in the image. Furthermore, overlap of different Bessel functions will occur at fairly low resolution. Attempts to use longer segments (at least 1000 pixels) were hampered due to curvature over this distance. The transforms of the reconstruction with imposed two-fold symmetry revealed layers line next to the meridian that had Bessel overlap. The helix has approximately 5 dimers per turn and this symmetry prevents separation of Bessel orders on the same layer line (that is n = -1 and n = 4). Consequently, only a limited number of views can be seen in the 2D projection of the reconstruction (figure 5.11 (b)).

The computed transform of a helix can be directly interpreted using the n,l plot, which is generated using the helical selection rule, l = um + tn (see section 5.2.1). This plot, a reciprocal lattice of a helical net, is used to index the diffraction pattern of a helix. The

convention is to use "1-start" to geometrically define helical symmetry. In the case of the computed transform of J1 nitrilase, n = -1 coincides with the left-handed reconstruction of J1 nitrilase and is reflected in the n,l plot (figure 5.12).

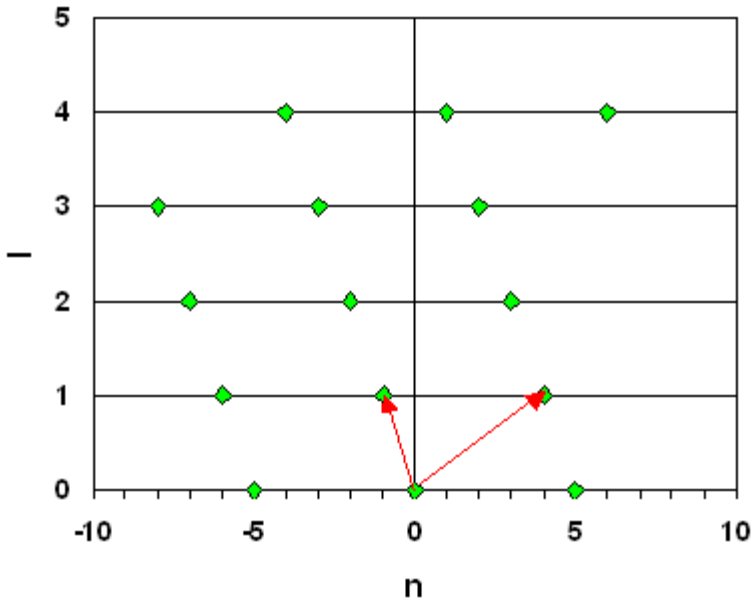


Figure 5.12. Representation of the computed transform of the reconstruction (figure 5.11 (b)) using the n,l plot. Bessel orders n = 4 and n = -1 occur on the same layer line (l) and cannot be distinguished in the computed transforms (figure 5.11). Red arrows represent vectors that can be used to generate all the other points in the plot.

It can be seen that Bessel functions n = -1 and n = 4 occur on the same layer line. These cannot be clearly distinguished in the computed transform (figure 5.11 (b)) of the reconstruction in which multiple Bessel orders can be seen to occur on the same layer line as a result of the symmetry of the helix. These results demonstrate the advantage of using IHRSR method (Egelman, 2000) over Fourier-Bessel methods because there is no need to index the computed transform prior to 3D reconstruction.

5.5 The Three-dimensional map

The negatively stained structure of J1 nitrilase on which two-fold symmetry was imposed has a resolution of 1.7nm. At this resolution, the shape and symmetry of the structure is well defined. The map reveals J1nitrilase forms a long fibre with a diameter of about 13nm and infinite length (figure 5.13). A slice through the 3D density reveals approximately 5 dimers per turn and a central core of zero density (figure 5.14). The structure is an open 5-start helix. Although the handedness was not determined *a priori*, the connectivity between the subunits shows the structure is clearly left-handed (see section 5.4.1). The reconstructed 3D volume encloses approximately 32 dimers of J1 nitrilase corresponding to a molecular envelope of ~3157nm^3.

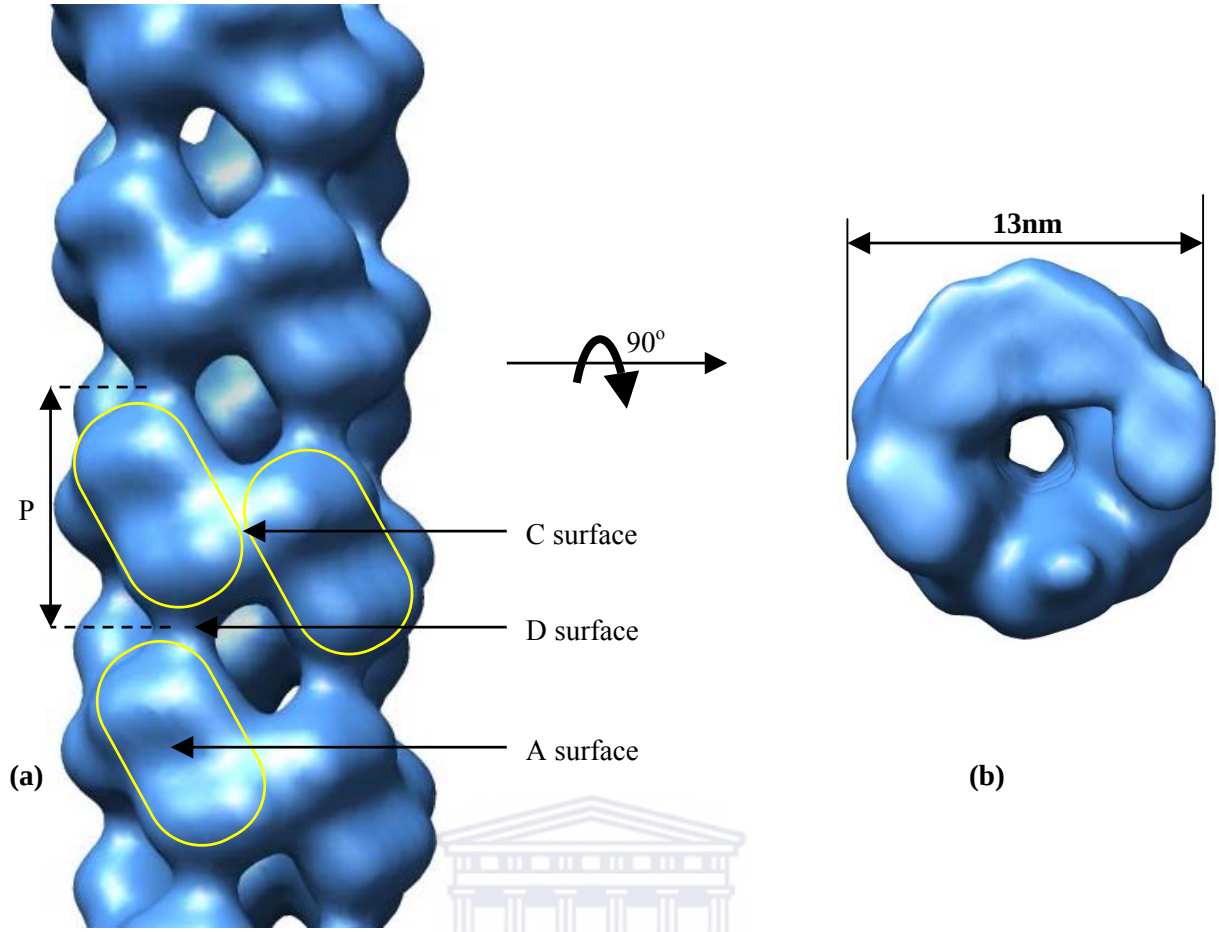


Figure 5.13. The left-handed helix of the nitrilase from *R.rhodochrous* J1. **(a)**. A section of the two-fold symmetric reconstruction. Each dimer (outlined in yellow) is rotated about a two-fold axis located at the A surface and perpendicular to the helix axis. The pitch (P) is 77.23 Å. Inter subunit contacts, namely A, C and D surfaces are also highlighted. **(b)**. Top view of the map showing the structure has approximately 5-fold symmetry.

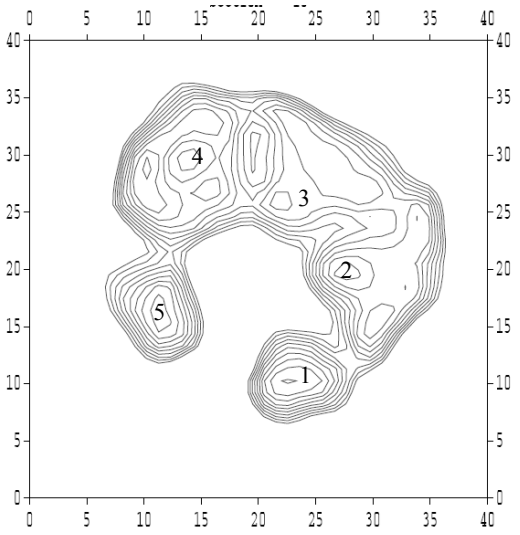


Figure 5.14. A cross-section of the reconstruction. There are ~5 dimers per turn. The central core is seen to be empty. The scale of the cross-section is 4 Å/pixel

As seen in figure 5.13 (a), there are three inter-subunit contacts, namely the A, C and D surfaces. The A surface is conserved in all the nitrilase atomic homologues whereas the C and

D surfaces are similar to those previously reported in the cyanide dihydratases of *B.pumilus* C1 and *P.stutzeri* AK61 (Sewell et al., 2003; 2005).

5.6 Interpretation of the map

The negatively stained structure of the J1 nitrilase has certain features, which allow the model previously constructed (see chapter 2) to be placed in the 3D map. Besides the coupled screw operations, namely delta phi and delta z, there exists another symmetry operation in the map. For any line group, the helical symmetry can be defined by either cyclic *n*-fold (C_n) or dihedral *n*-fold (D_n) coupled to S_σ rotational operation, where *n* is the rotational order of the point group whereas S_σ corresponds to the σ-fold screw symmetry (Makowski and Caspar, 1980). For instance, the symmetry of double-helical B-DNA that has a perpendicular dyad (designated by 1-fold dihedral point group D_1) in this notation is D_1S_{10}. No distinction is made between point groups with $n \geqslant 1$.

In the helical structure of J1 nitrilase, a two-fold axis exists about a line passing through the helical axis and perpendicular to it. This line coincides with the dyad axis of the dimer located at the A surface and comes out through the hole on the other side of the map. A second horizontal dyad is located at the D surface by about half a rise of the dimer. A line through this dyad passes through the C surface on the other side of the experimental map (figure 5.15). Therefore, the symmetry of the helical structure of J1 nitrilase is $D_1\ S_{\sim 4.89}$.

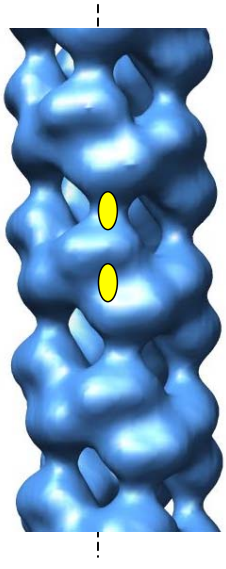


Figure 5.15. A section of the helical structure of J1 nitrilase showing two horizontal dyads (shown in black) located at the A and D surfaces and perpendicular to the axis of the helix (black dashed line). The line group symmetry of the structure is $D_1S_{\sim 4.89}$.

There are only two degrees of freedom per dimer, namely the azimuthal rotation and radial translation along the two-fold axis as a consequence of this symmetry. Therefore, automatic fitting with SITUS (Chacon and Wriggers, 2002) was fast and located the dimer with the C terminal facing the inside of the map (Table 5.1).

Table 5.1. Docking of different helical models comprising two turns into the 3D map of the nitrilase from *Rhodococcus rhodochrous* J1. The best fit corresponds to highest average correlation score.

Test Model (pdb)	Azimuthal rotation (degrees)	Radial translation (Å)	Average correlation score
J1fit01	20	35	0.13549
J1fit02	20	39	0.19047
J1fit03	23	36	0.16099
J1fit04	23	39	0.19649
J1fit05	24	39	0.19801
J1fit06	25	37	0.18102
J1fit07	25	38	0.19321
J1fit08	25	39	0.19925
J1fit09	26	39	0.19996
J1fit10	27	38	0.19420
J1fit11	27	39	0.20016
J1fit12	28	38	0.19383
J1fit13	28	39	0.19982
J1fit14	29	39	0.19862
J1fit15	30	39	0.19680

The result was similar to the 3D structural interpretation of the cyanide dihydratase from *P.stutzeri* AK61 (Sewell et al., 2003). As seen in Table 5.1, the model was best located in the 3D map when the centroid was translated 39 Å along the two-fold axis and rotated arbitrarily by 27^{o} along the same axis. The model was also lacking the significant insertions relative the solved structures (figure 5.16 (a)). The fitting confirmed the helical structure is built primarily from dimers which are mating across a global two-fold axis located at the A surface (Figure 5.16 (b)). As can be seen from the view in Figure 5.16 (c), there are empty regions of density on the outside of the helix corresponding to major insertions (residues 54-73 and 234-249) relative to the nitrilase atomic homologues. However, there is not enough defined density on the inside to accommodate 53 missing residues in the C terminal. This is probably because the enzyme loses about 39 residues in the C terminal tail prior to helix formation or may be because this region is highly disordered and flexible which lowers visible density at this resolution.

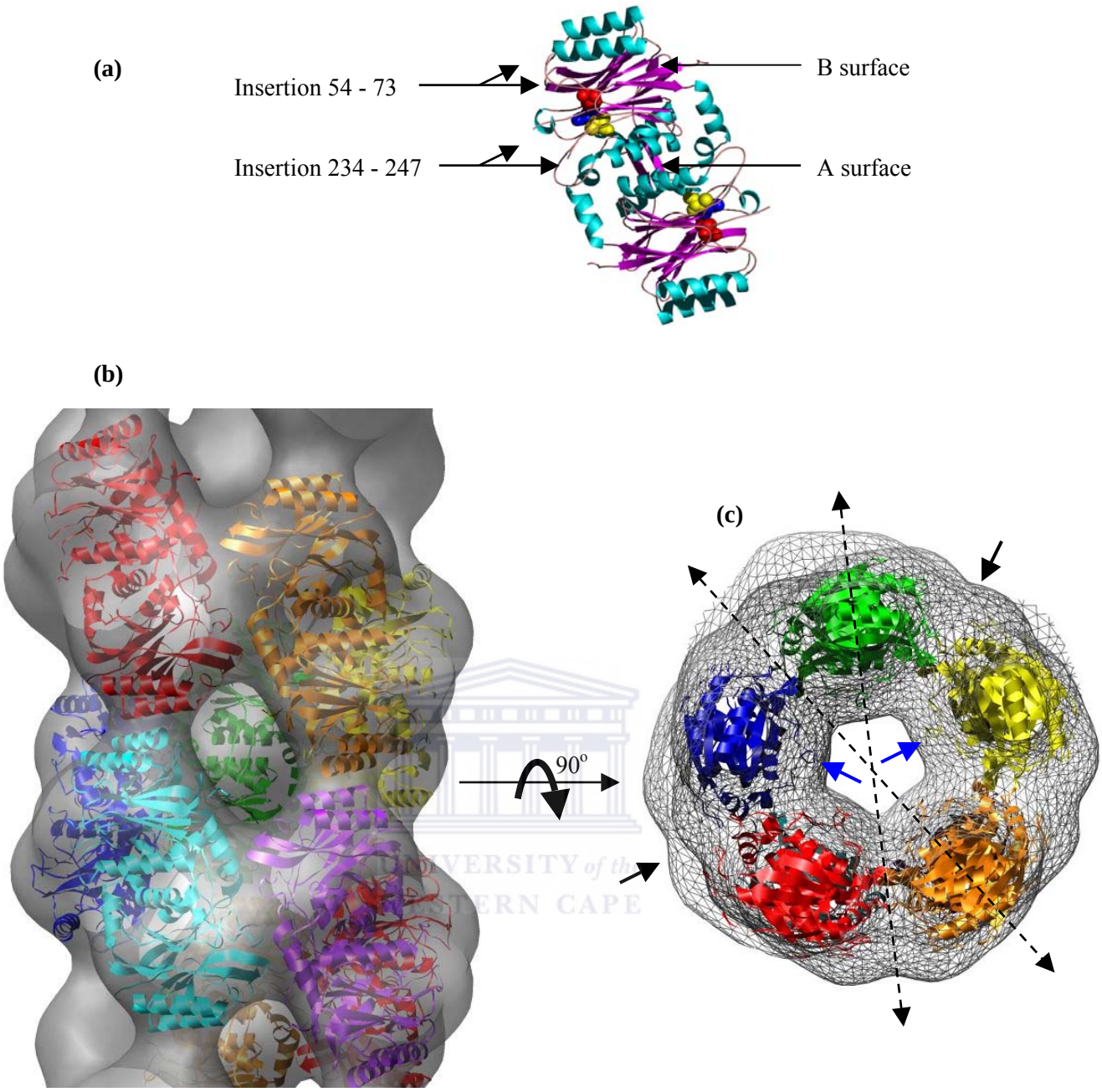


Figure 5.16. Fitting of dimer models of the nitrilase from *R.rhodochrous* J1 to the 3D map. (**a**). Ribbon representation of the model that was docked into the 3D map. Residues 54-73 and 234-247 have been deleted. The catalytic residues, E48 (red), K131 (Blue) and C165 (yellow) are shown in space filling representation **(b)**. Best fit of 9 dimers (ribbons) to the molecular envelope (translucent) of the reconstruction. (**c**). Top view (mesh) of the fitted molecular envelope. Regions of empty density are observed on the outside (black arrows) and the inside (blue arrows) corresponding to the C surface and missing residues in the C terminal respectively. The two-fold axes are represented using black dashed arrows. One dyad passes through the C and D surfaces whereas the other passes through the A surface and the hole as shown in (b) which is centered on the two-fold axis passing through the A surface.

5.7 Inter-subunit contacts

A view of the map in figure 5.13 reveals contacts between the subunits along its length. One of the contacts occurs at the A surface. This is located on the two-fold axis of the dimer as seen in figure 5.16(a) and corresponds to the A surface which is preserved in the atomic structures of 1ems, 1erz, 1f89 and 1j31 (see chapter 2). The other two contacts occur at the C and D surfaces where the two horizontal dyads are located. These interactions are absent in the solved structures, which do not form spirals and we propose that they are necessary for spiral formation and stability. There is clearly defined density showing there is contact at these surfaces (figure 5.17).

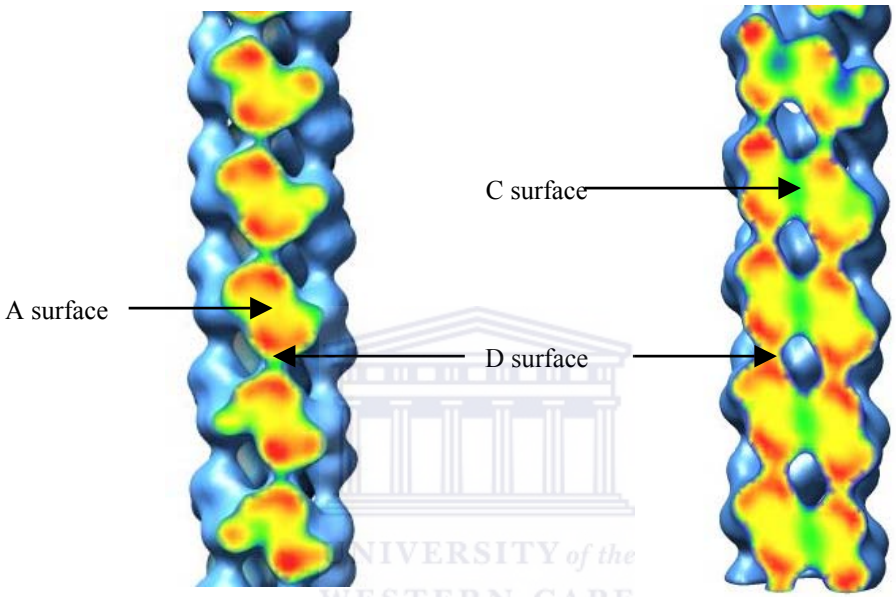


Figure 5.17. Clipped surface representations of density in the 3D map. Each subunit (highlighted in red) is seen to be mating across the A surface (highlighted in yellow) to form a dimer. Dimers next to each other are connected by C and D surfaces (highlighted in green). Regions of strongest density are coloured red whereas those of weakest density are coloured green

Since no attempt was made to model residues 54-73 and 234-249 located in the C surface (due to lack of a template), these residues were deleted in the model before docking and interactions across this surface cannot be located in the map. However, contacts at the D surface can be interpreted at atomic resolution. In section 2.18 of Chapter 2, it was predicted there are at least two salt bridges formed between residues 91DAARDHN97 and their two-fold related symmetry counterparts across the D surface (see section 2.18 of chapter 2). As seen from figure 5.18, there is possibility of four symmetric ion-pair interactions forming between residues 86VQRLLDAARDHN97 which are located near the end of an external helix corresponding to NH2 in Nit (see figure 2.10 of chapter 2). The density at the D surface is significantly less than at the C surface suggesting these residues may play a different role other than spiral formation.

It is predicted that these interactions confer stability to the spiral and would have little or no effect on the activity of this enzyme. A detailed study designed to investigate the effect on activity and structure upon modification of these residues is planned in the future.

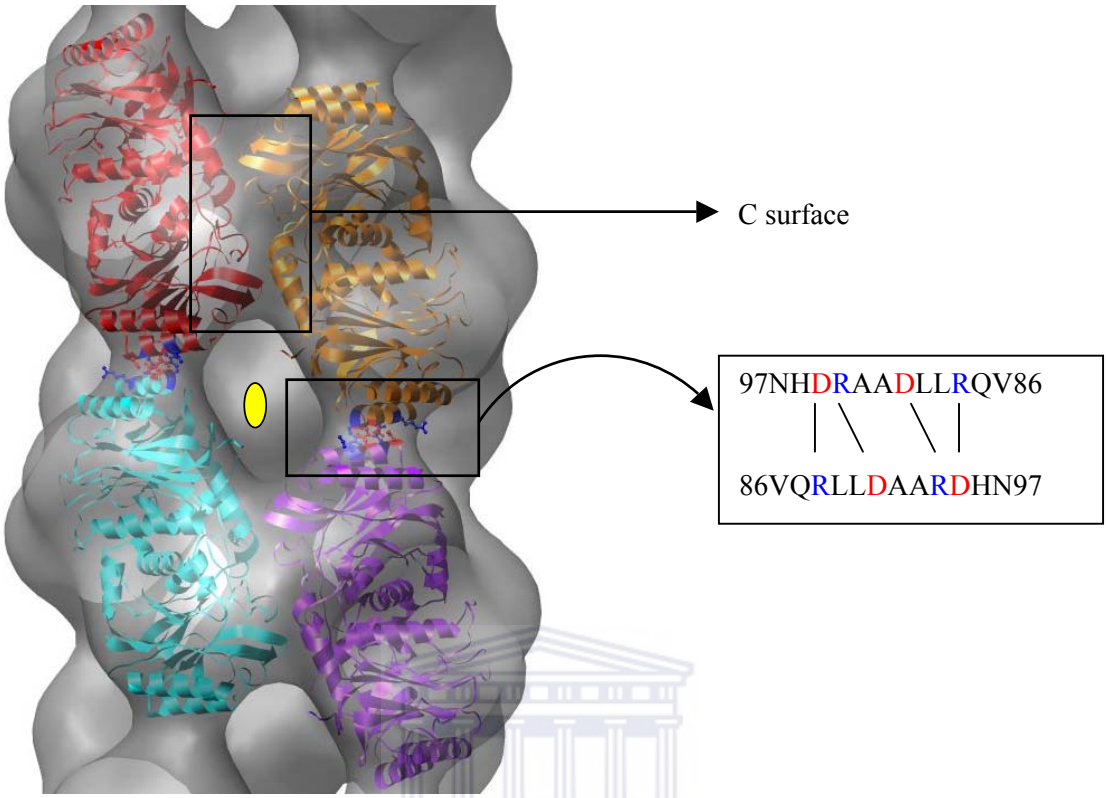


Figure 5.18. Symmetric interactions between highlighted residues (black rectangle) result in formation of four salt bridges (dashed lines) across the D surface. These residues may be involved in maintaining the helix, whereas those interacting at the C surface (deleted insertions not shown) form the helix. A horizontal dyad (shown in yellow) occurs at the hole passing through the A surface on the other side of the 3D map.

5.8 Discussion

The 3D map of the nitrilase from *Rhodococcus rhodochrous* J1 presented here is the first to be determined among the Rhodococcal nitrilases. Long regular filaments have previously been reported in the cyanide dihydratase from *B.pumilus* C1 (Sewell et al., 2002) and the cyanide hydratase from *G.sorghi* (Price et al. 2002). The structure is a regular, two-fold symmetric open helix made up of subunits interacting via the A, C and D surfaces. The handedness of the 1-start helix is based on the experimental results of Jandhyala et al (2003). This handedness is also coincident with the location of the insertion leading to the C surface as discussed by Sewell et al (2003, 2005).

Members in the nitrilase superfamily are widely reported to exist as inactive dimers, which must associate into a larger assembly and in the process become active. The association of

monomers occurs via the A surface, which is conserved in the crystal structures of 1ems, 1erz, 1f89 and 1j31. These dimers are primary building blocks towards higher oligomers via interactions in the C and D surfaces, which are not present in the solved structures. A comparative model of the dimer based on the structures of the nitrilase homologues was successfully docked into the 3D map. Model fitting was unambiguous and in excellent agreement with the structural interpretation of the cyanide dihydratase from *P.stutzeri* AK61 (Sewell et al., 2003). This revealed the two-fold axis of the dimer located at the A surface coincides with a horizontal dyad perpendicular to the helix axis, suggesting this surface is common among nitrilases. Some empty density is visible on the inside of the negatively stained map, which can accommodate an extra 14 residues (from each monomer) missing in the C terminal (that is, the homology model has 313 amino acids whereas the helix forming dimer has ~ 327 residues following the loss of 39 amino acids). As earlier suggested (see chapter 2), a beta sheet in the C terminal (similar to that present in 1ems) may be involved in interactions at the A surface. The loss of 39 amino acids in the C terminal, suggests J1 nitrilase does not require most of its C terminal for activity. Similarly, the cyanide dihydratase of *B.pumilus* C1 is reported to remain active after truncations of up to 40 amino acids in its C terminal (Sewell et al, 2005). This suggests there are other interactions responsible for maintaining a helix with a truncated C terminal (see below).

Helix formation is made possible by interactions at the C surface, which is observed to be the main point of connectivity forming an open left-handed 1-start helix (Figure 5.16(b)). A region of empty density coincident with the dyad axis is observed on the outside of the map (figure 5.16(c)). This density can accommodate amino acids (which were deleted in the model prior to docking) corresponding to significant residue insertions in its sequence between 54-73 and 234-249 relative to the solved structures. Although these residues are yet to be located at atomic resolution, deletion of the loop corresponding to 234-279 in the cyanide dihydratases rendered them inactive (Sewell et al., 2005). The symmetric interactions at the C surface are positioned about a line passing through a horizontal dyad located on the other side of the map at the D surface. There is possibility of four symmetric salt bridges across the groove at the D surface (figure 5.18) between residues located on an alpha helix corresponding to NH2 in Nit. These interactions are analogous to those observed in the cyanide dihydratases, which when modified did not affect the activity in the enzyme from *B.pumilus* C1 (Sewell at al., 2005). Although the specific residues that interact symmetrically across this surface are significantly different, it would seem they are required in maintaining the helix. Even though helix

formation occurs following loss of part of the C terminal residues and interactions across the C and D surfaces, it remains to be discovered what role the J1 nitrilase helix plays *in vivo*.

The location of the catalytic residues (E48, K131 and C165) is shown in figure 5.16(a) and these residues were seen to be buried when the model was docked into the 3D map. As seen in figure 5.16 (c), the central core of the helix is empty, suggesting there is a channel running through the entire structure. The purpose of this channel and how the substrates access or exit the active site pocket is not known. The crystal structures of 1ems (Pace et al, 2000) and 1erz (Nakai et al, 2000) form tetramers via interactions at the B surface which is comprised of two beta sheets corresponding to NS11 and NS12 in 1ems. This surface is seen to be located on the outside of the helix, where it could act as docking point for other enzymes forming a larger complex.

5.9. Conclusion

The helical structure of the nitrilase from *R.rhodochrous* J1 has been determined at a resolution 1.7nm. This enzyme forms a regular two-fold symmetric filament with a diameter of 13nm and a pitch of 77.23 Å. The symmetry of the helix is $D_1S_{\sim 4.89}$, a one-start left-handed helix with 44 dimers equally spaced along nine turns. The clarity of the 3D map shows clearly the symmetry and shape of subunits and has led to the identification of two surfaces involved in forming and maintaining the helix. These surfaces have previously been identified in other members of the nitrilase superfamily (Sewell et al., 2005), and suggest this may be a common oligomerization principle among microbial nitrilases. The functional significance of such a structure is yet unknown, however, biotechnological applications of artificially made helices are suggested. A longer enzyme suggests there are more active sites for hydrolyzing a whole range of substrates. Furthermore, these homo-oligomers are robust, easy to purify, can be stored for long periods and easily immobilized. More insights into the functional role of this structure among members in the superfamily could be gained through the determination of a high-resolution structure. The helix has approximately 5-fold symmetry, which has implications for X-ray studies because only proteins with 2, 3, 4 and 6 fold symmetry will crystallize. While crystallization and X-ray structure determination of J1 nitrilase remains elusive, a 3D map of the complex at higher resolution than that presented here will be determined using cryo-electron microscopy in the immediate future.

CHAPTER 6
SUMMARY AND CONCLUSIONS

6.1 Summary

6.1.1 Bioinformatics

Structural alignment of 20 sequences of biochemically characterized nitrilases against crystal structures of distant homologues, namely 1ems, 1erz, 1f89 and 1j31 revealed microbial nitrilases have extensions of 9-80 amino acids in their C terminal, which has been implicated in interactions at the dimer interface referred to as the A surface. Modeling based on homology of the J1 nitrilase sequence to atomic homologues revealed all structures have two-fold symmetry which conserves the $\alpha\beta\beta\alpha$-$\alpha\beta\beta\alpha$ fold across the A surface. The atomic homologues exist as dimers or tetramers whereas microbial nitrilases are found as high molecular weight homo-oligomers. The J1 nitrilase has two significant insertions of between 12-14 residues in its sequence relative to the solved structures, which possibly correspond to the C surface. In addition, intermolecular interactions previously identified in the spiral structures of the cyanide dihydratases from *B.pumilus* C1 and *P.stutzeri* AK61 (Sewell et al., 2003; 2005), which correspond to the C and D surfaces also occur in J1 nitrilase. The C and D surfaces, which appear conserved among microbial nitrilases, are responsible for spiral formation and absent in the atomic homologues. The dimer model of J1 nitrilase revealed the active site residues are buried, was useful in predicting residues that line the substrate-binding pocket and also in the design of mutants bearing modifications in the spiral forming interfaces.

6.1.2 Purification and negative stain EM

The recombinant J1 nitrilase was expressed in *E.coli* BL21 and subsequently purified by ion exchange and gel filtration chromatography. The enzyme was eluted in the absence of substrate as two distinct and apparently stable oligomeric forms corresponding to an active 480 kDa and an 80 kDa state. This observation contradicts those of Nagasawa et al (2000) and suggests the 480 kDa oligomer is not isolatable from the native organism. Reducing SDS-PAGE showed a single band of ~40 kDa. Preliminary circular dichroism spectroscopy showed the enzyme undergoes structural change in the presence of benzonitrile, thus confirming there is correspondence between oligomeric state and activity. Negative stain EM and image analysis revealed the active oligomer forms distinct particles of various shapes and sizes, including c-shaped particles (or open rings) that could not form spirals.

The chromatographic re-elution of an active fraction of 1-month old J1 nitrilase revealed an active form with a mass >1.5 MDa. Reducing SDS-PAGE and mass spectroscopy showed the subunit atomic mass was 36.5 (±0.6) kDa as a result of proteolysis whereas N-terminal sequencing confirmed that no amino acids were cleaved from this end. EM revealed a novel quaternary structure, that is, enzymatically active long fibres of J1 nitrilase.

This immediately raises the question as to why the c-shaped particles could not form a closed ring or a spiral structure. Previous structural studies have shown that the C-terminal tail of each dimer is located facing the inside of the spiral structure (Sewell et al., 2003). The J1 nitrilase has 53 extra residues in its C terminal relative to the crystal structures. The evidence presented suggests the loss of part of the C terminal leads to formation of long fibres. Therefore, inability of J1 nitrilase to close the ring and form at least one complete turn of the spiral arises from steric hindrances between these extra residues in its C terminal. Furthermore, failure to close the ring would prevent interactions at the D surface required to maintain a long regular helix.

6.1.3 3D reconstruction and interpretation of structure

Micrographs (109) were collected, short filament segments (13,506) were selected and subsequently, two independent 3D reconstructions were generated using the IHRSR algorithm (Egelman, 2000) and executed in SPIDER (Frank et al., 1996). Starting from different references, both independent reconstructions converged to the same structure and the resolution using FSC 0.5 criterion was 17 Å. The helix structure has $D_1S_{\sim 4.9}$ line group symmetry and a diameter of 13nm with ~5 dimers per turn in a pitch of 77.23 Å. The homology model of J1 nitrilase was located into the EM map for subsequent structural interpretation. This revealed the helix is built primarily from dimers that interact via the C and D surfaces. This interlocking arrangement between the dimers leading to a formation of a long spiral is also conserved in the helical structure of the enzyme from *B.pumilus* C1. The residues that potentially interact across the D surface in J1 nitrilase have been identified and these have previously been shown to confer stability to the helix (Sewell et al., 2005). Evidence suggests that conservation of the insertions and the possibility of salt bridge formation across the D surface is common among microbial nitrilases. In addition, it is predicted that removal of part of the C terminal would result in regular long helices, which have potential biotechnological applications.

6.2 Future work and conclusion

The key outcome is the helical form of the nitrilase from *Rhodococcus rhodochrous* J1 and evidence presented suggests this may be a common form among this important class of industrial enzymes. The J1 nitrilase helix structure was interpreted at an atomic level using a combination of homology modeling and 3D electron microscopy giving insights into the process of oligomerization. The active enzyme has ~5-fold symmetry along the helix axis, which suggests that the native enzyme cannot be crystallized. Constructs of this enzyme bearing modifications of residues in the A, C and D surfaces and truncations in its C terminal have already been prepared. The aim is to seek further structural insights into the activity and oligomerization process of this Rhodococcal nitrilase. A preliminary investigation has confirmed recombinant J1 nitrilase bearing truncations in its C terminal does indeed form long fibres. Furthermore, the presence of a distinct band on SDS PAGE suggests the cleavage of amino acids from its C terminal is not random but arises from a specific proteolytic or autolytic mechanism. While the crystal structure of J1 nitrilase remains elusive, its potential as an enhanced biocatalyst is feasible. This is because long active fibres are easy to purify, would possess higher activity, can be easily immobilized and stored for long periods. This could potentially contribute to the establishment of a competitive South African biotechnology industry.

REFERENCES

Alamatawah QA, Cramp R, and Cowan DA (1999). Characterization of an inducible nitrilase from a thermophilic bacillus. *Extremophiles* 3, 283-291.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17), 3889 – 3402.

Baker TS and Johnson JE (1997). Principles of virus structure determination. In: Chiu, Burnett, Garcia (Eds.). *Structural Biology of Viruses*. Oxford University Press, Oxford, 38-55.

Baker D and Sali A (2001). Protein structure prediction and structural genomics. *Science* 294, 93-96.

Bandyopadhyay AK, Nagasawa T, Asano Y, Fujishiro K, Tani Y, Yamada H (1986). Purification and characterization of benzonitrilases from *Arthrobacter* sp. Strain J-1. *Applied and Environmental Microbiology* 51, 302-306.

Banerjee A, Sharma R and Banerjee UC (2002). The nitrile-degrading enzymes: current status and future prospects. *Applied Microbiology and Biotechnology* 60, 30-44.

Barclay M, Tett VA and Knowles CJ (1998). Metabolism and enzymology of cyanide/metallocyanide biodegradation by *Fusarium solani* under neutral and acidic conditions. *Enzyme and Microbial Technology* 23, 321-330.

Barclay M, Day JC, Thompson IP, Knowles CJ, and Bailey MJ (2002). Substrate-regulated cyanide hydratase (chy) gene expression in *Fusarium solani*: the potential of a transcription-based assay for monitoring the biotransformation of cyanide complexes. *Environmental Microbiology* 4, 183-189.

Bartling D, Seedorf M, Mithofer A, and Weiler EW (1992). Cloning and expression of an *Arabidopsis* nitrilase which can convert Indole-3-acetonitrile to the plant hormone, Indole-3-acetic acid. *European Journal of Biochemistry* 205(1), 417-424.

Bartling D, Seedorf M, Schimdt RC, and Weiler EW (1994). Molecular characterization of two cloned nitrilases from *Arabidopsis thaliana*: key enzymes in biosynthesis of plant hormone indole -3- acetic acid. *Proceedings of the National Academy of Sciences* USA 91, 6021-5.

Bhalla TC, Miura A, Wakamoto A, Ohba Y, and Furuhasi K (1992). Asymmteric hydrolysis of α-aminonitriles to optically active aminoacids by a nitrilase of *Rhodococcus rhodochrous* PA-34. *Applied Microbiology and Biotechnology* 37, 184-190.

Bellinzoni M, Buroni S, Pasca MR, Guglierame P, Arcesi F, De Rossi E and Riccardi G (2005). Glutamine amidotransferance activity of NAD+ synthetase from *Mycobacterium tuberculosis* depends on an amino-terminal nitrilase domain. *Research in Microbiology* 156, 173-177.

Berman MN (2003). Quaternary structures of the cyanide dihydratases of *Bacillus pumilus* C1 and *Pseudomonas stutzeri* AK61. M.Sc. Thesis, University of Cape Town.

Bolleter WT, Bushman CJ, Tidwell PW (1961). Spectrophotometric determination of ammonia as indophenol. *Analytical Chemistry* 33, 592-594.

Bork P and Koonin EV (1994). A new family of Carbon-Nitrogen hydrolases. *Protein Science* 3(8), 1344-6.

Bottcher B, Wynne SA, Crowther RA (1997). Determination of the fold of the core protein of hepatitis B virus by electron microscopy. *Nature* 365(6620), 88-91

Bower JM, Cohen FE and Dunbrack Jr RL (1997). Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: A new homology modeling tool. *Journal of Molecular Biology* 267, 1268-1282.

Bradford MM (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72, 248-254.

Brady D, Beeton A, Zeevaart J, Kgaje C, van Rantwijk F and Sheldon RA (2004). Characterisation of nitrilase and nitrile hydratase biocatalytic systems. *Applied Microbiology and Biotechnology* 64, 76-85.

Brenner C (2002). Catalysis in the nitrilase superfamily. *Current Opinion in Structural Biology* 12, 775-782.

Burk MJ (2002). Discovery of novel enzymes from natural diversity: Development of nitrilases for production of pharmaceutical intermediates. *The Pharmaceutical Fine Chemicals and Manufacturing report* © PharmaVentures Ltd.

Canutescu AA, Shelenkov AA and Dunbrack Jr RL (2003). A graph-theory algorithm for rapid protein side-chain prediction. *Protein Science* 12, 2001-2014.

Chauhan S, Wu S, Blumerman WS, Fallon RD, Gavagan JE, DiCosimo R, and Payne MS (2003). Purification, cloning, sequencing and over-expression in *Escherichia coli* of a regioselective aliphatic nitrilase from *Acidovorax facilis* 72W. *Applied Microbiology and Biotechnology* 61, 118-122.

Chen C, Chiu W, Liu J, Hsu W, and Wang W (2003). Structural Basis for Catalysis and Substrate Specificity of *Agrobacterium radiobacter* N-Carbamoyl-D-amino Acid Amidohydrolase. *Journal of Biological Chemistry* 278(28), 26194–26201.

Cluness MJ, Turner PD, Clements E, Brown DT and O'Reilly C (1993). Purification and properties of cyanide hydratase from *Fusarium lateritium* and analysis of the corresponding chy1 gene. *Journal of General Microbiology* 139, 1807-1815.

Cohen GH (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *Journal of Applied Crystallography* 30, 1160-1161.

Conway JF, Cheng N, Zlotnick A, Wingfield PT, Stahl SJ and Steven AC (1997). Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature* 386(6620), 91-94.

DeLano WL (2002). The PyMOL molecular graphics system. DeLano Scientific LLC, San Francisco, USA (http://www.pymol.org).

DeRosier DJ and Klug A (1968). Reconstruction of three-dimensional structures from electron micrographs. *Nature* 217, 130-134.

DeSantis G, Zhu Z, Greenberg WA, Wong K, Chaplin J, Hanson SR, Farewell B, Nicholson LW, Rand CL, Weiner DP, Robertson DE, and Burk MJ (2002). An enzyme library approach to biocatalysis: Development of nitrilases for enantioselective production of carboxylic acid derivatives. *Journal of American Chemical Society* 124, 9024-9025.

Dhillon JK and Shivaraman N (1999a). Biodegradation of cyanide compounds by a *Pseudomonas* species S1. *Canadian Journal of Microbiology* 45(3), 201-208.

Dhillon JK, Chatre S, Shanker R and Shivaraman N (1999b). Transformation of aliphatic and aromatic nitriles by a nitrilase from *Pseudomonas* sp. *Canadian Journal of Microbiology* 45(10), 811-815.

Dubey SK and Holmes DS (1995). Biological cyanide destruction mediated by microorganisms. *World Journal of Microbiology and Biotechnology* 11, 257-265.

Ebbs S (2004). Biological degradation of cyanide compounds. *Current Opinion in Biotechnology* 15, 231-236.

Effenberger E and Osswald S (2001). Selective hydrolysis of aliphatic dinitriles to monocarboxylic acids by a nitrilase from *Arabidopsis thaliana*. *Synthesis* 12, 1866-1872.

Egelman EH, Francis N and DeRosier DJ (1982). F-actin is a helix with a random variable twist. *Nature* 298, 131-135.

Egelman EH (1986). An algorithm for straightening images of curved filamentous structures. *Ultramicroscopy* 19, 367-373.

Egelman EH and Stasiak A (1988). Structure of helical RecA-DNA complexes.II. Local conformational changes visualized in bundles of RecA-ATP-γ-S filaments. *Journal of Molecular Biology* 200(2), 329-349.

Egelman EH (2000). A robust algorithm for the reconstruction of helical filaments using single particle methods. *Ultramicroscopy* 85, 225-234.

Eisenberg D, Luthy R and Bowie JU (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in Enzymology* 277, 396-404.

Eswar N, Bino J, Mirkovic N, Fiser A, Ilyin VA, Pieper U, Stuart AC, Marti-Renom MA, Madhusudhan MS, Yerkovich B and Sali A (2003). Tools for comparative protein structure modeling and analysis. *Nucleic Acids Research* 31(13), 3375-3380.

Farnaud S, Tata R, Sohi MK, Wan T, Brown PR and Sutton BJ (1999). Evidence that cysteine-166 is the active site nucleophile of *Pseudomonas aeruginosa* amidase: crystallization and preliminary X-ray diffraction analysis of the enzyme. *Biochemical Journal* 340,711-714.

Fiser A (2004). Protein structure modeling in the proteomics era. *Expert Review of Proteomics* 1(1), 97-110.

Fiser A and Sali A (2003). MODELLER: generation and refinement of homology-based protein structure models. *Methods in Enzymology* 374, 463-493.

Frank J (1989). Images analysis of Single macromolecules. *Electron Microscopy Review* 2, 53-74.

Frank J, Radermacher M, Penczek P. Zhu J, Li Y, Ladjadj M and Leith A (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *Journal of Structural Biology* 116(1), 190-199.

Frank J (2002). Single – Particle imaging of macromolecules by cryo-electron microscopy. *Annual Reviews of Biophysics and Biomolecular Structure* 31, 303-319.

Fry WE and Millar RL (1972). Cyanide degradation by an enzyme from *Stemphylium loti*. *Archives of Biochemistry and Biophysics* 151, 468-474.

Fry WE and Munch DC (1975). Hydrogen cyanide detoxification by *Gloeocercospora sorghi*. *Physiological Plant Pathology* 7, 23-33.

Gagavan JE, DiCosimo R, Eisenberg A, Fager SK, Folsom PW, Hahn EC, Schneider KJ and Fallon RD (1999). A gram-negative bacterium producing a heat-stable nitrilase highly active on aliphatic dinitriles. *Applied Microbiology and Biotechnology* 52, 654-659.

Gradley ML and Knowles CJ (1994). Asymmetric hydrolysis of chiral nitriles by *Rhodococcus rhodochrous* NCIMB 11216 nitrilase. *Biotechnology Letters* 16(1), 41-46.

Goddard TD, Huang CC and Ferrin TE (2005). Software extensions to UCSF Chimera for interactive visualization of large molecular assemblies. *Structure* 13(3), 473-482.

Goldlust A and Bohak Z (1989). Induction, purification, and characterization of the nitrilase of *Fusarium oxysporum* f.sp *melonis*. *Biotechnology and Applied Biochemistry* 11, 581-601.

Harauz G and van Heel MG (1986). Exact filters for general geometry three-dimensional reconstructions. *Optik* 73, 146-156.

Harper DB (1977a). Fungal degradation of aromatic nitriles: enzymology of C-N cleavage by *Fusarium solani*. *Biochemical Journal* 167, 685-692.

Harper DB (1977b). Microbial metabolism of aromatic nitriles: enzymology of C-N cleavage by *Norcadia* sp. (*Rhodochrous* group) NCIB 11216. *Biochemical Journal* 165, 309-319.

Harper DB (1985). Characterization of a nitrilase from *Norcadia* sp. (*Rhodochrous* group) NCIB 11215, using *p*-hydroxybenzonitrile as sole carbon source. *International Journal of Biochemistry* 17(6), 677-683.

Hann EC, Sigmund AE, Fager SK, Cooling FB, Gavagan JE, Bramucci MG, Chauhan S, Payne MS and DiCosimo R (2004) Regioselective biocatalytic hydrolysis of (E, Z)-2-methyl-2-butenenitrile for production of (E)-2-methyl-2-butenoic acid. *Tetrahedron* 60, 577–581.

Heinemann U, Engels D, Burger S, Kiziak C, Mattes R, and Stolz A (2003). Cloning of a nitrilase gene from the cyanobacterium *Synechocystis* sp. Strain PCC6803 and heterologous expression and characterization of the encoded protein. *Applied and Environmental Microbiology* 69(8), 4359-4366.

Hiller A and Van slyke D (1933). Determination of ammonia in blood. Journal of Biological Chemistry 102, 499.

Hooft RW, Vriend G, Sander C and Abola EE (1996). Errors in protein structures. *Nature* 381 (6580), 272.

Hook RH and Robinson WG (1964). Ricinine Nitrilase II: Purification properties. *Journal of Biological Chemistry* 239(12), 4263-4267.

Hoyle AJ, Bunch AW, and Knowles CJ (1998). The nitrilases of *Rhodococcus rhodochrous* NCIMB 11216. *Enzyme and Microbial Technology* 23, 475-482.

Ingvorsen K, H∅jer-Pedersen B, and Godtfredsen SE (1991). Novel cyanide hydrolysiing enzyme from *Alcaligenes xylosoxidans* subsp. *denitrificans*. *Applied and Environmental Microbiology* 57(6), 1783-1789.

Jandhyala D, Berman M, Meyers PR, Sewell BT, Willson RC, and Benedik MJ (2003). Cyn D, the cyanide dihydratase from *Bacillus pumillus*: Gene cloning and structural studies. *Applied and Environmental Microbiology* 69(8), 4794-4805.

Jandhyala DM, Wilson RC, Sewell BT, and Benedik MJ (2005). Comparison of cyanide degrading nitrilases. *Applied Microbiology and Biotechnology* 68(3), 327-335.

Jeanmougin F, Thompson JD, Gouy M, Higgins DG and Gibson TJ (1998). Multiple sequence alignment with ClustalX. *Trends in Biochemical Sciences* 23, 403-405.

Jeng TW, Crowther RA, Stubbs G and Chiu W (1989). Visualization of alpha-helices in tobacco mosaic virus by cryo-electron microscopy. *Journal of Molecular Biology* 205 (1), 251-257.

Jiang W and Ludtke SJ (2005). Electron cryomicroscopy of single particles at subnanometer resolution. *Current Opinion In Structural Biology* 15, 571-577.

Johnson WC (1990). Protein Secondary Structure and Circular Dichroism: A Practical Guide. *PROTEINS: Structure, Function and Genetics* 7, 205-214.

Jones DT (1999). GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. *Journal of Molecular Biology* 287, 797-815.

Kato Y, Nakamura K, Sakiyama H, Mayhew SG, and Asano Y (2000). Novel Heme-containing Lyase, Phenylacetaldoxime Dehydratase from *Bacillus sp*. Strain OxB-1: Purification, Characterization, and Molecular cloning of the gene. *Biochemistry* 39, 800-809.

Kiziak C, Conradt D, Stolz A, Mattes R and Klein J (2005). Nitrilase from *Pseudomonas fluorescens* EBC191: cloning and heterologous expression of the gene and biochemical characterization of the recombinant enzyme. *Microbiology* 151, 3639-3648.

Kobayashi M, Nagasawa T, and Yamada H (1988). Regiospecific hydrolysis of dinitrile compounds by nitrilase from *Rhodococcus rhodochrous* J1. *Applied Microbiology and Biotechnology* 29, 231-233.

Kobayashi M, Nagasawa T, and Yamada H (1989). Nitrilase of *Rhodococcus rhodochrous* J1: Purification and characterization. *European Journal of Biochemistry* 182, 349-356.

Kobayashi M, Yanaka N, Nagasawa T, and Yamada H (1990). Purification and characterization of a novel nitrilase of *Rhodococcus rhodochrous* K22 that acts on aliphatic nitriles. *Journal of Bacteriology* 172 (9), 4807-4815.

Kobayashi M, Yanaka N, Nagasawa T and Yamada H (1992a). Primary structure of an aliphatic nitrile-degrading enzyme, aliphatic nitrilase, from *Rhodococcus rhodochrous* K22 and expression of its gene and identification of the active site residue. *Biochemistry* 31, 9000-9007.

Kobayashi M, Komeda H, Yanaka N, Nagasawa T and Yamada H (1992b). Nitrilase from *Rhodococcus rhodochrous* J1: Sequencing and overexpression of the gene and identification of an essential cysteine residue. *Journal of Biological Chemistry* 267(29), 20746-20751.

Kobayashi M, Izui H, Nagasawa T and Yamada H (1993). Nitrilase in the biosynthesis of the plant hormone indole-3-acetic acid from indole-3-acetonitrile: Cloning of the *Alcaligenes* gene and site-directed mutagenesis of cysteine residues *Proceedings of the National Academy of Sciences* USA 90, 247-251.

Kobayashi M, Suzuki T, Fujita T, Masuda M, and Shimizu S (1995). Occurrence of enzymes involved in biosynthesis of Indole-3-acetic acid from Indole-3-acetonitrile in plant-associated bacteria, *Agrobacterium* and *Rhizopium*. *Proceedings of the National Academy of Sciences* USA, 92(3), 714-718.

Kobayashi M, Fujiwara Y, Goda M, Komeda H, and Shimizu S (1997). Identification of active sites in amidase: Evolutionary relationship between amide bond- and peptide bond-cleaving enzymes. *Proceedings of the National Academy of Sciences* USA 94, 11986-11991.

Kobayashi M, Goda M, and Shimizu S (1998a). The catalytic mechanism of amidase also involves nitrile hydrolysis. *FEBS letters* 439, 325-328.

Kobayashi M, Goda M, and Shimizu S (1998b). Nitrilase catalyzes amide hydrolysis as well as nitrile nydrolysis. *Biochemical and Biophysical Research Communications* 253, 662-668.

Kobayashi M and Shimizu S (2000). Nitrile Hydrolases. *Current Opinion in Chemical Biology* 4, 95-102.

Kumaran D, Eswaramoorthy S, Gerchman SE, Kycia H, Studier FW, and Swaminathan S (2003). Crystal structure of putative CN hydrolase from yeast. *PROTEINS: Structure, Function, and Genetics* 52, 283–291.

Kunz DA, Wang CS, and Chen JL (1994). Alternative routes of enzymatic cyanide metabolism in *Pseudomonas fluorescens* NCIMB 11764. *Microbiology* 140, 1705-1712.

Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* 227(5259), 680-685.

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

Layh N, Paratt J and Willets A (1998). Characterization and partial purification of an enantioselective arylacetonitrilase from *Pseudomonas fluorescens* DSM7155. *Journal of Molecular Catalysis B: Enzymatic* 5(5-6), 467-474.

Levy-Schil S, Soubrier F, Crutz-le Coq AM, Faucher D, Crouzet J and Petre D (1995). Aliphatic nitrilase from a soil-isolated *Comamonas testoteroni* sp: gene cloning and over expression, purification and primary structure. *Gene* 161, 15-20.

Lovell SC, Davis IW, Arendall WB 3rd, deBakker PW, Word JM, Prisant MG, Richardson JS and Richardson DC (2003). Structure validation by Cα geometry: Phi, Psi and Cβ deviation. *PROTEINS: Structure, Function and Genetics* 50(3), 437-450.

Ludtke SJ, Baldwin PR and Chiu W (1999). EMAN: semi-automated software for high-resolution single particle reconstructions. *Journal of Structural Biology* 128(1), 82-96.

Ludtke SJ, Chen DH, Song JL, Chuang DT and Chiu W (2004). Seeing GroEL at 6Å resolution by single particle electron cryomicroscopy. *Structure* 12(7), 1129-1136.

Makowski L and Caspar DLD (1980). The symmetries of filamentous phage particles. Journal of Molecular Biology 145, 611-617.

Martinkova L and Kren V (2002). Nitrile- and Amide-converting microbial enzymes: Stereo-, Regio- and Chemoselectivity. *Biocatalysis and Biotransformation* 20(2), 73-93.

Marti-Renom MA, Madhusudhan MS, Fiser A and Sali A (2002). Reliability of assessment of protein structure prediction methods. *Structure* 10(3), 435-440.

Mathew CD, Nagasawa T, Kobayashi M and Yamada H (1998). Nitrilase catalyzed production of nicotinic acid from 3-cyanopyridine in *Rhodococcus rhodochrous* J1. *Applied Environmental Microbiology* 54 (3), 1030-1032.

McGuffin LJ, Bryson K and Jones DT (2000). The PSIPRED protein structure prediction server. *Bioinformatics* 16(4), 404-405.

McGuffin LJ and Jones DT (2003). Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* 19(7), 874-881.

Meyers PR, Rawlings DE, Woods DR, and Lindsey GG (1993). Isolation and characterization of a cyanide dihydratase from *Bacillus pumilus* C1. *Journal of Bacteriology* 175(19), 6105-6112.

Miyazawa A, Fujiyoshi Y, Stowell M and Unwin N (1999). Nicotinic acetylcholine receptor at 4.6Å resolution: transverse tunnels in the channel wall. *Journal of Molecular Biology* 288, 765-786.

Miyazawa A, Fujiyoshi Y, Stowell M and Unwin N (2003). Structure and gating mechanism of the acetylcholine receptor pore. *Nature* 424, 949-955.

Mizuguchi K, Deane CM, Blundell TL, Johnson MS and Overington JP (1998). JOY: protein sequence-structure alignments representation and analyis. *Bioinformatics* 14, 617-623.

Mylerova V and Martinkova L (2003). Synthetic applications of Nitrile-Converting enzymes. *Current Organic Chemistry* 7(14), 1-17.

Nagasawa T, Kobayashi M and Yamada H (1988). Optimum culture conditions for the production of benzonitrilase by *Rhodococcus rhodochrous* J1. *Archives of Microbiology* 150, 89-94.

Nagasawa T, Mauger J and Yamada H (1990). A novel nitrilase, arylacetonitrilase, of *Alcaligenes faecalis* JM3. Purification and characterization. *European Journal of Biochemistry* 194, 765-772.

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(3) 581-589.

Nagasawa T, Wieser M, Nakamura T, Iwahara H, Yoshida T, and Gekko K (2000). Nitrilase of *Rhodococcus rhodochrous* J1: Conversion into the active form by subunit association. *European Journal of Biochemistry* 267, 138-144.

Nakai T, Hasegawa T, Yamashita E, Yamamoto M, Kumasaka T, Ueki T, Nanba H, Ikenaka Y, Takahashi S, Sato M, and Tsukihara T (2000). Crystal structure of *N*-carbamyl-D-amino acid amidohydrolase with a novel catalytic framework common to amidohydrolases. *Structure* 8(7), 729–739.

Namba K, Pattanayek R and Stubbs G (1989). Visualization of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9Å resolution by X-ray fiber diffraction. *Journal of Molecular Biology* 208, 307-325

Nogales E and Grigorieff N (2001). Molecular machines: putting the pieces together. *The Journal Of Cell Biology* 152(1), F1-F10.

Nolan LM, Harnedy PA, Turner P, Hearne AB and O'Reilly C (2003). The cyanide hydratase enzyme of *Fusarium lateritium* also has nitrilase activity. *FEMS Microbiology Letters* 221, 161-165.

Novo C, Tata R, Clemente A and Brown PR (1995). *Pseudomonas aeruginosa* aliphatic amidase is related to the nitrlilase/cyanide hydrates enzyme family and Cys166 is predicted to be the active site nucleophile of the catalytic mechanism. *FEBS Letters* 367,275-279.

Novo C, Farnaud S, Tata R, Clemente A and Brown PR (2002). Support for three-dimensional structure predicting a Cys-Glu-Lys catalytic triad for *Pseudomonas aeruginosa* amidase comes from site-directed mutagenesis and mutations altering substrate specifity. *Biochemical Journal* 365, 731-738.

Ohi M, Li Y, Cheng Y and Walz T (2004). Negative staining and image classification – Powerful tools in modern electron microscopy. *Biological Procedures Online* 6(1), 23-34.

O'Reilly C and Turner PD (2003). The nitrilase family of CN hydrolyzing enzymes – a comparative study. *Journal of Applied Microbiology* 95, 1161-1174.

Osswald S, Wajant H, and Effenberger F (2002). Characterization and synthetic applications of recombinant AtNIT1 from *Arabidopsis thaliana*. *European Journal of Biochemistry* 269, 680-687.

Pace HC, Hodawadekar SC, Draganescu A, Huang J, Bieganowski P, Pekarsky Y, Croce C.M, and Brenner C (2000). Crystal structure of the worm NitFhit Rosetta Stone protein reveals a Nit tetramer binding two Fhit dimmers. *Current Biology* 10, 907–917.

Pace H and Brenner C (2001). The nitrilase superfamily: classification, structure and function. *Genome Biology* 2(1), reviews 001.1-0001-9.

Pearson WR (2000). Flexible sequence similarity searching with FASTA3 program package. *Methods in Molecular Biology* 132, 185-219.

Penczek P, Radermacher M and Frank J (1992). Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy* 40(1), 33-53.

Penczek P, Grassucii RA, Frank J (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* 53(3), 251-270.

Penczek P (1996). Appendix: Rotationally invariant K-means algorithm. *Ultramicroscopy* 63, 205-218.

Penczek P (1998). Appendix: Measures of resolution using Fourier Shell Correlation. *Journal of Molecular Biology* 280, 115-116.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE (2004). UCSF Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605-1612.

Piotrowski M, Schonfelder S, and Weiler EW (2001). The *Arabidopsis thaliana* isogene *NIT4* and its orthologs in tobacco encode β-cyano-L-alanine hydratase/Nitrilase. *The Journal of Biological Chemistry* 276(4), 2616-2621.

Price B, Chang P, Jandhyala DM, Benedik M and Sewell BT (2002). The quaternary structure of *Gloeocercospora sorghi* nitrilase (cyanide dihydratase) as revealed by negative staining. *Proceedings of the 15th International Congress on Electron Microscopy* 2, 565-566.

Radermacher M, WagenKnecht T, Verschoor A and Frank J (1987). Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *Journal of Microscopy* 146, 113-136.

Radermacher M (1998). Three-dimensional reconstruction of single particles from random and non-random tilt series. *Journal of Electron Microscopy and Technology* 9(4), 359-394.

Ramakrishna C, Dave H, and Ravindranathan M (1999). Microbial Matabolism of nitriles and its biotechnological potential. *Journal of Scientific and Industrial Research* 58 (12), 925-947.

Robertson DE, Chaplin JA, DeSantis G, Podar M, Madden M, Chi E, Richardson T, Milan A, Miller M, Weiner DP, Wong K, McQuaid J, Farwell B, Preston LA, Tan X, Snead MA, Keller

M, Mathur E, Kretz PL, Burk MJ, and Short JM (2004). Exploring nitrilase sequence space for anantioselective catalysis. *Applied and Environmental Microbiology* 70(4), 2429-2436.

Ruprecht J and Nield J (2001). Determining the structure of biological macromolecules by transmission electron microscopy, single particle analysis and 3D reconstruction. *Progress in Biophysics and Molecular Biology* 75, 121-164.

Sakai N, Tajika Y, Yao M, Watanabe N and Tanaka I (2004). Crystal structure of the hypothetical protein pH0642 from *Pyrococcus Horikoshii* and structure-based prediction of enzymatic reaction. *RCSB Protein Data Bank* (1J31).

Sali A and Blundell TL (1993). Comparative protein modeling by satisfaction of spatial restraints. *Journal of Molecular Biology* 234, 779-815.

Sali A, Sanchez R and Badretinov A (2000). MODELLER: A protein structure modeling program release 6. (http://salilab.org).

Sanchez R and Sali A (2000). Comparative protein structure modeling: Introduction and practical examples with MODELLER. *Methods in Molecular Biology*143, 97-129.

Sanchez R, Pieper U, Melo F, Eswar N, Marti-Renom MA, Madhusudhan MS, Mirkovic N and Sali A (2000). Protein structure modeling for structural genomics. *Nature Structural Biology* 7, 986-990.

Schoehn G, Mavrakis M, Aurelie A, Wade R, Hoenger A and Ruigrok RWH (2004). The 12Å structure of Trypsin-treated Measles virus N-RNA. *Journal of Molecular Biology* 339, 301-312.

Sewell BT, Berman MN, Meyers P, Jandhyala DM and Benedik M (2002). PH dependent quaternary structural transitions in the cyanide degrading nitrilase from *Bacillus pumilus*. *Proceedings of the 15th International Congress on Electron Microscopy* 2, 569-570.

Sewell BT, Berman MN, Meyers PR, Jandhyala D, and Benedik M (2003). The cyanide degrading nitrilase from *Pseudomonas stutzeri* AK61 is a two-fold symmetric, 14-subunit spiral. *Structure* 11, 1-20.

Sewell BT, Thuku RN, Zhang X and Benedik M (2005). The oligomeric structure of nitrilases: the effect of mutating interfacial residues on activity. *Annals of New York Academy of Sciences* 1056, 153-159.

Sexton AC and Howlett BJ (2000). Characterisation of a cyanide hydratase gene in the phytopathogenic fungus *Leptosphaeria maculans. Molecular and General Genetics* 263, 463-470.

Shi J, Blundell TL and Mizuguchi K (2001). FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology* 310, 243-257.

Sippl MJ (1993). Recognition of errors in three-dimensional structures of proteins. *PROTEINS: Structure, Function and Genetics* 17(4), 355-362.

Stevenson DE, Feng R, and Storer AC (1990). Detection of covalent enzyme- substrate complexes of nitrilase by ion-spray mass spectroscopy. *FEBS Letters* 277(1-2), 112-114.

Stevenson DE, Feng R, Dumas F, Groleau D, Mihoc A and Storer AC (1992). Mechanistsic and structural studies on *Rhodococcus* ATCC 39484 nitrilase. *Biotechnology and Applied Biochemistry* 15, 283-302.

Tao Y and Zhang W (2000). Recent developments in cryoelectron microscopy reconstruction of single particles. *Current Opinion in Structural Biology* 10, 616-622.

Thuku RN and Sewell BT (2005). The Quaternary Structure of *Rhodococcus rhodochrous* J1 as revealed by negative staining. *Proceedings of the Microscopy Society of Southern Africa* 35, 31.

Thumann-Commike PA (2001). Single particle macromolecular structure determination via electron microscopy. *FEBS Letters* 505, 199-205.

Trachtenberg A, Galkin VE and Egelman EH (2005). Refining the structure of the *Halobacterium salinarum* flagellar filament using Iterative Helical Real Space Reconstruction method: Insights in polymorphism. *Journal of Molecular Biology* 346, 665-676.

van Heel M (1987). Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy* 21, 111-123.

van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M and Padwardhan A (2000). Single-Particle electron cryomicroscopy: towards atomic resolution. *Quaterly Reviews of Biophysics* 33(4), 307-369.

Vorwerk S, Biernacki S, Hillebrand H, Janzik I, Muller A, Weiler EW and Piotrowski M (2001). Enzymatic characterization of the recombinant *Arabidopsis thaliana* nitrilase subfamily encoded by the *NIT2/NIT1/NIT3*-gene cluster. *Planta* 212, 508-516.

Wang P, Mathews DE and VanEtten HD (1992a). Cloning and properties of a cyanide hydratase gene from the phytopathogenic fungus *Gloeocercospora sorghi*. *Biochemical and Biophysical Research Communications* 187, 1048-1054.

Wang P, Mathews DE and VanEtten HD (1992b). Purification and characterization of cyanide hydratase from the phytopathogenic fungus *Gloeocercospora sorghi*. *Archives of Biochemistry and Biophysics* 298, 569-575.

Wang W, Hsu W, Chien F, and Chen C (2001). Crystal structure and site-directed mutagenesis studies of N-Carbamoyl-D-amino-acid amidohydrolase from *Agrobacterium radiobacter* reveals a homotetramer and insight into a catalytic cleft. *Journal of Molecular Biology* 306, 251-261.

Watanabe A, Yano K, Ikebukuro K, and Karube I (1998a). Cyanide hydrolysis in a cyanide-degrading bacterium, *Pseudomonas stutzeri* AK61, by cyanidase. *Microbiology* 144, 1677-1682.

Watanabe A, Yano K, Ikebukuro K, and Karube I (1998b). Cloning and expression of a gene encoding cyanidase from *Pseudomonas stutzeri* AK61. *Applied Microbiology and Biotechnology* 50, 93-97.

Watanabe A, Yano K, Ikebukuro K, and Karube I (1998c). Investigation of the potential active site of a cyanide dihydratase using site-directed mutagenesis. *Biochimica et Biophysica Acta* 1382, 1-4.

Wyatt J and Knowles C (1995). Microbial degradation of acrylonitrile waste effluents and condensates from manufacture of acrylonitrile. *Int Biodeteriro Biodegrad* 35, 227-248.

Yamamoto K, Ueno Y, Otsubo K, Fujimatsu I and Komatsu K (1990). Production of S-(+)-Ibuprofen from a nitrile compound by *Acinetobacter* sp. Strain AK 226. *Applied and Environmental Microbiology* 56(10), 3125-3129.

Yamamoto K, Oishi K, Kawakami K, and Komatsu K (1990). Production of R-(-)-Mandelic acid from a Mandelonitrile by *Alcaligenes faecalis* ATCC 8750. *Applied and Environmental Microbiology* 57(10), 3028-3032.

Yamamoto K and Komatsu K (1991). Purification and characterization of nitrilase responsible for enantioselective hydrolysis from *Acinetobacter sp*. AK 226. *Agricultural and Biological Chemistry* 55,1459-1466.

Yamamoto K, Fijimatsu I and Komatsu K (1992). Purification and characterization of the nitrilase from *Alcaligenes faecalis* ATCC 8750 responsible for enantioselective hydrolysis of mandelonitrile. *Journal of Fermentation and Bioengineering* 73, 425-430.

Yanase H, Sakamoto A, Okamoto K, Kita K, and Sato Y (2000). Degradation of the metal-cyano complex tetracyanonickelate (II) by *Fusarium oxysporum* N-10. *Applied Microbiology and Biotechnology* 53, 328-334.

Yang S, Yu X, Galkin E and Egelman EH (2003). Issue of resolution and polymorphism in single particle reconstruction. *Journal of Structural Biology* 144, 162-171.

Yonekura K, Maki-Yonekura S and Namba K (2003). Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* 424, 643-650