Factors involved in the oligomerisation of cyanide dihydratase from *Bacillus pumilus* C1

by

Andani Errol Mulelu

A thesis submitted in fulfilment of the requirements

for the degree of Master of Science

in Medical Biochemistry

In the Department of Clinical Laboratory Sciences

UNIVERSITY OF CAPETOWN

March 2013

Supervisor: Prof B.T. Sewell

DECLARATION

I declare that *Factors involved in the oligomerisation of cyanide dihydratase from Bacillus pumilus C1* is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Andani Errol Mulelu

March 2013

Signed: Signed by candidate

ABSTRACT

The cyanide dihydratase enzyme from *Bacillus pumilus* is a member of the nitrilase superfamily which comprises the ubiquitous thiol enzymes that perform a wide variety of non-peptide carbon-nitrogen hydrolysis reactions. The cyanide dihydratase enzyme from *Bacillus pumilus* is known to specifically catalyse the conversion of cyanide into formic acid and ammonia. This characteristic makes the enzyme a good candidate for bioremediation of cyanide found in industrial waste waters produced from processes such as leach mining, metal finishing and electroplating. Bacterial detoxification could provide a cheaper and safer solution than the costly conventional chemical methods. The high alkaline pH of the cyanide waste water poses a problem in that it inactivates the wild type enzyme and therefore improvement of stability is required to synthesize an effective enzyme. The purpose of this project is to understand the structure and function of the cyanide dihydratase enzyme from *Bacillus pumilus* to help engineer mutants with improved stability. Several mutant enzymes were designed using structural knowledge and purified using a three-step purification process consisting of ammonium sulphate fractional precipitation, anion exchange, and gel filtration chromatography. The pH stability of the mutants was assayed by measuring the cyanide degrading activity in various pH buffers and thermal stability was assayed by measuring the cyanide degrading activity at high temperature. Differential Scanning Fluorimetry was also carried out to measure the melting temperatures of the individual mutants. Using negative stain Transmission Electron Microscopy, gel filtration peak fractions of the mutants were visualized at various pH values and datasets of the long fiber forming enzymes were collected and then subjected to Iterative Helical Real-Space Reconstruction using the SPIDER software package to produce 3D models. All the enzymes formed long fibres at pH 5.4 except one, designated $CynD_{pum\text{-}stut}$, a mutant with the *B. pumilus* C-terminal 44 residues mutated to the 48 corresponding residues of the homologous, non-fibre-forming, cyanide dihydratase from *Pseudomonas stutzeri* which has no histidines in this region. Other mutants, namely: H305K+H308K+H323K, Q86R and Q86R+E96G+D254E+E327G, formed longer oligomeric structures at pH 5.4 and pH 8 than the wild type. This demonstrates that the C terminal region may be in part responsible for the pH dependent transition from short spirals to long fibres at pH 5.4 almost certainly because of the high density of histidine residues. In particular, the mutants Q86R+E96G+D254E+E327G and $CynD_{pum\text{-}stut}$ were identified as being stable at alkaline pH and elevated temperatures. The mutations had no effect on the helical symmetry of the enzyme's quaternary structure.

ACKNOWLEDGEMENTS

I wish to thank, first and foremost, my supervisor Professor Trevor Sewell for his knowledge, outstanding support and encouragement which helped me hurdle all the obstacles in the completion of this research work.

Special thanks to Dr. Michael Benedik from Texas A&M University (USA) for his advice, collaboration and providing me with the right DNA plasmids needed to start my project.

I would like to thank Mohammed Jaffer from the Electron Microscope Unit (UCT) for his time and patience while helping me find my way around the microscopes.

I would like to thank Brandon Webber for his shared valuable insights and advise while I was performing experiments in the lab.

I would like to thank Nathan Van Wyk from the Department of Chemical Engineering (UCT) for sharing his expertise and equipment during the thermostability experiments.

I am particularly grateful to Mare Vlok and Busiswa Kekana from the Centre for Proteomic and Genomic Research (Cape Town, SA) for their help with the mass spectrometry protein identification.

I would like to thank the Electron Microscope Unit (UCT) postdoctoral students Serah Kimani, Jeremy Woodward and Ndoria Thuku who despite their heavy schedules took time to share valuable insights and help with my project and thesis write-up.

I would like to also thank my family for their encouragement and support throughout the duration of my studies.

Lastly, I would like to thank the National Research Foundation and the Microscope Society of South Africa for providing me with the funds to help me carryout my research.

NOMENCLATURE

KCN	Potassium cyanide
CynD	Cyanide dihydratase
$CynD_{pum}$	Cyanide dihydratase from *Bacillus pumilus*
$CynD_{stut}$	Cyanide dihydratase from *Pseudomonas stutzeri* AK61
WT	Wild type
FSC	Fourier shell correlation
DPR	Differential phase residual
PDB	Protein data bank
DCase	*N*-carbamyl-*D*-amino acid amidohydrolase
EM	Electron microscopy/microscope
IHRSR	Iterative Helical Real Space Reconstruction
LA	Lurai agar
NAD^+	Nicotinamide adenine dinucleotide
TEM	Transmission electron microscopy/microscope
CCD	Charge coupled device
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SNR	Signal-to-noise-ratio
IPTG	isopropyl-β-D-thiogalactopyranoside
TRIS	Tris(hydroxymethyl)aminomethane
DSF	Differential scanning fluorimetry
MS	Mass spectrometry
MDFF	Molecular dynamics flexible fitting

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

LITERATURE REVIEW

1.1 Nitrilase classification

Nitrilases are commonly-occurring enzymes found in most species of plants, animals, fungi and many prokaryotes with more than 150 known members (Pace and Brenner, 2001). These enzymes belong to a large family of enzymes called the nitrilase superfamily. This superfamily comprises thiol enzymes that are mostly involved in natural product biosynthesis and post translational modification in most organisms and these carry out this functions by performing a wide variety of non-peptide carbon-nitrogen hydrolysis reactions (Pace and Brenner, 2001) (Fig. 1.1). Nitrilase-related sequences have also been found in phylogenetically isolated prokaryotes that show an ecological relationship to plants and animals. This shows evidence that the nitrilase superfamily may have emerged prior to the separation of plants, animals and fungi, radiated into families, and then spread laterally to bacteria and archea (Pace and Brenner, 2001).

Members of the nitrilase superfamily can be classified into 13 branches (Fig. 1.2) as proposed by Pace and Brenner (2001). The classification of nitrilases was achieved on the basis of global and structure-based sequence analysis, techniques that also lead to the identification of domain fusion events within the nitrilase superfamily members. Members of seven branches of the superfamily have demonstrated domain fusion capability where a nitrilase related domain is fused to at least one additional conserved domain (Fig. 1.2). The fused domains consist of separate polypeptides which appear to be fused in order to coordinate biochemical reactions, cellular functions and cellular localization (Pace and Brenner, 2001). Although nitrilase enzymes have been historically classified as nitrilase-related based on sequence analysis, only branch 1 members have been identified to carry out nitrile-hydrolyzing activity to form corresponding acids and ammonia (reaction a, Fig. 1.1). Members from eight of some of the remaining branches (branches 2–11) have apparent amidase or amide-condensation activities (reaction b or d, Fig. 1.1).

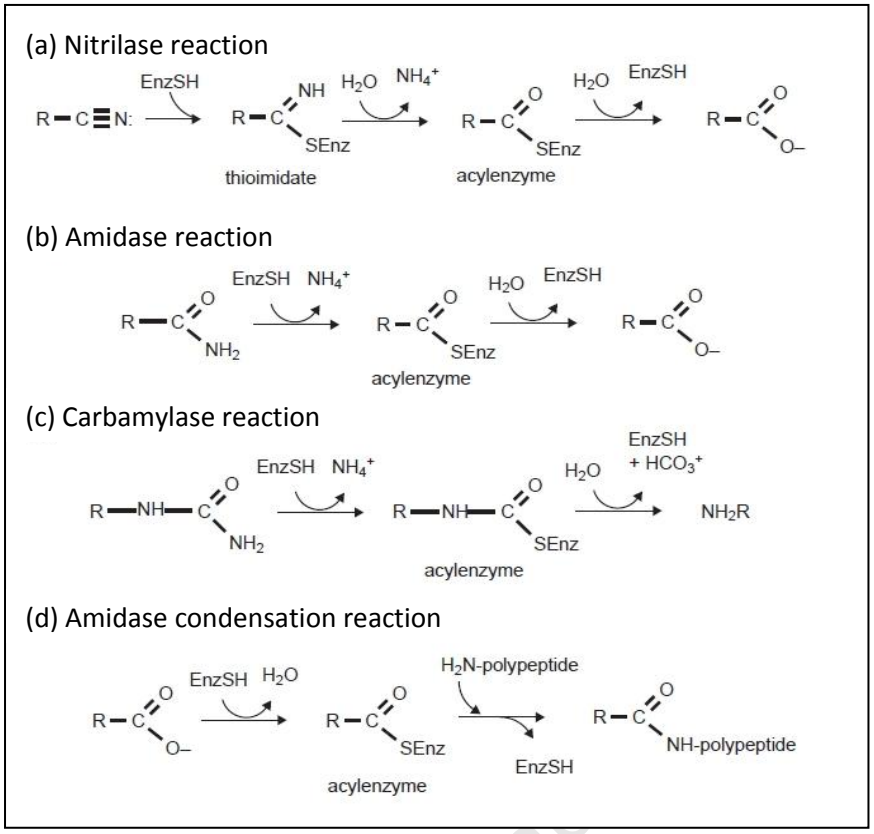


Figure 1.1. Four types of reaction carried out by nitrilase superfamily members (Pace and Brenner, 2001).

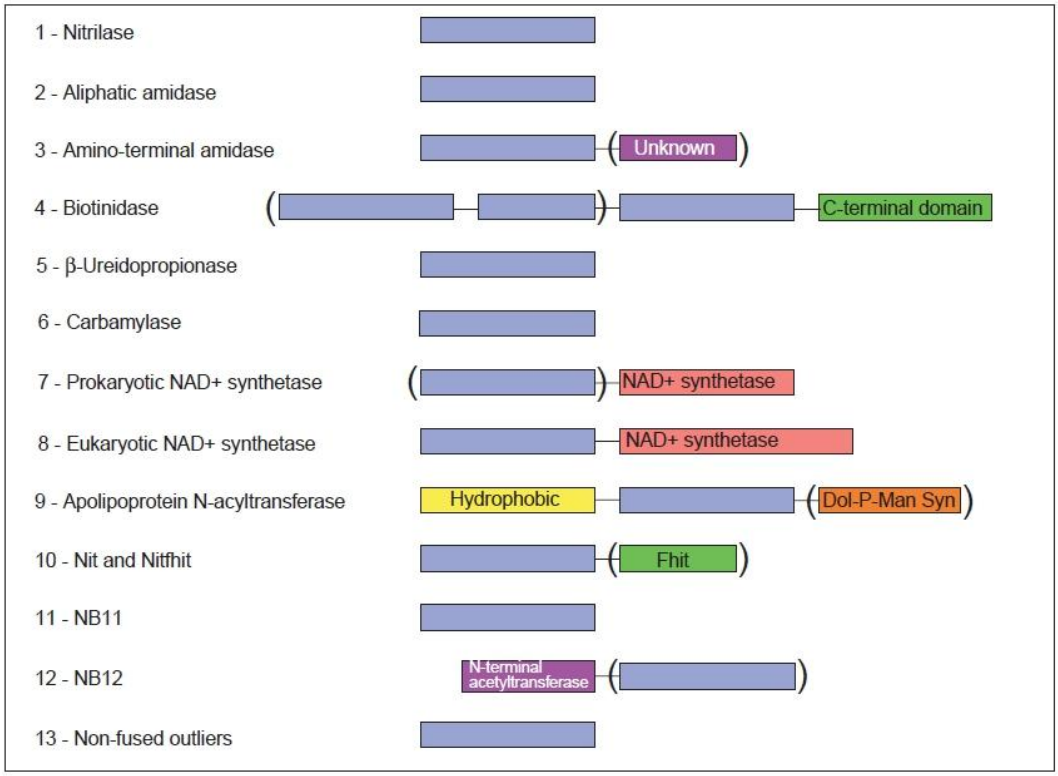


Figure 1.2. Domain structures of the 13 branches of the nitrilase superfamily. Parentheses denote domains found in only some members of the branch (Pace and Brenner, 2001).

Branch 1 consists of true nitrilases which are the only nitrilases that are known to carry out nitrile hydrolysis reactions as previously mentioned (reaction a, Fig. 1.1). These Branch 1 nitrilases are commonly found in bacterial species that are known to metabolise nitriles as a sole source of carbon and nitrogen (Banerjee *et al*., 2002). The physiological functions of these bacterial nitrilases are still unclear but in bacterial species like *Rhodococcus* and *Bacillus*, these nitrilases are capable of metabolizing aldoximes through nitrile intermediates, thus it was speculated that they could be components of complex pathways that control production and degradation of cyanogenic glycosides and related compounds where aldoximes are key intermediates (Banerjee *et al*., 2002; Kato *et al*., 1998; Kato *et al*., 2000). Some of the bacterial nitrilases are often used for biochemical syntheses and environmental remediation (Pace and Brenner, 2001). Members of this branch are also found in plants, animals and fungi. From the plant species *Arabidopsis thaliana*, three similar branch 1 nitrilases Nit1, Nit2 and Nit3 have been identified and shown to primarily act on nitriles (Bartling *et al*., 1992; Vorwerk *et al*., 2001). The primary function of nitrilases in plants is in the cyanide detoxification process and the accompanying recovery and recycling of essential nitrogen micronutrients (Piotrowski, 2008).

Branch 2 consists of aliphatic amidases, which are enzymes that hydrolyze short aliphatic amides such as the carboxamide sidechains of glutamine and asparagines with the aid of the conserved cysteine found within the nitrilase family (Pace and Brenner, 2001). These nitrilases perform amidase reactions which involve the hydrolysis of amides to corresponding acids and ammonia (reaction b, Fig. 1.1). The physiological function of branch 2 nitrilases is unclear, but they are postulated to be likely involved in the degradation of toxic nitriles (Banerjee *et al.,* 2002; Kato *et al*., 1998) in the same pathways as nitrile hydratases (Kobayashi *et al*., 1998; Yamada and Kobayashi, 1996), as well as in carbon and nitrogen fixation in bacterial cells (Sharma *et al*., 2009). This branch comprises of nearly identical proteins found in *Pseudomonas*, *Bacillus*, *Brevibacteria* and *Helicobacteria* (Pace and Brenner, 2001).

Branch 3 comprises of amino-terminal amidases which are involved in deamination of amino-terminal asparagines and glutamine residues to aspartate and glutamate, and they function in the N-end rule pathway of protein degradation (Mogk *et al*., 2007; Varshavsky, 1997) that is present in both prokaryotic and eukaryotic organisms for regulation of protein

concentrations during various cellular functions (Eisele and Wolf, 2008). Branch 3 enzymes include the Nta1 from *Saccharomyces cerevisiae* which has homologs found in fungi, but appear to be unrelated to mammalian amino-terminal amidases (Pace and Brenner, 2001). Like the branch 2 enzymes, the branch 3 enzymes also perform amidase reactions (reaction b, Fig. 1.1)

Branch 4 consists of biotinidases and pantetheinases which are secondary amidases, with similar specificities that are involved in vitamin recycling (Aurrand-Lions *et al.,* 1996; Maras *et al.,* 1999). Biotinidases are involved in the recycling of vitamin H while pantetheinases are involved in the recycling on vitamin B_5. . This branch is the only branch with amidases which are known to prefer secondary amine substrates of the form R-C=O(NHR') as opposed to simple acid amides (Brenner, 2002). Biotinidases and pantetheinases have been shown to have carboxy-terminal domains (Fig. 1.2). Branch 4 enzymes perform amidase reactions (reaction b, Fig. 1.1)

Branch 5 comprises of β-ureidopropionases which are enzymes involved in the catabolism of pyrimidine bases and the production of β-alanine (Kvalnes-Krick *et al*., 1993). These enzymes utilize carbamylase type substrate to produce a non-standard amino acid product (reaction c, Fig. 1.1).

Branch 6 consists of enzymes primarily known as bacterial N-carbamyl-D-amino acid amidohydrolases which are involved in the decarbamylation of D-amino acids (Nakai *et al.,* 2000). These enzymes are expressed in a variety of bacteria and have been exploited in the production of semisynthetic β-lactam antibiotics (Brenner*,* 2002). Like the branch 5 enzymes, the branch 6 enzymes perform carbamylase reactions (reaction c, Fig. 1.1)

Branches 7 and 8 comprise of glutamine-dependent NAD^+ synthetases which contain domain fusion events (Fig. 1.2) and function in the last step of the NAD^+ biosynthetic pathways (LaRonde-LeBlanc *et al*., 2009). These enzymes contain two domains, one with glutamine amidohydrolase (GAT) activity involved in glutamine hydrolysis to produce ammonia which is then transferred to the $NaAD^+$ substrate by the NAD^+ synthetase domain (LaRonde-LeBlanc *et al*., 2009; Zalkin and Smith*,* 1998). The reaction carried out by these NAD^+ synthetases can be generalised as an amidase reaction specific for glutamine (reaction b, Fig. 1.1). These enzymes occur in all eukaryotes and some prokaryotes, thus were place in

separate branches. Prokaryotic NAD^+ synthetases were placed in branch 7 and eukaryotic NAD^+ synthetases were placed in branch 8.

Branch 9 comprises of apolipoprotein N-acyltransferase which are involved in the modification and processing of Braun's lipoprotein. This lipoprotein is a major component of the outer membrane of some bacteria such as *E.coli* (Tokunaga *et al.* 1982; Rogers *et al.,* 1991). Branch 9 is the only branch known to catalyse the reverse amidase (condensation) reaction *in vivo* (reaction d, Fig. 1.1). Apolipoprotein N-acyltransferases always have a hydrophobic amino-terminal domain and one member is fused to an apparent dolichol phosphate mannose synthetase (Fig. 1.2), which contradicts the proposed function of branch 9 enzymes in post-translational modification.

Branch 10 consists of amidases from mammals, plants, bacteria and fungi, whose physiological function include nitrogen and sulphur metabolic and salvage pathways, among others (Cobzaru *et al*., 2011; Krasnikov *et al*., 2009a). Branch 10 includes two mammalian enzymes nitrilase-1 (Nit1) and nitrilase-2 (Nit2), whose expression in a wide range of tissues has been shown to have tumour suppression properties (Barglow *et al*., 2008; Krasnikov *et al*., 2009b; Semba *et al*., 2006).

Branch 11 contains bacterial and plant enzymes which are postulated to function as carbamylases for the production of putrescine or related amines, such as the N-carbamyl putrescine amidohydrolase product of the Pseudomonas AguB gene (Nakanda *et al.,* 2001) Branch 12 consists mostly of Rosetta Stone proteins, where distinctive nitrilase-related domains are fused to an amino terminal domain of approximately 201 amino acids (Fig. 1.2). Branch 12 enzymes have been implicated in protein post translational modifications because of their association with the RimI N-terminal acetyltransferases (Pace and Brenner, 2001). Branch 13 contains uncharacterized, non fused nitrilase-related proteins that are difficult to place in distinct groups.

1.2 Applications of the nitrilase enzymes

As demonstrated before, several branches of the nitrilase superfamily are found in mammalian genomes which encode eight nitrilase enzymes (Barglow *et al.,* 2008). These include the vanins/biotinidases/pantetheinases involved in vitamin synthesis, and the glutaminase domain of glutamine-dependent NAD^+ synthatase which generates ammonia from glutamine for NAD^+ synthesis (Barglow *et al.,* 2008). Studies done on these nitrilases could help in drug design and identification of potential drug targets. These potential drug targets include branch 7 glutamine-dependent NAD^+ synthetase and branch 2 amidases from *Helicobacter pylori*. In the glutamine-dependent NAD^+ synthetases, glutaminase domain converts L-glutamine into ammonia and directly transfers the ammonia to a synthetase domain where it is used to generate NAD^+ (LaRonde-LeBlanc *et al.,* 2009). NAD^+ is constantly recycled in organisms and the recycling pathways in some species are not dependent on NAD^+ synthetase. Human NAD^+ recycling pathways are not all dependent on NAD^+ synthetase while some prokaryotes such as *Mycobacterium tuberculosis* only have NAD^+ recycling pathways that are dependent on NAD^+ synthetase making their NAD^+ synthetase potential drug target for treatment of tuberculosis (LaRonde-LeBlanc *et al.,* 2009). *H.pylori* is a human gastric pathogen that colonises the mucius membrane of the stomach resulting in various stomach ailments including ulcers (Bury-Mone *et al.,* 2003). *H.pylori* has two branch 2 amidases, AmiE and AmiF, which play a role in protecting the bacterium from the harsh acidic stomach conditions making these amidases potential drug targets. β-Ureidopropionases and the prominent nitrilase-1 (Nit1) and nitrilase-2 (Nit2), which have been shown to exert effects on cell growth in mice and possibly serve as tumor suppressor genes, are furthermore examples of nitrilases found in the mammalian genome (Barglow *et al.,* 2008). More insights about these nitrilases could potentially lead to the design of cancer drugs.

In industry, nitrilases have the potential to be used as biocatalysts to convert cyano functionality to either acid or amide. Traditional chemical methods of converting nitriles to acids and amides have proven challenging due to the harsh conditions required for the reactions such strong acidity or basicity and high temperatures. Other drawbacks of these chemical methods include the formation of toxic by-products. The desirable features that make nitrilases good candidates for these reactions are their ability to carry out chemo,

regio, stereo and enatio-selective transformations, and effect reactions in a 'greener' manner (Banerjee *et al.,* 2002). There have been a number of successes in the application of nitrile-converting organisms as biocatalysts. These include the production of acrylamide using *Rhodococccus rhodochrous* J1 (Yamada and Nagasawa *et al.,* 1994), conversion of surfacial nitrile groups of acrylic fibres to the corresponding amides by *R. rhodochro*us NCIMB 11216 nitrile hydratase (Tauber *et al.,* 2000) and conversion of 3-cyanopyridine to nicotinamide as well as nicotinic acid using *R. rhodochrous* J1 (Matthew *et al.,* 1988). *R. rhodochrous* K22 is used to catalyse the conversion of adiponitrile to 5-cyanovaleric acid during the production of nylon-6 (Godtfredsen *et al.,* 1985), and *Acremonium sp*. is used to catalyse the selective mono-hydrolysis of trans 1,4-dicyano cyclohexane during the production of Tranexamic acid, a homeostatic drug (Nishise et al. 1987).

Some nitrilases could also be used for bioremediation of industrial wastewater that contains synthetic nitrile compounds. Most of the nitrile compounds are toxic, carcinogenic and mutagenic and thus their release into the environment needs to be controlled to prevent pollution. There have been a few successes were mixed cultures of bacteria containing different nitrile hydrolysing enzymes were used to metabolize effluent containing acrylonitrile, furmaronitrile, succinonitrile, etc by growing the bacteria on these components of waste (Wyatt and Knowles *et al.,* 1995). Cyanide dihydratase from *Bacillus pumilus* C1 is a good candidate for bioremediation and is later explained. Some soil micro-organisms have been reported to degrade nitrile-containing herbicides in soil such as *Agrobacterium radiobacter*, a soil bacterium capable of degrading bromoxynil herbicide (Freyssinet *et al*., 1996; Muller and Gabriel, 1999). These enzymes could potentially prevent be accumulation of the herbicides in foods.

1.3 Structure and homology of nitrilases

The nitrilase superfamily enzymes share a significant structural homology (Fig. 1.3) despite varying sequence conservation and differing substrate affinities (Pace and Brenner, 2001). In the early years only three nitrilase crystal structures were solved and these included two members of branch 10 enzymes; *Caenorhabditis elegans* NitFhit (PDB id, 1ems; Pace *et al.,* 2000) and *Saccharomyces cerevisiae* Nit3 (PDB id, 1f89; Kumaran *et al.,* 2003), and one

member from branch 6, *Agrobacterium N*-carbamyl-D-amino acid amidohydrolase (PDB id, 1erz; Nakai *et al.,* 2000). The nitrilase crystal structures indicated that these enzymes consist of monomers with an αββα sandwich architecture (Fig. 1.5). These αββα monomers were observed to stack together to form active dimer complexes. These dimers are basic building units that often further interact to form larger complexes which differ across members of the superfamily (Table 1.1).

C. elegans NitFhit and *Agrobacterium N*-carbamyl-D-amino acid amidohydrolase form active tetramer while *S. cerevisiae* Nit3 protein forms dimers (Table 1.1). After the *C. elegans* NitFhit, *S. cerevisiae* Nit3 protien and *Agrobacterium N*-carbamyl-D-amino acid amidohydrolase crystal structures were solved, more nitrilase structures began to appear in the Protein Data Bank (PDB). Among these were ligand bound and mutant structures which have provided further insight into the structure and the catalytic mechanism of some of these and other nitrilases (Table 1.1). Although there is an increase in the number of nitrilase structures, the crystal structure of a Branch 1 nitrilases are still elusive. One of the main reasons why branch 1 nitrilases cannot be crystallised readily is because of their inability to form uniform crystal lattices due to oligomerisation of monomers that result in the formation of long complex quaternary structures. Due to the lack of a branch 1 crystal structure, other methods of structure prediction such as homology modelling and electron microscopy (EM) had to be used. As mentioned previously, nitrilases demonstrate a significant structural homology despite their low sequence conservation (Fig. 1.3) and this characteristic is used in homology modelling where crystal structures of solved distant homologs are used to predict the unknown structures of branch 1 nitrilases. This method has successfully helped in the structural model prediction of some branch 1 nitrilases such as the cyanide dihydratase from *Pseudomonas stutzeri* AK61 (Sewell *et al.,* 2005), the nitrilase from *Rhodococcus rhodochrous J1* (Thuku *et al*., 2007) and the cyanide dihydratase from *Bacillus pumilus* C1 (Scheffer, 2006). These models were docked into low resolution 3D reconstruction EM maps to provide key insights into the oligomeric structure of the branch 1 nitrilases. These branch 1 nitrilases were shown to have a two fold symmetric structure which conserves a αββααββα and are homodimeric (Sewell *et al.,* 2005; Thuku *et al.,* 2007).

Figure 1.3. Alignment of the sequences of the cyanide dihydratases from B. pumilus C1 ($CynD_{pum}$) and P. stutzeri AK61 ($CynD_{stu}$) with the sequences of the four homologs: 1f89, Nit3 from *S. cerevisiae*; 1j31, PH0642 from *Pyrococcus horikoshii*; 1ems, NitFhit from *C. elegans*; DCase from *Agrobacterium,* for which the atomic structures have been determined. Pairwise alignment of 1f89, 1j31, 1ems, and 1erz was done with ALIGN (Cohen*,* 1997), and the CynD sequences were aligned with Gen-THREADER.13. The secondary structural elements referring to 1ems are indicated in the bottom line and use the notation of Pace *et al.* (2000). The approximate regions of the interacting surfaces are indicated in the top line with the A surface highlighted green, C surface highlighted yellow, D surface highlighted in pink. Charged residues that may be involved in these interactions are white on a black background. The residues mutated as indicated in TABLE 2 are underlined. The conserved active site residues are highlighted in grey (Sewell *et al.,* 2005).

Table 1.1 List of members of the nitrilase superfamily with solved atomic structures

PDB ID	Enzyme	Species	Superfamiliy Branch	Resolution	Oligomeric form	Reference
1ems	Nit-fragile histidine triad fusion protein	*Caenorhabditis elegans*	10	2.80 Å	Tetramer	Pace *et al* ., 2000
1erz	*N*-carbanyl-D-amino acid amidohydrolase	*Agrobacterium* sp. KNK712	6	1.70 Å	Tetramer	Nakai *et al.,* 2000
1f89	Nit 3	*Saccharomyces cerevisiae*	10	2.40 Å	Dimer	Kumaran *et al.,* 2003
1fo6	*N*-carbamyl-D-amino-acid amidohydrolase	*Agrobacterium radiobacter*	10	1.95 Å	Tetramer	Wang *et al.,* 2001
1j31	Hypothetical protein PH0642 from Pyrococcus horikoshii	*Pyrococcus horikoshii*	Unclassified	1.60 Å	Dimer	Sakai *et al.,* 2004
1uf4 1uf5 1uf7 1uf8	*N*-carbamyl-D-amino acid amidohydrolase	*Agrobacterium* sp.	10	2.15Å 1.6 0Å 1.9 0Å 1.80Å	Tetramer	Hashimoto, H (unpublished)
2dyu 2dyv 2e2k 2e2l	Formamidase (AmiF)	*Helicobacter pylori*	2	1.75 Å 2.00 Å 2.50 Å 2.29 Å	Hexamer	Hung *et al.,* 2007
2e11	XC1258	*Xanthomonas campestris*	Unclassified	2.30 Å	Dimer	Chin *et al.,* 2007
2ggk 2ggl	N-carbamoyl-D-amino-acid amidohydrolase (D-NCAase)	*Agrobacterium tumefaciens*	6	2.30 Å 2.40 Å	Tetramer	Chiu *et al.,* 2006
2plq	Aliphatic Amidase	*Geobacillus pallidus* RAPc8	2	1.90 Å	Hexamer	Agarkar *et al.,* 2006
2uxy	Aliphatic Amidase	*Pseudomonas aeruginosa*	2	1.25 Å	Hexaner	Andrade *et al.,* 2007
2vhh 2vhi	Beta-alanine synthase	*Drosophila melanogaster*	5	2.80 Å 3.30 Å	Octamer	Lundgren *et al.,* 2008
2w1v	Nitrilase-2	*Mus musculus*	10	1.49 Å	Dimer	Barglow *et al.,* 2008
3dla	Glutamine-dependent NAD^+ synthetase	*Mycobacterium tuberculosis*	7	2.35 Å	Octamer	LaRonde-LeBlanc *et al.,* 2009
3hkx	Amidase	*Nesterenkonia sp.*	2	1.66 Å	Dimer	Nel *et al*., 2011
3iw3 3ivz	Hyperthermophilic nitrilase	*Pyrococcus abyssi* ge5	Unclassified	1.80 Å 1.57 Å	Dimer	Raczynska *et al*., 2010
3ilv	Glutamine dependent NAD^+ synthetase	*Cytophaga hutchinsonii*	7 or 8	1.79 Å	Dimer	Palani K. (unpublished)

1.4 Spiral formation and interfacial regions

All the microbial nitrilases studied by our group form spiral quaternary structure (Thuku *et al*., 2009). These nitrilase include cyanide dihydratases and cyanide hydratases which either form helices of variable length or short, terminating spirals which have a specific number of subunits (Fig. 1.4). The cyanide dihydratase from *Psedomonus stutzeri AK61* ($CynD_{stut}$) forms a 14-subunit spiral (Sewell *et al*., 2003) while the homologous enzyme from *Bacillus pumilus* C1 ($CynD_{pum}$) and 8A3, form an 18- and a 22-subunit spirals, respectively (Jandhyala *et al*., 2003; Eicher, 2007). Short spirals are formed by nitrilases from *Rhodococcus rhodochrous* J1 (Thuku *et al*., 2007) and *Fusarium solani* O1 (Vejvoda *et al*., 2008), although the exact number of subunits has not been determined. The cyanide hydratases from *Gloecercospora sorghi* (Woodward *et al*., 2008), *Neurospora crassa* (Dent *et al*., 2008) and *Aspergillus niger* K10 (Vejvoda *et al*., 2008), occur as long, variable length helices.

Electron microscopy and modelling of branch 1 nitrilases based on nitrilase homologs with solved crystal structures has aided in the identification of interfacial regions or surfaces that are involved in the association of monomers to form oligomers (Sewell *et al*., 2005). On the basis of sequence alignment, model building and three-dimensional electron microscopy, it can be postulated that the interactions at these surfaces are generally electrostatic (Thuku *et al.,* 2009). Branch 1 nitrilases generally have six of these surfaces namely the A, B, C, D and F surfaces (Fig. 1.4–1.6). The association of subunits at different interfaces results in a variety of oligomeric shapes that are observed in the nitrilases superfamily (Fig. 1.4).

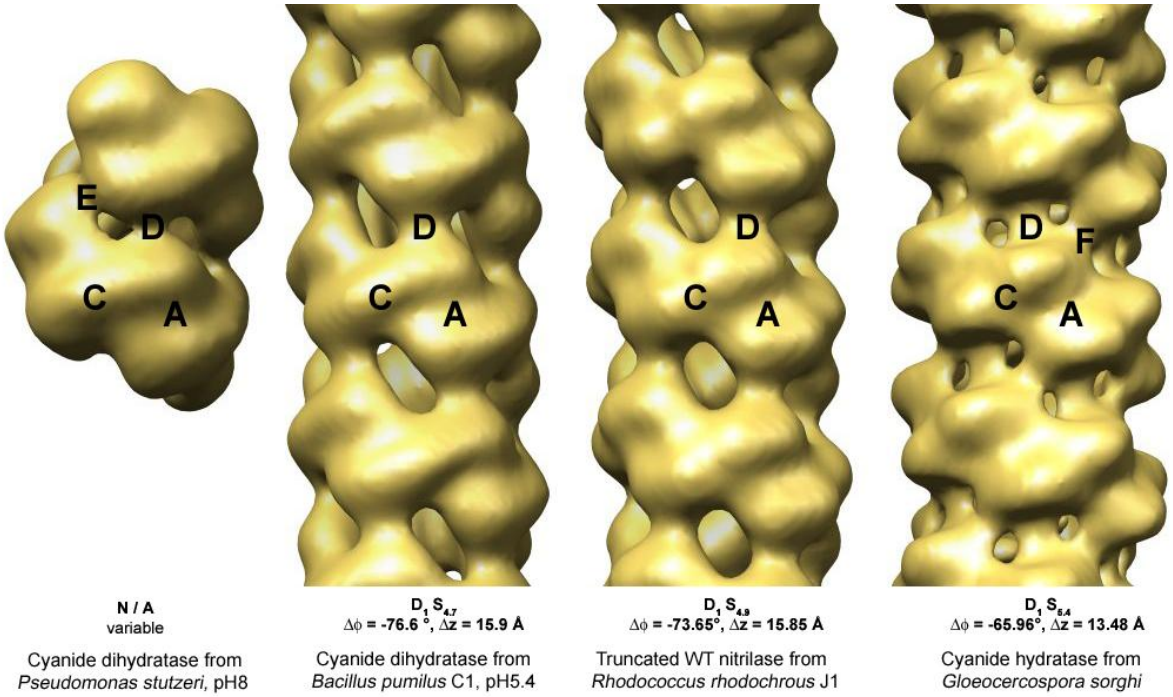


Figure 1.4. Three dimensional electron microscopic reconstuction of the microbial nitrilases studied at low resolution (unpublished)

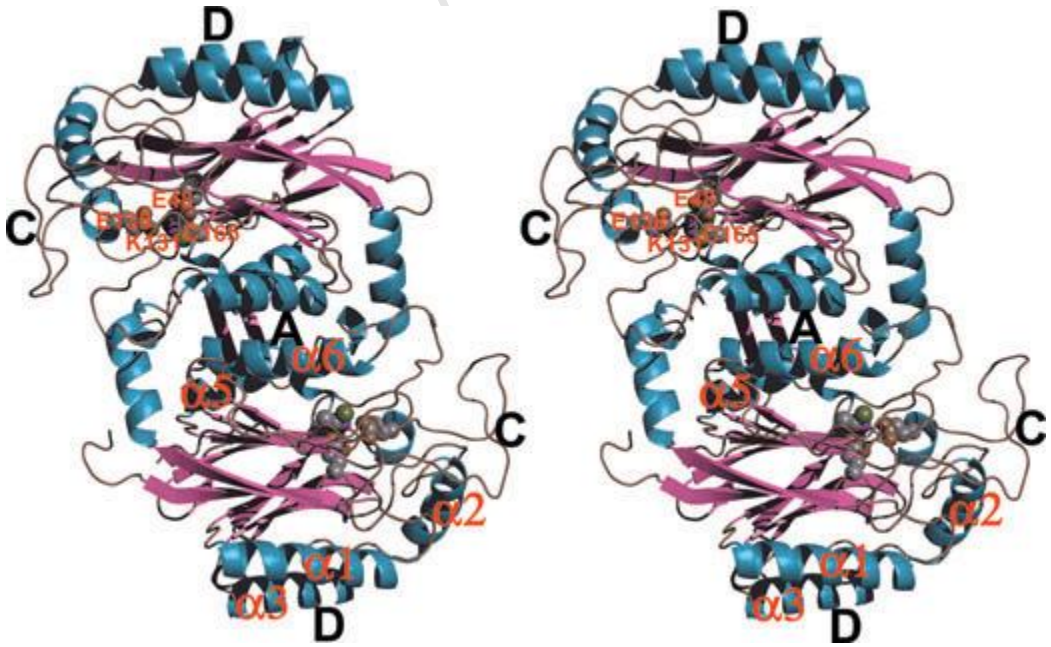


Figure 1.5. Stereo view of a dimer model of the nitrilase from *Rhodococccus rhodochrous* J1 with highlighted interfacial regions and conserved active site residues (Thuku *et al.,* 2009). The model was built based on structural homology of the nitrilase-related atomic structures of 1erz (Nakai *et al.,* 2000), 1ems (Pace *et al.,* 2000), 1j31 (Sakai *et al.,* 2004) and 2vhi (Lundgren *et al.,* 2008)

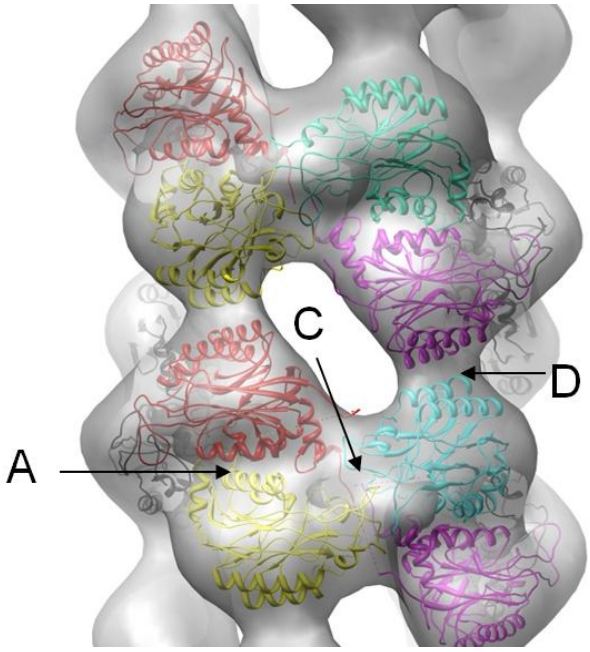


Figure 1.6. The docking of a homology model of the cyanide dihydratase from *B. pumilus* C1 into the negatively stained 3D electron microscopy reconstruction of the cyanide dihydratase from *B. pumilus* C1 (unpublished work).

The A surface and C surface are the only surfaces that have been visualized at atomic resolution in amidases. Details of the other surfaces have not been verified since they are obtained from docking homology models into low resolution 3D electron microscopic reconstructions (Thuku *et al.,* 2009). The A surface comprises α helices α5 and α6 (residues 170–179 and 192–205 respectively) as well as a β sheet (residues 289–296, Fig. 1.5) based on *R. rhodochrous* J1 homology model. The association of the two α helices occurs in all crystalline and spiral structures (Thuku *et al.,* 2009). The A surface is located between two monomers and interactions at this surface are postulated to be the basis of dimerisation, although the details of the intermolecular ionic and hydrophobic interactions of high resolution structures from different organisms vary (Wang *et al.,* 2001; Chin *et al.,* 2007; Kimani *et al.,* 2007).

The association of the C surface is postulated to be the key to the formation of the spiral quaternary structures. This surface is located approximately at right angles to the A surface and comprises two insertions, residues 63–74 and residues 216–233 based on *R. rhodochrous* J1 homology model (Fig. 1.5).

The D surface association occurs only when a spiral completes one turn. The interactions at this surface occur across the groove of the spiral or helix (Fig. 1.4 and 1.6). The residues that

form this surface have been shown to differ from organism to organism (Sewell *et al.,* 2005). Different residues that formed the D surface have been identified between the cyanide dihydratase from *B. pumilus* (residue 90–97) and the cyanide dihydratase from *P. stutzeri* (residue 82–87) and these residues formed α helices in both organisms (Sewell *et al.,* 2005). The D surface comprises two helices, α1 and α3 (Fig. 1.5), based on *R. rhodochrous* J1 homology model, and both have a mixture of positively and negatively charged residues. It is suggested that these charged residues form two-fold symmetric, electrostatic interactions that maintain the elongating assembly during oligomerisation (Sewell *et al.,* 2005).

The F surface is located near the D surface and is postulated to be involved in the further stabilization of the two fold symmetric interactions in certain enzymes such as cyanide hydratases from *Gloecercospora sorghi* (Woodward *et al.,* 2008). The E surface is different from the other surfaces in that its contributors are asymmetric and arise from different regions of the subunit (Fig. 1.4). The E surface only occurs in terminating spirals.

Bacterial nitrilases have an extended C-terminal sequence that is about 40–100 amino acid longer than that of other related nitrilase homologous structures. The C-terminal region is postulated to be located on the inside of the spirals formed by nitrilases (Thuku *et al*., 2007). This location was also predicted based on distinct electron densities, observed on the inner surface of the oligomeric structures of bacterial nitrilases, in 3D reconstruction models (Sewell *et al.,* 2003; Thuku *et al.,* 2007). The C-terminal tail has been shown to play an important role in oligomerisation and activity in branch 1 nitrilases (Sewell *et al.,* 2005; Thuku *et al.,* 2007). The C-terminal has been visualized on the inside rather than outside of several amidase structures although the size of the helical content does not match with that of branch 1 nitrilases (Andrade *et al.,* 2007; Hung *et al.,* 2007; Kimani *et al.,* 2007).

1.5 Catalysis

Members of the nitrilase superfamily were postulated to utilize a novel catalytic triad consisting of a glutamate, lysine and cysteine (Pace *et al.,* 2000; Nakai *et al.,* 2000) (Fig. 1.5). The cysteine was the first catalytic residue to be identified due to its high reactivity which could be specifically hindered by labelling with iodoacetate, resulting in an inactive form of

the *N*-carbamyl-D-amino acid amidohydrolase (D-NCAase) from *Agrobacterium* (Grifantini *et al.,* 1996). Other evidence that identified the cysteine as a catalytic residue includes site directed mutagenesis of the cysteine (Cys172 in D-NCAase) which produced a non-functional mutant (Grifantini *et al.,* 1996). A solvent-accessible cleft or pocket in the D-NCAase was later identified within the vicinity of the catalytic cysteine (Grifantini *et al.,* 1996). Around this cleft highly conserved residues were identified by comparison with aligned structures of different nitrilases. The comparison indicated five highly conserved residues, Glu47, Lys127, Glu146, Cys172 and in D-NCAase which were located close to each other and clustered near the floor of the pocket (Wang *et al.,* 2001). It was then postulated that the three conserved residues Cys172, Glu47 and Lys127 of D-NCAase (Cys164, Glu48 and Lys130 in $CynD_{pum}$) are most likely to participate in the amidohydrolytic reaction and recent evidence has suggested the participation of a second glutamate, Glu146 in D-NCAase (Glu137 in $CynD_{pum}$) (Kimani *et al.,* 2007). The essentiality of the first glutamate (Glu47 in D-NCAase) and lysine in catalysis was demonstrated by mutagenesis experiments of Pseudomonas aliphatic amidases where Glu59Gln and Lys134Asn mutations of the aliphatic amidase resulted in an inactive enzyme while Glu59Gln and Lys134Arg mutations resulted in an inactive and unstable enzyme which not only showed the importance of these residues in catalysis but also indicated their role in protein stability (Novo *et al.,* 2002).

A plausible catalytic mechanism was later proposed were the catalytic cysteine acts as a nucleophile, the glutamate acts as a general base catalyst and the lysine stabilizes a tetrahedral intermediate (Wang *et al.,* 2001). The two conserved glutamates (Glu47 and Glu164) have been proposed to play a role in positioning the substrate after visualization of substrate-bound crystal structures of D-NCAase and amidase (Chen *et al.,* 2003, Hung *et al.,* 2007) revealed that the amide nitrogen of the substrate is within hydrogen distance of the two glutamates. Using information of the catalytic site from homologous nitrilase structures, a cyanide dihydratase catalysis mechanism was proposed (Fig. 1.7; Jandhyala *et al*., 2005). The highly electronegative nitrogen from the substrate cyano group withdraws electrons away from the substrate cyano carbon thus making it more positive. This is followed by nucleophilic attack of the positive carbon by an activated cysteine which results in the formation of a thioimidate intermediate (1). The lysine, acting as a general base, then catalyses a hydrolysis, leading to the tetrahedral intermediate (2).The C-N bond then breaks

leading to the release of ammonia and formation of an acyl intermediate (3). Glutamate then activates a second hydrolysis leading to a second tetrahedral intermediate (4), which decomposes to release formic acid

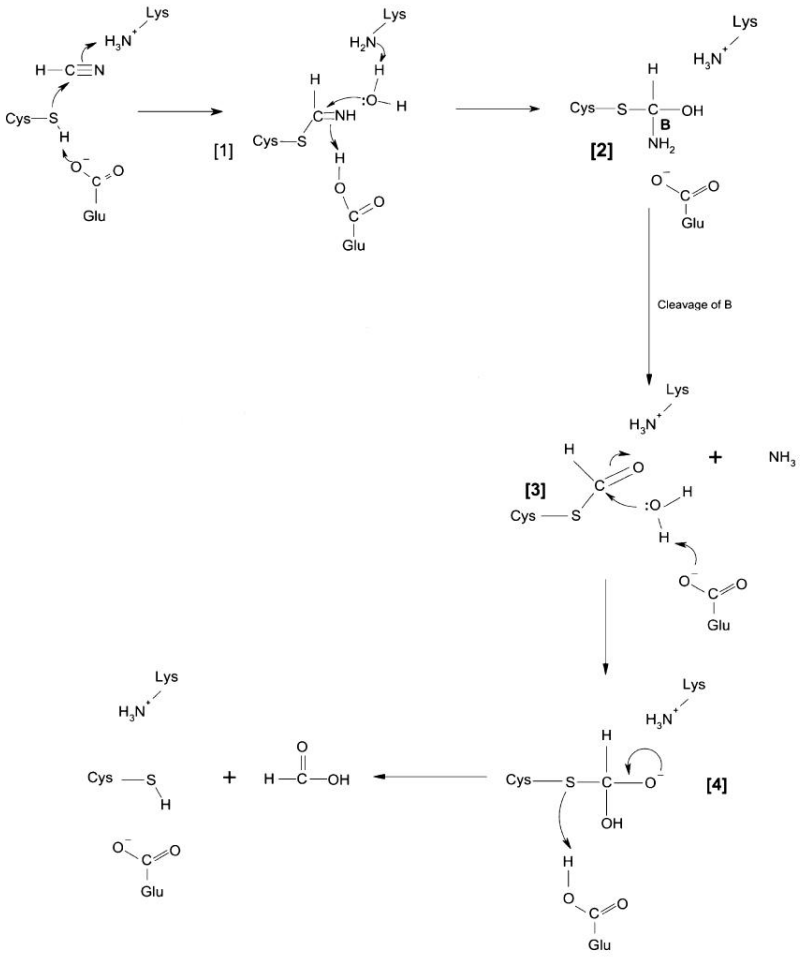


Figure 1.7. The proposed nitrilase catalytic mechanism of the *Bacillus pumilus C1* cyanide dihydratase (Jandhyala *et al*., 2005).

1.6 *Bacillus pumilus* C1 cyanide dihydratase

This study will focus on the cyanide dihydratase from *Bacillus pumilus* C1 ($CynD_{pum}$) which is a member of the branch 1 nitrilases. This branch includes cyanide and nitrile hydrolysing enzymes which have been studied in a wide range of microbial species (O' Relly and Tuner *et al.,* 2003). The *Bacillus pumilus* C1 was isolated from a soil sample near a cyanide-

containing wastewater dam in South Africa (Meyers *et al.,* 1991). $CynD_{pum}$ is an enzyme that is known to convert cyanide directly into formate and ammonia (reaction a, Fig. 1.1). The activity of the enzyme is optimal between 37°C and 42°C and its maximal activity is in the range of pH 7–8 (Jandhyala *et al.,* 2005). Like most nitrilases, $CynD_{pum}$ has a homodimeric building block with the αββα-αββα sandwich fold mentioned above. Over the pH range of 6–8 the enzyme exists as active short spirals approximately 18.9nm in length (Fig. 1.8). The spirals associate to form long active helical fibers at pH 5.4 (Fig. 1.8, Sewell *et al.,* 2005). These fibers have a diameter of 13nm and length of greater than 100nm (Jandhyala *et al.,* 2005). The onset of fiber formation has been shown to correspond to a small increase in activity for $CynD_{pum}$ (Jandhyala *et al.,* 2005), although most cyanide-degrading enzymes decrease steadily in activity as a function of pH below the optimum (Sewell *et al.,* 2005). Terminal subunits in the short spirals at pH8 are thought to be inactive and hence the structural transformation to long helices at pH 5.4 results in more active subunits in the population (Jandhyala *et al.,* 2005; Sewell *et al.,* 2005)

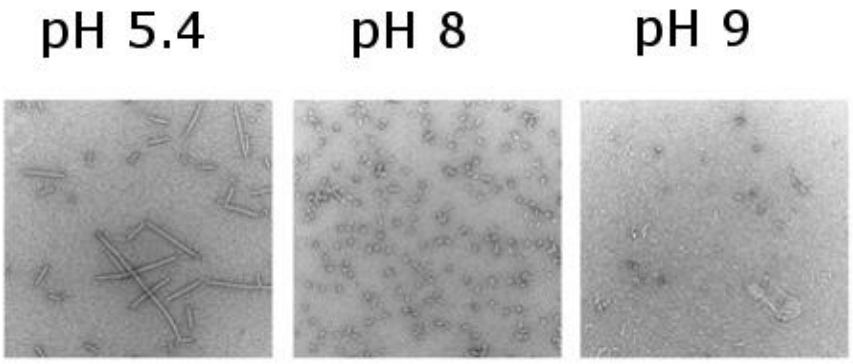


Figure 1.8. Negative stain electron micrographs of $CynD_{pum}$ wild type at pH 5.4, pH8 and pH 9.

The reason for the termination of the spiral at 18 subunits at pH 8 is not fully understood but it was initially hypothesized to be due to distortion of the helical symmetry (Sewell *et al.,* 2003). The theory at the time was that as subunits were added at the terminus of the spiral, new opportunities arose for interaction across the E surface, located across the groove of the spiral after the formation of a single turn (Fig. 1.5). Further elongation of the spiral by addition of more subunits would only occur if it is more energetically favourable to add a subunit than allowing the interaction across the E surface. Strong interactions across

the E surface could distort the helical symmetry and result in steric hindrance that prevents addition of further subunits (Sewell *et al.,* 2003). The transition from spirals to long fibers at pH 5.4 was therefore believed to be a result of environmental change which may affect the helical symmetry of the subunits by alterations of interactions responsible for the overall structure of the subunits.

Although it is still unclear whether or not the termination of the spirals at pH 8 is due to distortion of the helical symmetry, certain residues have been postulated to be involved in the pH-dependent transition from spirals to helices. The transition from short spirals to long helices occurs as the pH drops below 6.0 and this structural change implicates the participation of histidine which is the only amino acid charged at pH 5.4 (Histidine pKa of 6.0). There are 10 histidines in the amino acid sequence of $CynD_{pum}$, three of which are located on the C-terminal tail (Fig. 1.3). The involvement of histidines is further demonstrated by the fact that the highly homologous cyanide dihydratase from *Pseudomonas stutzeri* ($CynD_{stut}$) which lacks these C terminal histidines and does not undergo the transition from short spirals to long fibers at pH 6 (Sewell *et al.,* 2005). Eicher (2007) demonstrated that a similar strain of $CynD_{pum}$ from *Bacillus pumilus 8A3* which also lacks the histidines on the C-terminal tail, forms aggregated and irregular short fibres at pH below 6 instead of regular long fibers formed by the *Bacillus pumilus* C1 strain. We then proposed that transition from long fibers to short spirals at pH 5.4 is due to changes in the C-terminal tail histidines interaction which are caused by changes in charge of the residue. This may result in the distortion of the helical symmetry of the quaternary structure preventing more subunits from binding to form the long helices.

1.7 Previous studies on *Bacillus pumilus* C1 cyanide dihydratase

A number of experiments have been carried out to investigate the effects of the C-terminal residues and residues at the interfacial region on both activity and structure. This was mainly accomplished through mutagenesis experiments where specific residues were deleted or substituted and the effect on structure and activity was observed. Most of the $CynD_{pum}$ mutagenesis experiments were done by Sewell *et al.* (2005). A C surface mutation (mutant 5, Table 1.2; underlined in Fig. 1.3) with excised insertions which resulted in a shorter C surface β sheet, similar to those found in non-spiral forming homologs, showed inactivity or failure to express. This suggested that the C surface interactions may be involved in fiber formation and activation of the enzyme. $CynD_{pum}$ D surface mutations (mutant 6, Table 1.2) where $CynD_{pum}$ residues 90–97 (residues postulated to form salt bridges across the D surface that stabilized the spirals) were mutated to residues 89–96 from $CynD_{stut}$ residues that cannot form salt bridges across the D surface), did not affect activity (Sewell *et al.,* 2005). These results show that the spirals are still active in the absence of the salt bridge suggesting that the activity is due to other compensating interactions. Mutation of residues on the A surface (mutant 4, Table 1.2), which interact across the two fold axis to form a pair of salt bridges, resulted in the inactivation of the enzyme. This provides evidence of the involvement of the A surface in the activation of the enzyme (Sewell *et al.,* 2005). When a series of C-terminal truncations were generated for both $CynD_{pum}$ (mutant 1–3, Table 1.2) and $CynD_{stut}$ (mutant 10, Table 1.2) it was demonstrated that the $CynD_{pum}$ retained activity even after a larger truncation than that of $CynD_{stut}$. Because these enzymes only differ in residues on the D surface and C-terminus, it was suggested that the D surface may be involved in the stabilization of the spirals which contain truncated C-terminal tail (Sewell *et al.,* 2005). These mutation studies of the interfacial regions suggest that the A surface, C surface, D surface and the C terminus are involved in activation on the enzyme.

Some of the $CynD_{pum}$ mutagenesis experiments were done by one of our collaborators (Wang *et al*., 2011), who showed that the D surface and the C-terminal tail are involved in the oligomerisation on of the $CynD_{pum}$ enzyme. Through random mutagenesis, Wang *et al*. (2011) indicated that mutations on the D surface and the C-terminal tail (mutant 7–9, Table

1.2) result in an enhancement of oligomerisation that was displayed by the formation of oligomers significantly longer than those of the wild type at pH 5.4–pH9. Wang *et al*. (2011) study also showed a correlation between improved fiber formation and increase in alkaline pH stability and thermal stability of the $CynD_{pum}$. Hence the key to engineer better industrial enzymes is to first understand the interactions at the interfacial region and the C-terminal tail.

Table 1.2. Effect of mutating interfacial residues and C-terminal tail of $CynD_{pum}$ and $CynD_{stut}$

Mutant	Location	Change	Observed effect	References
B. pumilus				
1. Delta 303	C-terminal tail	V303→stop[a]	Full activity	Sewell *et al.,* 2005
2. Delta 293	C-terminal tail	M293→stop[a]	Partial activity	Sewell *et al.,* 2005
3. Delta 279	C-terminal tail	Y276→stop[a]	Inactive	Sewell *et al.,* 2005
4. Y201D/A204D	A surface	Y201→D, A204→D	Inactive	Sewell *et al.,* 2005
5. Delta 219-233	C surface	219MKEMICLTQEQRDYF233 was deleted	Inactive	Sewell *et al.,* 2005
6. 90	D surface	90EAAKRNE97→AAARKNK	Full activity	Sewell *et al.,* 2005
7. Q86R	D surface	Q86→R substitution	Full activity and improved stability with formation of long fibre at pH 5.4	Wang *et al.,* 2011 (part of this work)
8. C5	D surface	Q86→R,E96→G,D254→E substitution	Full activity and improved stability with formation of long fibre at pH 5.4	Wang *et al.,* 2011 (part of this work)
9. C5+E327G	D surface + C-terminal tail	Q86→R,E96→G,D254→E, E327→G substitution	Full activity and improved stability with formation of long fibre at pH 5.4	Wang *et al.,* 2011 (part of this work)
P. stutzeri				
10. Delta 310	C-terminal tail	310→stop[a]	Inactive	Sewell *et al.,* 2005
Hybrids				
11. $CynD_{pum-stut}$	C-terminal tail	Residues 1–286 from *B. pumilus*→ 287–end from *P. stutzeri*	Full activity and improved stable	Sewell *et al.,* 2005 (part of this work)
12. $CynD_{stut-pum}$	C-terminal tail	Residues 1–286 from *P.stutzeri*→ 287–end from *B.pumilus*	Inactive	Sewell *et al.,* 2005

[a] Truncation at the specified residue position

1.8 Motivation and study objectives

Motivation

Cyanide dihydratase enzymes have potential to be used in detoxification of cyanide waste produced during industrial processes, such as leach mining, metal finishing, and electroplating. These processes produce cyanide-containing wastewater which must be detoxified before it's released in to the environment. Cyanide is a potent inhibitor of cytochrome oxidase, making it toxic to aerobic organisms, and extended exposure of humans to cyanide has been linked to certain neuropathies. Conventional methods used for cyanide detoxification, such as alkaline chlorination, use hazardous chemicals that are both costly and generate toxic by-products and therefore the use of biological based methods is desirable (Young and Jordan, 1995). Bacterial detoxification could provide a cheap, on-site enzymatic solution. $CynD_{pum}$ is a good candidate because it does not need additional cofactors or substrates and because of its functionality over a wide substrate concentration (Jandhyala *et al.,* 2005). The drawback for this cyanide dihydratase enzyme is that cyanide wastes are alkaline and inactivate the wild type enzymes therefore enhanced stability and activity at alkaline pH would allow for a more effective use of this enzyme.

So far crystallisation of the branch 1 nitrilase members has not been accomplished as previously mentioned. This has made it difficult to obtain accurate 3D coordinates of $CynD_{pum}$ hence making further studies difficult due to the reliance on predicted structures. $CynD_{pum}$ monomers have been shown to associate and form complex helical quaternary structures. Although these helical structures have been successfully solved using the 3D Electron Microscopy method, the structures' resolution obtained so far are surpassed by the resolution of other biological macromolecules such as viruses, intracellular actin filaments, myosin filaments, microtubules, extracellular collagen, keratin and bacterial flagella whose structures were obtained using the same method. One possible factor preventing the acquisition of high resolution 3D Electron Microscopy structures of $CynD_{pum}$ could be the proteins low stability which results in structural variability of the protein in solution. The measure of thermostability has been successfully used to assess the stability of various proteins and has helped optimise stability of proteins for use in subsequent structural characterisation techniques such as X-ray crystallisation (Ericsson *et al*., 2006, Krajewski *et*

al., 2008). The thermostability of $CynD_{pum}$ has been demonstrated to improve in the presence of certain mutations (Wang *et al*., 2011) thus measuring thermostability could directly help identify protein stabilizing mutations. The results will not only help engineer alkaline pH stable $CynD_{pum}$ variants but will also help in the identification of mutations that result in a more stable enzyme with reduced structural variability thus enable the acquisition of a high resolution Electron Microscopy maps of $CynD_{pum}$ mutant which will be useful in future structural studies.

Study objectives

The purpose of this project is to understand the structure and function of $CynD_{pum}$ to help engineer mutants with improved stability. This was to be accomplished by addressing three main objectives:

- **To design, purify and visualise $CynD_{pum}$ mutants with improved structural stability (Chapter II).** Based on previous mutagenesis and structural studies, various D surface and C-terminal $CynD_{pum}$ mutants with improved stability will be designed and purified. To confirm structural stability, the mutants will also be visualized under a transmission microscope at various pH to observe their oligomeric pattern at acidic and alkali pH.
- **To identify $CynD_{pum}$ mutations that confer alkaline pH stability and thermostability (Chapter III).** Enzyme activity assays will be used to characterise the pH stability and thermostability stability of the $CynD_{pum}$ mutants. The melting temperatures of the $CynD_{pum}$ mutants will also be assayed using differential scanning fluorimetry to confirm stability.
- **To investigate the structural role of the $CynD_{pum}$ mutations (Chapter IV)**. Electron microscopy 3D reconstruction techniques will be used to visualise structural differences in the quaternary structure between the $CynD_{pum}$ mutants and wild type at various pH. A homology model will be docked into the 3D maps constructed to help understand the mechanism behind the stability at an atomic level.

This investigation will give us a better understanding of the basis for the transition from short to long helices and better understanding of factors that determine the high pH and thermal stability leading to the design of useful industrial enzymes. This study can potentially give more insights into other nitrilases and related enzymes due to the high structural homology of the nitrilase superfamily.

CHAPTER II

PURIFICATION AND MICROSCOPY

2.1 Introduction

$CynD_{pum}$ has potential applications in bioremediation of industrial cyanide wastewater but to order to accomplish this, stable variants of this enzyme have to be engineered to prevent denaturation by the alkaline pH conditions of the wastewater. Variants of $CynD_{pum}$ were constructed through mutagenesis and the mutants produced were purified and visualised by transmission electron microscopy. Correlation between oligomerisation and stability of $CynD_{pum}$ has been demonstrated (Wang *et al.,* 2011) and thus visualisation of the degree of oligomerisation by the introduced mutations may provide a preliminary indication of stability. Described in this chapter are the methods and results of the expression, purification and visualisation of $CynD_{pum}$ wild type and mutants. The mutants were at various pH to observe their oligomeric pattern at acidic and alkaline pH. The mutants include two with D-surface mutations (Q86R and Q86R+E96G +D254E+E327G), six C-terminal histidine mutants (H305K, H308K, H323K, H305K+H308K, H305K+H323K, H308K+H323K and H308K+H305K+H323K) and one hybrid mutant were the *B. pumilus* C-terminal tail is mutated to the tail of a non-fibre-forming homolog cyanide dihydratase from *Pseudomonas stutzeri* ($CynD_{pum\text{-}stut}$).

2.2 Origin of mutations

A series of stable mutants were identified by our collaborator Dr Michael J. Benedik (Texas A&M University, Department of Biology, USA) using a random mutagenesis technique that enables rapid generation of random mutant libraries (Abou-Nader and Benedik, 2010). During the random mutagenesis experiments Dr Michael J. Benedik identified Q86R+E96G +D254E+E327G (Table 2.1, strain 9) as a alkaline pH and thermostable $CynD_{pum}$ mutant, and from this insight we designed the Q86R mutant (Table 2.1, strain 8) which based on

structural studies appeared likely to be the main mutation to confer these new properties. Literature and previous work (Thuku *et al.*, 2007; Kiziak *et al.,* 2007) have demonstrated the participation of the C-terminal tail in oligomerisation, activity and stability, and this led to the design of the C-terminal tail mutants. In view of the fact that the transition from short spirals to long helices occurs at pH 6, the participation of histidine was proposed as it is the only amino acid charged at pH 5.4 (histidine pKa = 6.0). There are 3 histidines located in the C-terminal tail of $CynD_{pum}$. This led to the design of mutants in which the C-terminal tail histidines were mutated to lysines in a way that covers all possible permutations (7 possible arrangements). These mutants include H305K, H308K, H323K, H305K+H308K, H305K+H323K, H308K+H323K and H308K+H305K+H323K. The H323K mutant was omitted from this study since it was previously characterised and showed no effect on stability (Mulelu*,* 2010).The histidines were converted to lysine to allow for the retention of the positive charge above pH 6 (Lys pKa = 10.5) as this will allow us to observe any effects due to changes in the residue charges. The genes for the $CynD_{pum}$ mutants cloned into pET28b expression plasmids were provided by our collaborator Dr Michael J. Benedik (Texas A&M University, Department of Biology, USA).

Table 2.1. List of mutants used in this study. The location of the mutations were based on the wild type homology model.

Strains	**Mutations**	**Location of mutations on $CynD_{pum}$ structure**
H305K	H305K	C-terminal tail
H308K	H308K	C-terminal tail
H305K+H308K	H305K and H308K	C-terminal tail
H305K+H323K	H305K and H323K	C-terminal tail
H308K+H323K	H308K and H323K	C-terminal tail
H308K+H305K+H323K	H308K, H305K and H323K	C-terminal tail
Q86R	Q86R	D-surface
Q86R+E96G +D254E+E327G	Q86R, E96G, D254E and E327G	D-surface and C-terminal tail
$CynD_{pum\text{-}stut}$	C-terminal tail residues 287-end from *B. pumilus* (SNQSLSMNFNQQPTPVVKHLNHQKNEVFTYEDIQYQHGILEEKV) mutated to C-terminal tail residues 287-end from *P. stutzeri* (NQSLSMNFNQSPNPVVRKIGERDSTVFTYDDLNLSVSDEEPVVRSLR K)	C-terminal tail

2.3 Materials and Methods

Transformation and Expression

$CynD_{pum}$ plasmid DNA (100 ng) was added to ice thawed *Escherichia coli* BL21 (DE3) chemically competent cells, and the mixture was incubated on ice for 30 min. The mixture was then heat-shocked for 45 seconds at 42°C followed by a 2 min chill on ice. Nutrient broth recovery medium was added to the mixture and incubated at 37°C with shaking. The cells were then plated on LA + kanamycin plates (25 µg/ ml kanamycin) and incubated at 37°C overnight. One of the *E. coli* BL21 transformed colonies was used to grow 5 ml of overnight starter culture which was then added to 1 litre of nutrient broth (25 µg/ ml kanamycin) and cultured at 37°C to an $OD_{600\ nm}$ of 0.4–0.6. Expression was then induced by adding 1 mM IPTG to the culture. After the addition of IPTG the culture were incubated overnight at 30°. The culture was then centrifuged at 4000 x g for 15 min and the cell pellet was stored at -20°C.

Cell lysis

The stored cell pellets were thawed and resuspended in 50 mM Tris-HCl, pH 8 buffer with EDTA-free protease inhibitors (Roche Ltd, USA) on ice and then sonicated for 6 min (15sec pulsar on/off) using Sonicator® 3000 (Misonix, USA). An ethanol and ice water bath was used to cool the sample while sonicating. After sonication samples were centrifuged at 20 000 x g for 30 min and ammonium sulphate precipitation was immediately carried out.

Ammonium sulphate $(NH_4)_2SO_4$ precipitation (salting out)

Increasing amounts of ammonium sulphate were added using 10% saturation increments from 20%–50%. After each addition of ammonium sulphate, the sample was incubated for 20 min on ice to allow proteins to precipitate out of solution. After incubation, the suspension was centrifuged at 10000 x g for 10 min at 4^0C. The supernatant was then collected and its volume was measured in order to calculate the mass of ammonium sulphate needed for the next 10% saturation increase. The ammonium sulphate was then added to the supernatant while the pellet was resuspended in 50 mM Tris-HCl, pH 8 b0uffer. The process was repeated until there was no further apparent protein precipitation.

Samples were taken for analysis by SDS PAGE gel (described below) to determine which fraction contained the $CynD_{pum}$.

Anion exchange chromatography

The ammonium sulphate precipitation fraction with the $CynD_{pum}$ was run through a Q Sepharose XL anion exchange column (Amersham Biosciences, Sweden) for further protein separation. The proteins were eluted with a linear gradient from low salt buffer (50 mM Tris-HCl pH 8, 100 mM NaCl) to high salt buffer (50 mM Tris-HCl pH 8, 1M NaCl) at a flow rate of 2.5 ml/min. SDS PAGE gels (described below) were run for all fractions that produced a A_{280nm} peak and the results were used to identify the fraction that contained the $CynD_{pum}$.

Gel filtration

The protein solution from the anion exchange chromatography was then run through a TSKgel PWXL4000, HPLC gel filtration column (Tosoh Corporation, Tokyo) pre-equilibrated and run with 150 mM NaCl in 50 mM Tris-HCl pH 8 buffer, at a flow rate of 0.5 ml/min. $CynD_{pum}$ fractions with the highest A_{280nm} peak were collected and tested for activity using the picric acid activity assay (described below). Active fractions were buffer exchanged into appropriate pH buffers and then visualised by electron microscopy.

Enzyme assay

The picric acid activity assay (Fisher and Brown, 1952) was used to measure nitrilase activity of the purified protein. 70 µl of 50 mM Tris-HCl pH 8 buffer was added to 10 µl of protein and then 20 µl of 25 mM KCN was added to this protein mixture (5 mM final [KCN] in 100 µl reaction mixture). The reaction was allowed to continue for 20 min at 37°C and stopped by addition of 80 µl of the picric acid/carbonate mixture (50% picric acid, 50% carbonate). The samples were then boiled for 5 minutes to allow the colour complex to develop. Room temperature distilled water (1 ml) of was added to stop the colour reaction. Absorbance of 100 µl of this solution was measured at 520 nm on a Multiskan spectrophotometer (Titertek, Helsinki). One unit of activity is defined as the degradation of 1 µmol of cyanide per minute.

Protein concentration determination

The concentration of protein was determined using the Bradford assay (Bradford, 1976). Bovine serum albumin was used to prepare 0.1–0.5 mg/ml protein standards. The protein samples (10 µl) were mixed with 200 µl of the Bio-Rad protein assay solution (Bio-Rad, USA) and allowed to incubate for 10 min at room temperature. After 10 min the absorbance was measured at 595 nm in a Multiskan spectrophotometer (Titertek, Helsinki)

SDS-PAGE

SDS-PAGE was performed on BIORAD Mini-PROTEAN® 3 Cell (BIORAD, UK) gel apparatus. 15 µl of protein samples were mixed with 5 µl of 4× Sample Loading Buffer. 10 µl of the mixture was boiled for 10 min and then run on a 10% SDS polyacrylamide gel. Page Ruler Prestained protein ladder (Fermentas, Canada) was used to determine the molecular weight of the protein. After the run, the SDS gels were stained with Coomassie Brilliant Blue R-250.

Negative stain electron microscopy

Purified samples from gel filtration chromatography were buffer exchanged in buffers of desired pH using Nanosep, 10kDa molecular weight cut-off, spin columns (Microsep Pty Ltd, South Africa). The samples were buffer exchanged with 50 mM Citrate/ Na_2HPO_4 buffer at pH 5.4, 50 mM glycine/NaOH buffer at pH 9 and left in 50 mM Tris-HCl for pH 8 visualisation. Carbon-coated copper grids were glow-discharged at a plasma current of 25 mA for 30 seconds using a EMS100× Glow Discharge Unit (Electron Microscopy Sciences, USA). 3 µl of protein solution was then placed on the grid and allowed to stand for 30 seconds. After the 30 seconds the grid was washed with 2 drops of water then with 3 drops of uranyl acetate while blotting with filter paper between each wash. The grids were blotted and air dried then loaded into the LEO 912 OMEGA transmission electron microscope with a Proscan 2048 x 2048 CCD camera. Snapshot micrographs were collected at 50 000 x calibrated magnification. The uranyl acetate stain coats the molecules on the grid and produces a 'shell' of contrast which allows us to reconstruct the outer shape of the molecule when viewing under the electron microscope. For each enzyme and pH condition, 20 oligomers

were selected and measured using the BOXER program of the EMAN suite (Ludtke *et al.,* 1999), and then average lengths of the oligomers were calculated.

2.4 Results

Purification

A three step purification process was used to purify the $CynD_{pum}$ enzyme. All the $CynD_{pum}$ precipitated out at the 40–50% saturation during the ammonium sulphate fractional precipitation. Cyanide degrading activity was measured for each of these steps and purification tables were constructed (Table 2.2–2.11). Total activity and yield values were omitted because of an incorrect measurement method. Cyanide degrading activity of the ammonium sulphate precipitation fractions could not be measured because the ammonium sulphate interfered with the activity assay. The purification tables of $CynD_{pum}$ mutants and wild type demonstrated a general increase in the purification fold and specific activity as the purification steps were carried out, with the gel filtration step showing the most purity except in the case of H305K, H308K, H305K+H308K and Q86R which showed a slight decrease in specific activity compared to the previous purification step. This decrease in specific activity was mainly due to the freezing of these particular mutants before running them through gel filtration resulting in slight denaturation of the protein. The purification was also demonstrated by the appearance of a more pronounced and distinct band, with fewer background bands visible, of about 37kDa in the SDS-PAGE gels as the purification steps were performed (Fig. 2.1b–2.10b). Comparing the SDS-PAGE gels and the purification tables it was evident that the purification fold presented by the tables was lower than what one would expect based on the gels. This demonstrates that there may be loss of activity which could be due to oxidation of the active site cysteine or the denaturation caused by the freezing and thawing of the protein during storage as previously mentioned. Activity assay were carried out on gel filtration chromatography fractions with protein (Fig. 2.1a–2.10a). The gel filtration elution profiles of each of the enzymes consisted of a distinct sharp peak around fraction 50 which was confirmed to have the most activity (Fig. 2.1a–2.10a). The H305K+H308K+H323K mutant gel filtration elution profile demonstrated a broader peak than the other mutants, with elution of the $CynD_{pum}$ starting as early as fraction 40. This broad peak was confirmed to be due to the high heterogeneity of $CynD_{pum}$ oligomers.

Table 2.2. Wild Type purification table

Purification Step	Protein Concentration (mg)	Activity (U)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.76	13.26	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.70	29.78	2.25
Gel filtration	0.057	2.33	40.90	3.08

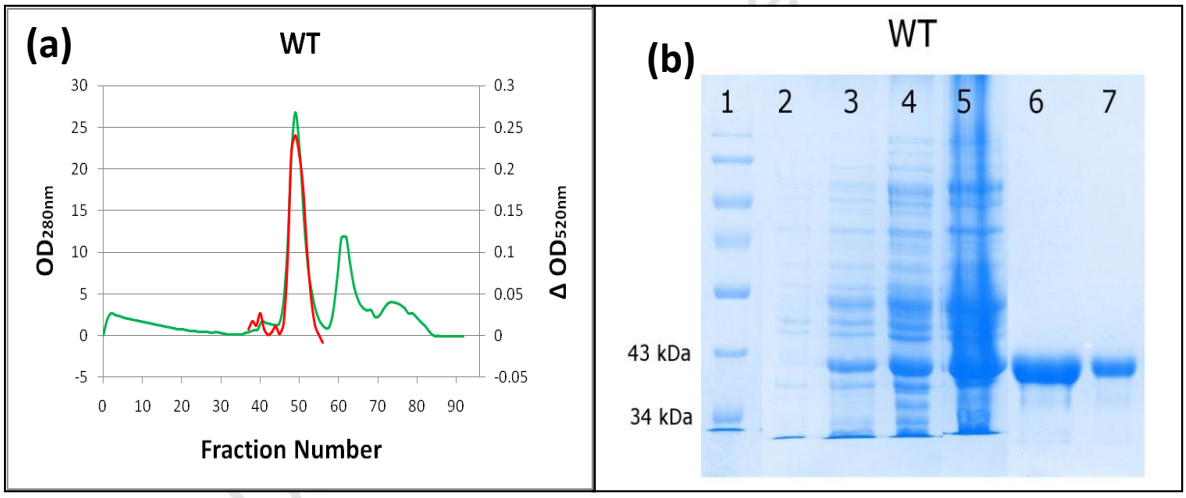


Figure 2.1. (a) The gel filtration elution profile of the wild type. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the wild type. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.3. H305K purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.37	6.43	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.77	31.01	4.83
Gel filtration	0.057	0.99	17.34	2.70

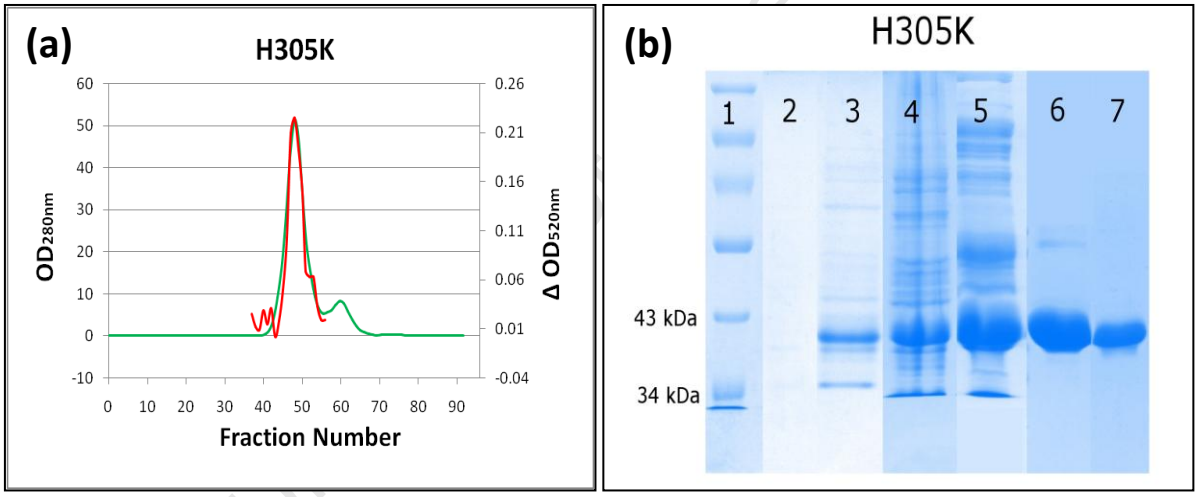


Figure 2.2. (a) The gel filtration elution profile of the H305K mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the H305K mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.4. H308K purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.56	9.89	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.53	26.93	2.72
Gel filtration	0.057	0.99	17.34	1.75

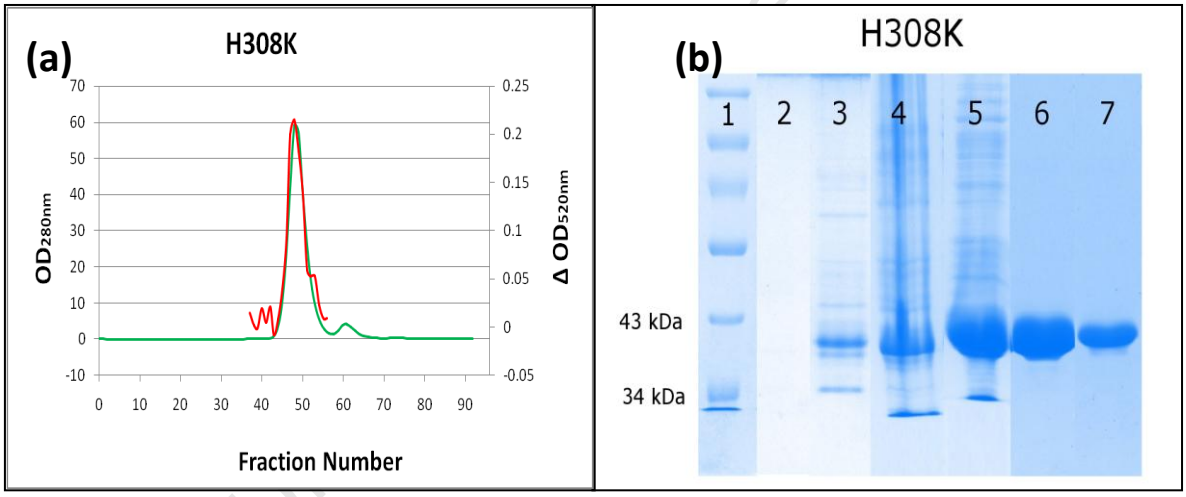


Figure 2.3. (a) The gel filtration elution profile of the H308K mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the H308K mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.5. H305K+H308K purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.39	6.83	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.35	23.77	3.48
Gel filtration	0.057	1.13	19.89	2.91

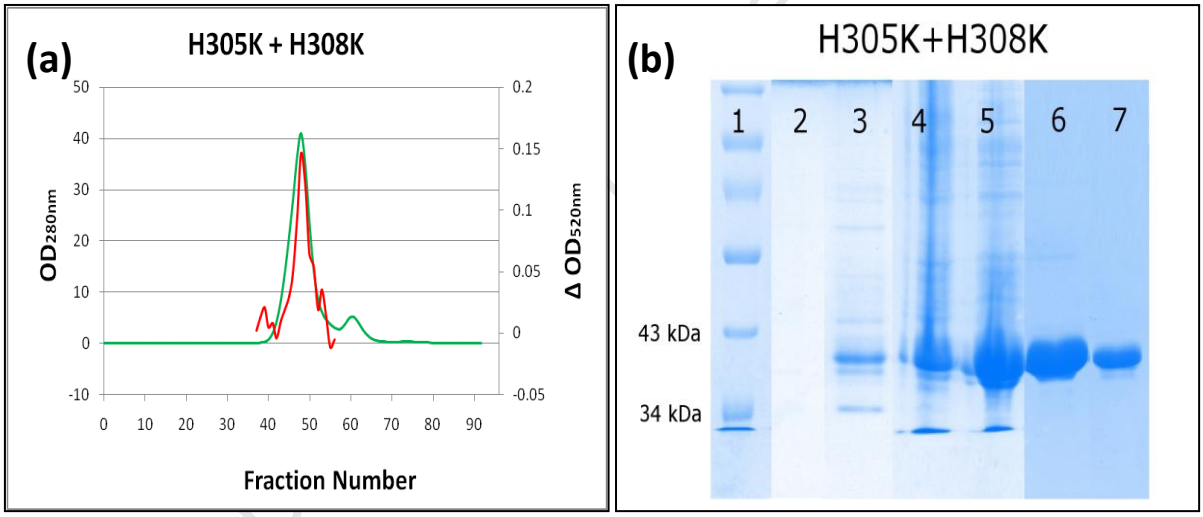


Figure 2.4. (a) The gel filtration elution profile of the H305K+H308K mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the H305K+H308K mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.6. H305K+H323K purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.35	6.13	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.69	29.72	4.84
Gel filtration	0.057	2.00	35.09	5.72

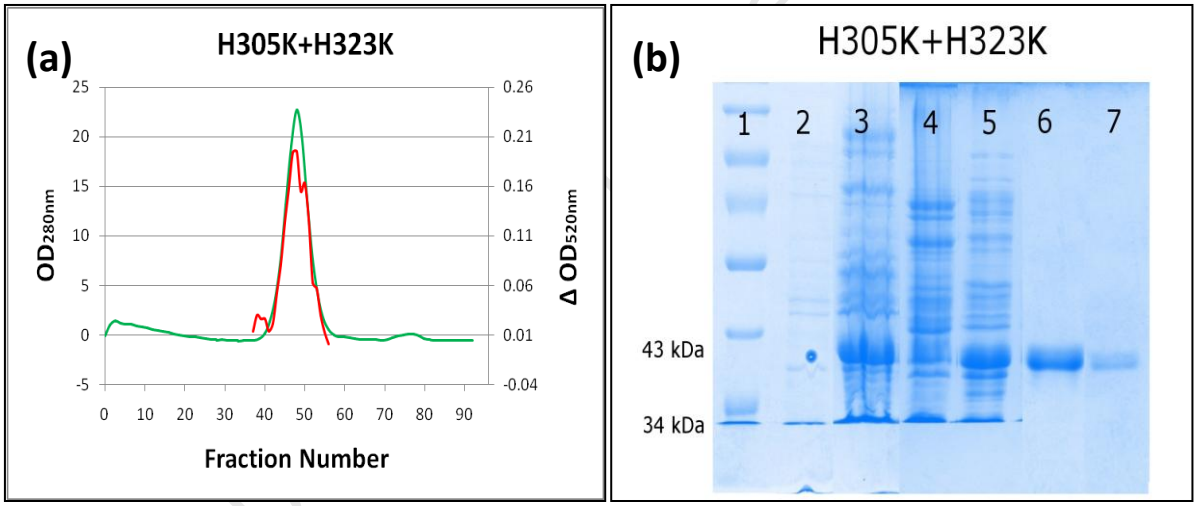


Figure 2.5. (a) The gel filtration elution profile of the H305K+H323K mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the H305K+H323K mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.7. H308K+H323K purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.44	7.68	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.54	27.00	3.51
Gel filtration	0.057	2.19	38.42	5.00

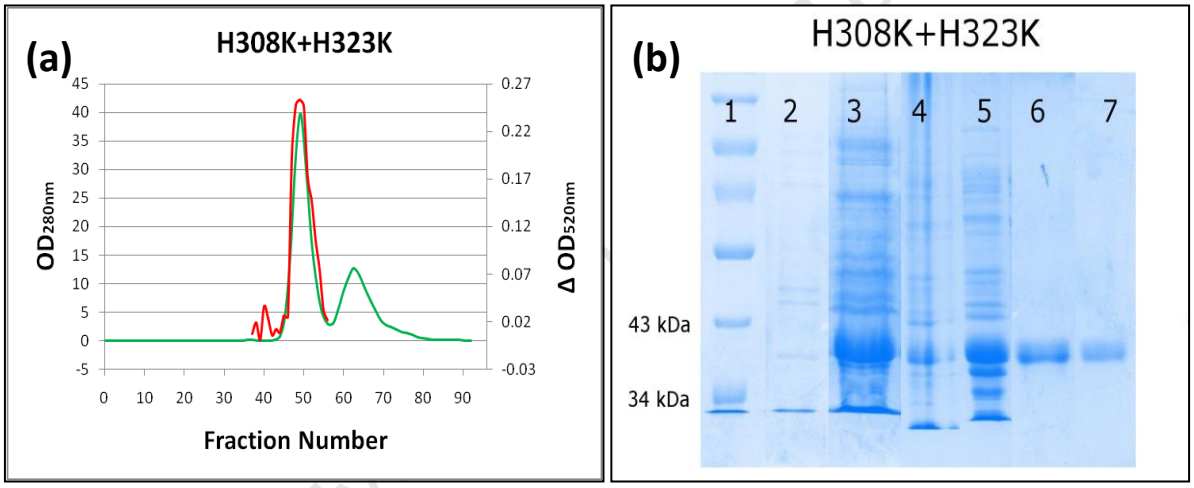


Figure 2.6. (a) The gel filtration elution profile of the H308K+H323K mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the H308K+H323K mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.8. H305K+H308K+H323K purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.50	8.71	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.94	34.00	3.90
Gel filtration	0.057	2.08	36.45	4.18

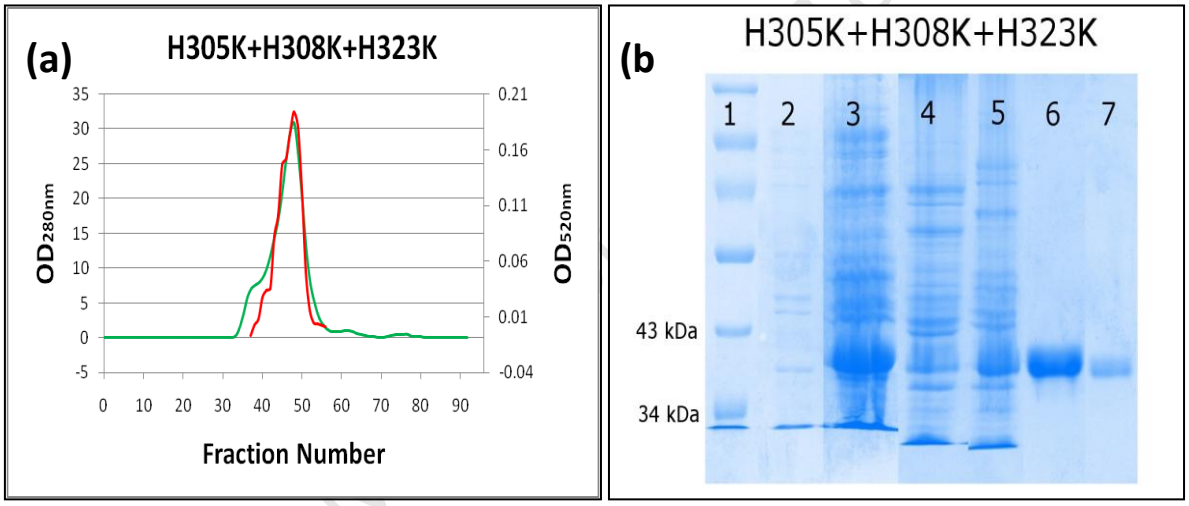


Figure 2.7. (a) The gel filtration elution profile of the H305K+H308K+H323K mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the H305K+H308K+H323K mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.9. Q86R purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	1.60	28.15	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	2.67	46.85	1.66
Gel filtration	0.057	2.62	45.90	1.63

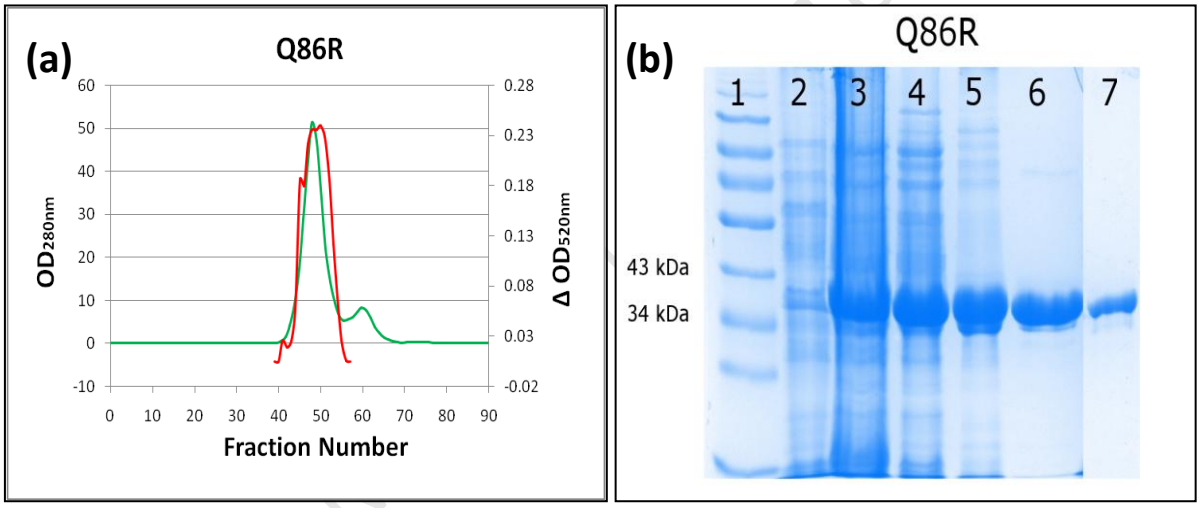


Figure 2.8. (a) The gel filtration elution profile of the Q86R mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the Q86R mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.10. Q86R+E96G+D254E+E327G purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	1.57	27.61	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	4.67	82.01	2.97
Gel filtration	0.057	3.79	66.50	2.41

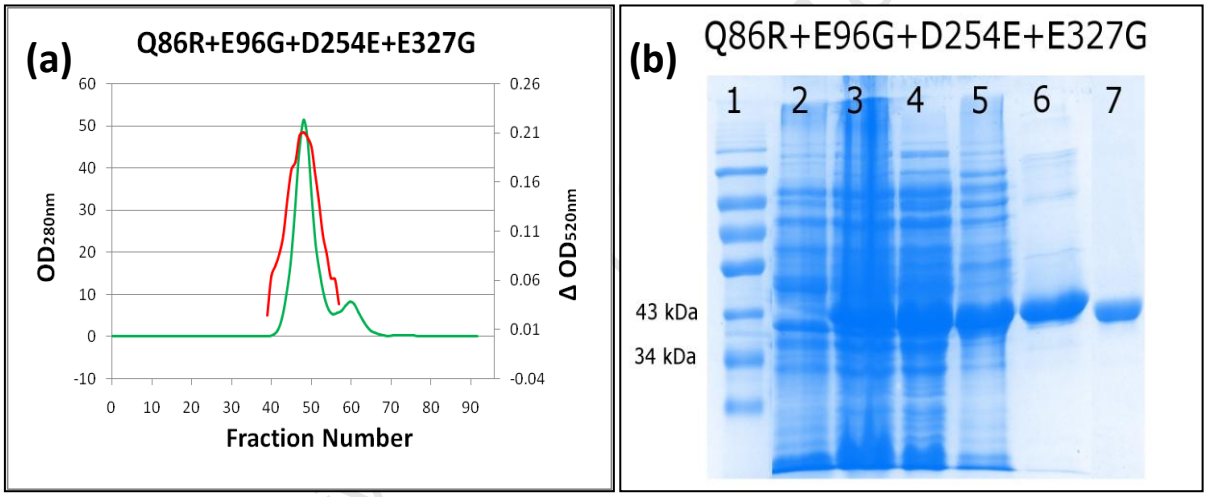


Figure 2.9. (a) The gel filtration elution profile of the Q86R+E96G+D254E+E327G mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the Q86R+E96G+D254E+E327G mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Table 2.11. $CynD_{pum\text{-}stut}$ purification table

Purification Step	Protein Concentration (mg)	Total Activity(U/ml)	Specific Activity (U/mg)	Fold Purification
Lysate	0.057	0.65	11.42	1.00
$(NH_4)_2SO_4$ precipitation	0.057	–	–	–
Anion exchange	0.057	1.59	27.98	2.45
Gel filtration	0.057	1.60	28.15	2.46

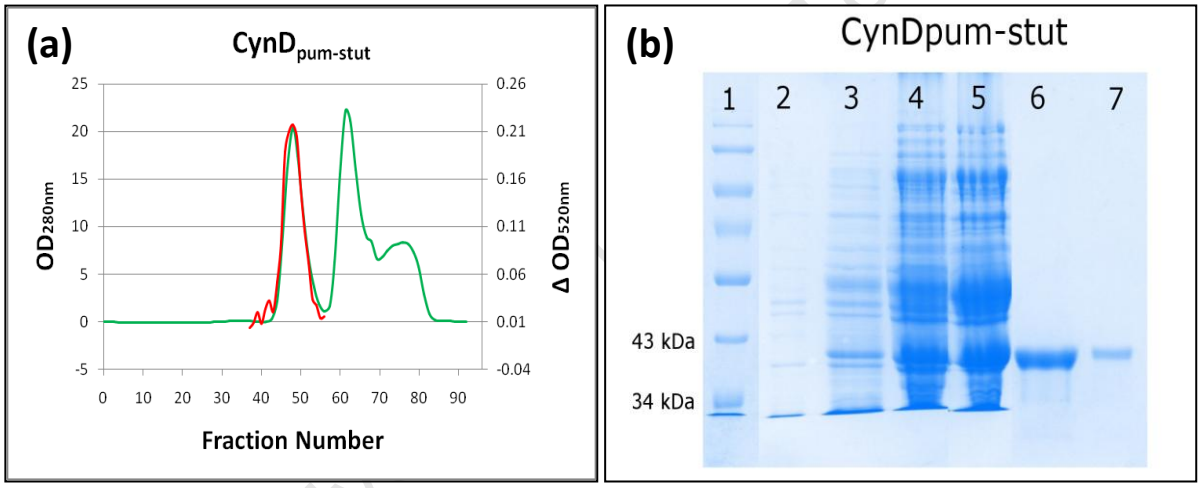


Figure 2.10. (a) The gel filtration elution profile of the $CynD_{pum\text{-}stut}$ mutant. Protein content (green) of each fraction was obtained by measuring the absorbance at 280 nm and cyanide degrading activity of the peak fractions (red) was obtained by measuring absorbance of the cyanide-picrate complex at 520 nm. **(b)** 12% SDS PAGE gels were run for each purification step of the $CynD_{pum\text{-}stut}$ mutant. Lanes: 1, molecular mass ladder; 2, uninduced cell extract; 3, induced cell extract; 4, lysate; 5, ammonium sulphate precipitation fraction; 6, anion exchange eluate; 7, gel filtration eluate.

Microscopy

Electron microscopy confirmed that the material eluted from the gel filtration column contained pure $CynD_{pum}$ protein (Fig. 2.11 and 2.12). Although some micrographs contained a heterogeneous mixture of participles of different sizes, a predominant size could be observed within the population. The oligomers formed at the various pH were measured and the averages were calculated (Table 2.12). The wild type formed long fibres at pH 5.4 (89 nm), short spirals at pH 8 (23 nm) and what appeared to be partly denatured short spirals at pH 9 (12 nm) (Fig. 2.11 and 2.12). Mutants H305K, H308K, H305K+H308K, H305K+H323K and H308K+H323K formed slightly longer fibres than the wild type at pH 5.4 (about two fold increase in length for H308K, H305K+H308K and H305K+H323K) but displayed oligomers of similar size to the wild type at pH 8 and pH 9, with the exception of H308K+H323K which formed oligomers half the size of the wild type at pH 9 (Fig. 2.11 and 2.12). The triple histidine mutant, H305K+H308K+H323K formed extremely long fibres at pH 5.4 which were about 5 times longer than the wild type (Fig. 2.12). Shorter fibres were observed at pH8 for this triple histidine mutant and short spirals similar to those formed by the wild type at pH 8 were observed at pH 9 (Fig. 2.12). The Q86R and Q86R+E96G+D254E+E327G, also formed significantly longer fibres (4–6 fold longer) than the wild type at pH 5.4 (Fig. 2.12). These mutants retained their helical form at pH 8 and still showed evidence of short helices at pH 9 (Fig. 2.12). All the tail mutants formed long fibres at pH 5.4 except the $CynD_{pum\text{-}stut}$ hybrid $CynD_{pum\text{-}stut}$ which appeared to exist as short spirals at pH 5.4, pH 8 and pH 9 (Fig. 2.12).

Table 2.12. Effect of pH on the oligomerisation state of $CynD_{pum}$ wild type and mutants. Values were calculated from 20 selected oligomers per pH condition.

Strain	Length at pH 5.4 (nm)		Length at pH 8 (nm)		Length at pH 9 (nm)	
	Average	(Range)	Average	(Range)	Average	(Range)
Wild type	89	(29–238)	23	(12–42)	12	(6–18)
H305K	131	(51–284)	30	(14–53)	11	(6–17)
H308K	204	(62–837)	25	(11–37)	15	(9–24)
H305K+H308K	241	(72–554)	34	(21–53)	15	(8–28)
H305K+H323K	196	(55–490)	36	(19–81)	10	(6–16)
H308K+H323K	111	(40–312)	30	(15–47)	7	(5–9)
H308K+H305K+H323K	335	(46–724)	94	(49–166)	29	(10–89)
Q86R	>547	(31–>1455)	73	(30–171)	16	(10–31)
Q86R+E96G +D254E+E327G	324	(24–1139)	51	(25–121)	60	(29–100
$CynD_{pum–stut}$	32	(15–54)	29	(19–37)	26	(16–43)

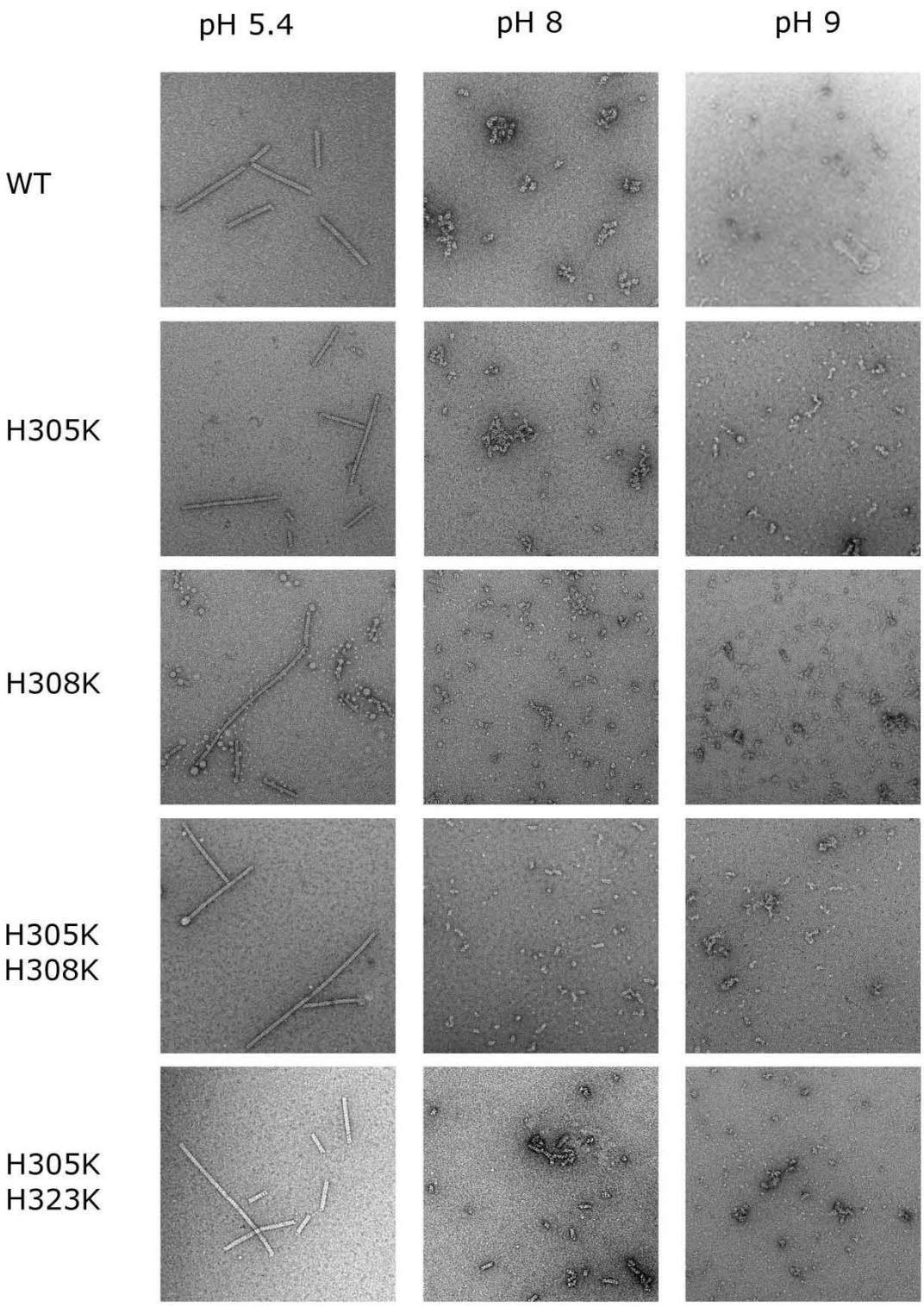


Figure 2.11. Negative-stain electron micrographs of $CynD_{pum}$ mutants H305K, H308K, H305K+H308K, H305K+H323K and wild type at pH 5.4 (1st column), pH 8 (2nd column) and pH 9 (3rd column)

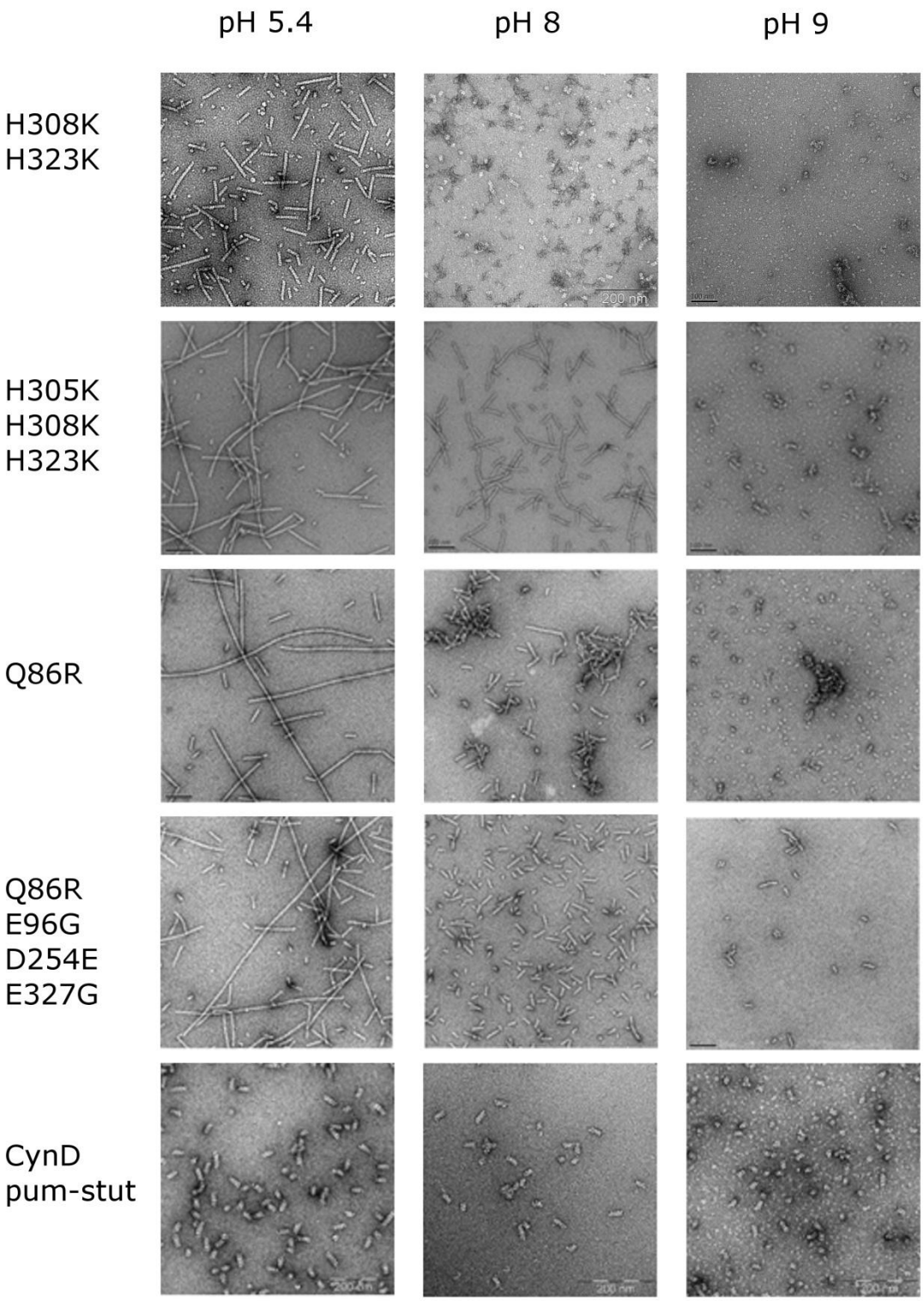


Figure 2.12. Negative-stain electron micrographs of $CynD_{pum}$ mutants H308+H323K, H305K+H308K+H323K, Q86R, Q86R+E96G +D254E+E327G and $CynD_{pum–stut}$ at pH 5.4 (1st column), pH 8 (2nd column) and pH 9 (3rdcolumn)

2.5 Discussion

The $CynD_{pum}$ wild type and mutants were successfully purified and were confirmed by the presence of activity in the gel filtration peak fractions and purification table demonstrating the high purity of the enzyme preparations. SDS PAGE gel also confirmed the purity of the enzyme preparations with the observed 37kDa band corresponds to the size of a $CynD_{pum}$ monomer.

Any possible changes in the oligomerisation of the $CynD_{pum}$ enzymes were visualised using electron microscopy. The full length of a functional $CynD_{pum}$ wild type at pH 8 has been reported to have maximum dimension of 21.2 nm and exists as 18 subunit oligomer (Jandhyala *et al.,* 2003). Given that the observed size distribution of wild type spirals at pH 8 (23 nm) is entirely consistent with the previously reported dimension and we propose that the enzyme is primarily in the octadecameric state. The wild type formed short spirals at pH 8 which associated to from long fibres at pH 5.4. This observed oligomerisation pattern of the wild type is consistent with previous observations (Jandhyala *et al.,* 2003).

The micrographs clearly demonstrated that the H305K+H308K+H323K, Q86R, Q86R+E96G +D254E+E327G and the $CynD_{pum\text{-}stut}$ mutations had dramatic effect on the oligomerisation when compared to the $CynD_{pum}$ wild type. At this point we can only speculate that these effects on the oligomerisation are due to the changes in charges of the mutated residue. The mutations may have introduced or eliminated charges resulting in the formation of new interactions that affected oligomerisation by altering the folding of the $CynD_{pum}$ enzyme. Q86R and Q86R+E96G +D254E+E327G displayed a similar pH-dependent transition (from short oligomers at pH 8 to longer oligomers at pH 5.4) as observed for the wild type although there was a higher degree of oligomerisation. This demonstrates that the same residues responsible for the pH-dependent transition of the wild type may still be present in the mutants hence the pH-dependent transition may not be due to the mutated residues. Q86R's increase in length of fibres at pH 5.4 and pH 8 compared to the wild type, demonstrates improved stability relative to wild type and that the Q86R may solely be responsible for this increase in the degree of oligomerisation. The Q86R mutation introduces a positive charge which may result in the formation of new electrostatic interactions which stabilise the enzyme during oligomerisation. The Q86R+E96G

+D254E+E327G displayed the same degree of oligomerisation, at pH 5.4 and pH 8, as Q86R but retained some of its fibres at pH 9 demonstrating more stability than Q86R. In addition to the Q86R mutation, the Q86R+E96G +D254E+E327G mutant has an E96G mutation which eliminates a positive charge. The elimination of this charge may allow other previously hindered interactions to occur which offer better stabilization of the enzyme. The D254E mutation increases the side chain length of a negatively charged residue and this may allow new electrostatic interactions with previous inaccessible residues. The E327G mutation is located on the C-terminal tail of $CynD_{pum}$ (based on the amino acid sequence) and it eliminates a negative charge and this may allow previously hindered interaction, which better stabilize the enzyme to occur. The E327G single mutation has been shown to be involved in the stabilisation of the quaternary structure of $CynD_{pum}$ but to a less extent compared to Q86R (Wang *et al.,* 2011). The extent at which the individual Q86R+E96G +D254E+E327G mutations contribute to stability may be difficult to assess but due to the similar degree of oligomerisation between Q86R+E96G +D254E+E327G and Q86R, it is likely that the Q86R mutation is largely responsible for the observed increase in stability.

Histidine residues have previously been speculated to be involved in the pH-dependent transition from short spirals at pH 8 to long fibres at pH 5.4 (Sewell *et al.,* 2005). Three histidines are located on the C-terminal tail of $CynD_{pum}$ (H305, H308 and H323) and these were mutated to lysine and visualised by electron microscopy at different pH. H305K+H308K+H323K displayed considerably longer fibres than the wild type at pH 5.4 and still retained the fibres at pH 8 and a few at pH 9, thus demonstrating an increase in structural stability. The H305K+H308K+H323K mutant displayed the most pronounced effects on the oligomerisation of $CynD_{pum}$ compared to the other histidine mutants, demonstrating that all three histidine to lysine mutations are required for the increase in structural stability. The observed effect of the mutations on the oligomerisation of the $CynD_{pum}$ enzyme at pH 5.4 may due to the presence of more favourable electrostatic interaction at the C-terminus produced by the lysines rather than the histidines. The presence of fibers at pH 8 and pH9 may be due to the retention of the positive charge, which may be responsible for the stabilizing the long fibres at pH 5.4, at higher pH due the lysines higher pKa (Lys pKa=10.5) rather than that of the histidines. Although one would expect H305K+H308K+H323K to form long fibres at all the pH (due to the retention of the

positive charge at pH above 6), to demonstrate the involvement of the C-terminal histidines in the pH-dependent oligomerisation, we could still propose that these C-terminal histidines are in part involved in the pH-dependent oligomerisation and also confer structural stability at alkaline pH.

To further investigate the influence of the tail on the oligomerisation of $CynD_{pum}$, a $CynD_{pum\text{-}stut}$ mutant was constructed were the *B. pumilus* C-terminal tail was mutated to the C-terminal tail of a non-fibre-forming homolog cyanide dihydratase from *Pseudomonas stutzeri* ($CynD_{stut}$). $CynD_{stut}$ has been shown to exist as short spirals at pH 8 and does not undergo the pH-dependent transition to long fibres at pH 5.4 (Sewell *et al*., 2003). The absence of the pH-dependent transition has been proposed to be due to the absence of histidines on the C-terminal tail of $CynD_{stut}$ (Fig. 2.13). $CynD_{pum\text{-}stut}$ existed as short spirals at pH 5.4, pH 8 and pH 9, an oligomerisation pattern similar to the one displayed by $CynD_{stut}$ (Sewell *et al*., 2003). This shows evidence that the C-terminal tail may be required for the pH-dependent oligomeric transition of $CynD_{pum}$. This hypothesis is further verified by the cyanide dihydratase from *Bacillus pumilus* 8A3, a strain similar to that of *Bacillus pumilus* C1. The cyanide dihydratase from *B. pumilus* 8A3 lacks the histidines on the C- terminal tail (Fig. 2.13) and has been shown to form aggregated and irregular strands at pH below 6 instead of the regular fibres formed by the *B. pumilus* C1 strain (Eicher, 2007). Although our previous experiment did not clearly show that the C-terminal histidines are exclusively responsible for the pH-dependent oligomerisation, the lack of fibre formation at pH 5.4 by $CynD_{pum\text{-}stut}$ mutant and the *B. pumilus* 8A3 cyanide dihydratase clearly demonstrate the involvement of the C-terminal tail on the pH-dependent oligomerisation of $CynD_{pum}$. The pH-dependent transition may involve other residues on the C-terminal tail and not exclusively the histidines.

```
BpumC1     1:  MTSIYPKFRA AAVQAAPIYL NLEASVEKSC ELIDEAASNG AKLVAFPEAF LPGYPWFAFI GHPEYTRKFY
Pstut      1:  -MAHYPKFKA AAVQAAPVYL NLDATVEKSV KLIEEAASNG AKLVAFPEAF IPGYPWFAFL GHPEYTRRFY
Bpum8A3    1:  MTSIYPKFRA AAVQAAPIYL NLEATVQKSC ELIDEAASNG AKLVAFPEAF LPGYPWFAFI GHPEYTRKFY

BpumC1    71:  HELYKNAVEI PSLAIQKISE AAKRNETYVC ISCSEKDGGS LYLAQLWFNP NGDLIGKHRK MRASVAERLI
Pstut     70:  HTLYLNAVEI PSEAVQKISA AARKNKIYVC ISCSEKDGGS LYLAQLWFNP EGDLIGKHRK MRVSVAERLC
Bpum8A3   71:  HELYKNAVEI PSLAIQKISE AAKRNETYVC ISCSEKDGGS LYLAQLWFNP NGDLIGKHRK MRASVAERLI

BpumC1   141:  WGDGSG-SMM PVFQTEIGNL GGLMCWEHQV PLDLMAMNAQ NEQVHVASWP GYFD------ ----------
Pstut    140:  WGDGNG-SMM PVFETEIGNL GGLMCWEHNV PLDIAAMNSQ NEQVHVAAWP GFFD------ ----------
Bpum8A3  141:  WGDGSG-SMM PVFQTEIGNL GGLMCWEHQV PLDLMAMNAQ NEQVHVASWP GYFD------ ----------

BpumC1   194:  -DEISSRYYA IATQTFVLMT SSIYTEEMKE MICLTQEQRD YFETFKSGHT CIYGPDGEPI SDMVPAETEG
Pstut    193:  -DETASSHYA ICNQAFVLMT SSIYSEEMKD MLCETQEERD YFNTFKSGHT RIYGPDGEPI SDLVPAETEG
Bpum8A3  194:  -DEISSRYYA IATQTFVLMT SSIYTEEMKE MICLTQEQRD YFETFKSGHT CIYGPDGEPI SDMVPAETEG

BpumC1   263:  IAYAEIDVER VIDYKYYIDP AGHYSNQS-L SMNFNQQPTP VVKHLNHQKN EVFTYEDIQY QHGILEEKV-
Pstut    262:  IAYAEIDIEK IIDFKYYIDP VGHYSNQS-L SMNFNQSPNP VVRKIGERDS TVFTYDDLNL SVSDEEPVVR
Bpum8A3  263:  IAYADIDVER VIDYKYYIDP AGHYSNQS-L SMNFNQQPTP VVKQLNDNKN EVLTYEAIQY QNG MLEEKV-

BpumC1   331:  ----------
Pstut    331:  SLRK------
Bpum8A3  331:  ----------
```

Figure 2.13. Sequence alignment of the cyanide dihydratases from *Bacillus pumilus* C1 (BpumC1), *Pseudomonas stutzeri* (Pstut) and *Bacillus pumilus* 8A3 (Bpum8A3). Histidine residues are highlighted in red. The C-terminal tail is highlighted in yellow.

2.6 Conclusion

We demonstrated that Q86R and Q86R+E96G+D254E+E327G mutations improve the structural stability of $CynD_{pum}$. The $CynD_{pum\text{-}stut}$ inability to form long fibres at pH 5.4 demonstrated how the C-terminal tail is required for the pH-dependent oligomerisation of $CynD_{pum}$. The H305K+H308K+H323K mutant demonstrated how the C-terminal histidines are in part involved in the pH-dependent oligomerisation and also how they can be modified to confer structural stability.

CHAPTER III

BIOCHEMICAL CHARACTERISATION

3.1 Introduction

As previously mentioned, $CynD_{pum}$ has potential industrial applications but properties such as high alkaline pH and thermostability have to be introduced to make $CynD_{pum}$ a viable industrial enzyme. The $CynD_{pum}$ enzymes discussed in the previous chapter were biochemically characterized. In this chapter we discuss results obtain from different assays used to characterise the stability of the $CynD_{pum}$ enzymes as well as the basis of the methods used to acquire them. The pH stability and thermostability of the $CynD_{pum}$ enzymes were first measured using the picric acid activity assay. The thermostability was later measured based on the melting temperature of the $CynD_{pum}$ enzymes by differential scanning fluorimetry.

3.2 Review of enzyme assays

The picric acid assay

The picric acid activity assay measures the concentration of cyanide in a solution. It was developed based on a standard colorimetric assay for cyanide (Fisher and Brown, 1952). On addition of the yellow picric acid, the cyanide in solution reacts with the picric acid to from a red cyanide-picrate complex which can be quantified by measuring its absorbance at 520 nm. By using this assay to measure the amount of cyanide at the start and end of the reaction, one can measure the amount of cyanide degraded in a solution hence the activity can be calculated. To construct pH activity profiles for the $CynD_{pum}$ enzymes, the picric acid activity assay was carried out in the presence of different pH buffers thus the activity at the different pH could be obtained and compared. For the thermostability assays, the picric assay was carried after incubation of the $CynD_{pum}$ enzymes at high temperature.

Thermostability assay

Enzyme thermostability incorporates both thermodynamic and kinetic stability factors (Vieille and Zeikus, 2001). The two parameters can be used to describe different aspects of temperature-dependent protein unfolding. Thermodynamic stability is easily measured for proteins that undergo reversible unfolding but impossible to measure for proteins that undergo irreversible unfolding (Sterner and Liebl, 2001). For the proteins that undergo irreversible unfolding, the kinetic stability or the rate of thermal inactivation is used as the determinant of thermostability of an enzyme. In industrial bioprocesses, kinetic stability is a key determinant of biocatalyst functionality. Thermodynamic stability is dependent on the difference in the free energies between the folded and unfolded states (ΔG) whereas kinetic stability is dependent on activation energy (Ea) of unfolding (ΔG*) from folded to unfolded states (Fig. 3.1).

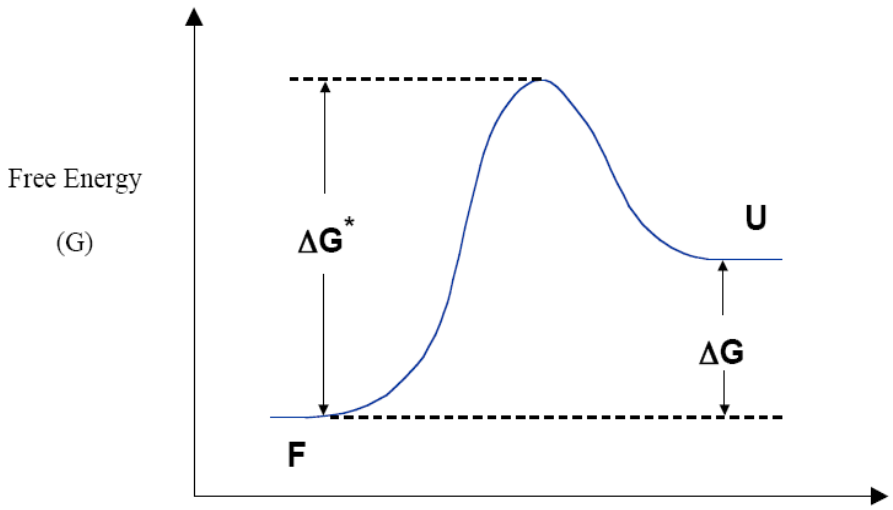


Figure 3.1. Schematic of the free energy between folded (F) and unfolded (U) states of a protein (Van Wyk, 2001).

The measurement of kinetic stability of proteins usually involves assaying for enzyme activity (Polizzi *et al*., 2007) and is most commonly expressed as the overall observed deactivation constant (k_d). $CynD_{pum}$ enzymes have been shown to exhibit first order thermal inactivation kinetics (Wang *et al*., 2011), which fit to equation (1), thus the residual enzyme activity after incubation at a constant temperature can be used as an indicator of the thermostability.

ln (% residual activity) = $-k_d$ x t **(1)** (Van Wyk, 2001)

Using equation (1), the rate of thermal inactivation (k_d) can be derived from the slope of a ln (% residual activity) versus t plot where t is the time of incubation. Mutants displaying increased thermostability have greater activity after a corresponding time period than the wild type. The ratio of the two rate constants (k_d) can be used to compute the change in ΔG (ΔΔG) required for unfolding relative to the wild-type enzyme. Using the Arrhenius equation (equation 2) we can obtain the relationship between the rate of thermal inactivation and the activation energy to yield equation 3 for the wild type and equation 4 for the mutants.

$$k_d = Ae^{-(\mathrm{Ea}/RT)} \quad \textbf{(2)}$$

$$k_{wt} = Ae^{-(\Delta G^*_{wt}/RT)} \quad \textbf{(3)} \qquad k_{mut} = Ae^{-(\Delta G^*_{mut}/RT)} \quad \textbf{(4)}$$ (Van Wyk, 2001)

Since ΔG* is the activation energy of thermal unfolding, the differences in ΔG* between the wild type and the mutants (ΔΔG) can be obtained by subtracting equation 4 from equation (3) to give equation (5). The ΔΔG of the different mutants were calculated using equation 5 and the values were used to compare thermostability of the mutants.

$$\Delta\Delta G^* = \Delta G^*_{wt} - \Delta G^*_{mut} = \mathrm{RT}\ln(k_{wt}/k_{mut}) \quad \textbf{(5)}$$ (Van Wyk, 2001)

Differential scanning fluorimetry

To measure the unfolding temperatures of the $CynD_{pum}$ enzymes and to confirm the thermostability results obtained using the picric acid assay, we carried out differential scanning fluorimetry (DSF). DSF is a fluorescence-based thermostability assay developed by Pantoliano *et al.* (2001). It is a high throughput screen method used for buffer optimisation

and ligand-induced stabilisation which was initially developed for drug discovery to allow rapid identification of ligands of target protein from compound libraries (Pantoliano *et al.,* 2001). The principle behind DSF is that folded and unfolded proteins can be distinguished through exposure to a hydrophobic fluoro probe. The fluoro probe is quenched in aqueous solution but will have an affinity to bind to the exposed hydrophobic interior of an unfolding protein resulting in a sharp decrease in quenching which produces a readily detected fluorescence emission that can be studied as a function of temperature (Errisson *et al.,* 2006). In DSF, the temperature of the protein-probe solution is gradually increased as a function of time leading to the unfolding of the protein. As the protein unfolds, the fluoro probe binds to its exposed hydrophobic region resulting in a significant increase in fluorescence emission (Fig. 3.2). As the temperature continues to increase more hydrophobic regions are exposed and an exponential increase of the fluorescence is observed until a maxima is reached which is followed by a steady decrease (Fig. 3.2). The steady decrease in fluorescence has been postulated to be due to precipitation and aggregation of the protein. Thermally induced unfolding is an irreversible unfolding process which follows a typical two-state model with a sharp transition between the folded and unfolded states. T_m (Fig. 3.2) is defined as the midpoint of temperature of the protein – unfolding transition (Errisson *et al.,* 2006). The simplest way to calculate T_m values is to determine the maximum of the first derivative of the fluorescence versus temperature plot. T_m values are used to calculate changes in the thermostability of the protein. The most commonly used fluoro probe is SYPRO Orange dye, due to its high signal to noise ratio. Another property that makes SYPRO Orange a favourable dye is its relatively high wavelength of excitation (465 nm excitation, 500 nm emission) which decreases the like hood that any small molecule would interfere with the optical properties of the dye which usually result in the quenching of the fluorescence intensity (Niesen *et al.,* 2009). DSF has been successfully applied in buffer screening (Phillips *et al.,* 2011), kinetic studies (Matulis *et al.,* 2005), ligand screening (Caver *et al.,* 2005) and testing of the stability of different functional protein variants (Lavinder *et al.,* 2009), a method which we adopted to identify $CynD_{pum}$ mutants with enhanced stability. DSF is an excellent method to screen for conditions or mutations that stabilise proteins, owing to the very small quantities of protein needed, reproducible results, quick simultaneous screening and no required prior knowledge proteins (Matulis *et al.,* 2005).

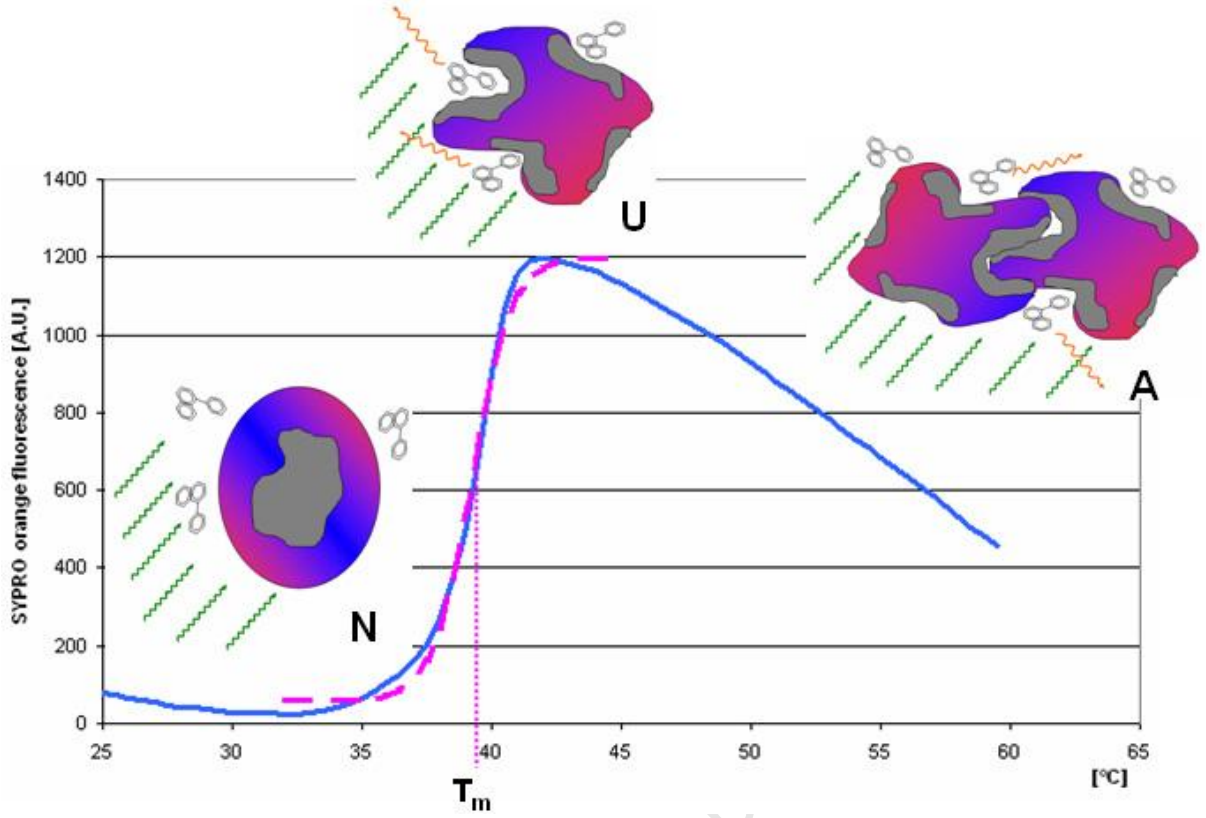


Figure 3.2. Typical recording of fluorescence intensity versus temperature for the unfolding of a protein in the presence of SYPRO orange. The dye is symbolized as a three-ring aromatic molecule. In the presence of a globular protein (spherical shape at the baseline of the curve), a basic fluorescence intensity is excited by light of 492 nm (depicted schematically by green curved arrows). Through unfolding of the protein, hydrophobic patches (in grey) become exposed, and strong fluorescent light of 610 nm (depicted by orange curved arrows) is emitted by the dye molecules bound to them. Following the peak in the intensity, a gradual decrease is observed, which is mainly explained by protein being removed from solution owing to precipitation and aggregation. The midpoint of temperature of the protein –unfolding transition T_m is used as a measure of stability (Errisson *et al.,* 2006).

3.3 Materials and Methods

pH activity profiles

The activity of purified enzymes was tested over the pH range 5.0–12.0 using a picric acid assay as follows. Citrate/Na_2HPO_4 buffer was used for pH 5.0 and 6.0, phosphate buffer for pH 7.0 and 8.0, glycine/NaOH for pH 9.0, sodium carbonate for pH 10.0, NaH_2PO_4 /NaOH for pH 11.0 and KCl/NaOH for pH 12.0. All buffers were used at a concentration of 50 mM and contained 150 mM NaCl. The mutants were assayed in parallel using the wild type enzyme as an internal control, and reactions were carried out in triplicate. Protein concentration was determined in triplicate using the Bradford method (described in section 2.3) and

protein samples were diluted to 0.057 mg/ml. Of this sample, 10 µl was mixed with 70 µl of buffer and pre-incubated for 1 h at 37°C, prior to addition of 20 µl of 25 mM KCN (in appropriate buffer). The reaction was allowed to continue for 20 min at 37°C in a water bath, and then stopped by addition of 80 µl of alkaline picric acid (50% picric acid, 50% carbonate). Colour was developed by heating samples at 100°C for 5 min in a heating block, then removing them to room temperature. Room temperature distilled water (1 ml) was added to stop the colour reaction, and the absorbance of this solution was measured at 520 nm on a Multiskan spectrophotometer (Titertek, Helsinki). These readings were converted to KCN concentration using standard curves determined in triplicate, separately at each pH condition. One unit of activity is defined as the degradation of 1 µmol of cyanide per minute per microgram of protein.

Thermostability assay

Thermostability assays of the $CynD_{pum}$ were carried out in order to test if the observed pH stability results were due to improved stability (slowed unfolding at unfavourable pH) on the protein rather than changes in the catalytic properties which may have been caused by the mutations. From a previous study in which the stability of $CynD_{pum}$ was tested by prolonged incubation at 23°C, 37°C, 42°C and 55°C, we identified 42°C as an appropriate condition to assess enzyme stability (Jandhyala *et al.,* 2005). To determine thermostability, separate tubes containing 10 µl of purified enzyme (0.057 mg/ml) in 70 µl of 50mM Tris-HCl pH 8 buffer were made up and stored at -4°C. These tubes were placed on a heating block at different times for incubation at 42°C. The tubes were placed on the heating block in such a way that the time difference from the start of the incubation to the end time of the incubation of all the tubes is equivalent to a desired time point, thus all the tubes reached the incubation endpoint at the same time. The samples were incubated for times ranging from 0 to 72 h. At the end of the incubation, activities of all the tubes were cooled to room temperature and activity was measured simultaneously using the picric acid assay as described above but using 50mM Tris-HCl pH 8 (the $CynD_{pum}$ pH optimum). The activities were measured simultaneously to prevent any errors that may be caused by minor differences in the picric acid assay conditions which were observed when using separate assays. From the results, the thermostability of each mutant was calculated as described above.

Differential Scanning Fluorimetry

Solutions containing 10 µl of SYPRO Orange (Sigma-Aldrich, USA) dilution and 30 µl of protein both in the same buffer were made up in transparent 0.2 ml PCR tubes. The tubes were placed in a Real Time PCR machine (info). The tubes were heated in the Real Time PCR from 30°C to 95 °C in increments of 0.5°C, while the fluorescence changes were monitored and recorded for each tube with wavelengths of excitation and emission at 490 nm and 575 nm respectively. To obtain the temperature midpoint for the protein unfolding transition (T_m), the maximum value of the first derivative of the fluorescence versus temperature curve was calculated as previously discussed. Pre optimization experiments were initially carried out to optimise the concentration of $CynD_{pum}$ and SYPRO Orange dye. The Protein concentration conditions used for optimisation were 0, 0.01, 0.1and 0.5mg/ml in the presence of 1X, 5X and 50X SYPRO Orange. The optimisation was carried out using wild type enzyme in 50 mM Tris-HCl at pH 8. Two separate DSF experiment were then carried out using the different $CynD_{pum}$ strains at the optimised concentrations, with one set run in the presence of 50 mM Citrate/Na2HPO4 buffer at pH 5.4 and the other in 50mM Tris-HCl at pH 8.

3.4 Results

pH activity profiles

Activity of the purified $CynD_{pum}$ mutants was measured at different pH values and the results were compared with that of the wild type (Fig. 3.3). At pH 5 to pH 8 all the $CynD_{pum}$ mutants displayed similar or less activity relative to the wild type. At pH 8 the mutants either showed elevated activity (H305K, H308K, H323K, H305K+H308K, H305K+H323K, H308K+H323K and H308K+H305K+H323K and Q86R) or similar activity (Q86R+E96G+D254E+E327G) with respect to the wild type. At pH 9 all the mutant displayed similar activities relative to the wild type, except the Q86R+E96G +D254E+E327G mutant and the $CynD_{pum\text{-}stut}$ mutant which demonstrated pronounced increase in activity of approximately five and four fold higher than the wild type, respectively. No activity was observed at pH 10–12 even when using higher enzyme concentrations (not shown).

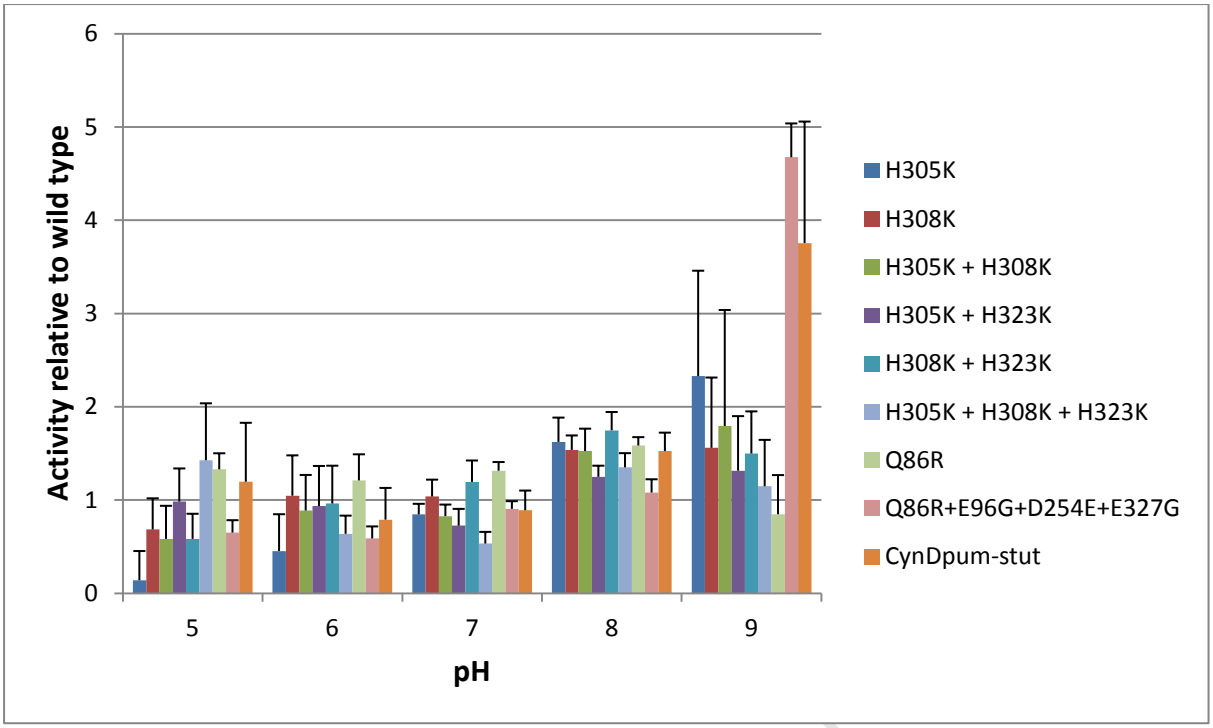


Figure 3.3. pH profiles of $CynD_{pum}$ mutants relative to the wild type. Enzymatic activities of the $CynD_{pum}$ mutants were compared to the wild type enzyme at different pH values, using the picric acid assay. The values are expressed as an average of triplicate experiments, with error bars representing standard deviations.

Thermostability

We investigated the thermostability of the $CynD_{pum}$ wild type and mutants by incubating them at 42°C for 0 to 72hrs at pH 8. The natural log of the residual activity was calculated and plotted versus time (Fig. 3.4). The graphs showed that there was more retention of activity by the mutants (represented by the slopes of the graphs), even after 72hrs, than the wild type which lost 80 % of activity after 4hrs (Fig. 3.4). The slopes obtained from the Figure 3.4 graphs were then used to calculate ΔΔG (Fig. 3.5). All the $CynD_{pum}$ mutants showed a higher ΔΔG relative to the wild type. H305K, H308K, H323K, H305K+H323K, Q86R and Q86R+E96G +D254E+E327G showed the highest ΔΔG relative to the wild type. $CynD_{pum\text{-}stut}$ displayed one of the lowest ΔΔG relative to the wild type, although it previously displayed one of the highest alkaline pH stability in the pH profile. There was no significant difference between the $CynD_{pum}$ mutants according to the error bars (Fig. 3.5).

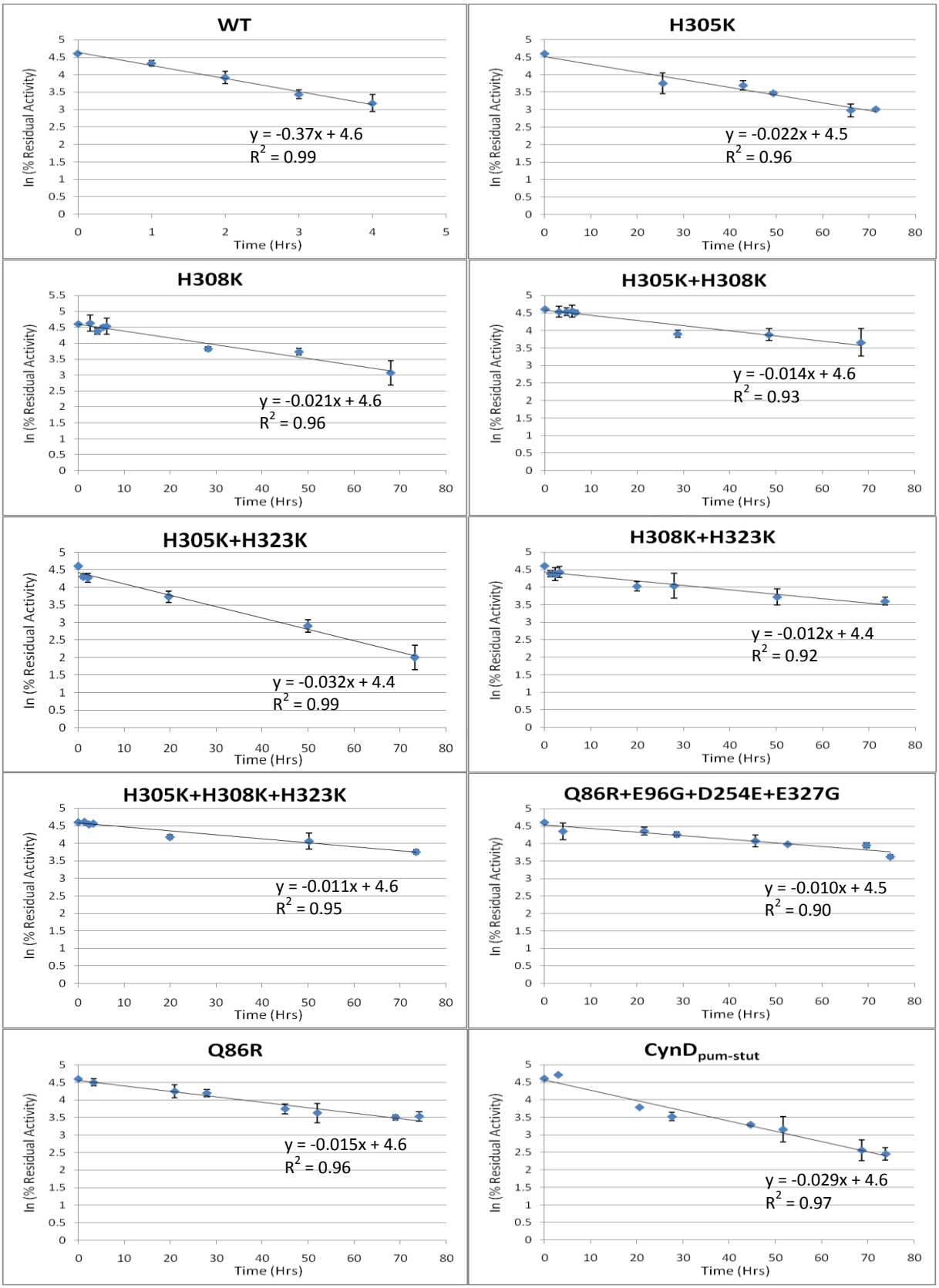


Figure 3.4. Thermostability of the $CynD_{pum}$ wild type and mutants at pH 8. The natural logarithm of the percentage residual activity after the stated time periods was calculated and plotted. The values of the percentage residual activity were calculated relative to the initial activity and are expressed as the average of duplicate experiments, with error bars representing standard deviations.

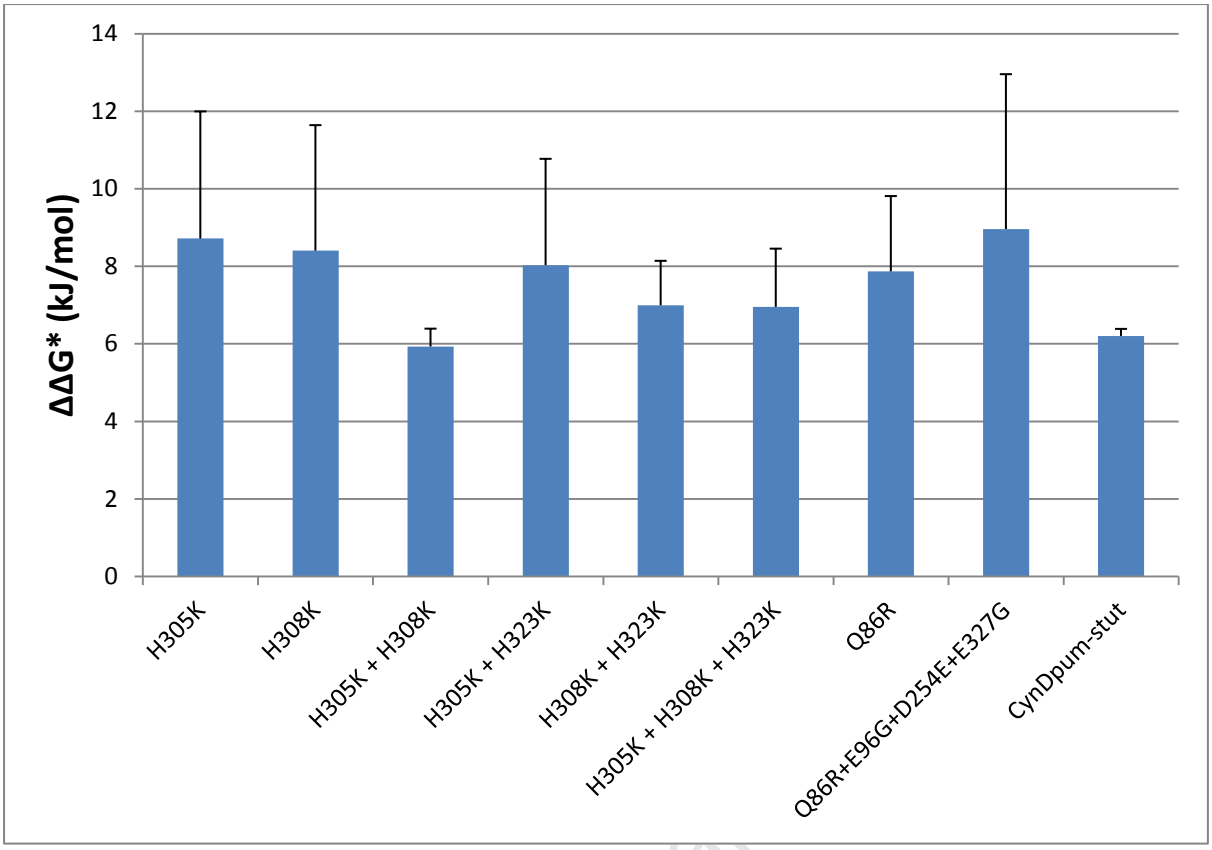


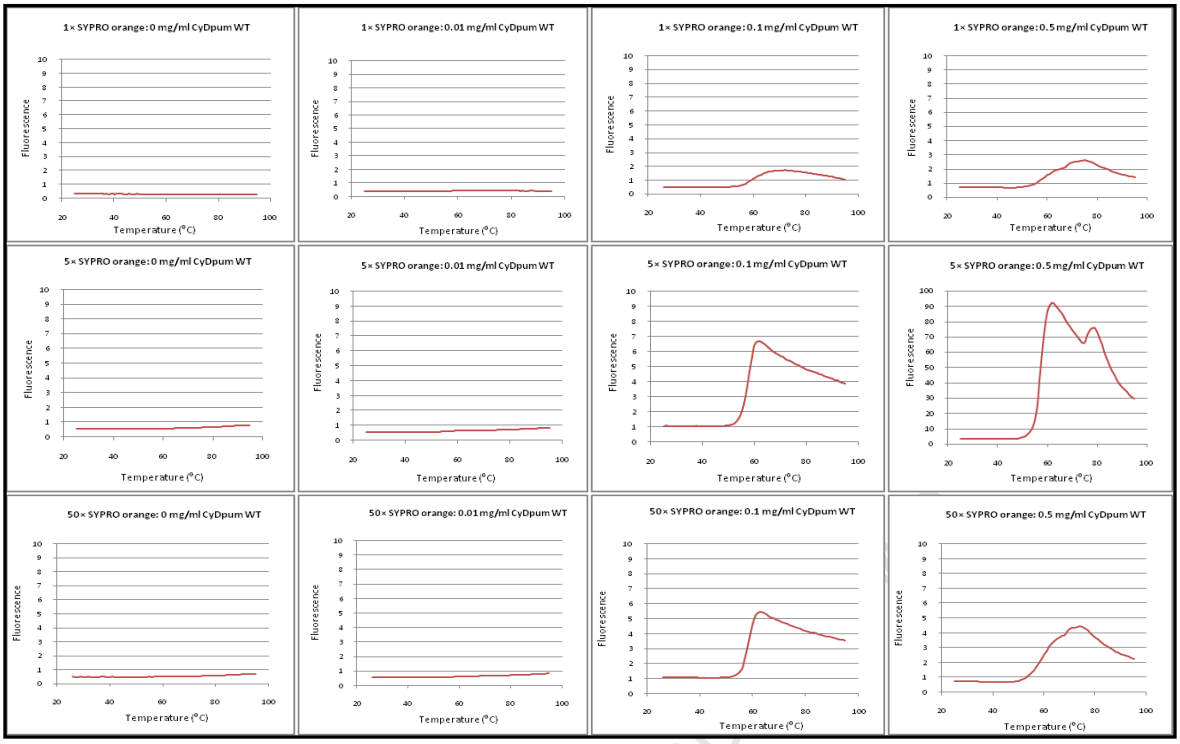
Figure 3.5. Calculated ΔΔG* (kJ/mol) for the $CynD_{pum}$ mutants relative to wild type. Error bars represent standard deviations calculated from the error of the slopes of the graphs in Fig. 3.4.

Differential Scanning Fluorimetry

Optimum DSF conditions were obtained by carrying out the DSF experiment at different wild type $CynD_{pum}$ and dye concentrations. From the optimization experiment we identified 0.1mg/ml of enzyme in the presence of 5X SYPRO Orange as the optimum condition for the experiment, as these conditions gave us the sharpest transition and high quantum yield in the presence of the least protein concentration (Fig. 3.6). The midpoint of temperature of the protein –unfolding transition temperatures (T_m) of $CynD_{pum}$ wild type and mutants were measured in both pH 5.4 (Fig. 3.7) and pH8 buffers (Fig. 3.8). The T_m of the enzymes were calculated (Fig. 3.9a) and they displayed a general decrease as the pH changed from pH 5.4 to pH 8. The T_m values of the mutants were compared with that of the wild type (Fig. 3.9b). At pH 8 all the mutants demonstrated an increase in T_m relative to the wild type. $CynD_{pum\text{-}stut}$ demonstrated the highest T_m relative to the wild type at pH8. H305K+H308K+H323K, Q86R, Q86R +E96G +E327G also showed a significantly higher T_m relative to the wild type at pH8. At pH 5.4 most of the mutants demonstrated similar T_m to that of the wild type with the

exception of $CynD_{pum\text{-}stut}$ which displayed the highest T_m at this pH. H305K+H308K+H323K, Q86R and Q86R+E96G+D254E+E327G showed a slightly higher T_m compared to the wild type at pH 5.4. Interestingly H308K+H323K demonstrated a lower T_m than the wild type, at pH 5.4, shown by the negative shift which may be due to partial destabilisation or aggregation of the protein.

(a)



(b)

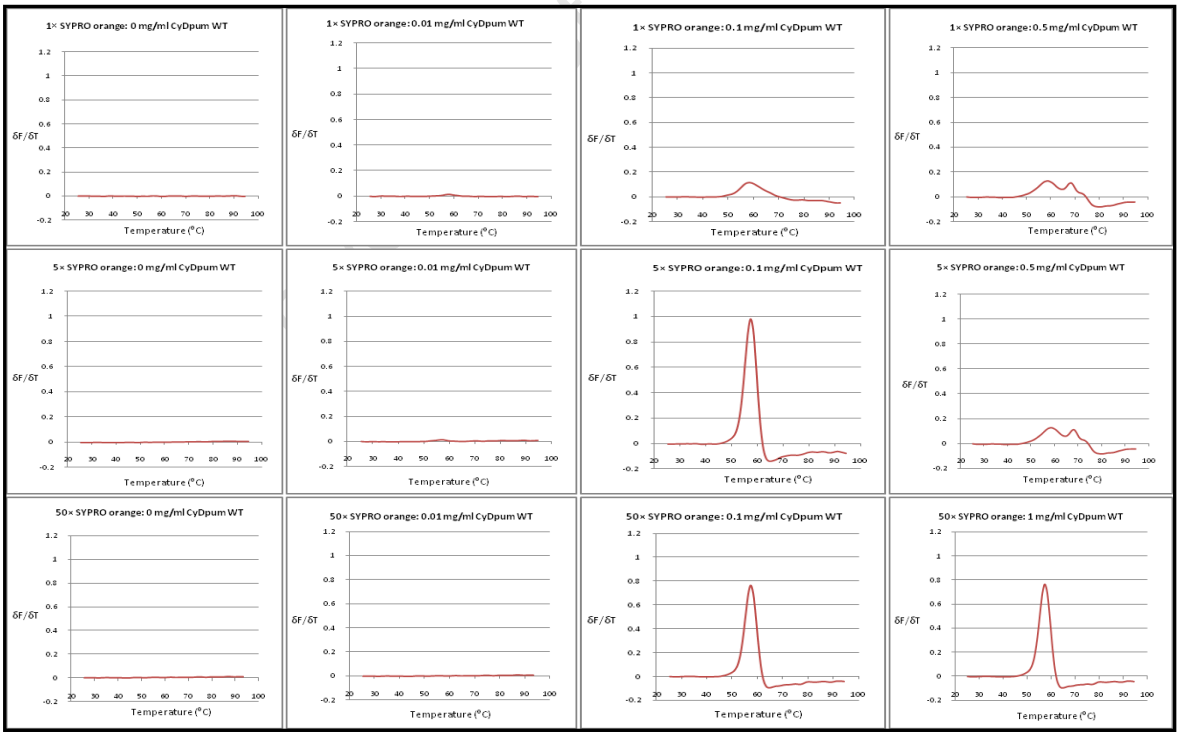


Figure 3.6. DSF optimisation results. **(a)** The thermal denaturation of various concentrations of $CynD_{pum}$ wild type in the presence of different SYPRO Orange concentrations. **(b)** The first derivative of graph (a) were maximum value determines T_m.

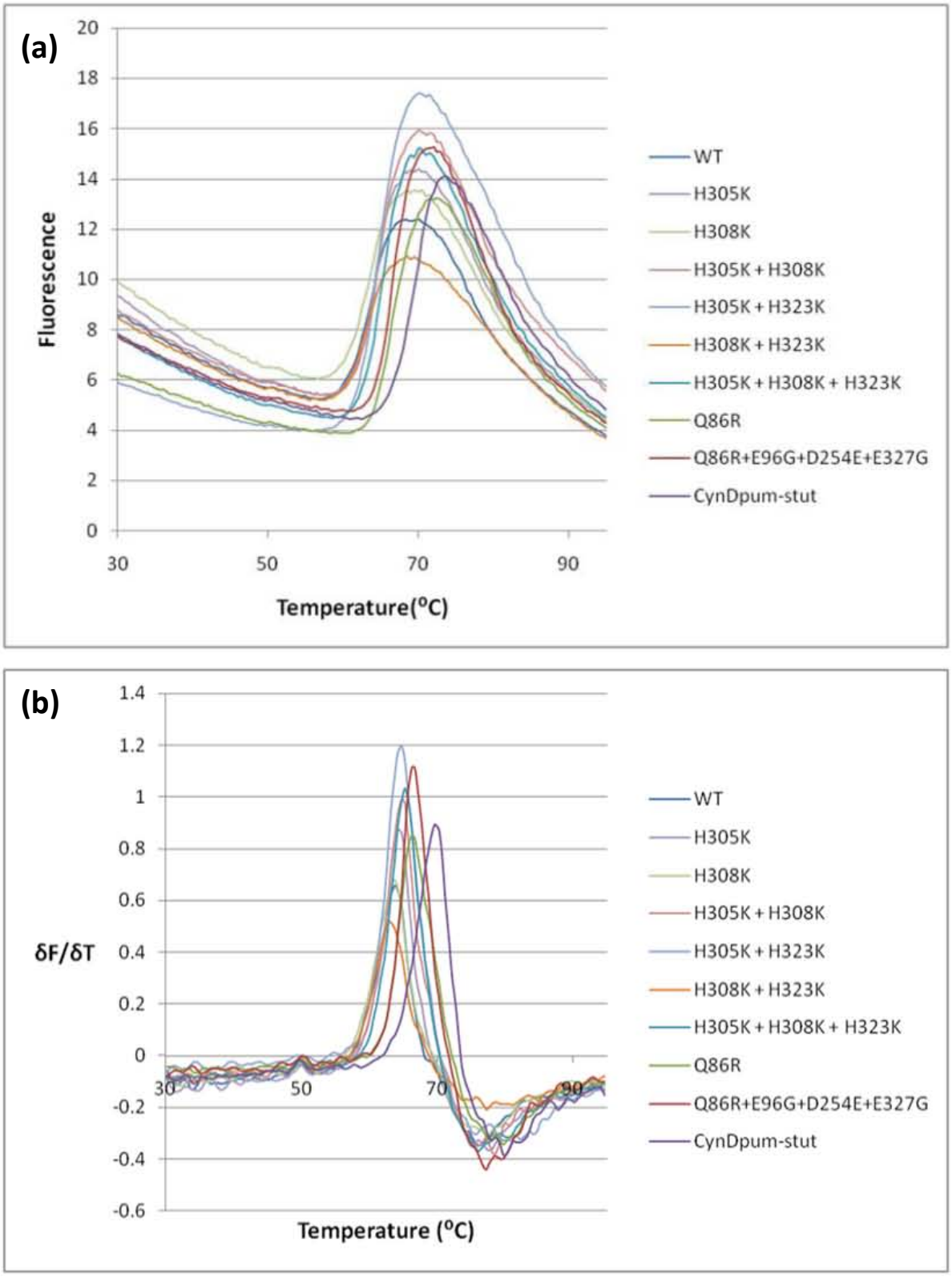


Figure 3.7. (a) DSF results showing the thermal denaturation of the $CynD_{pum}$ enzymes at pH 5.4. **(b)** The first derivative of graph (a) were maximum value determines T_m. The individual T_m values were extrapolated and summarised in fig. 3.9a.

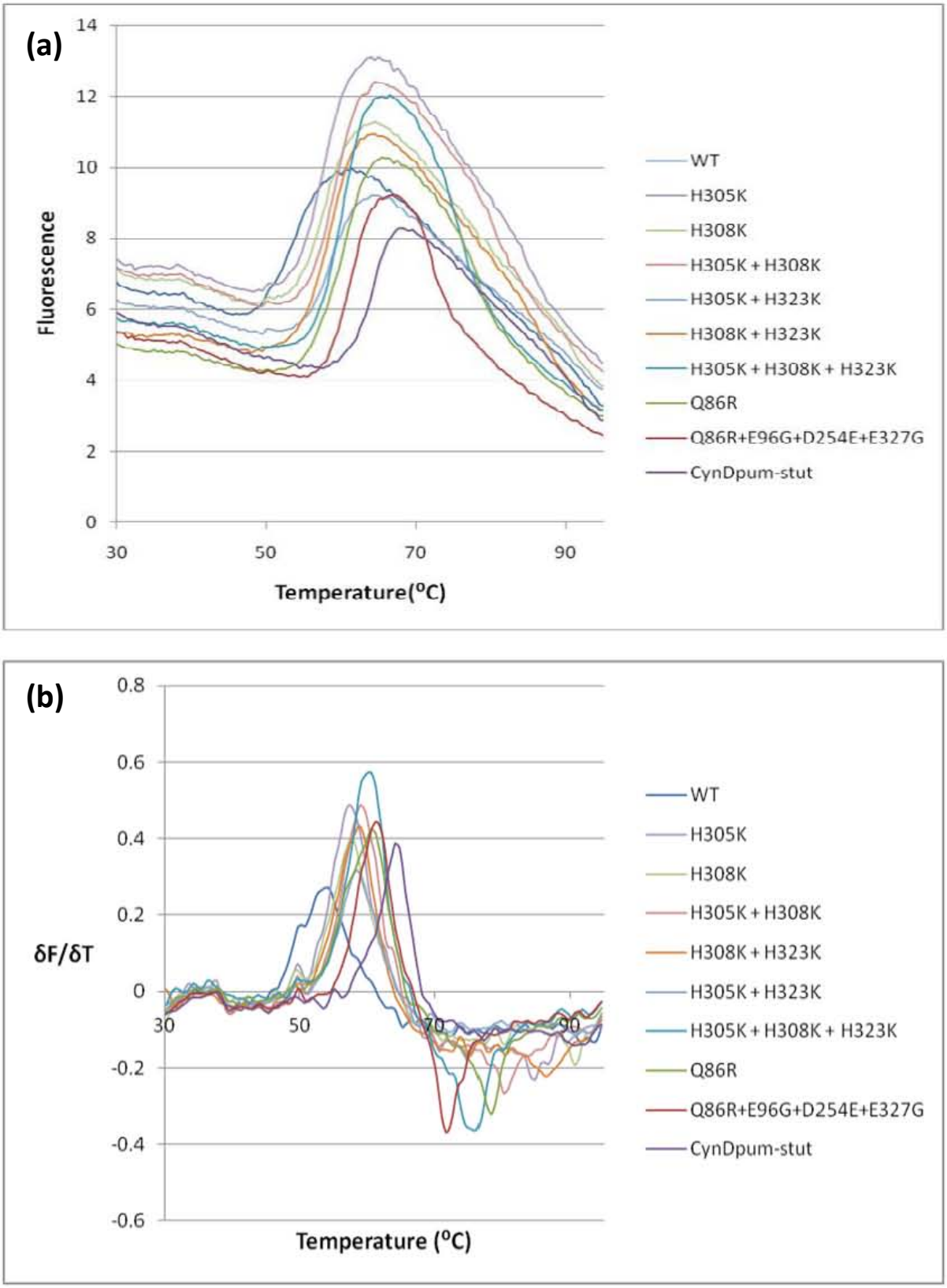


Figure 3.8. (a) DSF results showing the thermal denaturation of the $CynD_{pum}$ enzymes at pH 8. **(b)** The first derivative of graph (a) were maximum value determines T_m. The individual T_m values were extrapolated and summarised in fig. 3.9a.

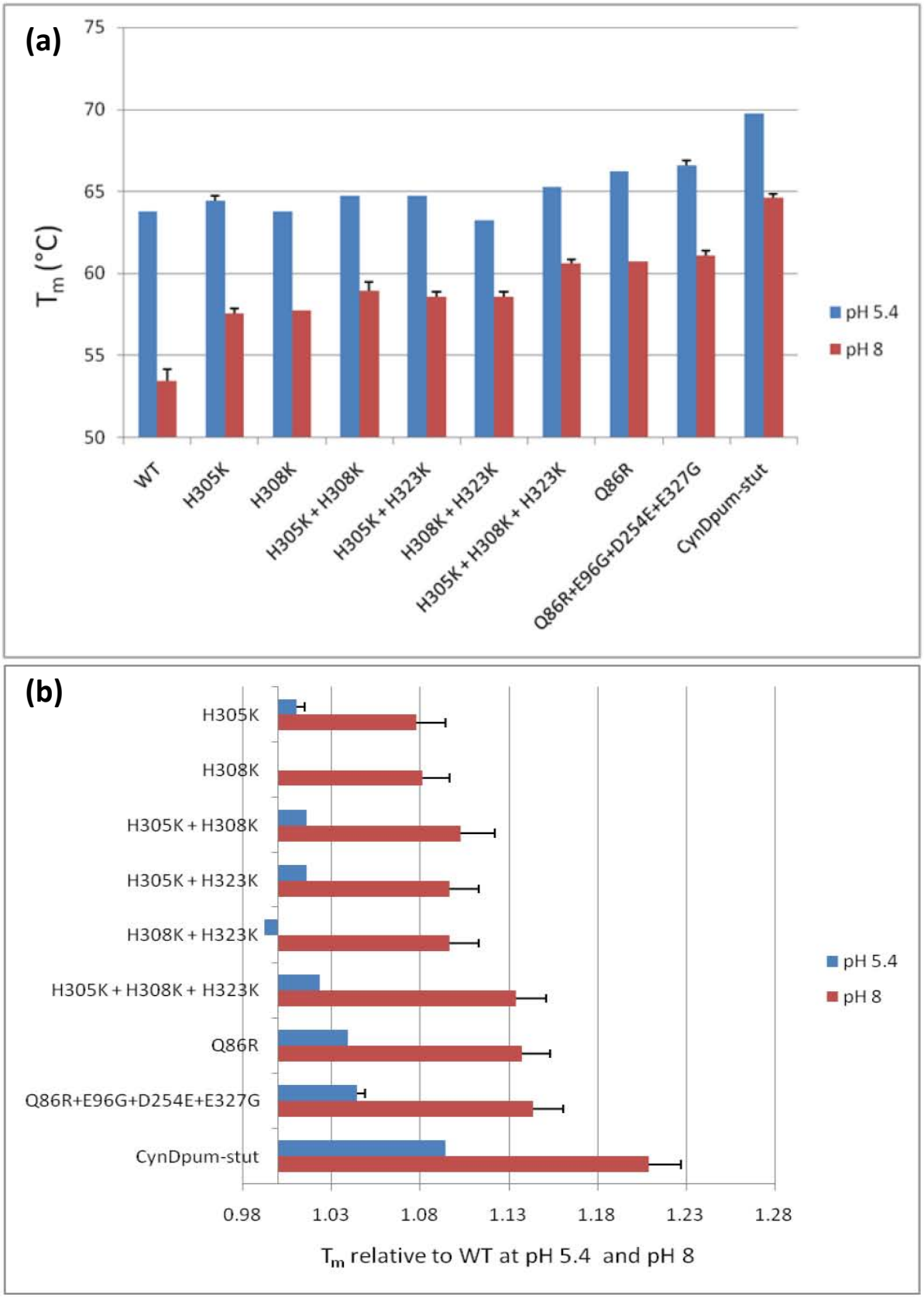


Figure 3.9. (a) Midpoint temperatures of the protein-unfolding transition (T_m) for the $CynD_{pum}$ wild type and mutants at pH 5.4 and pH 8. **(b)** The T_m of the mutants relative to the wild type at pH 5.4 and pH 8.

3.5 Discussion

The picric acid activity assay provides a powerful tool to measure activity of the $CynD_{pum}$. Coupled with various conditions such as different pH or temperature, this assay can be successfully used to identify pH and thermostable mutant strains of $CynD_{pum}$. It was previously proposed that terminal subunits in the helices may be inactive and that the structural transformation to long spirals results in fewer terminal subunits resulting in more active subunits in the population (Jandhyala *et al.,* 2005; Sewell *et al.,* 2005). Based on this assumption, one would expect the long fibres formed by H308K+H305K+H323K, Q86R and Q86R+E96G +D254E+E327G at pH 5.4 to be more active than the wild type since the wild type exits as shorter oligomers at the various pH as observed in the previous chapter (Fig. 2.12). This was not the case as the pH activity profiles showed no noticeable differences in the mutant's activity relative to the wild type at pH 5 to pH 7. The results indicate that the terminal subunits may not be inactive and that the difference in the population of the terminal subunits (in equally concentrated solution) does not affect activity. Another explanation could be that the terminal subunits a too few in comparison to the active subunits that no noticeable differences in activity would be observed. The degree of the increase in activity at pH 8 was lower than expected for H308K+H305K+H323K, Q86R and Q86R+E96G +D254E+E327G since they still displayed longer fibres than the wild type at this pH, shown in the previous chapter (Fig. 2.12), and this may also be due to the existence of active terminal subunits in wild type's short spiral population. At pH 9, Q86R+E96G+D254E+E327G and the $CynD_{pum\text{-}stut}$ mutants displayed higher activity relative to the wild type. The high activity of the Q86R+E96G+D254E+E327G mutant was expected as it showed signs of structural stability at pH 9 when viewed by electron microscopy (Fig. 2.12). Higher activity relative to the wild type was not expected for $CynD_{pum\text{-}stut}$ mutant since this mutant did not display any change in oligomerisation at pH 5.4, pH 8, and pH 9 (Fig. 2.12). Thus at pH 9 two scenarios were observed, in one case the higher activity correlated with the fibre stability but in the other case it did not indicating evidence of two different effects. It was evident that the $CynD_{pum\text{-}stut}$'s 44 residues mutations not only gave this hybrid the oligomeric characteristics of $CynD_{stut}$ but also conferred alkaline pH stability. It is difficult to provide a mechanism for the alkaline pH stability observed for the $CynD_{pum\text{-}stut}$ due to the multiple amino acid mutations present in the mutant which make it challenging to predict the key amino acid responsible for the observed effect. These results nevertheless

demonstrate the importance of the C-terminal tail in stabilizing the $CynD_{pum}$ at alkaline pH. Q86R+E96G +D254E+E327G and $CynD_{pum-stut}$ displayed the most pronounced alkaline pH stability with the presence of high activity at pH 9, relative to the wild type. This acquired property makes Q86R+E96G +D254E+E327G and $CynD_{pum-stut}$ good candidates for use in bioremediation of alkaline cyanide waste.

The improvement in activity at the different pH observed above could either be due to changes in the stability of the enzyme or changes in the catalytic properties of the enzyme caused by the mutations. Thermostability assays were carried out to test which of the two mechanisms is responsible for the pH stability. Since there was a general loss in enzyme activity after samples were cooled down from 42°C to room temperature prior the picric acid assay, we can conclude that the thermal unfolding of the $CynD_{pum}$ enzymes is an irreversible process. All the $CynD_{pum}$ mutants that were tested appeared to be more thermostable than the wild type indicating that all the mutations confer thermostability. To further assess the thermostability the ΔΔG of the mutant was calculated relative to the wild type and the results confirmed the enhancement of thermostability of the mutants. The thermostability of the $CynD_{pum}$ mutants showed a different trend from that observed for the previous pH stability results. Q86R+E96G +D254E+E327G displayed the highest thermostability and this agreed with the pH profile results which demonstrated that the mutants also have the highest alkaline pH stability. This indicates that the alkaline pH stability may be due to improved stability rather than a change in catalytic properties of this $CynD_{pum}$ mutant. On the other hand $CynD_{pum-stut}$ displayed the second lowest thermostability compared to the other mutants even though it displayed the second highest alkaline pH stability in the pH profiles. This illustrates that the conferred alkaline pH stability of the $CynD_{pum}$ may not be only due to the improved stability of the enzyme. Changes in the catalytic properties may partially contribute to the alkaline pH stability or a completely different, yet to be discovered mechanism may be involved.

Using DSF, the midpoint of temperature of the protein-unfolding transition (T_m) of the $CynD_{pum}$ wild type and mutants was calculated at pH 5.4 and pH8. The measured T_m showed a different stability trend compared to the previous thermostability results. This was not completely unexpected as the previous thermostability assay measured activity at a fixed temperature while DSF measures the temperature required to unfold a protein regardless of

the activity. With the exception of $CynD_{pum\text{-}stut}$, the T_m of the other $CynD_{pum}$ (Fig. 3.9a) showed a similar trend to the structural stability observed on the micrographs in chapter II (Fig. 2.11 and 2.12), were a high T_m corresponded to a high level of oligomerisation and vice versa. This provides evidence of a general correlation between improved oligomerisation and improved structural stability as previously described by Wang *et al.* (2011). The observed change in stabilization of the protein may be due to the different intrinsic properties of the pH buffers, although this is unlikely as this would not account for the individual mutant's difference in the change in T_m (ΔT_m) across the different pH. $CynD_{pum\text{-}stut}$ demonstrated the highest stability across pH 5.4 and pH 8, followed by Q86R+E96G+D254E+E327G, Q86R then H305K+H308K+H323K in order of decreasing stability. When comparing the DSF results to the pH profiles, it is evident that $CynD_{pum\text{-}stut}$ and Q86R+E96G+D254E+E327G displayed both the highest alkaline pH stability and highest T_m illustrating how a $CynD_{pum}$ enzyme has to be highly thermodynamically stable to be alkaline pH stable. $CynD_{pum\text{-}stut}$'s low $\Delta\Delta G$ in the thermostability assay demonstrates how kinetic stability may not be a determinant of alkaline pH stability.

DSF provides a very reliable and high throughput method to measure the stability of the enzymes. The experiments were done in triplicates and the standard error was minimal which demonstrates the accuracy and reproducibility of the DSF method (Fig. 3.9). The minimal discrepancy of the results may have been due to differences in excitation intensities across the different PCR tubes, or due to optical edge effects from the detector. DSF is a more reliable method of measuring thermostability of enzymes compared to the picric acid thermostability assay but its drawback is that it does not factor in activity when measuring thermostability.

There is no prior evidence in general that pH stability is related to thermostability. However in the case of $CynD_{pum}$ activity is lost at pHs in which lysines become deprotonated and thus lose their charge. These lysines may be responsible for the interactions that stabilise the enzyme against thermal denaturation.

3.6 Conclusion

We have successfully identified $CynD_{pum}$ mutations with alkaline pH stability namely $CynD_{pum\text{-}stut}$ and Q86R+E96G+D254E+E327G. These mutants were also shown to be very thermostable based on DSF were the $CynD_{pum\text{-}stut}$ and Q86R+E96G+D254E+E327G mutants displayed T_m shifts of 11.2° and 7.2° respectively compared to the T_m of the wild type at pH 8. The thermostability of these mutants was confirmed by the thermostability assay which measured the ΔΔG of the enzymes, but in the case of $CynD_{pum\text{-}stut}$ the thermostability was low indicating that $CynD_{pum\text{-}stut}$ loses its activity before it completely unfolds, thus thermodynamic stability and not kinetic stability is responsible for the observed increase in stability at alkaline pH.

CHAPTER IV

THREE –DIMENSIONAL RECONSTRUCTION

4.1 Introduction

Electron microscopes are powerful and versatile tool that allow visualisation of microscopic object such as biological macromolecules. The transmission electron microscope in particular, uses high-energy electrons to form an image of very thin objects. When a beam of electrons is passed through matter, the electrons are strongly scattered and the scattered electron's paths can be bent by magnetic lenses so that an image is formed on a fluorescent screen, photographic film, or scintillator of a charge-coupled device (CCD camera). Electrons have a much shorter wavelength compared to light and hence allow transmission electron microscope to produce higher resolution images. 2D electron micrographs can be used to computationally produce 3D reconstructions of objects. 3D reconstructions have improved over the years due to advances of computation systems and refinement of algorithms thus allowing one to generate 3D reconstructions of high resolution at reduced time. In this chapter we describe how we used electron microscopy 3D reconstruction techniques to visualise structural differences in the quaternary structure between the $CynD_{pum}$ wild type and the long fibre forming mutants at various pH.

4.2 Review of helical reconstruction

A number of protein complexes found in prokaryotic, eukaryotic and archaeal cells exist as helical polymers. These include assemblies such as intracellular actin filaments, myosin filaments, microtubules, extracellular collagen, keratin and bacterial flagella. Despite having significant roles in cells, only a minority of these assemblies have solved atomic or near-atomic structures. This is primarily due to limitation of conventional methods of atomic structure determination such as X- ray crystallography. X-ray crystallography requires a crystal of the molecule to be diffracted, and the only filaments that have been demonstrated to pack with crystalline symmetry are those that have exactly two, three, four or six subunits per turn (Egelman, 2007). Most helical polymers fail to crystallize in their native polymeric state often due to disruption of the crystal lattice by irregularly shaped

surfaces of the polymer. Electron microscopy techniques have proven successful in elucidating the structures of helical polymers, primarily 3D reconstruction from electron microscopic images. Some of the first application of 3D reconstruction techniques was to a helical polymer (DeRosier and Klug, 1968) and was later adopted to single particle specimens. Over the years there have been advances and refinements to the method and it is now even possible to get images of very high resolution. Problems such as flexibility and disorder of the specimen's filaments have presented complications in 3D reconstruction (Egelman, 2007). Real space approaches have provided solutions to these problems such as the algorithm known as Iterative Helical Real Space Reconstruction (IHRSR). Unlike single particle reconstruction, which is generally performed using microscopy datasets that consists of randomly oriented electron microscopy images to obtain angular views of the molecule, IHRSR further takes advantage of the helical symmetry of the helical filament which it imposes on an asymmetric reconstruction model (Egelman, 2000). Helical symmetry is symmetry of objects related by a coupled rotation about an axis with a translation parallel to the axis, and is defined by two parameters, the rise between adjacent subunits (Δz) and the rotation per subunits ($\Delta\phi$).

Iterative Helical Real Space Reconstruction (IHRSR)

IHRSR is a robust algorithm for the reconstruction of helical filaments using single-particle methods and was developed by Egelman (2000). This technique overcomes some of the difficulties previously encountered in helical image analysis such as indexing, unbending and the need to find long helically symmetric filament segments. In this method, helical filaments are boxed out from the electron micrograph to produce individual helical filaments which are further divided into segments of appropriate length. These smaller segments are treated as individual particles in subsequent steps. The box size of these segments is calculated and optimized empirically for different structures by using different factors such as the mass per unit length, the signal-to –noise ratio and flexibility. To obtain an optimal box size, the size should not be too long so as to encounter problems with filament bending and internal disorder, and should not be too short that cross correlation would be ambiguous. An initial model then has to be inserted into the algorithm (Fig. 4.1). This model could be generated computationally using the microscopy datasets or sourced from 3D models of homologous specimens although this might introduce model bias and

hence a commonly used featureless cylinder is preferable. Projections of the starting model are then generated by azimuthally rotating the model at appropriate angle increments about the filament axis, generating 2D reference projections for each rotation.

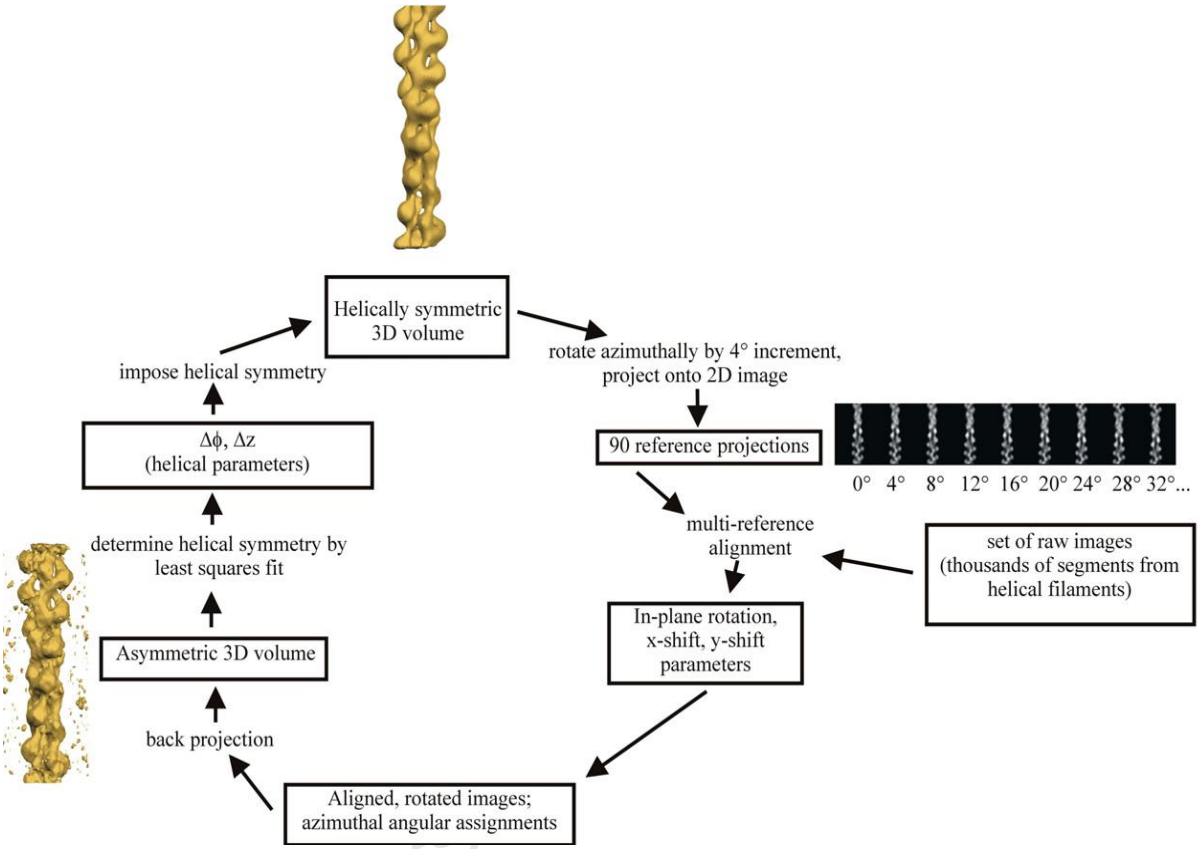


Figure 4.1. Schematic diagram of the cycle used the IHRSR algorithm (Egelman, 2000)

The 2D references are then used for multi-reference alignment against all the segments of the helical filament (Fig. 4.1). Multi reference alignment is achieved by computationally calculating the best cross-correlation of each segment image to each reference projection and assigning that segment image to a reference projection that gives the highest cross-correlation. The multi reference alignment step yields x-shifts, y-shifts and in-plane rotation parameters which are used to bring the aligned images into register with their assigned reference projection in order to generate an asymmetric 3D volume by a process called back projection (Fig. 4.1). A least-squares search algorithm (Egelman, 2000) then calculates the helical symmetry (Δz and Δϕ) of the asymmetric 3D volume. The least-square search algorithm calculates the mean-squared deviation in density between densities at symmetry-related positions in the asymmetric volume. An initial Δϕ guess is input into the

IHRSR algorithm and is obtained either from previous reconstruction or calculated using Fourier-Bessel methods (DeRosier and Klug, 1968). The calculated helical symmetry parameters are imposed to the asymmetric 3D volume to produce a symmetrized volume which is then used as a new reference for the next cycle (Fig. 4.1). The whole process is then iterated until a stable helical symmetry solution is obtained hence until the helical symmetry shows no further change and converges to a single solution (Fig. 4.1).

Resolution criteria

One of the difficulties in 3D reconstruction is that there are limited means of validating a reconstruction. The most common way to achieve this is to quantitatively measure the resolution of the reconstruction. The concept of resolution is defined as the minimal distance between two points in an image at which they can be still be distinguished from each other. Over the years a number of methods to calculate resolution of 3D reconstructions have been developed. The most popular Half-Set method is based on the comparison of two independent averages from two randomly drawn subsets of equal size, in the Fourier domain. In this method the Fourier transforms of the two subsets are compared and a measure of discrepancy is calculated as a function of spatial frequency. The discrepancy is commonly measured as either Differential Phase Residual (DPR, Frank *et al*., 1981) or Fourier Shell Correlation (FSC, Saxton and Baumeister, 1982; van Heel *et al*., 1983). DPR measures the phase difference between the two Fourier transforms weighted by the average Fourier amplitude while FSC (or Fourier ring correlation if in 2D) measures the correlation as a function of resolution shells between Fourier coefficients computed from the reconstructions, generated by the two datasets. Unlike DPR, FSC directly compares the two Fourier transforms. Over time the use of DPR has been abandoned in favour of the FSC generally because of DRF necessity to use a fixed threshold value to determine resolution (Orlova *et al*., 2007; van Heel, 1987) although in recent years authors have again started to use fixed-value resolution thresholds but in conjunction with FSC curves. The common resolution-defining threshold value is 0.5 (Beck-Mann *et al*., 1997; Bottcher *et al*., 1997). The traditional measure of resolution of single-particle reconstruction using FSC as a measure of discrepancy has raised issues over the years and perhaps the most considered is based on the fact that reconstructions are susceptible to model bias (Grigorieff, 2000; Shaikh *et al*., 2003).

Consequently, noise that is present can become correlated during the alignment process (Boekema *et al*., 1986) due to the fact that the same reference is repeatedly used for aligning two sets of images. Noise has been shown to lead to an artefactual resolution extention due to the strong size dependent correlation of Fourier coefficients in adjacent shells of a 3D Fourier transform thus small complexes deviate more from the true resolution than larger complexes being reconstructed (Yang *et al.,* 2003). It has been suggested that for the FSC method to be really valid, the two independent reconstructions have to be generated from different initial reference structures (Yang *et al.,* 2003). Previously the FSC method involved aligning two sets of images against a common reference resulting in two reconstructions which were not truly independent. It is proposed that when different initial structures are used for iterative cycles of reconstruction, it will lead to convergence of the structure but will diverge for the noise hence uncorrelated the noise. Yang *et al.* (2003) proposed a simple method of obtaining the independent reconstructions using IHRSR. By using different starting helical symmetries for the IHRSR process one can eliminate all correlations in the starting structures expect at the very lowest spatial frequencies.

Other methods of determine resolution include the Spectral Signal-to-Noise Ratio (SSNR) method which measures Signal-to-Noise ratio (SNR) as a function of spatial frequency using the whole data set. It was first introduced by Unser *et al.* (1987) and was initially applicable to 2D forms but was later adopted to 3D forms by Penczek (2002) and Unser *et al.* (2005). The SSNR method has been demonstrated to have better statistical performance than DRF and the FSC due to the evaluation of the whole data set (Liao*,* 2010), but with this advantage SSNR is still less commonly used than the FSC method.

4.3 Materials and Methods

Sample preparation and data collection

Negative stained protein preps on carbon coated grids were prepared as described in section 2.3. Electron microscopy was performed at low dose using the LEO 912 TEM operated at 120kV. Micrographs were recorded on a Proscan 2048 x 2048 CCD camera at 50 000 × magnification.The calibrated sampling on the CCD was 2.17 Å/pixel and defocus of

0.8–1.0 µm was used to manually capture data sets of micrograph. Electron micrographs of the fibres formed by the $CynD_{pum}$ wild type enzyme and $CynD_{pum}$ mutants that formed fibres at pH 5.4 and pH 8 (H308K+H305K+H323K and Q86R; Fig. 2.12) were collected. Although Q86R+E96G +D254E+E327G formed fibres, micrographs were not collected as it was concluded in chapter II that Q86R is the predominant mutation in this mutant.

Image pre-processing

The Helixboxer program of the EMAN suite (Ludtke *et al*., 1999) was used to pick and extract straight helical filaments from the micrographs (Table 4.1). Power spectra were calculated for each of the boxed helices with the removal of helices that failed to diffract. All the successful power spectra were then averaged together and the resultant power spectrum was used to evaluate the data. The boxed filaments were then further cut into several overlapping segments of 160 × 160 pixels with a shift of 10 pixels between adjacent segments, and all the segments were compiled into a stack (Table 4.1). The raw images were centred by applying x shifts only using SPIDER routines (Frank *et al*., 1996) and IHRSR. The new stack was further visually inspected using V2 of the EMAN suite (Ludtke *et al*., 1999) and bad segments were deleted from the stack.

Table 4.1. Image pre-processing summary

Strain	pH	# of micrographs	# of boxed helical filaments	# of segmented images in stack
Wild type	**pH 5.4**	105	401	8474
H305K+ H308K+ H323K	**pH 5.4**	100	357	9790
	pH 8	160	687	8914
Q86R	**pH 5.4**	100	516	15260
	pH 8	128	847	11770

Three-dimensional reconstruction

3D reconstruction was computed using the IHRSR++ software package (Parent *et al.,* 2010) which is a modified version of the IHRSR method of Egelman (2000). The package consist of a variety of individual programs but for our purpose the programs where run through one SPIDER script. IHRSR++ package includes added features such as imposing of dihedral symmetry, search for out of plane tilt and generation of random models. The random model

construction feature was used to generate model and either the resultant model or a featureless cylinder (130 Å in diameter) was used as starting volumes for IHRSR. Reference projections were constructed using 4° azimuthal angle increments to produce 90 2D projections of the starting model. The processed images were aligned to the 2D projections using the multi-reference alignment procedure. The helical parameters for the wild-type $CynD_{pum}$ ($\Delta\phi$ = -75.0°, Δz = 16.0 Å) were obtained from previous studies and are based on approximate indexing of the average power spectrum of an aligned stack and the helical parameters of the homologous nitrilase from *Rhodococcus rhodochrous* J1 (Thuku *et al*., 2007). These helical symmetry parameters were entered into the IHRSR algorithim and a 3D symmetric reconstruction was generated at the end of each IHRSR cycle. Each process was iterated for 20 cycles for each enzyme. Two separate IHRSR algorithms were run with different initial helical symmetries to clearly observe any convergence of the helical symmetry, thus validating the models. A power spectrum of the reconstructions was finally constructed to validate the 3D models. This was achieved by generating 2D projection of an extended fibre constructed from an un-symmetrized reconstruction volume and then calculating the power spectrum using IHRSR.

Resolution determination

The FSC method was used for resolution determination. As suggest by Yang *et al.* (2003), the image stack was split into two subsets and two separate IHRSR 3D reconstructions were performed using different starting helical symmetries and a featureless cylinder as a starting model. The models were iterated 20 times to allow convergence of the helical symmetry and the resultant asymmetric volumes were used in subsequent steps. It has been observed that a common mask between two volumes extends the FSC out further than it should by creating some high frequencies that arise from the well correlated common mask function and not from the two structures hence to obtain an accurate FSC one has to uncorrelate the radial masks. To accomplish this, the himpose program (IHRSR) was used to change the maximum radius value of one of the asymmetric volumes. The two volumes were aligned (using SPIDER) in z and ϕ to bring them into register with each other. The two volumes were windowed and padded (using SPIDER) to get rid of the vertical shift needed to align them. The FSC of the volumes was then calculated and plotted as a function of spatial frequency.

The 'conservative' resolution-defining threshold value "0.5" (Beck-Mann *et al.,* 1997; Bottcher *et al.,* 1997) was used to interpolate and calculate the resolution of the models.

Visualisation and docking of homology model

Visualisations and measurements of 2D images were done using BOXER and V2 both of the EMAN suite (Ludtke *et al*., 1999). UCSF Chimera (Petterson *et al*., 2004) was used to view the 3D reconstruction volume and to docking the homology model into the 3D reconstruction volumes. Docking was done manually and aided by the 'Fit model in map' function in Chimera to optimise the docking.

Due to the absence of a $CynD_{pum}$ crystal structure, methods such as homology modelling had to be implemented in order to generate a structure prediction of the enzyme. Over the years modelling of $CynD_{pum}$ structural models that have gone through a number of refinements based on insights obtained about the $CynD_{pum}$ enzymes and other related nitrilase structures, has been attempted. In this section we describe the methods used to construct the most resent $CynD_{pum}$ model, which was docked into the 3D reconstruction density maps. This model was generated by Dr Jean M. Watermeyer and Professor B. Trevor Sewell (University of Cape Town, South Africa) in collaboration with Dr James Gumbart and Dr Kwok-Yan Chan (University of Illinois at Urbana-Champaign). Six X-ray crystal structures were used as templates for homology modelling of CynDpum, based on their moderate sequence homology and high resolution: hypothetical nitrilase from *Pyrococcus horikoshii* (PDB code 1j31; (Sakai *et al*., 2004), N-carbamyl-D-amino acid amidohydrolase from *Agrobacterium sp*. strain KNK712 (PDB code 1erz; (Nakai *et al*., 2000), amidase from *Geobacillus pallidus* RAPc8 (PDB code 2plq; (Kimani *et al*., 2007), β-alanine synthase from *D. melanogaster* (PDB code 2vhi; (Lundgren *et al*., 2007), aliphatic amidase from *Nesterenkonia sp*. strain AN1 (PDB code 3hkx; (Nel *et al*., 2011) and formamidase AmiF from *Helicobacter pylori* (PDB code 2dyu; (Hung *et al*., 2007)). Once a homology was generated a structure-based multiple sequence alignment was prepared according to (Thuku *et al*., 2009). Secondary structure elements were predicted using PSIPRED (McGuffin *et al*., 2000), and elements predicted with high confidence scores were used as an aid to alignment and as a restraint during model building. Dimer homology models were generated using MODELLER 9 (Sali and Blundell, 1993). Side chain positions were refined using Scwrl3 (Canutescu *et al*.,

2003), keeping the positions of the four catalytic residues fixed. The model was initially docked into the 3D-TEM reconstructions using CoLoReS of the SITUS suite (Chacon & Wriggers, 2002). Dimer models were used to generate helical models using a Fortran script. In order to obtain the best interpretation of the intermolecular interfaces at high resolution, further rounds of fitting were carried out using a newly-implemented symmetrical flexible fitting route in NAMD (Chan *et al*., 2011; Trabuco *et al*., 2008).

4.4 Results

Data collection and image pre-processing

The collected micrographs contained straight intact fibre similar to images collected in section 2.4 (Fig. 4.2a). Boxed fibres were used to construct an average power spectrum which had visible layer lines corresponding to axial spacings of about 234, 110, 73, 55, 44 and 37 $Å^{-1}$ (Fig. 4.3a and 4.4a). The average power spectrum of the individual samples displayed similar axial spacing. The boxed fibres were further cut to produce to a stack of images which consisted of centred helical segments of the fibres (Fig. 4.2b).

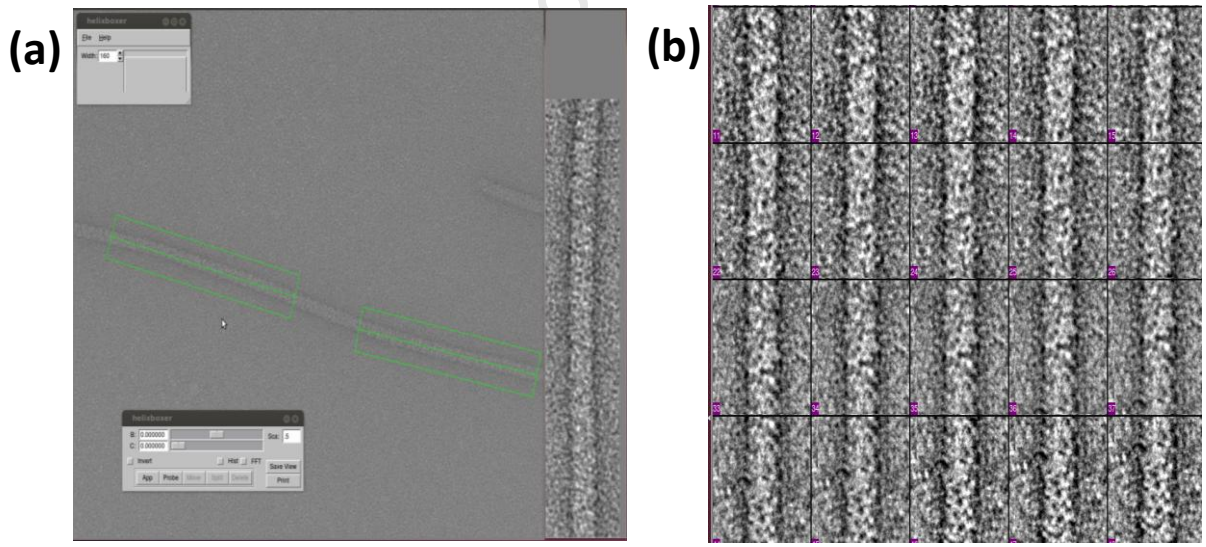


Figure 4.2. Representative image illustrating the pre-processing of EM micrographs of the $CynD_{pum}$ wild type. **(a)** Selection of helical fibres by boxing using the helixboxer program. **(b)** Image stack which consists of centred helical segments view using the V2 program.

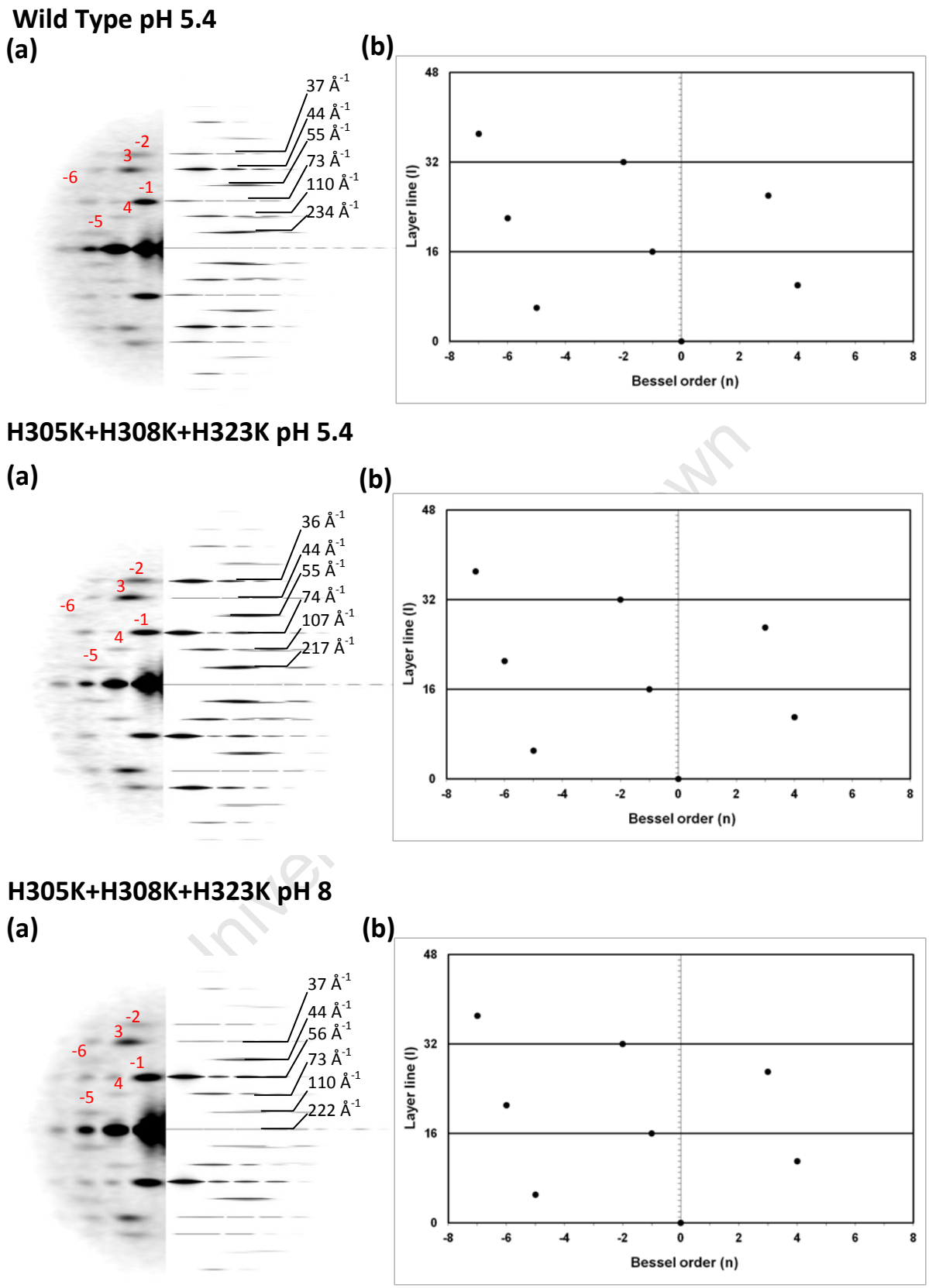


Figure 4.3. (a) Comparison of the average power spectrum obtained from the raw images (left) and from reconstructed maps (right) both demonstrating layer lines corresponding to the axial spacing shown in black which were interpreted as Bessel function of order shown in red. (b) n,l plot which enables indexing of the corresponding power spectrum of $CynD_{pum}$ wild type at pH 5.4, H305K+H308K+H323K at pH 5.4, H305K+H308K+H323K at pH8.

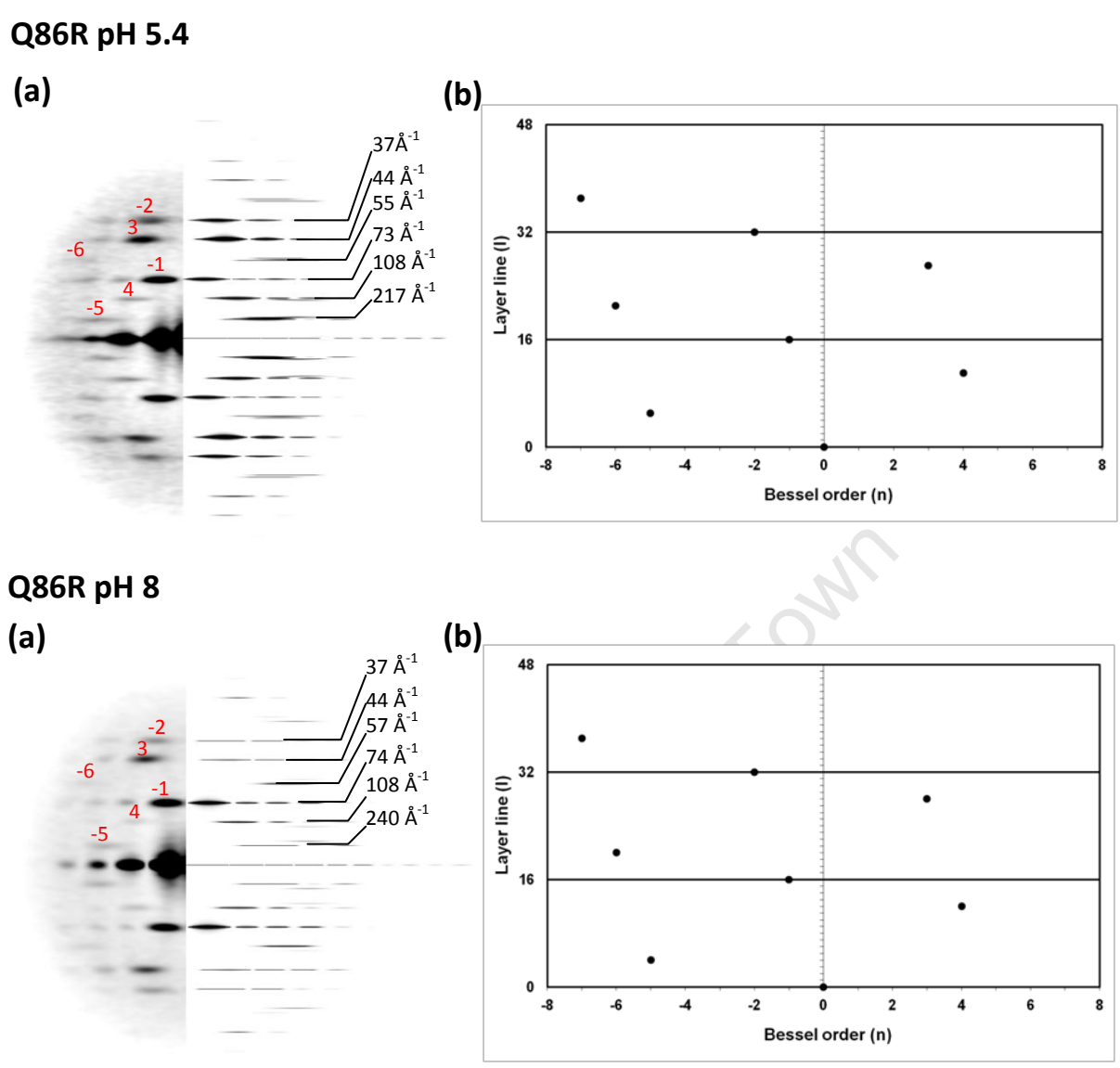


Figure 4.4. (a) Comparison of the average power spectrum obtained from the raw images (left) and from reconstructed maps (right) both demonstrating layer lines corresponding to the axial spacing shown in black which were interpreted as Bessel function of order shown in red. (b) n,l plot which enables indexing of the corresponding power spectrum of $CynD_{pum}$ Q86R at pH 5.4, Q86R at pH 8.

Three–dimensional reconstruction

IHRSR was carried out using the pre-processed image stacks to generate a 3D model and calculate the models' individual helical symmetry parameters in order to observe any quaternary structural changes. The frequency distribution of the reference projections of the individual reconstructions were evenly spread and most of the fluctuations could be explained by a random statistical model (Fig. 4.5a–4.9a). 3D maps of the $CynD_{pum}$ enzymes were produced and displayed a global 2-fold axis of symmetry perpendicular to the spiral axis which is due to the fibres being constructed from dimers (Fig. 4.10). The reconstructions all converged on stable models and this was demonstrated by the convergence of the helical symmetry parameters, from two independent reconstructions after 20 iteration of IHRSR (Fig. 4.5–4.9, b and c). The wild type at pH 5.4 converged on a helical symmetry of $\Delta\phi$ = -76.98°, Δz = 15.71. H305K+H308K+H323K displayed a similar $\Delta\phi$ to the wild type at both pH 5.4 and 8 but a minute increase in Δz (Δz = 15.80) was observed at both pH 5.4 and pH 8. Q86R displayed a similar $\Delta\phi$ to the wild type at pH 5.4 but this decreased to -76.70° at pH 8. The Δz of Q86R was lower than the wild type at pH 5.4 (Δz = 15.61) but increased to a Δz (Δz = 15.86) higher than the wild type at pH 8. The estimated resolution of the models was calculated from the spatial frequency obtained at FSC=0.5 (Fig. 4.11). The wild type displayed the highest resolution (17 Å) at pH 5.4. The power spectra produced by the reconstructions displayed the similar axial spacing to the averaged power spectrum produced from the raw data (Fig. 4.3 and 4.4). Using the helical symmetry values calculated by the IHRSR, n,l plots was constructed to enable indexing of the power spectra shown in Fig. 4.3 and 4.4. The visible layer lines observed on the average power spectra, corresponding to axial spacings of about 234, 110, 73, 55, 44 and 37 $Å^{-1}$ axial spacings were interpreted as having Bessel orders of -5, 4, -1, -6, 3 and -2 respectively based on there being 75 dimers in 16 turns (Fig. 4.3 and 4.4). The average power spectrum of the individual samples displayed similar axial spacing with similar Bessel orders.

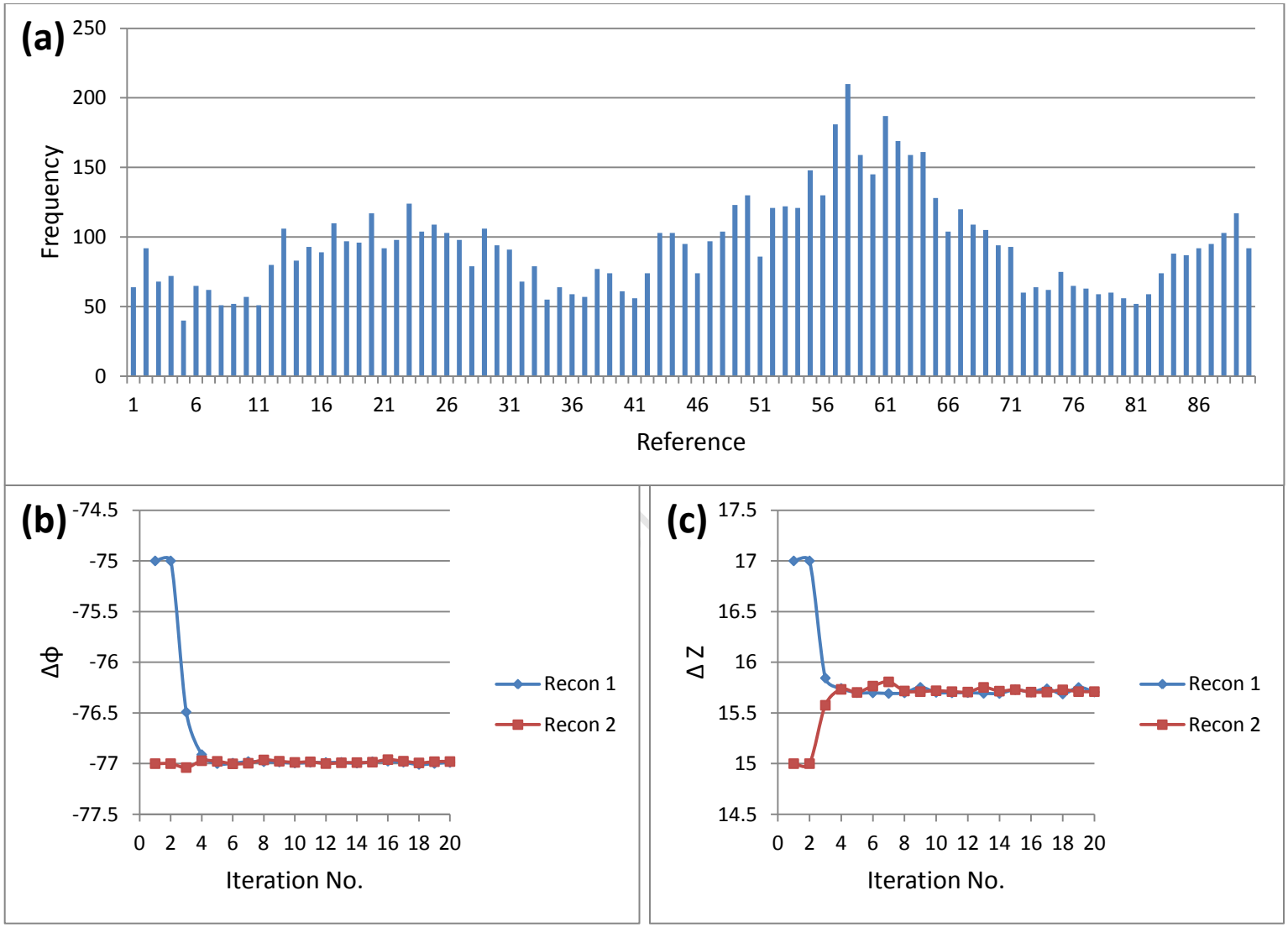


Figure 4.5. (a) Frequency distribution of reference projections of the 3D reconstruction of the wild type at pH 5.4. **(b)** and **(c)**, Convergence of two separate reconstructions (reconstruction 1 and 2) with different initial helical symmetry parameters . Iteration number was plotted against Δϕ (**b**), and Δ z**(c)**.

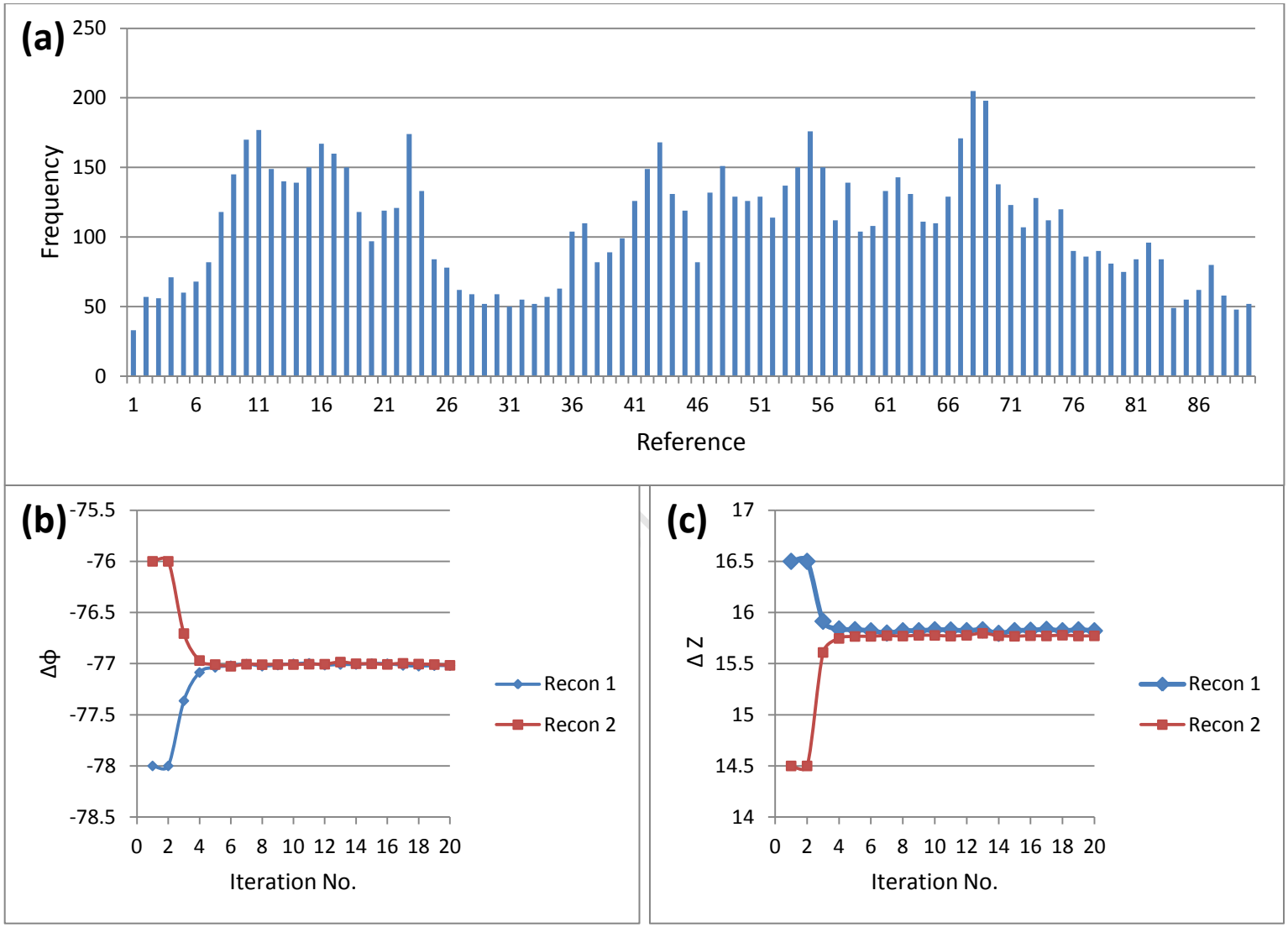


Figure 4.6. (a) Frequency distribution of reference projections of the 3D reconstruction of the H305K+H308K+H323K mutant at pH 5.4. **(b)** and **(c)**, Convergence of two separate reconstructions (reconstruction 1 and 2) with different initial helical symmetry parameters. Iteration number is plotted against Δϕ **(b)**, and Δ z **(c)**.

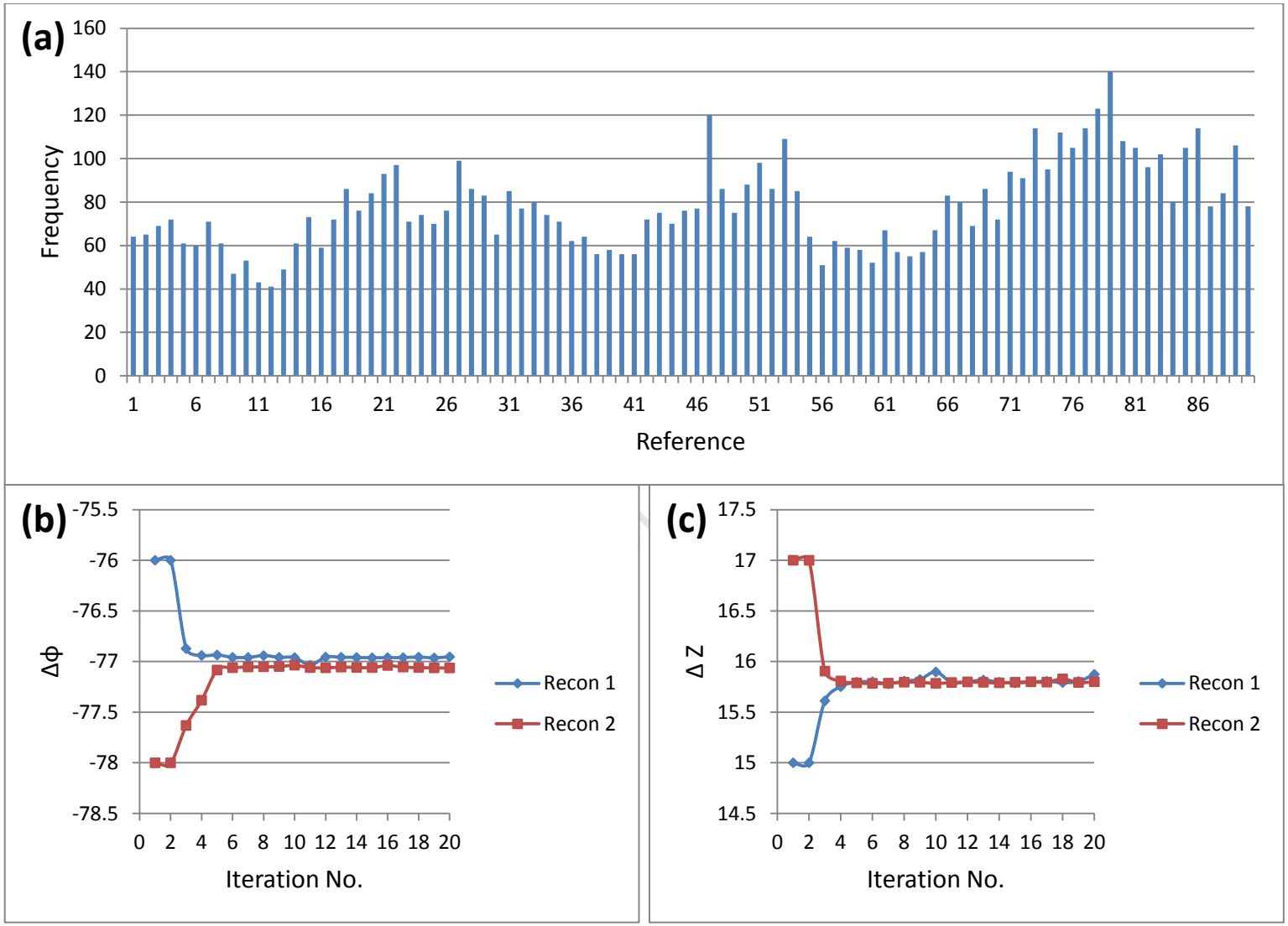


Figure 4.7. (a) Frequency distribution of reference projections of the 3D reconstruction of the H305K+H308K+H323K mutant at pH 8. **(b)** and **(c)**, Convergence of two separate reconstructions (reconstruction 1 and 2) with different initial helical symmetry parameters. Iteration number is plotted against Δϕ **(b)**, and Δ z **(c)**.

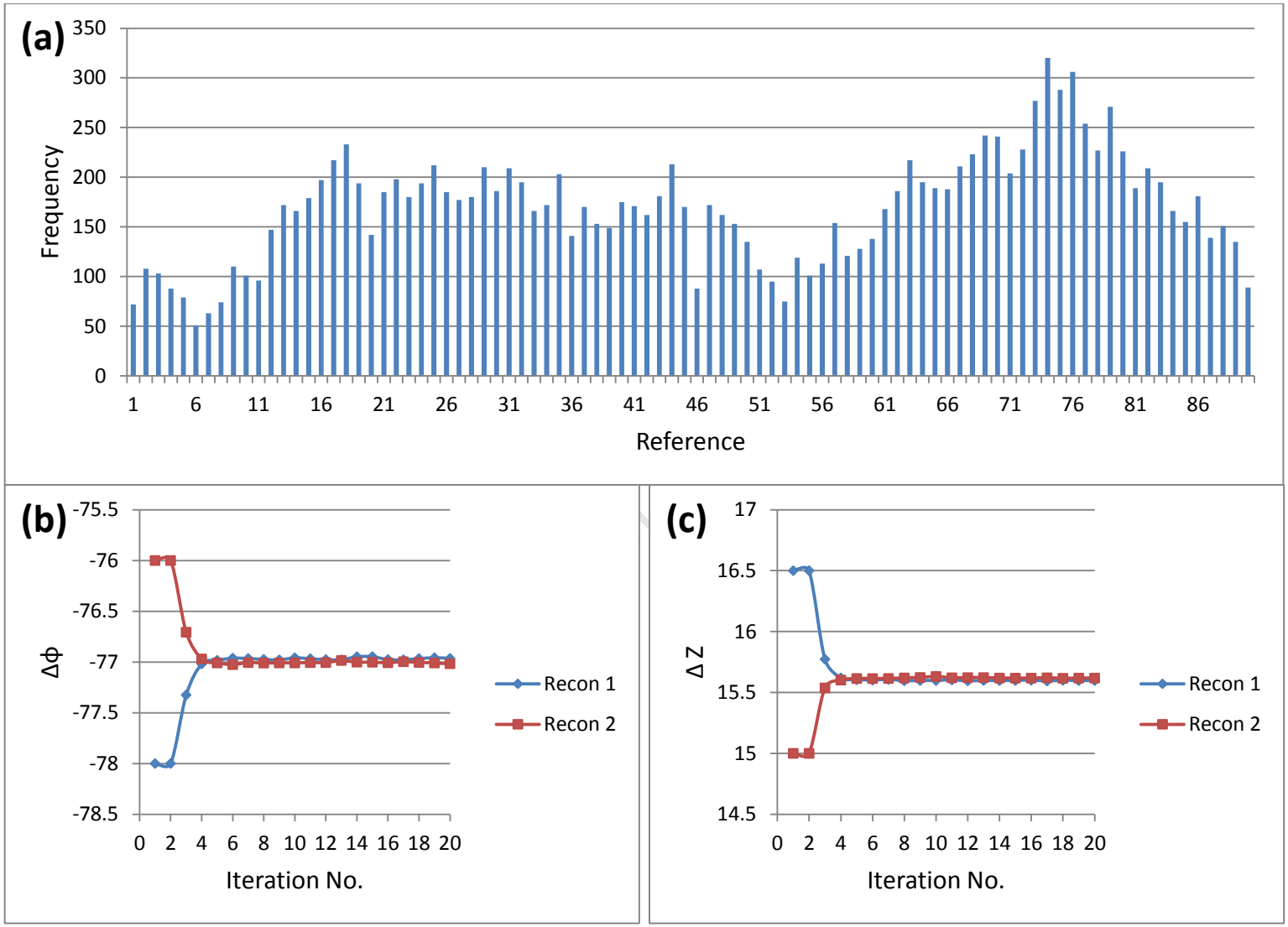


Figure 4.8. (a) Frequency distribution of reference projections of the 3D reconstruction of the Q86R mutant at pH 5.4. **(b)** and **(c)**, Convergence of two separate reconstructions (reconstruction 1 and 2) with different initial helical symmetry parameters. Iteration number is plotted against Δϕ **(b)**, and Δ z **(c)**.

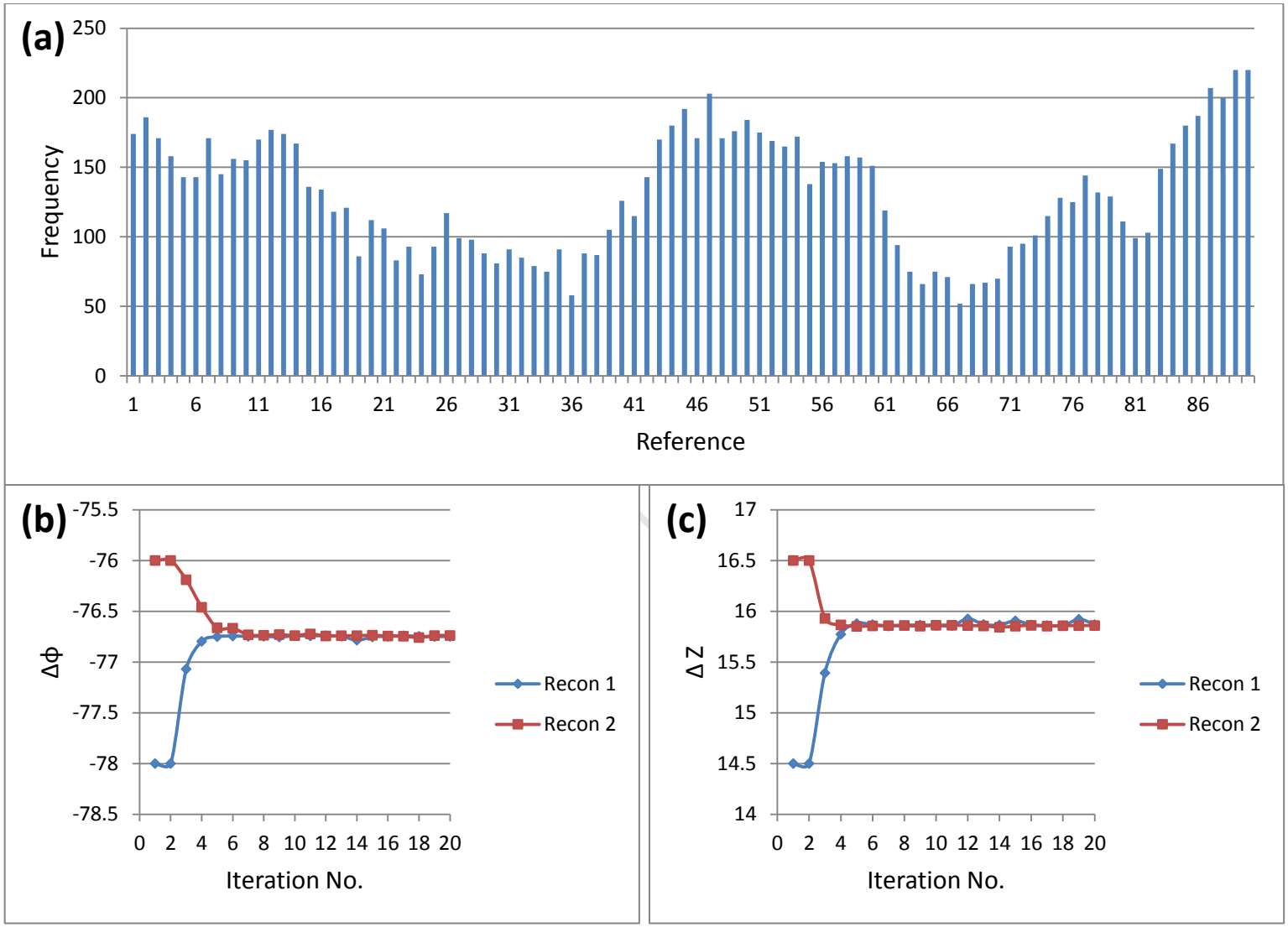


Figure 4.9. (a) Frequency distribution of reference projections of the 3D reconstruction of the Q86R mutant at pH 8. **(b)** and **(c)**, Convergence of two separate reconstructions (reconstruction 1 and 2) with different initial helical symmetry parameters. Iteration number is plotted against Δϕ **(b)**, and Δ z **(c)**

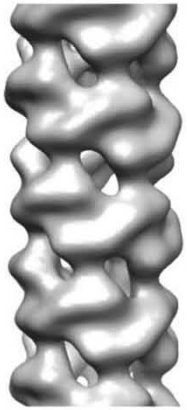
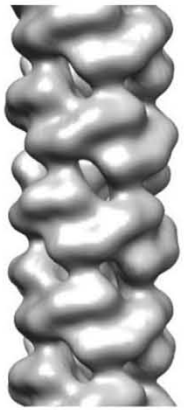
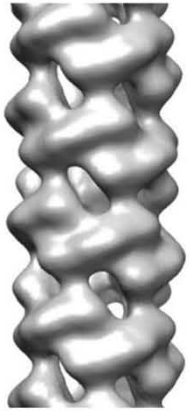
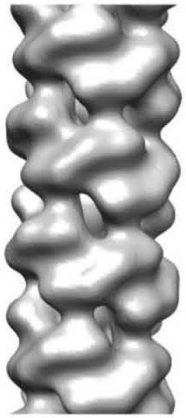
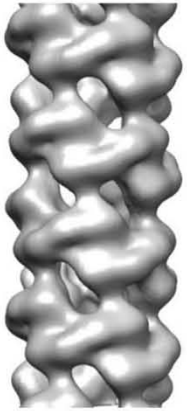
	WT	H305K, H308K, H323K	H305K, H308K, H323K	Q86R	Q86R
	pH 5.4	pH 5.4	pH 8	pH 5.4	pH 8
					
Helical Symmetry					
Δ φ	-76.98°	-77.02°	-77.00°	-77.00°	-76.70°
Δ z	15.71	15.80	15.80	15.61	15.86
Resolution	17 Å	20 Å	21 Å	20 Å	19 Å

Figure 4.10. Three-dimensional reconstructions and the converged helical symmetry parameters of the $CynD_{pum}$ mutants and wild type

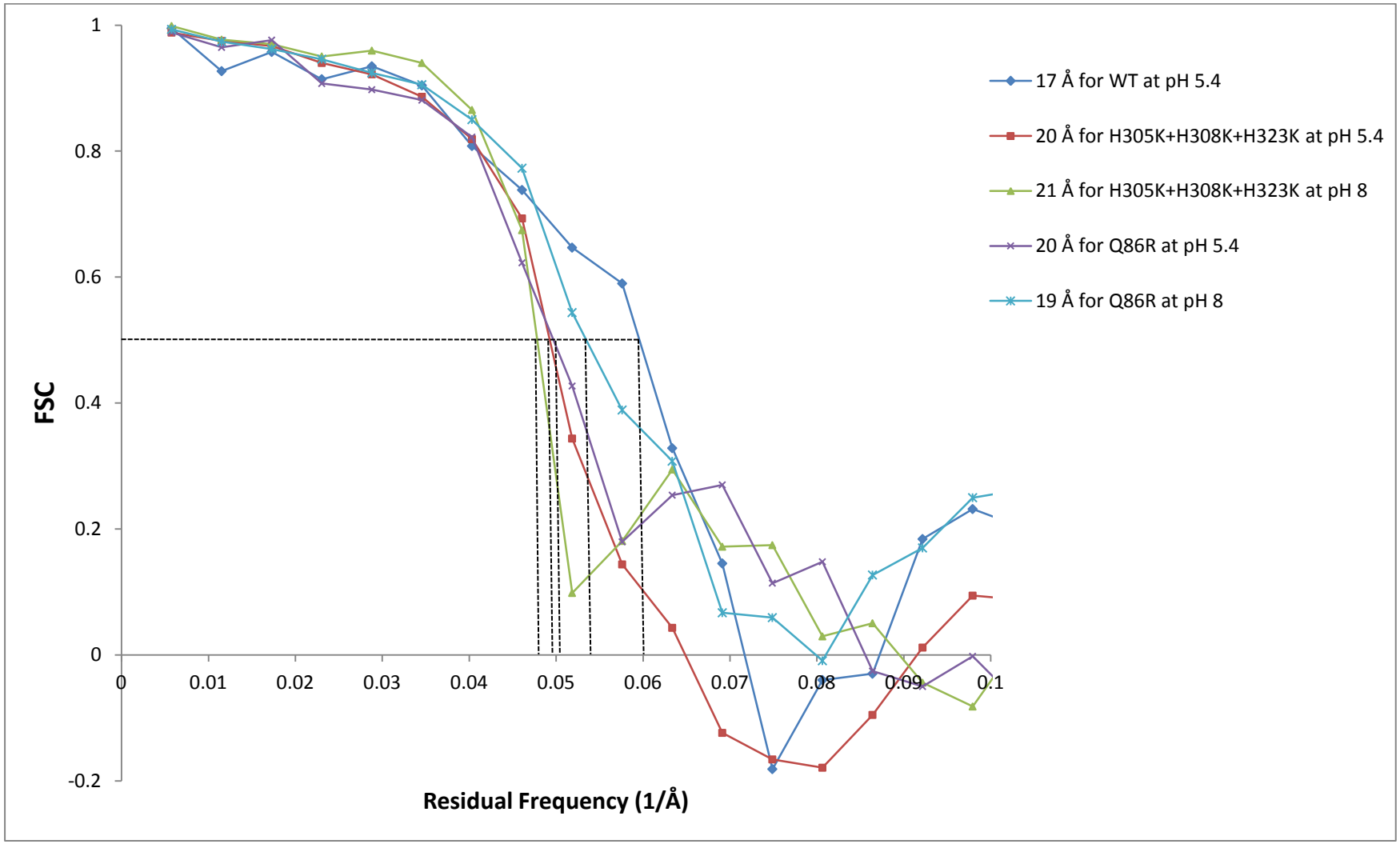


Figure 4.11. Estimated resolution of the 3-D reconstruction of the $CynD_{pum}$ mutants and wild type at the various pH. The resolution estimates were calculated from the residual frequency obtained at the 0.5 fourier shell correlation cut-off indicated on the graph.

Docking

The homology model mentioned in section 4.3 was docked into the 3D reconstruction density map of the wild type at pH 5.4 which had the highest resolution compared to the other 3D volumes generated. From the model we can clearly observe individual dimers along with the interfacial regions. The A surface is located between monomers of the same dimer, the D surface is located between adjacent dimers and the C surface is located is also located between adjacent dimers but at right angles to the A surface. The model also displayed the vacant densities located at the inside of the spiral. These densities have been proposed to be the location of the C-terminus (Thuku *et al.*, 2007). The C-terminus was omitted from the homology model as it is difficult to model and caused clashes during the flexible fitting procedure.

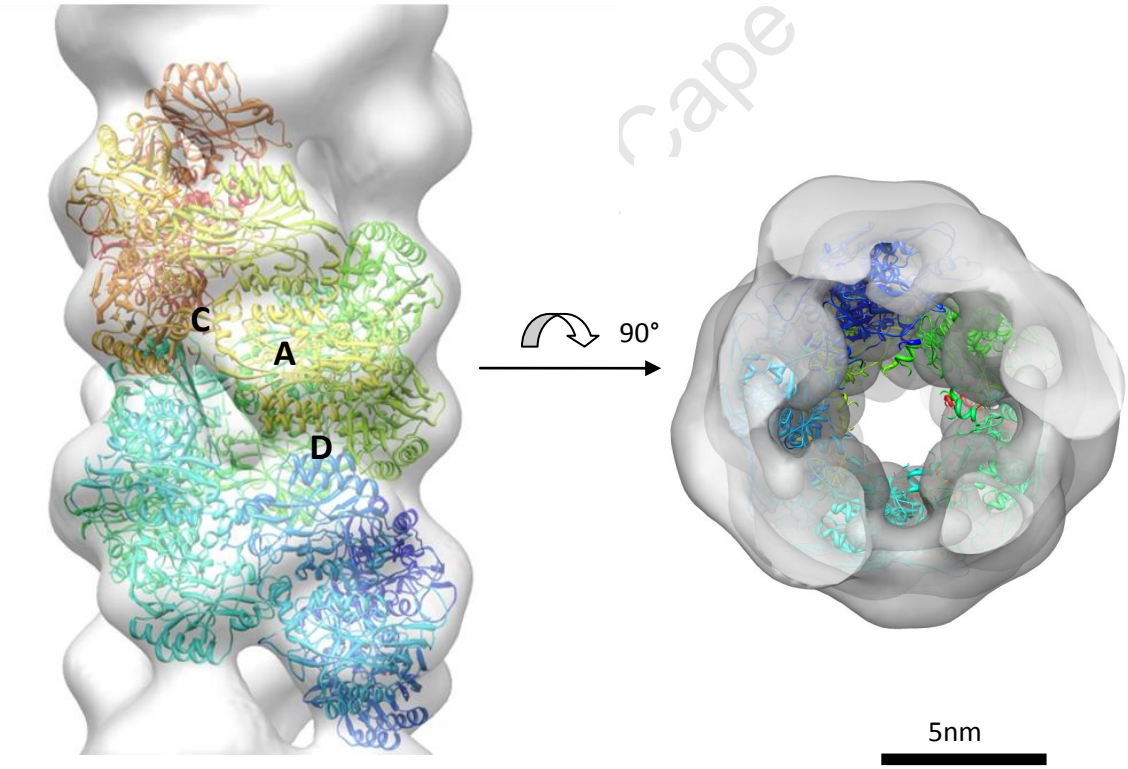


Figure 4.12. The docking of a homology model of $CynD_{pum}$ into a 3D reconstruction volume of $CynD_{pum}$ at pH 5.4. The location of the A, C and D surfaces are indicated in letters. A 90° rotation of the model shows the vacant densities proposed to be the location of the C-terminus.

4.5 Discussion

IHRSR produced the correct 3D maps of the $CynD_{pum}$ enzymes and these were validated by the convergence of the helical symmetries of two independent reconstruction. Similarities of the axial spacing between the reconstruction power spectra and the average power spectrum from the raw data also validated the 3D maps as it is highly unlikely that an incorrect structure would produce a correct power spectrum. It was previously proposed that the termination of the spirals during oligomerisation is due to deviations from perfect helical symmetry which results in structures with narrower ends preventing further addition of subunits (Sewell *et al*., 2003), thus the formation of the different sized $CynD_{pum}$ oligomers is due to changes in the helical symmetry of the quaternary structure. There were no significant changes in the helical symmetry across the different pH, but Q86R demonstrated a pH-dependent change in helical symmetry where slight changes in $\Delta\phi$ and Δz were observed as the pH changed from pH 5.4 to pH 8. Small changes in symmetry were observed at the different pH values. Furthermore the power spectra produced by both the raw data and the reconstructions displayed similar axial spacing across all the samples demonstrating no significant change in the quaternary structure of the $CynD_{pum}$ enzymes. The observed difference in helical symmetry for the Q86R mutant at the different pH may be due to heterogeneity which we have previously observed in both negative stain EM and Cryo-EM reconstructions of this particular mutant (not published). This heterogeneity has made it difficult to construct a high resolution 3D map even when using high resolution techniques such as Cryo-EM. There were also no significant differences in the helical symmetry across the different strains indicating that the mutations did not affect the helical symmetry of the quaternary structure of the mutants when compared to the wild type. This was also demonstrated by the similarities in the raw data and simulated power spectrum. Based on the absence of environmentally and mutationally induced changes in helical symmetry, we postulate that there may be factors other than change in helical symmetry, that result in the different fibre length of the $CynD_{pum}$ enzymes across various pH and across different strains. The 3D models we reconstructed were of low resolution and this may have masked any differences in helical symmetry, thus high resolution maps of the $CynD_{pum}$ enzymes have to be acquired before we can conclusively determine if there are any changes in helical symmetry at the different pH and between mutants.

In chapter II and III we observed an increase in stability by the Q86R, Q86R+E96G +D254E+E327G, H305K+H308K+H323K and $CynD_{pum-stut}$ mutants. In order to understand the structural mechanism of the effects of these mutations we first had to visualise them at an atomic level. Although no crystal structure of $CynD_{pum}$ is available, it has been possible to construct a homology model as mentioned in section 4.3. By docking this model into the 3D reconstruction density maps we generated, we were able to predict the positions of the mutations in the structure of $CynD_{pum}$ (Fig. 4.12). To better visualise any possible new interactions caused by the mutation, we further introduced the specific mutations to the model (Fig. 4.13 and Fig. 4.14). The Q86R mutant consists of a single mutation located on the D surface of the $CynD_{pum}$ oligomeric structure (Fig. 4.13). Q86R+E96G+D254E+E327G consists of two D surface mutations Q86R and E96G, together with a D254E mutation located on the outside of the $CynD_{pum}$ oligomer and a E327G mutation located at the C-terminal tail (Fig. 4.12 and Fig. 4.14). The mutation model of $CynD_{pum}$ demonstrates how the D surface mutation, Q86R, could be responsible for the observed increase in fibre length at the various pH (Table 2.12 and Fig. 2.12). This mutation is located on one of the D-surface α helices which has been proposed to form interactions between neighbouring subunits during oligomerisation (Sewell *et al.,* 2005). The Q86R mutation introduces a positive charge which may result in the formation of a new salt bridge across the neighbouring subunits and hence help stabilise the oligomeric quaternary structure of $CynD_{pum}$ (Fig. 4.13). In the presence of the Q86R mutation, one possible ionic interaction across the D surface may be between the R86 and E90 (oppositely charged residues) which are in close proximity on adjacent subunits (Fig. 4.13b).

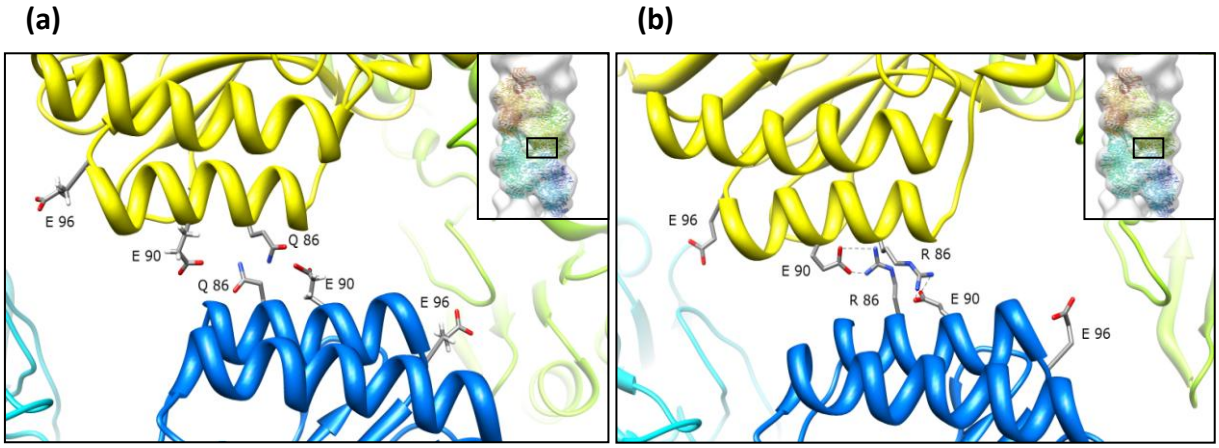


Figure 4.13. (a) Model of the D surface of $CynD_{pum}$ wild type. **(b)** Model of the D surface of $CynD_{pum}$ Q86R mutant.

On the other hand the additional E96G mutation in Q86R+E96G +D254E+E327G removes a negative charge on the D surface and is located in a remote location making it difficult to predict a mechanism of $CynD_{pum}$ stabilisation by this mutation (Fig. 4.14b). The location of D254E does not demonstrate the involvement of this mutation in stabilization, as it is located on the outside of the $CynD_{pum}$ structure which we propose is not involved in oligomerisation (Fig. 4.14, c and d). The remote locations of the E96G and D254E mutation in E96G +D254E+E327G, and the similar oligomerisation pattern between Q86R+E96G +D254E+E327G and Q86R observed in chapter II (Table 2.12 and Fig. 2.12) demonstrate that the Q86R mutation is largely responsible for the high stability in the Q86R+E96G +D254E+E327G mutant.

On the other hand we speculate that all the mutations of Q86R+E96G +D254E+E327G are required for a further increase stability due to the increase in alkaline pH stability (Fig. 3.3) and thermostability (Fig. 3.5 and Fig. 3.9) of Q86R+E96G +D254E+E327G relative to Q86R observed in chapter III. These results demonstrate the structural importance of the D surface interactions during oligomerisation and how these interactions can be modified to give a $CynD_{pum}$ that is structurally more stable than the wild-type.

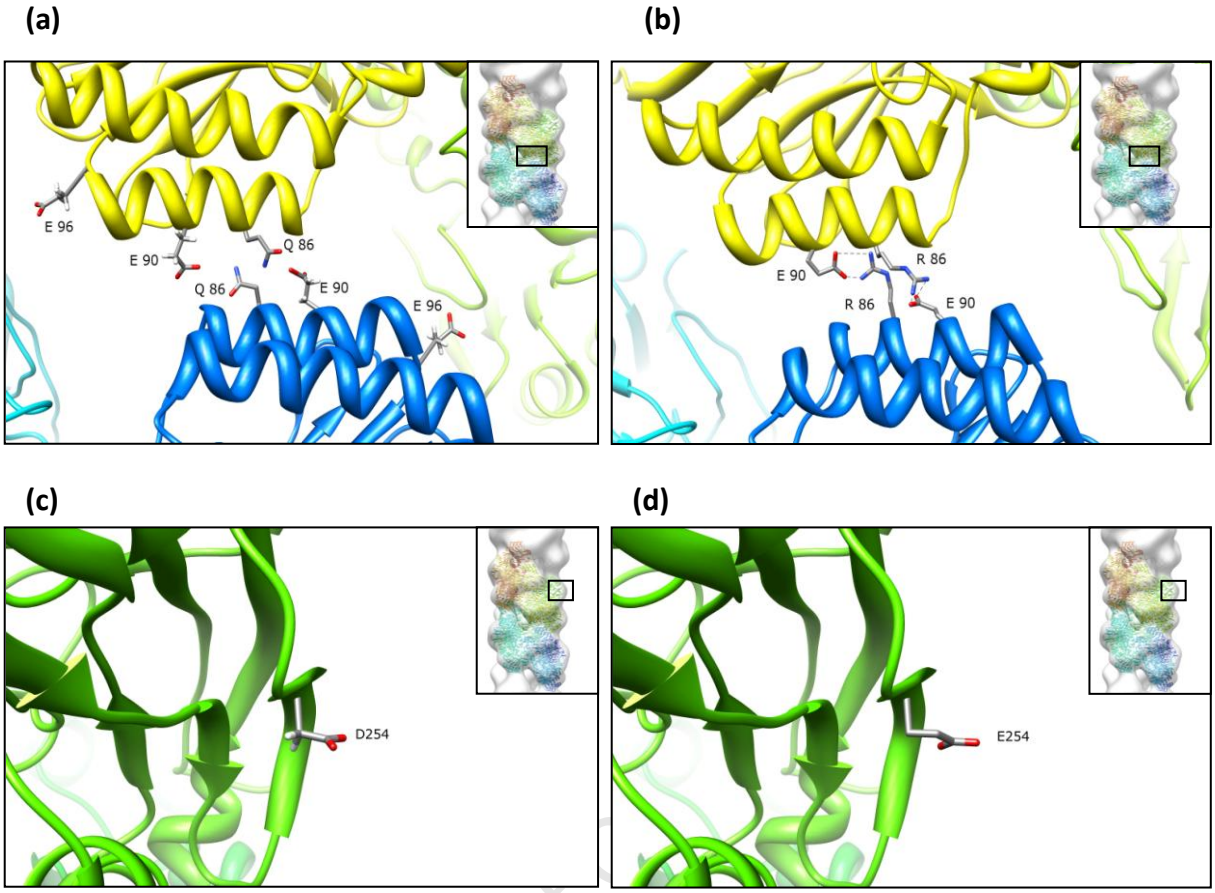


Figure 4.14. (a) Model of the D surface of $CynD_{pum}$ wild type. **(b)** Model of the D surface of $CynD_{pum}$ Q86R+E96G +D254E+E327G mutant **(c)** The location of D254 in $CynD_{pum}$ wild type. **(d)** The location of E254 in $CynD_{pum}$ Q86R+E96G +D254E+E327G mutant

The C-terminal tail has proven difficult to model accurately due to very low sequence identity with any of the template structures but has been predicted to be located on the inside of the $CynD_{pum}$ oligomeric structure (Fig. 4.12). The E327G mutation from Q86R+E96G +D254E+E327G is located in the C-terminal tail and this mutation removes a negative charge at the C- terminal tail (Fig. 4.12). This may result in a more positive inner surface or core of the $CynD_{pum}$ fibres. It was previously proposed that the C-terminal histidines are responsible for the pH-dependent transition observed in chapter II (Table 2.12 and Fig. 2.12) by contributing an ionic force either by repelling or attracting to stabilise the helical form of the enzymes as they make the transition below pH 6 and become positively charged. The E327G mutation may increase of the positive ionic force of the histidines at the core of the helical enzyme resulting in a more stable helical form of $CynD_{pum}$. The H305K+H308K+H323K mutant's retention of this ionic force at higher pH (lysine pKa=10.5) contradicts this

assumption since the presence of a pH-dependent transition from short to long helical forms between pH 8 and pH 5.4 was still observed for the H305K+H308K+H323K mutant (Table 2.12 and Fig. 2.12) and no change in helical symmetry was observed between pH 5.4 and pH 8 for the H305K+H308K+H323K mutant.

It is difficult to provide a structural mechanism for the increase in stability observed for the $CynD_{pum-stut}$ due to the multiple amino acid mutations present in the mutant which make it challenging to predict the key amino acid responsible for the observed effect. In addition, the mutations are located on the C-terminal tail which has proven difficult to model as mentioned previously.

Based on the homology model of the wild type $CynD_{pum}$ (Fig. 4.12), none of the mutations are located in proximity to the active site thus any changes in activity observed in the pH profile and thermostability assays (Fig. 3.3, 3.5 and 3.9) are unlikely to be due to the interruption of the active site. The theoretical pKa's of the active site residues has been calculated for *At*NIT4, a homologous plant nitrilase from *Arabidopsis thaliana* (Gordon *et al*., 2005; Myers *et al*., 2006) where the pKa's of the active site residue were found to be out of the range of the pH range we used for pH profiling (chapter III), thus the pH unlikely changed the protonation states of the active site residues. Therefore the observed increase in alkali pH stability relative to the wild type of mutants Q86R+E96G +D254E+E327G and $CynD_{pum-stut}$ are likely due to increase in stability (slowed unfolding of the enzyme) and not the change in catalytic properties of the enzymes.

4.6 Conclusion

The mutations on the $CynD_{pum}$ strains resulted in different oligomerisation patterns but the mechanism for the formation of these different oligomers may not be due to changes in the helical symmetry as previously proposed. The transition from short spirals at pH 8 to long fibres at pH 5.4 may be due to other mechanisms and not the environmentally induced changes in helical symmetry. The lack of a high resolution 3D map of $CynD_{pum}$ makes it difficult to validate these results. The Q86R mutation improves the stability of $CynD_{pum}$ by introducing a new salt bridge interaction across the D surface. Although the location of the

other mutations that increase $CynD_{pum}$ stability can be visualised, the mechanism behind their ability to confer these changes is still unclear.

CHAPTER V

GENERAL DISCUSSION AND FUTUREWORK

The $CynD_{pum}$ has the potential to provide a cheap and safe bioremediation solution in detoxification of industrial cyanide wastewater but it lacks the alkaline pH stability to accomplish this. Before this enzyme can be used for this task, structural and functional insights have to be acquired in order to engineer a more stable enzyme. The D surface and C-terminal tail mutants we contracted have given us some insights of the structure and function of $CynD_{pum}$ including residue modifications that can improve stability. We previously thought that an increase in the degree of oligomerisation of the $CynD_{pum}$ enzyme resulted in an increase in stability and although this was true for the D surface and C-terminal histidine mutants, the $CynD_{pum\text{-}stut}$ displayed an increase in stability independent to its oligomerisation. This demonstrates how improving the stability of $CynD_{pum}$ does not necessarily have to change the oligomerisation state of the enzymes thus stable $CynD_{pum}$ enzymes that do not form long helical fibres do exist and our simple model (Fig. 4.12) is not adequate. Stability is not limited to the D surface and the C-terminal tail. The effects of mutations at other interfacial region such as the A and C surface still have to be studied as there have been reports that these surfaces are involved in oligomerisation and activity (Sewell *et al*., 2005).

The D surface mutant Q86R+E96G +D254E+E327G and the C terminal tail hybrid $CynD_{pum\text{-}stut}$ showed dramatic increase in alkaline pH stability relative to the wild type. These mutants have potential to be used in bioremediation of cyanide wastewaters. Further studies on the reaction of these mutants to metals that contaminate some industrial cyanide wastewaters have to be carried out since these metals have been shown to inhibit activity of the wild type enzyme (Jandhyala *et al*., 2005). In vivo experiments have shown that there is a greater pH tolerance when the mutant enzymes are enclosed in cells rather than in vitro (Wang *et al*., 2011). The increased tolerance was presumed to be due to the cells ability to maintain partial pH homeostasis. This demonstrates that the use of bacteria whole cell cultures may provide an even more efficient method for bioremediation of cyanide wastewaters.

In this study we demonstrated the importance of the C-terminal tail in oligomerisation and stability of the $CynD_{pum}$ enzyme, although its role is not clear. The importance of the C-terminal tail has been demonstrated in other microbial nitrilase such as the *R. rhodochrous* J1 nitrilase. This nitrilase displayed loss of activity when 55 C-terminal residues were deleted. However, the enzyme performs an auto-cleavage of 39 amino acids from its tail in order to form active oligomers (Thuku *et al.*, 2007). It was then speculated that this region might be interacting with the A and C surface and is important for the oligomerisation of the protein (Thuku *et al.*, 2007). In a different study, the C-terminal region of the arylacetonitrilase from *Pseudomonas fluorescens* EBC191 was speculated to be necessary for activity and stability of this enzyme, as well as its enantioselectivity (Kiziak *et al.*, 2007). In the case of $CynD_{pum\text{-}stut}$ we identified the group of C-terminal residues that increase stability but the key residues have to be identified as it is unlikely that all the residues mutated are responsible for the conferred stability. Further work has to be done on the C-terminal region of other nitrilases to gain more insight on the possible role of this region in nitrilases.

There is an increase need for a branch 1 nitrilase high-resolution crystal structure to help validate the recent discoveries made about these enzymes. A high-resolution structure could provide more insight into the interactions involved at the interfacial regions thus allowing the design of more stable $CynD_{pum}$ enzymes. Until such a high-resolution structure is obtained, we have to rely on structure prediction methods, modelling and mutational studies. Although homology modelling of $CynD_{pum}$ has given insights into the structure of the enzyme, new methods have been developed to refine the generated models. One successful method for this refinement is the molecular dynamic flexible fitting (MDFF) developed by Trabuco *et al*. (2008). This method employs molecular dynamics simulations to perform flexible fitting of atomic models into EM maps. The method was further improved to incorporate knowledge of structural symmetry from EM maps into the fitting procedure thus the symmetry-restrained MDFF was developed (Chan *et al*., 2011). Symmetry-restrained MDFF will add EM map-derived potential further restrains into the molecular dynamic flexible fitting procedure to help improve the quality of our homology models. Besides structural symmetry other restrains can be added during the molecular dynamic flexible fitting procedure such as active site constrains, force field constrains and

interatomic distance constrains. One way to obtain knowledge of interatomic distances is to carry out protein cross-linking experiments coupled to mass spectrometry (MS). In this method the structure of proteins or protein complexes is ascertained by identifying amino acid pairs that are positioned in close proximity to each other (Rappsilber, 2011). More restraints could be obtained as more knowledge about the structure of branch 1 nitrilases is acquired.

REFERENCES

Abou-Nader, M. and Benedik, M.J. (2010) Rapid generation of random mutant libraries. *Bioeng Bugs* **1,** 337-340.

Agarkar, V.B., Kimani, S.W., Cowan, D.A., Sayed, M.F. and Sewell, B.T. **(**2006) The quaternary structure of the amidase from *Geobacillus pallidus* RAPc8 is revealed by its crystal packing. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62,** 1174-1178.

Andrade, J., Karmali, A., Carrondo, M.A. and Frazao, C. (2007) Structure of amidase from *Pseudomonas aeruginosa* showing a trapped acyl transfer reaction intermediate state. *J Biol Chem* **282**, 19598–19605.

Aurrand-Lions, M., Galland, F., Bazin, H., Zakharyev, V.M., Imhof, B.A. and Naquet, P. (1996) Vanin-1, a novel GPI-linked perivascular molecule involved in thymus homing. *Immunity* **5**, 391-405.

Banerjee, A., Sharma, R. and Banerjee, U.C. (2002) The nitrile-degrading enzymes: current status and future prospects. *Appl Microbiol Biotechnol* **60**, 30–44.

Barglow K.T., Saikatendu K.S., Bracey M.H., Huey R., Morris G.M., Olson A.J., Stevens R.C. and Cravatt, B.F.(2008) Fuctional proteomic and structural insights into molecular recognition in the nitrilase family enzymes. *Biochemistry* **47**, 13514-13523.

Bartling, D., Seedorf, M., Mithofer, A. and Weiler, E.W. (1992) Cloning and expression of an *Arabidopsis* nitrilase which can convert indole-3-acetonitrile to the plant hormone, indole-3-acetic acid. *Eur J Biochem* **205**, 417–424.

Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G. and Frank, J. (1997) Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* **278**, 2123–2126.

Bottcher, B., Wynne, S.A. and Crowther, R.A. (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature* **386**, 88–91.

Bradford, M.M. (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. *Anal Biochem* **72**, 248–54.

Brenner, C. (2002) Catalysis in the nitrilase superfamily. *Curr Opin Struct Biol* **12**, 775–782.

Carver, T.E., Bordeau, B., Cummings, M.D., Petrella, E.C., Pucci, M.J. and Zawadzke, L.E. (2005) Decrypting the biochemical function of an essential gene from *Streptococcus pneumoniae* using ThermoFluor technology. *J Biol Chem* **280**, 11704–11712.

Chan K.Y., Gumbart J., McGreevy R., Watermeyer J.M., Sewell B.T. and Schulten, K. (2011) Symmetry-restrained flexible fitting for symmetric EM maps. *Structure* **19**, 1211-1218.

Chen, C., Chiu, W., Liu, J., Hsu, W. and Wang, W.C. (2003) Structural basis for catalysis and substrate specificity of *Agrobacterium radiobacter* N-Carbamoyl-D-amino acid amidohydrolase. *J Biol Chem* **278**, 26194–26201.

Chin, K.H., Tsai, Y.D., Chan, N.L., Huang, K.F., Wang, A.H.J. and Chou, S.H. (2007) The crystal structure of XC1258 from *Xanthomonas campestris*: a putative prokaryotic Nit protein with an arsenic adduct in the active site. *Proteins: Struct Funct Bioinform* **69**, 665–671.

Cobzaru, C., Ganas, P., Mihasan, M., Schleberger, P. and Brandsch, R. (2011) Homologous gene clusters of nicotine catabolism, including a new omega-amidase for alpha-ketoglutaramate, in species of three genera of Gram-positive bacteria. *Res Microbiol* **162**, 285–291.

Cohen, G.E. (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J Appl Cryst* **30**, 1160–1161.

Curnow, A.W., Kw, H., Yuan, R., Si, K., Martins, O., Winkler, W., Henkin, T.M. and Soll, D. (1997) Glu-tRNAGln amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc Natl Acad Sci USA* **94**, 11819–11826.

Dent, K.C., Weber, B.W., Benedik, M.J. and Sewell, B.T. (2008) The cyanide hydratase from *Neurospora crassa* forms a helix which has a dimeric repeat. *Appl Microbiol Biotechn*ol doi: 10.1007 ⁄ s00253-008-1735-4.

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

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

Eicher, J. J. (2007) pH-dependence of the quaternary structure of the cyanide dihydratase from *Bacillus pumilus*. M. Sc. Thesis, University of Cape Town, Rondebosch, South Africa.

Eisele, F. and Wolf, D.H. (2008) Degradation of misfolded protein in the cytoplasm is mediated by the ubiquitin ligase Ubr1. *FEBS Lett* **582**, 4143–4146.

Ericsson, U.B., Hallberg, M.B., DeTitta, G.T., Dekker, N. and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289–298.

Fisher, F.B. and Brown, J.S. (1952) Colorimetric determination of cyanide in stack gas and waste water. *Anal Chem* **24***,* 1440–1444.

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*. J Struct Biol* **116**, 190–199.

Frank, J., Verschoor, A. and Boublik, M. (1981) Computer averaging of electron micrographs of 40S ribosomal subunits. *Science* **214**, 1353–1355.

Freyssinet, G., Peleissier, B., Freyssinet, M. and Delon, R. (1996) Crops resistant to oxynils: from the laboratory to the market. *Field Crops Res* **45**, 125–133.

Godtfredsen, S.E., Ingvorsen, K., Yde, B. and Anderson, O. (1985) In: Tramper J, Vanderplas HC, Linko P (eds) Biocatalysis in organic syntheses. Elsevier, Amsterdam, pp 3–18.

Gordon, J.C., Myers, J.B., Folta, T., Shoja, V., Heath, L.S. and Onufriev, A. (2005) H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res* **33**, 368–371.

Grifantini, R., Pratesi, C., Galli, G. and Grandi, G. (1996) Topological mapping of the cysteine residues of Ncarbamyl-D-amino-acid amidohydrolase and their role in enzymatic activity. *J Biol Chem* **271**, 9326–9331.

Grigorieff, N. (2000) Resolution measurement in structures derived from single particles. *Acta Crystallogr D Biol Crystallogr* **56**, 1270–1277.

Hung, C.L., Liu, J.H., Chiu, W.C., Huang, S.W. and Wang, W.C. (2007) Crystal structure of *Helicobacter pylori* Formamidase AmiF reveals a cysteine-glutamate-lysine catalytic triad. *J Biol Chem* **282**, 12220–12229.

Jandhyala, D.M., Berman, M., Meyers, P.R., Sewell, B.T., Willson, R.C. and Benedik, M.J. (2003) Cyn D, the cyanide dihydratase from *Bacillus pumilus*: gene cloning and structural studies. *Appl Environ Microbiol* **69**, 4794–4805.

Jandhyala, D.M., Willson, R.C., Sewell, B.T. and Benedik, M.J. (2005) Comparison of cyanide degrading nitrilases. *Appl Microbiol Biotechnol* **68**, 327–335.

Jeng, T.W., Crowther, R.A., Stubbs, G. and Chiu, W. (1989) Visualization of alpha-helices in tobacco mosaic virus by cryo-electron microscopy. *J Mol Biol* **205**, 251–257.

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

Kato, Y., Ooi, R. and Asano, Y. (2000) Distribution of aldoxime dehydratase in microorganisms. *Appl Environ Microbiol* **66**, 2290–2296.

Kimani, S.W., Agarkar, V.B., Cowan, D.A., Sayed, M.F.R. and Sewell, B.T. (2007) Structure of an aliphatic amidase from *Geobacillus pallidus* RAPc8. *Acta Crystallogr Sect D: Biol Crystallogr* **63**, 1048–1048.

Kiziak, C., Klein, J. and Stoltz, A. (2007) Influence of different carboxy-terminal mutations on the substrate-, reaction and enantiospecificity of the arylacetonitrilase from *Pseudomonas fluorescens* EBC191. *Prot Eng Des Sel* **20**, 385–396.

Kobayashi, M., Goda, M. and Shimizu, S. (1998) Nitrilase catalyzes amide hydrolysis as well as nitrile hydrolysis. *Biochem Biophys Res Commun* **253**, 662–666.

Krajewski, W.W., Collins, R., Holmberg-Schiavone, L., Jones, T.A., Karlberg, T. and Mowbray, S.L. (2008) Crystal structures of mammalian glutamine synthetases illustrate substrate-induced conformational changes and provide opportunities for drug and herbicide design. *J Mol Biol* **375**, 217-228.

Krasnikov, B.F., Chien, C.H., Nostramo, R., Pinto, J.T., Nieves, E., Callaway, M., Sun, J., Huebner, K. and Cooper, A.J. (2009a) Identification of the putative tumor suppressor Nit2 as omega-amidase, an enzyme metabolically linked to glutamine and asparagine transamination. *Biochimie* **91**, 1072–1080.

Krasnikov, B.F., Nostramo, R., Pinto, J.T. and Cooper, A.J. (2009b) Assay and purification of omegaamidase/ Nit2, a ubiquitously expressed putative tumor suppressor, that catalyzes the deamidation of the alpha-keto acid analogues of glutamine and asparagine. *Anal Biochem* **391**, 144–150.

Kumaran, D., Eswaramoorthy, S., Gerchman, S.E., Kycia, H., Studier, F.W. and Swaminathan, S. (2003) Crystal structure of a putative CN hydrolase from yeast. *Proteins* **52**, 283–291.

Kvalnes-Krick, K.L. and Traut, T.W. (1993) Cloning, sequencing, and expression of a cDNA encoding beta-alanine synthase from rat liver. *J Biol Chem* **268**, 5686–5693.

LaRonde-LeBlanc, N., Resto, M. and Gerratana, B. (2009) Regulation of active site coupling in glutamine-dependent NAD^+ synthetase. *Nat Struct Mol Biol* **16**, 421–429.

Lavinder, J.J., Hari, S.B., Sullivan, B.J. and Magliery, T.J. (2009) High-throughput thermal scanning: a general, rapid dye-binding thermal shift screen for protein engineering. *J Am Chem Soc* **131**, 3794–3795.

Liao, H.Y. and Frank, J. (2010) Definition and estimation of resolution in single-particle reconstructions. *Structure* **18**, 768–775.

Lundgren, S., Lohkamp, B., Andersen, B., Piskur, J. and Dobritzsch, D. (2008) The crystal structure of B-alanine synthase from *Drosophila melanogaster* reveals a homoactameric helical turn-like assembly. *J Mol Biol* **377**, 1544–1559.

Ludtke S.J., Baldwin, P.R. and Chiu, W. (1999) EMAN: semi-automated software for high-resolution single particle reconstructions. *J Struct Biol* **128**, 82–96.

Maras, B., Barra, D., Dupre, S. and Pitari, G. (1999) Is pantetheinase the actual identity of mouse and human vanin-1 proteins? *FEBS Lett* **461**, 149–152.

Mathew, C., Nagasawa, T., Kobayashi, M. and Yamada, H. (1988) Nitrilase catalyzed production of nicotinic acid from 3-cyanopyridine in *Rhodococcus rhodochrous* J1. *Appl Environ Microbiol* **54**, 1030–1032.

Matulis, D., Kranz, J.K., Salemme, F.R. and Todd, M.J. (2005) Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* **44**, 5258–5266.

Meyers, P.R., Rawlings, D.E., Woods, D.R. and Lindsey, G.G. (1993) Isolation and characterization of a cyanide dihydratase from *Bacillus pumilus* C1. *J Bacteriol* **175**, 6105–6112.

Miyazawa, A., Fujiyoshi, Y., Stowell, M. and Unwin, N. (1999) Nicotinic acetylcholine receptor at 4.6 A resolution: transverse tunnels in the channel wall. *J Mol Biol* **288**, 765–786.

Mogk, A., Schmidt, R. and Bukau, B. (2007) The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol* **17**, 165–172.

Mulelu, A.E. (2010) The Structure of high pH tolerant mutants of the cyanide dihydratase from *Bacillus pumilus* C1. Honours Thesis, University of Cape Town, Rondebosch, South Africa.

Muller, D. and Gabriel, J. (1999) Bacterial degradation of the herbicide bromoxynil by *Agrobacterium radiobacter* in biofilm. *Folia Microbiol (Praha)* **44**, 377–379.

Myers, J., Grothaus, G., Narayanan, S. and Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pK_as in macromolecules. *Proteins* **63**, 928–938.

Nakada, Y., Jiang, Y., Nishijyo, T., Itoh, Y. and Lu, C.D. (2001) Molecular characterization and regulation of the aguBA operon, responsible for agmatine utilization in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **183**, 6517–6524.

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

Nel, A.J., Tuffin, I.M., Sewell, B.T. and Cowan, D.A. (2011) Unique aliphatic amidase from a psychrotrophic and haloalkaliphilic nesterenkonia isolate. *Appl Environ Microbiol.* **77**, 3696–3702.

Niesen, F.H., Berglund, H. and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols* **2**, 2212–2221.

Nishise H, Kurihara M, Tani Y (1987) Microbial synthesis of tran-examic acid intermediate from dinitrile. *Agric Biol Chem* **51**, 2613–2616

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

Orlova, E.V., Dube, P., Harris, J.R., Beckmann, E., Zemlin, F., Markl, J. and van Heel, M. (1997) Structure of Keyhole Limpet Hemocyanin Type 1 (KLH1) at 15 Å resolution by electron cryomicroscopy and angular reconstitution. *J Mol Biol* **271**, 417–437.

Pace, H.C., Hodawadekar, S.C., 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 dimers. *Curr Biol* **10**, 907–917.

Pace, H. and Brenner, C. (2001) The nitrilase superfamily: classification, structure and function. *Genome Biol* **2**, 1–9.

Pantoliano M.W., Petrella E.C., Kwasnoski J.D., Lobanov V.S., Myslik J., Graf E., Carver T., Asel E., Springer B.A., Lane P. and Salemme F.R. (2001) High-density miniaturized thermal shift assays as a general strategy for drug discovery, *J Biomol Screen* **6**, 429–440.

Parent, K.N., Sinkovits, R.S., Suhanovsky, M.M., Teschke, C.M., Egelman, E.H. and Baker, T.S. (2010) Cryo-reconstructions of P22 polyheads suggest that phage assembly is nucleated by trimeric interactions among coat proteins. *Phys Biol* **7**, 045004.

Pekarsky, Y., Campiglio, M., Siprashvili, Z., Druck, T., Sedkov, Y., Tillib, S., Draganescu, A., Wermuth, P., Rothman, J.H. and Huebner, K. (1998) Nitrilase and Fhit homologs are encoded as fusion proteins in *Drosophila melanogaster* and *Caenorhabditis elegans. Proc Natl Acad Sci USA* **95**, 8744–8749.

Penczek, P.A. (2002) Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms. *J Struct Biol* **138**, 34–46.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.

Piotrowski, M. (2008) Primary or secondary? Versatile nitrilases in plant metabolism. *Phytochemistry* **69**, 2655–67.

Polizzi, K. M., Bommarius, A. S., Broering, J. M. and Chaparro-Riggers, J. F. (2007) Stability of biocatalysts. *Curr Opin Chem Biol* **11**, 220–225.

Raczynska, J.E., Vorgias, C.E., Antranikian, G. and Rypniewski, W. (2011) Crystallographic analysis of a thermoactive nitrilase. *J Struct Biol* **173**, 294–302.

Rappsilber, J. (2011) The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J Struct Biol* **173**, 530–540.

Rogers, S.D., Bhave, M.R., Mercer, J.F.B., Camakaris, J. and Lee, B.T.O. (1991) Cloning and characterization of *cutE*, a gene involved in copper transport in *Escherichia coli. J Bacteriol* **173**, 6742–6748.

Sakai, N., Tajika, Y., Yao, M., Watanabe, N. and Tanaka, I. (2004) Crystal structure of hypothetical protein PH0642 from *Pyrococcus horikoshii* at 1.6 Å resolution. *Proteins: Struct Funct Bioinform* **57**, 869–873.

Saxton, W.O. and Baumeister, W. (1982) The correlation averaging of a regularly arranged bacterial cell envelope protein. *J Microsc* **127**, 127–138.

Scheffer, M.P. (2006) Helical structures of the cyanide degrading enzymes from *Gloeocercospora sorghi* and *Bacillus pumilus* providing insights into nitrilase quaternary interactions. MSc Thesis, University of Cape Town, Rondebosch, South Africa.

Semba, S., Han, S.Y., Qin, H.R., McCorkell, K.A., Iliopoulos, D., Pekarsky, Y., Druck, T., Trapasso, F., Croce, C.M. and Huebner, K. (2006) Biological functions of mammalian Nit1, the counterpart of the invertebrate NitFhit Rosetta stone protein, a possible tumor suppressor. *J Biol Chem* **281**, 28244–28253.

Sewell, B.T., Thuku, R.N., Zhang, X. and Benedik, M.J. (2005) The oligomeric structure of nitrilases: the effect of mutating interfacial residues on activity. *Ann NY Acad Sci* **1056**, 153–159.

Sewell, B.T., Berman, M.N., Meyers, P.R., Jandhyala, D. and Benedik, M.J. (2003) The cyanide degrading nitrilase from *Pseudomonas stutzeri* AK61 is a two-fold symmetric, 14-subunit spiral. *Structure* **11**, 1–20.

Shaikh, T.R., Hegerl, R. and Frank, J. (2003) An approach to examining model dependence in EM reconstructions using cross-validation. *J Struct Biol* **142**, 301–310.

Sharma, M., Sharma, N. and Bhalla, T. (2009) Amidases: versatile enzymes in nature. *Rev Environ Sci Biotechnol* **8**, 343–366.

Sosedov, O., Baum. S., Bürger S., Kathrin Matzer, K., Kiziak, C. and Andreas Stolz, A. (2010) Construction and Application of Variants of the *Pseudomonas fluorescens* EBC191 Arylacetonitrilase for Increased Production of Acids or Amides*. Appl Environ Microbiol* **76***,* 3668–3674.

Sterner, R. and Liebl W. (2001) Thermophilic adaptation of proteins. *Crit Rev Biochem Mol Biol* **36**, 39–106.

Tauber, M., Cavaco-Paulo, A., Robra, K. and Gubitz, G. (2000) Nitrile hydratase and amidase from *Rhodococcus rhodochrous* hydrolyse acrylic fibers and granular polyacrylonitriles. *Appl Environ Microbiol* **66**, 1634–1638.

Thuku R.N., Brady D., Benedik M.J. and Sewell B.T. (2009) Microbial nitrilases: versatile, spiral forming, industrial enzymes. *J Appl Microbiol* **106**, 703–727.

Thuku, R.N., Weber, B.W., Varsani, A. and Sewell B.T. (2007) Post-translational cleavage of recombinantly expressed nitrilase from *Rhodococcus rhodochrous* J1 yields a stable, active helical form. *FEBS J* **274**, 2099-2108.

Tokunaga, M., Tokunaga, H. and Wu, H.C. (1982) Post-translational modification and processing of *Escherichia coli* prolipoprotein *in vitro. Proc Natl Acad Sci USA* **79**, 2255–2259.

Trachtenberg, S., DeRosier, D.J., Zemlin, F. and Beckmann, E. (1998) Non-helical perturbations of the flagellar filament: *Salmonella typhimurium* SJW117 at 9.6 A resolution. *J Mol Biol* **276**, 759–773.

Unser, M., Trus, B.L. and Steven, A.C. (1987) A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy* **23**, 39–51.

Unser, M., Sorzano, C.O., Thevenaz, P., Jonic, S., El-Bez, C., De Carlo, S., Conway, J.F. and Trus, B.L. (2005) Spectral signal-to-noise ratio and resolution assessment of 3D reconstructions. *J Struct Biol* **149**, 243–255.

van Heel, M., Keegstra, W., Schutter, W.G. and van Bruggen, E.F.J. (1983). Arthropod hemocyanin studied by image analysis. The EMBO Workshop on Invertebrate Respiratory Proteins Leeds, Structure Life Chemistry Reports. pp.69–73.

Van Heel, M. (1987) Similarity measures between images. *Ultramicroscopy* **21**, 95–100.

Van Wyk, J.C. (2008) The relationship between structure and thermostability of a nitrile hydratase from *Geobacillus pallidus* RAPc8. PhD Thesis, University of Western Cape, Cape Town, South Africa

Varshavsky, A. (1997) The N-end rule pathway of protein degradation. *Genes Cells* **2**, 13–28.

Vejvoda, V., Kaplan, O., Bezouska, K., Pompach, P., Sulc, M., Cantarella, M., Benada, O., Uhnakova, B. *et al.* (2008) Purification and characterization of a nitrilase from *Fusarium solani* O1. *J Mol Catal B Enzym* **50**, 99–106.

Vieille, C. and Zeikus, G. J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol R* **65**, 1–43.

Vorwerk, S., Biernacki, S., Hillebrand, H., Janzik, I., Muller, A., Weiler, E.W. 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 L., Watermeyer J.M., Mulelu A.E., Sewell B.T., Benedik M.J., 2011. Engineering pH-tolerant mutants of a cyanide dihydratase. *Appl Microbiol Biotechnol* **94**, 131-140.

Wang, W.C., Hsu, W.H., Chien, F.T. and Chen, C.Y. (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. *J Mol Biol* **306**, 251–261.

Woodward, J.D., Weber, B.W., Scheffer, M.P., Benedik, M.J., Hoenger, A. and Sewell, B.T. (2008) Helical structure of unidirectionally shadowed metal replicas of cyanide hydratase from *Gloeocercospora sorghi*. *J Struct Biol* **161**, 111–119.

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

Yamada, H. and Nagasawa, T. (1994) Process for biological production of amides with *R. rhodochrous* J1. US Patent 5334519.

Yamada, H. and Kobayashi, M. (1996) Nitrile hydratase and its application to industrial production of acrylamide. *Biosci Biotechnol Biochem* **60**, 1391–1400.

Yang, S., Yu, X., Galkin, V.E. and Egelman, E.H. (2003) Issues of resolution and polymorphism in single-particle reconstruction. *J Struct Biol* **144**, 162–171.

Young, C.A. and Jordan, T.S. (1995) Cyanide remediation: current and past technologies. Proceedings 10th Annual Conference on Hazard Waste Research, Manhattan, pp 104–128.

Zalkin, H. and Smith, J.L. (1998) Enzymes utilizing glutamine as an amide donor. *Adv Enzymol RAMB* **72**, 87–144.

Zhang, P., Toyoshima, C., Yonekura, K., Green, N.M. and Stokes, D.L. (1998) Structure of the calcium pump from sarcoplasmic reticulum at 8-A resolution. *Nature* **392**, 835–839.