2-1-2014

Cloning, expression and preliminary characterization of a novel nitrilase from the Red Sea, Atlantis II Deep brine pool using a metagenomic approach

Sarah Ali Sonbol

THE AMERICAN UNIVERSITY IN CAIRO
الجامعة الأمريكية بالقاهرة

School of Sciences and Engineering

Cloning, expression and preliminary characterization of a novel nitrilase from the Red Sea, Atlantis II Deep brine pool using a metagenomic approach

A Thesis Submitted to

The Biotechnology Master's Program

In partial fulfillment of the requirements for

the degree of Master of Science

By: Sarah Ali Ahmed Sonbol

Under the supervision of:

Dr. Ahmed Moustafa

And

Dr. Ari José Scattone Ferreira

December / 2013

The American University in Cairo

Cloning, expression and preliminary characterization of a novel nitrilase from the Red Sea, Atlantis II Deep brine pool using a metagenomic approach

A Thesis Submitted by
Sarah Ali Ahmed Sonbol

To the Biotechnology Graduate Program
December / 2013

In partial fulfillment of the requirements for
The degree of Master of Science
Has been approved by

Thesis Committee Supervisor/Chair
_____
Affiliation _____

Thesis Committee Reader/Examiner
_____
Affiliation _____

Thesis Committee Reader/Examiner
_____
Affiliation _____

Thesis Committee Reader/External Examiner
_____
Affiliation_____

_____ _____ _____
_____
Dept. Chair/Director Date Dean Date

DEDICATION

To my wonderful Mom and Dad who taught me to have faith, love science and never lose hope.

To Ahmad, the martyr, the marvellous and the best brother ever, who always believed that I can, taught me to live and die for a proper cause and to dedicate my life to God.

To Rokaya, Asmaa, Huda and Basma, my amazing sisters, who are always beside me.

To Yahia and Yosouf, my nephews, who brought great smiles and laughter to our life and to whom I wish to be as great as their uncle.

ACKNOWLEDGEMENTS

This work could have never been done without the help, support and guidance of Dr. Ari Ferreira. I sincerely thank him for all he has done.

Special thanks to Dr. Ahmed Moustafa, the director of the Biotechnology Graduate Program at AUC; I am extremely indebted to him for his extreme support and advising.

I am also grateful to all my lab-mates who supported me during my work and to Dr. Ari's team (Aya, Hadeel, Rama, Nahla and Salma).

Words can never describe my gratitude towards Nahla Hussein for her sincere help and useful remarks.

I owe Mariam RizkAllah and Mustafa Adel special thanks for their help in the bioinformatics part and Bothaina El Laimoni for her help in the lab work.

Great thanks to all professors who taught me during my courses (Dr. Hamza El Dorry, Dr. Walid Fouad, Dr. Ahmed Sayed, Dr. Asma Amleh, Dr. Ahmed Mustafa, Dr. Edwin Rivera, Dr. Rania Siam & Dr. Ari Ferreira).

I should also acknowledge Seung Hye Hong from Dr. Oh lab, Konkuk University, Korea, for providing me with a recombinant expression vector, with a nitrilase gene, as a positive control for my work.

I have to acknowledge the spring 2010 expedition team, especially members from AUC (Dr. Rania Siam, Dr. Mohamed Ghazy and Mr. Amgad Ouf) and all lab members who shared in collecting the samples, extracting the DNA and creating the database that I used in my work.

Finally, thanks to King Abdullah University of Science and Technology (KAUST) for funding my work and granting me a fellowship and thanks to the AUC for granting me a Merit fellowship.

ABSTRACT

The American University in Cairo

Cloning, expression and preliminary characterization of a novel nitrilase from the Red Sea, Atlantis II Deep brine pool using a metagenomic approach

Sarah Ali Ahmed Sonbol

Supervisor: Dr. Ahmed Moustafa and Dr. Ari José Scattone Ferreira

Summary:

With the abundance of hazardous nitrile compounds in nature and with the extensive use of them in industries of fine chemicals, the attention towards nitrilases has profoundly increased as biocatalysts in different industries and in bioremediation. The use of nitrilases, which hydrolyze nitriles into carboxylic acids and ammonia, was introduced as a biological method superior to conventional chemical means.

In contrast to conventional microbiological techniques using cultured microorganisms, and with the loads of genomic sequences available in the databases, identification of novel nitrilases using metagenomic approaches seems more promising.

In this study, a nitrilase was isolated from the microbial metagenomic DNA obtained from the Lower Convective Layer (LCL) of Atlantis II Deep Brine Pool in the Red Sea. Sequences of putative nitrilases were retrieved from the LCL database, followed by PCR amplification and gene synthesis of the chosen sequences. Amplified genes were cloned into pET SUMO expression vector, whereas a synthesized gene, with optimized codons for *E. coli,* was cloned into pET-28b+ vector flanked by *Sac*I and *Hind*III restriction sites. The recombinant proteins were then expressed in *Escherichia coli* BL21 (DE3) cells. The His-tagged protein (Nitra-S), from the recombinant pET-28b+ transformed cells, was successfully purified using Ni-NTA affinity chromatography columns under native conditions. Preliminary studies on the induced cells, crude extracts and the purified enzyme showed that Nitra-S exhibited nitrilase activity towards dinitriles (succinonitrile and gluteronitrile) rather than other nitrile classes. The used quantitative activity assay was based on measuring the amounts of released ammonia upon the action of nitrilases on their nitrile substrates. Full characterization of the purified nitrilase is still to be done; however, the isolation of this protein from the LCL with its unique characteristics, increase the odds towards finding exceptional properties of the newly identified nitrilase.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIR: Aminoimidazole Ribonucleotide

BCA: Bicinchoninic Acid

BLAST: Basic Local Alignment Search Tool

BSA: Bovine Serum Albumin

CNS: Central Nervous System

CTD: Conductivity, Temperature and Depth

DNA: Deoxyribonucleic Acid

DTT: Dithiothreitol

FAD: Flavin Adenine Dinucleotide

HMM: Hidden Markov Model

HPLC: High Performance Liquid Chromatography

IPTG: Isopropyl β-D-1-Thiogalactopyranoside

KAUST: King Abdullah University of Science and Technology

LB: Luria-Bertani

LCL: Lower Convective Layer

NCBI: National Center for Biotechnology Information

NTA: Nitrilotriacetic Acid

OD: Optical Density

ORF: Open Reading Frame

PBS: Phosphate Buffered Saline

PCR: Polymerase Chain Reaction

pI: Isoelectric point

PMSF: Phenylmethylsulfonyl fluoride

PVA: Poly Vinyl Alcohol

RBS: Ribosomal Binding Site

SA: Sodium Alginate

SAM: S-Adenosyl Methionine

SDS-PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SOD: Superoxide dismutase

Ta: Annealing temperature

UCL: Upper Convective Layer

X-gal : 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER 1: REVIEW OF LITERATURE AND STUDY OBJECTIVES

1. Introduction:

In this biotechnological era, it is hard to ignore the plethora of genomic data obtained from known and unknown organisms. The availability of different resources renders the task of identification of novel enzymes easier than before. Nitrilases have not received much attention except recently, after the realization of their potential uses in industry and bioremediation processes. The number of characterized nitrilases from cultured microorganisms is small, which makes it important to shift towards the uncultured world of microorganisms for the identification and isolation of new nitrilases. Exploring genes from extreme environments such as Atlantis II Deep brine pool in the Red Sea, and identification of the potential uses of the newly identified genes might be a step forward towards better understanding of the unique world of microbes and for better biotechnological future.

2. Nitrilases and their nitrile substrates:

2.1. Nitriles:

Nitriles are organic compounds that contain a cyano (CN) functional group. These compounds are ubiquitous in the environment. Different nitriles are produced by certain plants and animals (1), in addition to some bacterial, fungal and algal species (2). Cyanoglycosides, cyanolipids, aminonitriles, cyanohydrin and phenylacetonitrile are examples of naturally occurring nitrile compounds(1)(3).On the other hand, some nitriles are produced as industrial wastes or as valuable industrial products (1).

The high toxicity of most nitriles has raised awareness towards the nitrile waste products discharged from different industries. Cyanide containing compounds can cause gastric disturbances, respiratory distress, skeletal deformities, convulsions and other CNS symptoms. In addition, CN^- ions released from nitriles could be lethal by binding to the cytochrome-c-oxidase in the respiratory chain resulting in inactivation of respiration (1). In spite of the hazards that cyanide containing compounds could cause to humans and ecosystem, they are considered of great value from an industrial point of view (1)(2)(4).Various nitrile compounds are synthesized as herbicides, plastics, polymers and solvents (5). Moreover, nitriles are used as starting materials or intermediates in fine chemicals, pharmaceuticals and other industries(1).

2.2. Nitrilases and related enzymes:

Nitrilases (EC 3.5.5.1) are nitrile hydrolyzing enzymes that hydrolyze nitrile compounds (R-CN) directly into their corresponding carboxylic acids (R-COOH) and ammonia (NH_3) (1)(4)(6). In this aspect, they differ from nitrile hydratases (NHases) that convert nitriles into amides, which could then be converted into ammonia by the action of amidases (1)(4). Nitrilases are a subgroup of the carbon-nitrogen (C-N) hydrolase superfamily. Members of this superfamily are classified into 13 subgroups that are involved in non-peptide C-N hydrolysis; thus, amidases, carbamylases, amidohydrolases and N-acetyltransferases are members of this superfamily as well (3). This superfamily was first classified by Brenner (2002) as the nitrilase superfamily, even though one branch only exhibited true nitrilase function (1). All members of the C-N hydrolase superfamily are characterized by the presence of a catalytic triad of glutamate-lysine-cysteine and α-β-β-α sandwich fold (3).Data concerning structure and active sites were mainly derived from two members of the superfamily. However, a thermoactive nitrilase was crystallized recently by Raczynska *et al* (2011) confirming the previous findings (7).

According to their substrate specificity, nitrilases are classified into three major classes: aromatic, aliphatic and arylacetonitrile nitrilases. Nevertheless, some nitrilases have broad substrate specificity; thus accepting nitriles from different classes as substrates (1)(4).

2.3. Mechanism of action of nitrilases:

As the number of known structures of nitrilase enzymes is relatively small, a detailed mechanism describing the catalytic function of nitrilases is still unavailable. However, several mechanisms were suggested to explain the activity of nitrilases. Putting into consideration that some nitrilases have minor nitrile hydratase and/or amidase activity beside their main function as nitrilases, some studies suggested that in a process that requires the addition of two molecules of water, the cysteine residue initiates a nucleophilic attack on the carbon atom of the cyano group, forming a tetrahedral intermediate through an enzyme-thioimidate pathway. With some nitrilases, a small percentage of the tetrahedral intermediate can breakdown forming an amide, while the majority is transformed into an acylenzyme complex when the ammonia group leaves being a better leaving group. Finally, carboxylic acid can be formed after the addition of another water molecule (figure1) (1). In this process, the lysine residue could be important for electrostatic stabilization of the tetrahedral intermediate, donating protons to the substrate, while the glutamate might act as a catalytic base for the hydrolytic water (7). Mutations at any residue of the catalytic triad lead to inactivation of the nitrilase function. Several studies have also confirmed that thiol

complexing agents inactivate nitrilases as they could bind to the catalytic cysteine thiol group, whereas, thiol reducing agents enhance the activity (8). Thus, using thiol reducing agents such as Dithiothreitol (DTT) or β-mercaptoethanol would be a necessity during the reactions.

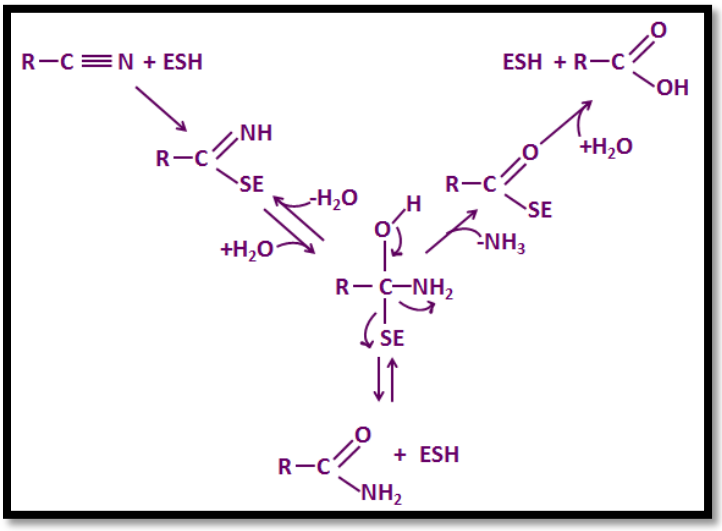

Figure 1: Reaction mechanism of nitrilase catalysis.

2.4. Nitrilases applications and advantages:

As nitrile compounds are extensively used in several industries and in the synthesis of a variety of fine chemicals and pharmaceutical compounds, interest towards the use of nitrilases has risen as a better alternative to conventional chemical methods. Chemical processing of nitrile intermediates necessitates the use of high temperatures, strong acidic or basic conditions and in some cases, the use of protective groups. Moreover, harmful byproducts such as toxic HCN could be produced during these processes (1)(9)(10)(11).

On the other hand, the use of nitrilases offers a number of advantages. First, reactions could be carried in mild reaction conditions. Besides, no toxic byproducts are to be formed, which means a reduction in industrial pollution (1)(9)(10)(11). Furthermore, some nitrilases are stereo- and/or regioselective, allowing the synthesis of specific isomers that could not be synthesized using traditional chemical means. One of the interesting findings, which makes nitrilases superior to chemical methods, is the ability of some nitrilases to hydrolyze a single cyano-group in dinitriles or polynitriles producing cyanocarboxylic acids that could be used in different industries (1) (11). Nitrilases isolated from *Acidovorax facilis* 72W and *Rhodococcus rhodochrous* K22 are examples of these highly selective nitrilases, which specifically hydrolyze single cyano-group in aliphatic dinitriles. According to their substrate specificity, nitrilases can accept substrates of

different complexities and can hydrolyze nitriles without affecting other labile functional groups (11).

Nitrilases are extremely useful catalysts in different industries. They were successfully utilized in manufacturing different types of polymers such as nylon-6, and other chemicals such as nicotinic and acrylic acids. They were also used in pharmaceutical industries as in the manufacturing of (*S*)-ibuprofen, a widely used non-steroidal anti-inflammatory drug. In a recent study done by Liu *et al* (2012), iminodiacetic acid was synthesized from immobilized nitrilase-producing recombinant *E. coli* BL21 (DE3) cells that were entrapped in polyvinyl alcohol/sodium alginate (PVA/SA) copolymer (12). Moreover, nitrilases were found to be of extreme importance in bioremediation processes as they were utilized in detoxification of cyanide containing wastes and in nitrile herbicides degradation (1)(4). The nitrilase isolated from *Klebsiella ozaenae* was found to effectively degrade the nitrile-herbicide bromoxynil (13). Bioremediation using nitrilases is faster, more economical and more efficient than other physical and chemical methods(1) (2) (4).

2.5. Distribution and physiological role of nitrilases:

The presence of nitrilases is infrequent in nature. They are relatively rare in plants and fungi; however, their frequency is considered higher in the bacterial kingdom. Nitrilases were found in bacterial species from different genera such as *Pseudomonas*, *Klebsiella*, *Nocardia*, and *Rhodococcus*.

Plant nitrilases were found to have a role in the production of Indole-3-acetic acid, which plays a role in the elongation and development of plants (14) (1). In addition, it was reported that these nitrilases are important in cyanide detoxification, as cyanide can be produced in plants as a by-product during the synthesis of ethylene hormone or by the degradation of cyanogenic glycosides (14).

The role of nitrilases in bacterial metabolism is still unclear; however, some bacterial strains that possess nitrilases are able to utilize nitriles as a sole source of carbon and nitrogen. Other studies suggest roles in complex pathways which direct the production and degradation of cyanogenic compounds(1)(9). Nitrilases are generally inducible enzymes(1)(15) that might have an important role in defense mechanisms and detoxification of xenobiotics or cyanogenic compounds produced by plants or other microorganisms(16).

2.6. Nitrilase gene family neighborhood:

In bacterial genomes, genes are usually found in clusters such as operons. Generally, genes that are found in one cluster are involved in a common metabolic pathway. In a phylogenetic study done by Robertson *et al* (2004), nitrilase sequences from completed genomes and environmental samples were studied and clustered into 6 different clades (3). Studying the neighborhood of one of the major clades (subfamily 1) by Podar *et al* (2005) revealed that these nitrilase sequences are part of a conserved gene cluster (Nit1C) composed of seven genes. The gene cluster is composed of [1] hypothetical protein,[2] nitrilase, [3] radical SAM superfamily member, [4] acetyltransferase, [5] AIR synthase, [6] hypothetical protein, and finally [7] a predicted flavoprotein which is either located at the end of the cluster or at the beginning but on the opposite strand. Slight variations in the arrangement of these genes in the cluster were observed which correlate with the taxonomic groups from which these genes were isolated (cyanobacteria, verrucomicrobia, beta and gamma-proteobacteria). The presence of Nit1C gene cluster in the virulence plasmid pLVPK of *Klebsiella pneumoniae* suggests a possible transfer of these genes to other bacterial species. For nitrilases that belong to other clades, conservation of the gene sequences in the neighborhood was found to be less prominent than that in Nit1C (16).

2.7. Nitrilase subunit size and molecular structure:

Sequencing and SDS-PAGE experiments studying different nitrilases have shown that most nitrilases have a molecular mass that range between 30-45 kDa (1)(17). Robertson *et al* (2004) have studied 137 nitrilase sequences reporting that sequences were between 304 and 385 amino acids in length. These variations reflected the variation in the C-termini lengths rather than the N-termini (3).

Different studies using gel filtration chromatography, light scattering or electron microscopy revealed that functional nitrilases form aggregates ranging from 6 to 26 subunits (1)(17). However, some were also found as smaller aggregates: dimers and trimers, as in nitrilases from *Klebsiella ozaenae* and *Pseudomonas fluorescens* DSM 7155, and in rare cases as monomers such as in *Rhodococcus rhodochrous* PA-34 (17). The activation and association of these subunits were found to be dependent on temperature, pH, substrate availability, concentrations of salt and organic solvents and the enzyme concentrations (1)(15).

2.8. Limitations of the utilization of nitrilases:

In spite of their presence as potential biocatalysts in the synthesis of a great number of fine chemicals; low activity, low stability and unsuitable substrate specificity of many available

nitrilases have limited their use (18). Substrate tolerance is another limitation as nitriles in high concentrations are toxic to nitrilases, limiting their activity. To date the number of characterized nitrilases is considered small and those already used in different applications are much smaller in number (19).

Overcoming these limitations would open the locked gates in fine chemicals industry and other potential applications. The discovery of novel nitrilases with different substrate specificities, higher activities and stabilities could make myriads of applications possible. Interestingly, few studies have reported the isolation of thermophilic nitrilases that were found to be more stable than their mesophilic homologues(9)(15). This indicates the possibility of finding robust nitrilases that can withstand harsh conditions.

2.9. Determination of nitrilase activity:

Different methods were developed for determination of nitrilase activity by either measuring the amount of produced carboxylic acid or ammonia. HPLC and infrared spectroscopy could be used for direct estimation of the produced acid; however, these methods require long preparation time and special instruments that might not be available (1).

A rapid, colorimetric method for the screening of nitrilase-producing strains were developed by Banerjee *et al* (2003); this method is based on the drop in pH due to acid formation, resulting in a change in the color of an acid-base indicator. The choice of the indicator is crucial and the difference between the pKa of the indicator and buffer should be within 0.1 units to ensure that the relative amount of protonated indicator to protonated buffer is constant, thus any change in the indicator color would reflect the number of protons released. This method is qualitative due to the presence of cells which interfere with spectrophotometric assays(20).

Many other methods were developed, measuring the amount of produced ammonia. Ninhydrin and Nesslerization methods are no longer used because of their low sensitivity, large sample volumes needed and the interference of some organic solvents and inorganic ions. Berthelot assay is a precise, simple colorimetric assay which depends on the reaction of ammonia with sodium phenate and hypochlorite to produce a blue-colored complex, in presence of nitroprusside as a catalyst (21). Nevertheless, the use of corrosive reagents is considered a serious drawback of this method (1).

A more sensitive, fluorometric assay method can be used which is based on the reaction

of ammonia with *o*-phthaldialdehyde and β-mercaptoethanol to form an isoindole derivative that emits fluorescence at 467nm when exited at 412nm (1)(22). The formed complex is colored, so it can also be measured spectrophotometrically at 410 with reliable results (23).

Finally, a recent metal ion-based method was developed by Yazbeck *et al* (2006) for the screening of nitrilases. This method is also based on the amount of released ammonia, which can form a complex with cobalt ions changing the color from light pink to yellow which can be measured spectrophotometrically at 375nm. Unlike other fluorometric and colorimetric assays, the color develops instantly in this assay (24).

3. Metagenomic approaches in the search of novel nitrilases:

Conventional screening methods by culturing different microorganisms on selective media with nitrile substrates were the main used techniques to isolate and characterize nitrilases (19). In spite of being tedious and time consuming, different nitrilases were identified using this strategy. Some of these nitrilases have already been utilized in different industries as mentioned previously. However, as cultured and characterized microorganisms represent less than 0.1% of the microbial world (3), the trend towards studying the unfathomable world of uncultured microorganisms seems promising. Metagenomic studies have provided us with a plethora of microbial genomic sequences from which novel enzymes, with promising potential in different industries, could be characterized. Extracting microbial genomes from different environments increases the possibility of the identification of enzymes that can adapt to these environments.

Microbial metagenomic studies are based on studying the genomes of the uncultured microbial world by extracting the DNA directly from the environment, bypassing the steps of isolation and cultivation of individual microorganisms. Metagenomic libraries of environmental DNA sequences can be constructed after mechanical shearing or enzymatic digestion of the DNA. Gene-identification from these libraries is done using two different approaches, either functional-based or sequence-based screening. In the former approach, the library clones are screened for specific protein function or enzyme activity (19). Such strategy was used by Robertson *et al* (3) and Bayer *et al* (11) to identify novel nitrilases. In sequence-based screening, determination of putative nitrilases is based on sequence homology with available sequences in the existing databases. This strategy combines bioinformatics, sequence analysis, sequencing techniques, PCR amplification and/or DNA synthesis (19). As the number of available sequences in the databases over-exceeds the number of biochemically tested ones, determination of the exact functions of the majority of these sequences is still lacking. Most nitrilase sequences, as well as enzymes with different functions, retrieved from available databases are annotated based

on sequence similarities with their homologs. These homologs could be distantly related to experimentally tested enzymes, which sometimes make the available annotations misleading. Different softwares such as Pfam can determine known patterns in gene sequences and thus assign these sequences to groups with major functions. Nevertheless, specific functions and substrate specificity of different enzymes can only be determined by experimental testing (25). In a study done by Seffernick *et al* (2009), several nitrilases were identified based on sequence homology followed by expression of identified sequences in suitable bacterial hosts and characterization of purified nitrilases (25).

4. Red Sea Atlantis II Deep brine pool:

Studies on the Red Sea, which is located between Africa and the Arabian Peninsula, revealed the presence of about 25 brine pools in its depth (26). Those are geothermal salt-enriched waters that are found in topographic depressions in the seafloor of the central and northern Red Sea (27). High density of these hypersaline waters stabilizes them and makes them poorly mixed with the above water (28).

The largest and one of the most interesting brine pools in the Red Sea is Atlantis II Deep. It is 2,194 m deep from the seawater surface (28), and located at latitude 21° 20’ N and longitude 38° 04’ (29) near the central rift of the Red Sea. This brine pool is a unique ecosystem which is characterized by elevated temperatures, salinity and pressure (26). Studies at different time intervals revealed that Atlantis II Deep is still dynamic and considered a part of an active hydrothermal system, as its temperatures have elevated since 1966 until recent years from 56.5°C (30) to 68°C (26).

Atlantis II Deep is characterized by four brine layers, which differ in a steep gradient in their temperatures, salinity and O_2 content. Those are three upper convective layers (UCL1, UCL2, UCL3) and one lower convective layer (LCL) which shows the highest temperature (68°C) and salinity (250 part per thousand or 7.5 times that of the normal sea water)(26)(27).

The anoxic conditions, elevated temperature, salinity and pressure of the LCL (26), increase the likelihood of finding unique novel enzymes that are produced by extremophiles which inhabit the LCL. Nitrilases are not an exception; nitrilases with higher stability and different substrate specificities could be discovered in the LCL. Extracting microbial DNA of the LCL inhabitants allows further investigations and retrieval of valuable genes from the extracted environmental DNA.

5. Study objectives and experimental design:

The first objective of this study was to identify a novel nitrilase from the LCL of Atlantis II Deep brine pool in the Red Sea, using a sequence-based metagenomic approach. To achieve this goal, extensive bioinformatics and computational work was done on the LCL microbial sequences database. Confirmation of the presence of the identified putative nitrilase gene sequences required amplification from environmental DNA using specific primers. The obtained amplicons were then sequenced to verify the sequences retrieved from the database.

The second objective was to express the identified nitrilase. In accordance with this aim, the amplified sequences were cloned into a proper expression vector and the same sequence was synthesized with optimization of codon usage before being cloned into another expression vector. The used expression vectors were chosen to express the desired sequence with either a C or an N-terminal His-tag to allow easier purification of the protein in later steps. The recombinant expression vectors were transformed into *Escherichia coli* BL21 (DE3) host cells and expression was induced by adding IPTG to the medium followed by visualization of the induced protein on SDS-PAGE gels.

The third objective was the purification of the expressed protein. This was done using Ni affinity chromatography columns, where the His-tagged protein can bind to the Ni^{+} ions, followed by elution of the bound protein with a suitable elution buffer that contains high concentrations of imidazole.

The fourth objective was activity determination of the purified potential nitrilase. Both qualitative and quantitative assays were done. The qualitative colorimetric assay was dependant on the change of color of the acid base indicator, bromothymol blue, upon the decrease in pH due to the release of acid. In contrast, in the quantitative assay, the amount of released ammonia was measured, based on its reaction with a buffered *o*-phthaldialdehyde/β-mercaptoethanol reagent, forming a colored complex that can be quantified spectrophotometrically.

The final objective was to preliminary characterize the purified nitrilase in terms of thermal stability and halotolerance. This was done by incubating the enzyme for a definite time interval at different temperatures and salt concentrations before measuring the residual activity of the enzyme.

CHAPTER 2: MATERIALS AND METHODS

1. Sampling:

On board the R/V Aegaeo research vessel, water samples were obtained during the King Abdullah University of Science and Technology (KAUST) expedition on the spring of 2010. Samples were taken from the four Atlantis II Deep layers including the LCL, in a *rosette* of Niskin bottles, supplied with CTD (Conductivity, Temperature and Depth) meter.

2. DNA extraction and sequencing:

Water samples were filtered through mixed cellulose ester filters of pore sizes of 3µm, 0.8µm and 0.1µm sequentially and the prokaryotic DNA was extracted from the cells retained on the 0.1µm filter. Extracted DNA was sequenced using 454 shotgun sequencing approach with GS FLX Titanium pyrosequencer (454 Life Sciences), creating an enormous database with raw sequences for all obtained reads. The expedition team and other lab-mates performed these steps.

3. Computational analysis:

A sequence-based metagenomic approach was used for finding new genes in the LCL database. The Pfam (PF00795) for the desired functional domain (Carbon-Nitrogen hydrolase) was obtained from Pfam 26.0 (31) and used to extract reads with this domain from the LCL database using an HMM scan. The reads were assembled and visualized using the Phred-Phrap-Consed tool (32)(33)(34)(35)(36) followed by BLASTing the consensus sequence of the contig, with the greatest number of reads, against a database of assembled LCL contigs with annotated open reading frames (ORF's). This step was done to identify the ORF of the desired gene, its regulatory elements and neighboring ORFs. The promoter region was then identified using the bprom tool (37) and the Shine Dalgarno sequence was detected manually. The sequence of the ORF of interest was BLASTed (BLASTx) (38) against the non redundant protein database in the National Center for Biotechnology Information (NCBI) for identification of the new nitrilase.

4. Amplification of the complete gene of interest:

Forward primer Nitr1_F (GATACCTGGACGGTTTAAGG) and reverse primer Nitr1_R (GAAAGCTGATGTTCGCAAT) were designed using Primer3Plus software (39) to amplify the desired sequence from the environmental DNA using an annealing temperature (Ta) of 54°C. The amplified sequence (A:1336bp fragment) included the desired ORF plus 76 bp upstream and

243 bp downstream. Another primer Nitr6_F (CAGATCGAAGTCACCATCC) with Nitr1_R were used to amplify part of the gene (D: 665bp fragment) from the environmental DNA using a Ta of 54°C. The amplified sequences were then extracted from the agarose gel using QIAquick® gel extraction kit (Qiagen).

5. Sequencing of the amplified sequence:

TA cloning was used to clone the amplicons into pGEM®-T Easy cloning vector (Promega). Recombinant plasmids were transformed into *E. coli* Top10 strains, which were grown on LB agar containing 0.5mM IPTG, 40ug/ml X-gal and 100ug/ml ampicillin. Colony PCR was performed for several positive white clones using the same primers used in amplification to confirm the presence of the amplified sequences. Plasmids were extracted from cultures obtained from colonies with positive results using PureYield ™ Plasmid Miniprep System (Promega). Extracted plasmids were then sequenced in ABI 3730xl DNA Analyzer using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems®) with the vector T7 forward and Sp6 reverse primers in addition to Nitr1_F, Nitr1_R and Nitr6_F primers. Direct sequencing of the purified amplicons using the amplicon primers was done as well. Sequencing results were visualized using BioEdit Sequence Alignment Editor version 7.1.3.0 (40).

6. Amplification and sequencing of the regulatory elements and upstream parts of the operon:

Several primers were designed to sequence the upstream part of the operon containing the ORF of interest with its regulatory elements. Five different PCR reactions were done using the following pairs of forward and reverse primers, all with a Ta of 54°C: Nitr3_F (GTCTCAGGCTACCACAATGA) and Nitr3_R (GTGATGAGCGACATATCCAG), Nitr3_F and Nitr4_R (TCGGTATCAAAGATGAGCTG), Nitr7_F (CAACAAGCAGATCACCAAGT) and Nitr3_R, Nitr4_F (ATCAGGTAAGCGGGTTTG) and Nitr4_R, and finally Nitr4_F and Nitr5_R (CTGCTTGTTGAAGTTGAGATG). Positions of used primers in reference to the assembled contig (contig00026) are shown in figure2. The amplified sequences were extracted from the agarose gel using QIAquick® gel extraction kit. Direct sequencing was done for all purified amplicons using the same primers used in amplification.

7. Amplification of the putative nitrilase gene using pET SUMO and pET-28b+ expression vectors:

In a PCR reaction at a Ta of 58°C, the expression primers Nitr_exp1_F (CGCGGATCCGGCATCACCATGTCACAGAA) and Nitr_exp2_R (CCCAAGCTTTCGGTCCGAAGCTCAGTTAC), with *Bam*HI and *Hind*III restriction sites, respectively, were used for the amplification of the desired ORF. Positions of the primers are shown in figure 2.

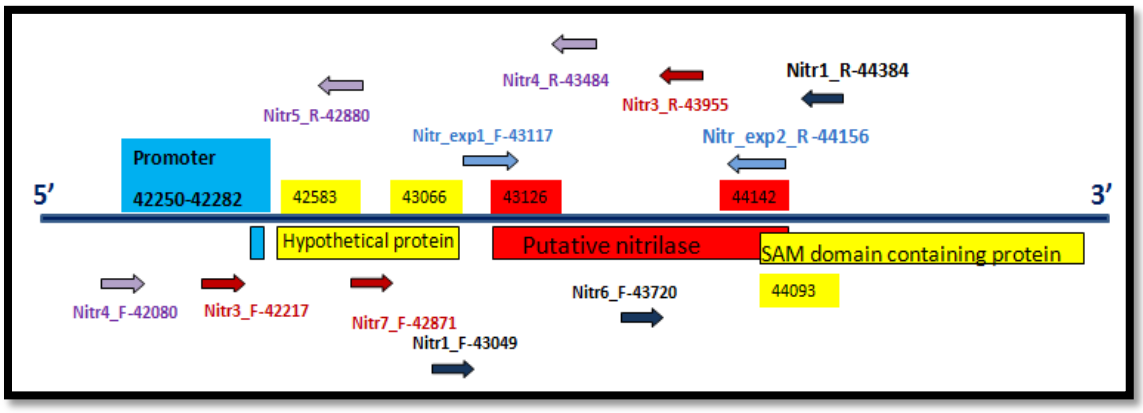

Figure 2: Schematic diagram showing the positions of all primers used in contig00026. Rectangles represent the putative nitrilase ORF, surrounding ORFs and the promoter region with numbers indicating their positions in the contig. Arrows represent forward and reverse primers. Numbers next to each primer indicate the position of the primer's first base in the contig.

8. Expression of the gene of interest using pET SUMO expression vector:

The amplified sequences obtained from the PCR reaction with the expression primers (Nitr_exp1_F and Nitr_exp2_R) were extracted from the agarose gel using QIAquick® gel extraction kit (Qiagen). The purified amplicons were then cloned into the pET SUMO expression vector, which has an N-terminal His-tag, using pET SUMO TA Cloning® Reagents provided with the Champion™ pET SUMO Protein Expression System (Invitrogen™). The recombinant plasmids were transformed into *E.coli* Top10 by Electroporation. Transformed cells were allowed to grow on LB agar plates with kanamycin (50ug/ml). For determination of the clones with inserts in the correct direction, a number of colonies were subjected to colony PCR using the vector Sumo_F (AGATTCTTGTACGACGGTATTAG) primer with Nitr3_R primer and Nitr6_F with the vector T7_R (TAGTTATTGCTCAGCGGTGG) primer using a Ta of 55°C. Plasmids were

extracted from cultures obtained from a colony with positive results using PureYield ™ Plasmid Miniprep System (Promega) followed by their transformation into *E.coli* BL21 (DE3) for expression and sequencing. Ten millilitres of LB broth with kanamycin (50ug/ml) were inoculated with a transformed colony and incubated for an overnight at 37°C and 225 rpm. A proper volume of the overnight culture was added to a fresh LB broth with kanamycin. Cultures were allowed to grow at 37°C and 225 rpm to an OD_{600} of about 0.6 and then induction of expression started at 37°C with different IPTG concentrations (0.1-0.8mM) and for different time intervals. Cultures were centrifuged and pellets were preserved at -80°C for analysis. All cell lysates were analyzed using Coumassie-stained 12% SDS-PAGE according to the method of Laemmli (1970) (41). Any further analysis was done using pellets obtained from 25 ml cultures, which were induced for 2 hours using 0.1mM IPTG.

9. Expression using a codon-optimized putative nitrilase gene in pET-28b+ expression vector:

For optimized expression into *E. coli* BL21 (DE3), a synthesized gene (based on the sequence amplified from the metagenomic DNA) with optimized codons for *E.coli* expression, was obtained from GenScript in pET-28b+ with a C-terminal His-tag. The gene sequence was synthesized into pUC57 with *Sac*I and *Hind*III restriction sites, and the sequence was then cloned into the expression vector between the previously mentioned restriction sites. This process was done by GenScript, Inc. The recombinant plasmids were transformed into *E. coli* Top10 that were kept as glycerol stocks and from which the plasmids were extracted using PureYield ™ Plasmid Miniprep System (Promega) to be transformed into *E. coli* BL21 (DE3). All transformed colonies were cultured on LB agar plates with kanamycin (50ug/ml). Transformation and induction of protein expression were performed as described in section 8 except for using a range of IPTG concentrations from 0.1 to 1 mM and temperatures of either 16°C or 37°C. Culture volumes of 25 or 50 ml, induced by 0.1mM IPTG for 2 hours at 37°C were used for further analysis.

10. Protein purification:

Pellets of induced cells were subjected to freezing (in ice cold ethanol) and thawing (42°C) 3 times, followed by re-suspension in binding buffer, pH 8 (20mM sodium phosphate buffer, 20mM or 40mM imidazole, 500mM NaCl and 5% glycerol). Lysozyme (1mg/ml) and 1mM of phenylmethylsulfonyl fluoride (PMSF) were added and the suspension was incubated on ice for 30 minutes with occasional shaking. Cells were subjected to sonication followed by centrifugation for 25 minutes to separate the supernatant from the cell debris. The protein (Nitra-S) was purified using Ni-NTA affinity chromatography with Ni-NTA Purification System (Invitrogen™). Purification

was done according to manufacturer's native condition specifications at 4°C. The His-tagged Nitra-S was eluted with 20mM Sodium phosphate buffer (pH 8) containing 500mM imidazole, 500mM NaCl and 50% glycerol. DTT (1mM) was added to the eluted fractions. Coumassie-stained 12% SDS-PAGE gel was used to visualize the protein during different purification stages according to the method of Laemmli (1970) (41).

11. Determination of protein concentration:

Protein concentration was determined using the Pierce™ BCA Protein Assay Kit (Thermo Scientific). The purified Nitra-S was 10-fold diluted in binding buffer, pH 8 (20mM imidazole, 500mM NaCl and 5% glycerol) before measuring its concentration. A standard curve was obtained using known concentrations of standard Bovine Serum Albumin (BSA) in the same buffer (figure 3) and the color produced was measured at 595nm using the FLUOstar OPTIMA microplate reader (BMG LABTECH). All measures were done at least in triplicates.

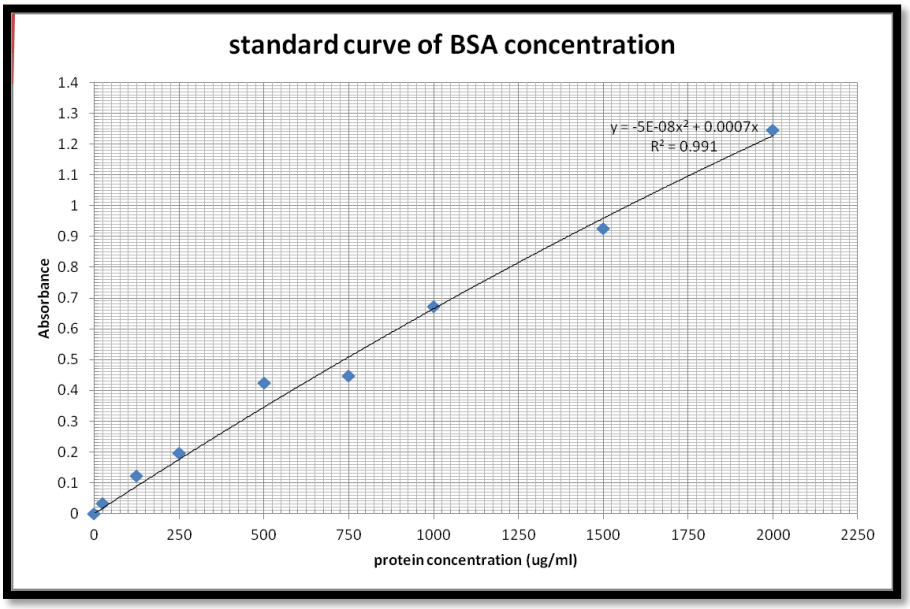

Figure 3: Standard curve of a standard BSA concentration *vs* absorbance at 595nm.

12. Nitrilase activity assay:

A preliminary colorimetric activity assay, developed by Banerjee *et al* (20) was done with minor modifications. Harvested cells from a 25ml overnight culture of transformed cells (both induced and uninduced) were washed with 10mM phosphate buffer (pH 7.2), and suspended in

the same buffer according to its OD (cultures with OD~1 were re-suspended in 1ml buffer). In a 230ul reaction, 6ul of nitrile substrate (from 1M stock solution) and 23ul of prepared cell suspension were added to 201ul of phosphate buffer (10 mM, pH 7.2) containing 0.01% bromothymol blue. The reaction mixture was incubated at 50°C for 6 hours. A color change of the bromothymol blue indicator from bluish-green to yellow was considered a positive result for nitrilase due to liberation of acid into the media. As a positive control for the assay, I used induced cells, transformed with recombinant pET-28a+ that carries a nitrilase gene from *Rhodococcus rhodochrous* ATCC 33278 (obtained from Dr. Oh D-K, Konkuk University, Korea). Unlike all other experiments, testing the activity of Nitra-S qualitatively with benzonitrile, glutaronitrile and succinonitrile was done only in one biological experiment with three technical replicates.

For quantitative activity assay, a spectrophotometric-based method was exploited in assaying the nitrilase activity (23). The method is based on measuring the amount of liberated ammonia by the action of the nitrilase on its substrate. In a 100ul reaction, 10ul of cell or debris suspension (in binding buffer), cell free extract (supernatant) or purified protein (100ug/ml) were added to 50mM potassium phosphate buffer (pH 8), 2mM DTT and 400mM substrate. The reaction was kept at 30°C/100rpm for 30 minutes, and then it was stopped with an equal volume of 100mM HCl. Centrifugation of the stopped reaction mixture was done at 5000rpm for 10 minutes. To detect liberated ammonia, 10ul of the reaction mixture were added to 140ul of buffered alcoholic *o*-phthaldialdehyde/β-mercaptoethanol reagent. The isoindole derivative was allowed to develop for 30 minutes at 30°C/100rpm, and the color intensity was measured at 405nm using the FLUOstar OPTIMA microplate reader (BMG LABTECH).

The reagent used in the assay was prepared one day before the assay as described by Banerjee *et al* (2003) (22). For the preparation of alcoholic *o*-phthaldialdehyde, 100mg *o*-phthaldialdehyde were dissolved in 10ml of absolute ethanol, while alcoholic β-mercaptoethanol was prepared by adding 50ul of β-mercaptoethanol to 10ml of absolute ethanol. For preparing the working reagent, 2.25ml of both alcoholic *o*-phthaldialdehyde and alcoholic β-mercaptoethanol were added to 45.5ml of 200mM potassium phosphate buffer (pH7.4). A standard curve using NH_4Cl was utilized for determination of the ammonia concentration (figure 4). Negative controls were done using non-transformed and uninduced cells.

13. Substrate specificity:

For determination of substrate specificity, the assay was done using several nitrilases: acetonitrile, glutaronitrile, succinonitrile, mandelonitrile and benzonitrile (Sigma-Aldrich). In

addition, for the detection of any possible amidase activity, acetamide and benzamide (Sigma-Aldrich) were used as substrates as well.

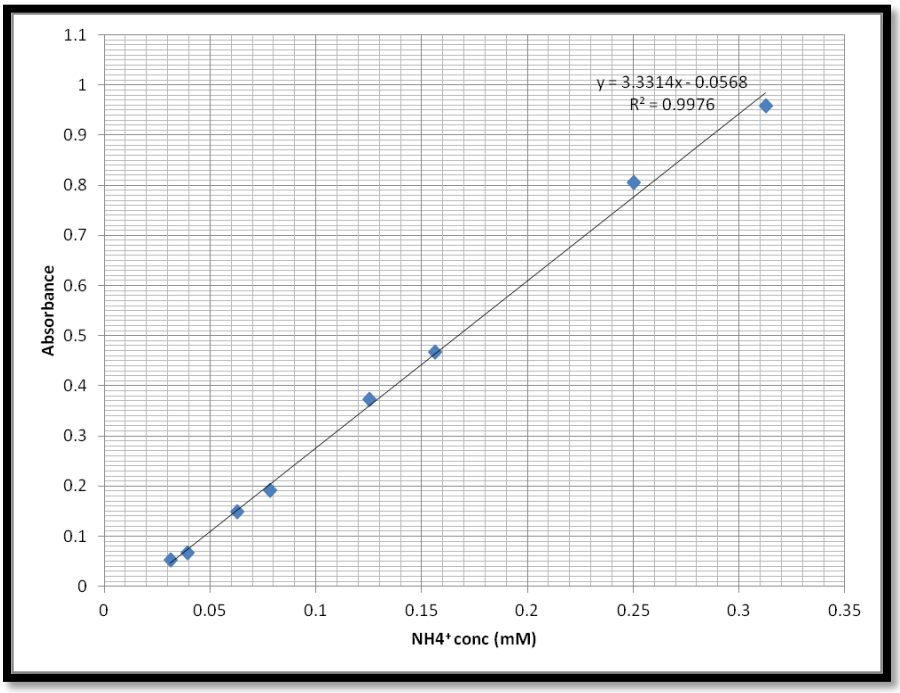

Figure 4: Standard curve for standard NH_4^+ concentrations *vs* absorbance at 405nm.

14. Enzyme characterization:

14.1. Enzyme kinetics:

For determination of Nitra-S kinetics for succinonitrile, the initial reaction rate was determined in a 10-minute reaction with a fixed concentration of the purified protein (100ug/ml) and different concentrations of succinonitrile (0-600mM). Trials were done to calculate the kinetic parameters, Km and Vmax to determine whether Nitra-S obeys Michaelis-Menten kinetics.

14.2. Enzyme thermal stability and thermosensitivity:

For determination of the enzyme thermal stability, the purified enzyme (Nitra-S) was incubated for an hour at different temperatures (30-70°C) before assaying the residual activity as mentioned previously in the quantitative spectrophotometric assay. For determination of the thermosensitivity of the enzyme, the activity assay was hold at different temperatures (30-70°C) as described before.

14.3. Enzyme halotolerance:

For examination of the salt tolerance for the identified nitrilase, the Nitra-S enzyme was incubated at 30°C for an hour at different NaCl concentrations (0.167-3.5M) before assaying the residual activity spectrophotometrically using 66.7ug/ml as final protein concentration.

All experiments in this study were done at least in triplicates.

CHAPTER 3: RESULTS AND DISCUSSION

1. Identification of nitrilase sequences in the LCL database:

Sequences containing the Pfam functional domain CN_hydrolase (PF00795) were extracted from the LCL database using an HMM scan to 8,025,636 reads. The obtained sequences were assembled giving rise to 15 contigs from which the contig with the largest number of reads (contig15: 40 reads, 1209bp) was chosen for further analysis. Upon aligning the obtained consensus sequence with assembled LCL contigs, with annotated ORFs, it aligned perfectly with one of the assembled contigs (contig00026: 66.091Kb) with a query coverage and fraction identity of 99.9%. The alignment of contig15 with contig00026 covered the bases starting from 42805 to 44011 positions in the later contig. Contig00026 with its ORFs was visualized using Artemis Release 13.2.0 (42). The results revealed that the sequence of interest was part of an operon of 9 ORFs (figure5). Regulatory regions, identified by bprom tool (37), were found upstream to the first ORF; a CGCAATGAT sequence was identified as the -10 box, while a TTAAAG sequence comprised the -35 box (figure6). In addition, using the same tool (37), two binding sites for two known transcriptional factors were found upstream: purR: CGTTTTTT and rpoD15: TTTTGTTT. The ORFs in the operon were annotated, after being BLASTed (BLASTx) using NCBI blast2.2.27+ (38), as follows: conserved hypothetical protein, putative nitrilase, putative radical SAM-domain containing protein, putative acetyltransferase, selenophosphate synthetase-related protein (AIR synthetase related protein), conserved hypothetical protein, putative FAD-dependant oxidoreductase, conserved hypothetical protein and putative methyl malonyl Co-A mutase. Analysis of the different ORFs of the operon, revealed that in a study done by Podar *et al* (2005), the first seven ORFs were found with the same arrangement in several bacterial phyla and in microorganisms that inhabit diverse environments in a cluster known as Nit1C (16). The study was done on previously characterized nitrilase genes, obtained from different environments (3). Different bacterial species belonging to beta and gamma proteobacteria have the same arrangement, while in some cyanobacterial species the 7^{th} ORF which is a putative FAD-dependant oxidoreductase involved in K^+ transport, is found on the other end of the cluster on the opposite strand (16). The high conservation of this cluster suggests its involvement in a conserved metabolic pathway; however, the lack of biochemical testing of the genes in this cluster and the presence of hypothetical proteins with unknown activities hinder the prediction of the function of this operon. Comparing the identified nitrilase sequence with sequences used by Podar *et al* (16) revealed that some sequences were closely related to the sequence identified herein. One of the genes which were studied by Podar *et al*, 1B7 nitrilase gene (Genbank accession number: AY487530.1) was found to be in a cluster with the same

arrangement mentioned before. By aligning the amino acid sequences for both 1B7 protein and the putative nitrilase under study, which was named “Nitra-S”, I found 77% identity with 85% coverage. In the study done by Robertson *et al* (2004), 1B7 nitrilase was found to be active towards *S*- nitrile enatiomers (3). In this study, however, no experiments were done to study selectivity towards different isomers.

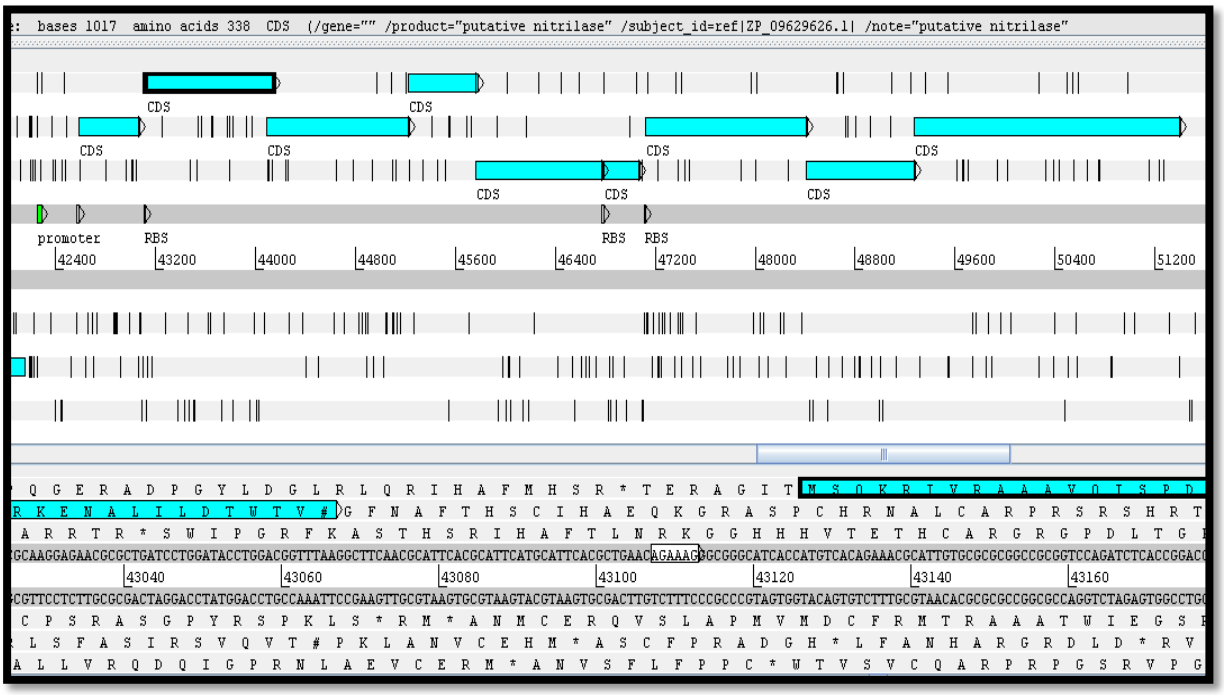

Figure 5: Artemis visualization of the ORFs in the operon at which the putative nitrilase gene is the second ORF. ORFs are shown as cyan arrows at different reading frames in the forward strand. The promoter region is indicated with a green arrow starting from 42250 position, while the putative nitrilase gene starts from 43126 position and ends at 44142 position.

Figure 6: DNA sequence of the putative nitrilase gene within its genomic context and showing its regulatory elements.

The encoded protein sequence of the putative nitrilase gene consists of 338 amino acids. Pfam (31) identified a catalytic triad of glutamate, lysine and cysteine at positions 47, 129 and 163, respectively (figure 7). This catalytic triad is conserved among all nitrilases. This high conservation is shown by aligning the putative nitrilase under study (Nitra-S) with several identified nitrilases using ClustalW2 tool (43) as shown in figure 8. The presence of upstream regulatory regions in addition to the catalytic triad, which is characteristic to all nitrilases, increases the possibility for this gene to encode for an active nitrilase. Using the BLASTp tool (38) to compare the identified protein sequence to the non-redundant protein database in the NCBI, hits with high identities were found; all were annotated as carbon-nitrogen hydrolases or aliphatic nitrilases. Although the best hit (carbon-nitrogen hydrolase in *Cupriavidus basilensis* OR16) was found to have 98% identity with the submitted sequence, this hit and all hits with high identities were annotated based on sequence similarities rather than experimental testing.

MSQKRIVRAAAVQISPDLEHGEGTLGKVCEAIDRAAREGVQLIVFPETFLPYYPYFSFVRPPVQSGSDHMRLYEQAVVV
PGPVTHAVSERARRHAMVVVLGVNERDHGSLYNTQLIFDTDGRLVLKRRKITPTFHERMIWGQGDAAGLKVADTAIG
RVGALACWEHYNPLARYALMTQHEEIHCSQFPGSLVGPIFADQIEVTIRHHALESGCFVVNATGWLTDEQIASVTTDPA
LQKALRGGCNTAIVSPEGQHLAPPLREGEGMVIADLDMSLITKRKRMMDSVGHYARPELLSLAINDRPAATASPMATA
FSNYHGSTHHELQRDDAGLEPVVGN

Figure 7: Amino acid sequence of the native Nitra-S with the catalytic triad of glutamate, lysine and cysteine shown in red.

In an attempt for determination of the 3D structure of Nitra-S, the amino acid sequence was submitted to Swiss-Model tool (44)(45)(46); however, I did not succeed in obtaining a reliable 3D structure of the nitrilase under study. The obtained QMEAN Z-Score was -5.705 and the sequence identity to the used template (2uxy) was 20.455%. The QMEAN Z-score measures the quality of the studied protein model compared to a template formed from reference structures obtained by x-ray crystallography (46). The obtained QMEAN Z-score means that the obtained model deviates from the template by 5.705 standard deviations, rendering this model unreliable. As these results are based on a fully automated mode, the target and the template should have at least 50% identity for the results to be reliable. Such a result was expected due to the scarcity of available nitrilase crystal structures.

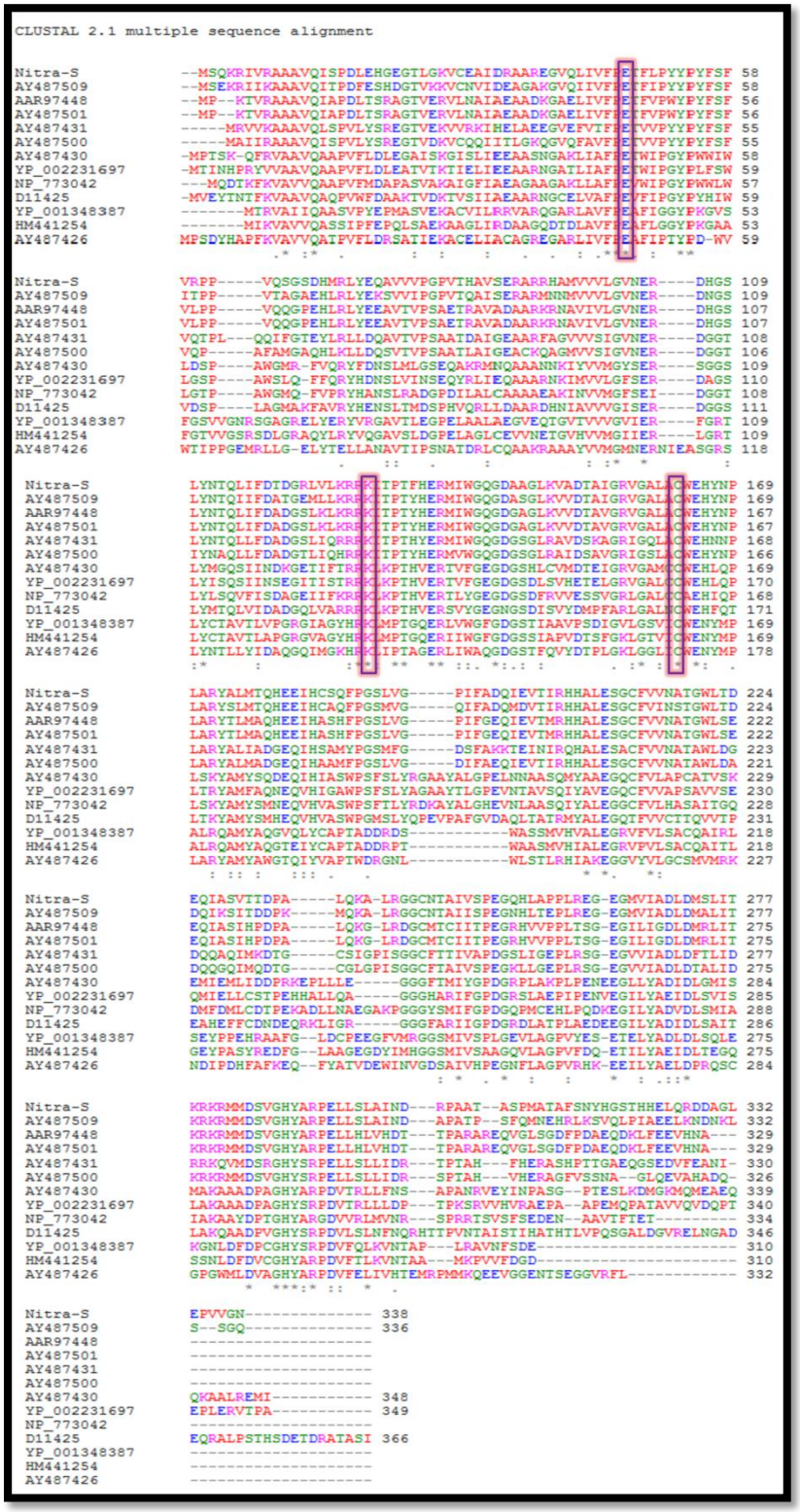

Figure 8: Multiple sequence alignment of Nitra-S with different nitrilases (identified by their NCBI accession numbers on the left panel). Residues of the catalytic triad are shown in rectangles.

2. Amplification and sequencing of the putative nitrilase gene and regulatory elements:

Different primers were used to amplify the nitrilase ORF and surrounding regions. The position of different primers used and the expected amplicon sizes were shown in figure 2. Either direct sequencing or sequencing of the recombinant pGEM vectors with the amplified sequences were done as mentioned in the materials and methods. Ligation into pGEM vector was based on TA cloning. In this type of cloning, the amplicons with the 3'-A overhangs, which is a characteristic of the used *taq* polymerase, can bind to the 3'-T overhangs in the open vector. The LCL microbial DNA was used in different PCR reactions with the following sets of primers: Nitr1_F with Nitr1_R, Nitr6_F with Nitr1_R, Nitr3_F with Nitr3_R, Nitr3_F with Nitr4_R, Nitr7_F with Nitr3_R, Nitr4_F with Nitr4_R or Nitr4_F with Nitr5_R. Amplicons of 1336bp (A fragment) and 665bp (D fragment) were obtained using the first two sets of primers as shown on 0.8% agarose gels (figures 9 &10). Figure 11 shows recombinant pGEM plasmids with either A or D fragments on 0.8% agarose gel before sequencing. Using the other sets of primers, amplicons of 1739bp, 1268bp, 1085bp, 1405bp and 801bp were obtained, respectively, as shown in figure 12. Sequencing confirmed the consensus sequence of the nitrilase ORF and regulatory regions obtained from the LCL database.

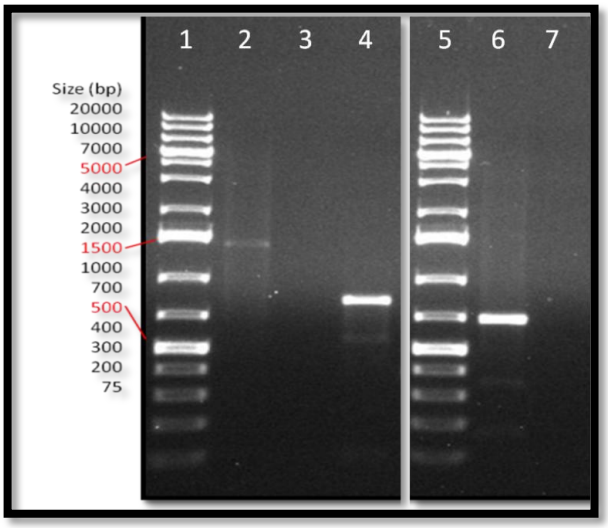

Figure 9: PCR amplification of the putative nitrilase gene and neighboring regions from environmental DNA. lane1: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific), lane2: gene amplification of the 1336 bp fragment (A) of the putative nitrilase gene using Nitr1_F and Nitr1_R primers, lane3: Negative control using Nitr1_F and Nitr1_R primers, lane4: positive control SOD with *E.coli* genomic DNA, lane5: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific), lane6: gene amplification of the 665 bp (D) fragment of the putative nitrilase gene using Nitr6_F and Nitr1_R primers, lane7: negative control using Nitr6_F and Nitr1_R primers.

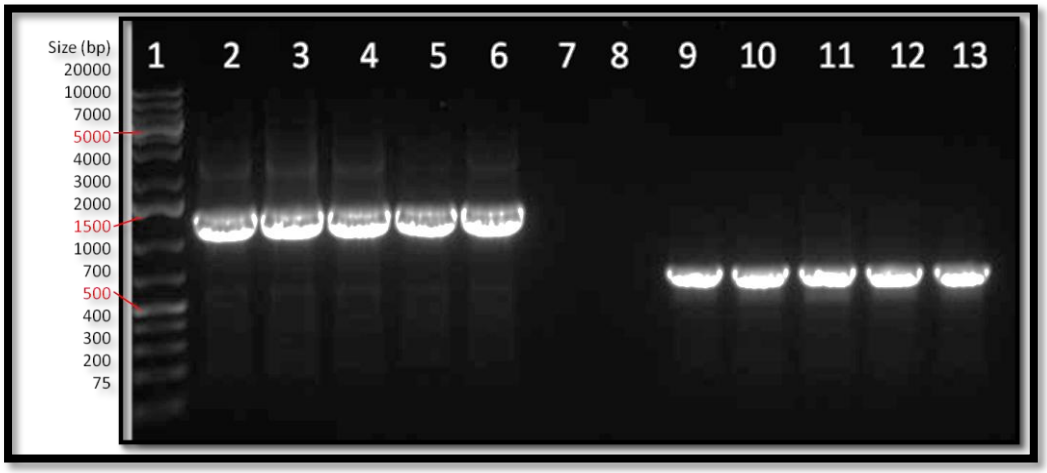

Figure 10: Colony PCR from 5 *E.coli* Top10 colonies transformed with either the (A: 1336bp) fragment or the (D: 665bp) fragment cloned in pGEM vector. lane1: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific), lanes 2-6: (A) amplified fragment using Nitr1_F and Nitr1_R primers, lane7: negative control using Nitr1_F and Nitr1_R primers, lane8: gap, lanes 9-13: (D) amplified fragment using Nitr6_F and Nitr1_R primers.

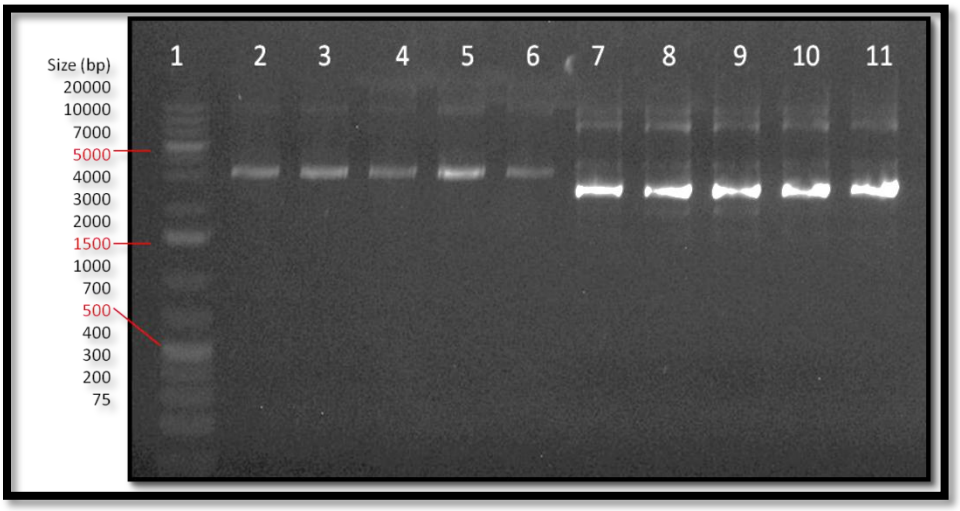

Figure 11: Purified recombinant pGEM plasmid with (A) insert or (D) insert. lane1: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific), lanes 2-6: recombinant pGEM plasmid with (A: 1336bp) insert, lanes 7-11: recombinant pGEM plasmid with (D: 665bp) insert.

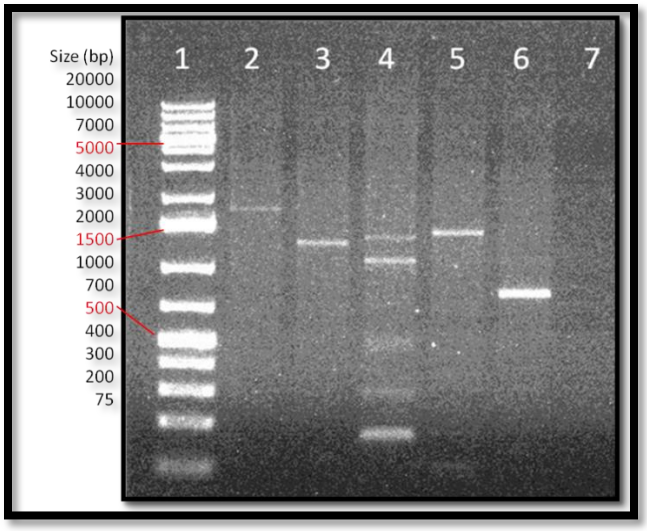

Figure 12: PCR amplification products from environmental DNA using different sets of primers for the region surrounding the nitrilase ORF. lane1: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific), lane2: 1739bp product using Nitr3_F and Nitr3_R primers, lane3: 1268bp product using Nitr3_F and Nitr4_R primers, lane4: 1085bp product using Nitr7_F and Nitr3_R primers, lane5: 1405bp product using Nitr4_F and Nitr4_R primers, lane6: 801bp product using Nitr4_F and Nitr5_R primers, lane7: negative control.

3. Expression of Nitra-S from the recombinant pET SUMO vector and optimization of the expression conditions:

In order to express the putative nitrilase, Nitr_exp1_F and Nitr_exp2_R primers were used to amplify the complete ORF using the recombinant pGEM-insert (A) purified plasmid as a template. An amplicon of 1058bp was obtained (figure 13), and subsequently ligated to the pET SUMO vector by TA cloning, followed by transformation into *E. coli* Top10. After overnight culturing of the transformed cells, 24 clones were subjected to colony PCR using both Sumo_F with Nitr3_R and Nitr6_F with T7_R primers to identify clones with the proper orientation of the insert. Sumo_F with Nitr3_R were used to amplify the 5' fragment (953bp), while Nitr6_F with T7_R were used to amplify the 3' fragment (581bp). Bands at the expected size (953bp with the first set of primers and 581bp with the second set) were obtained with only one colony among the colonies screened (figure 14). The corresponding recombinant plasmid was extracted, sequenced and transformed into *E. coli* BL21 (DE3) for expression.

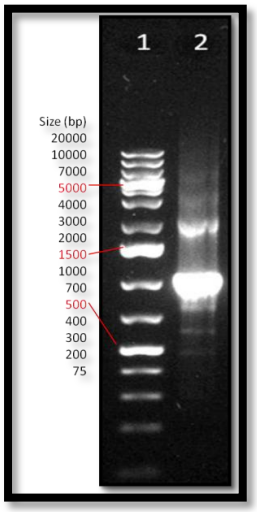

Figure 13: Amplified nitrilase ORF sequence using expression primers (Nitr_exp1_F and Nitr_exp2_R). lane 1: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific), lane2: 1058bp product amplified with Nitr_exp1_F and Nitr_exp2_R primers using recombinant pGEM with the previously amplified (A) fragment as a template.

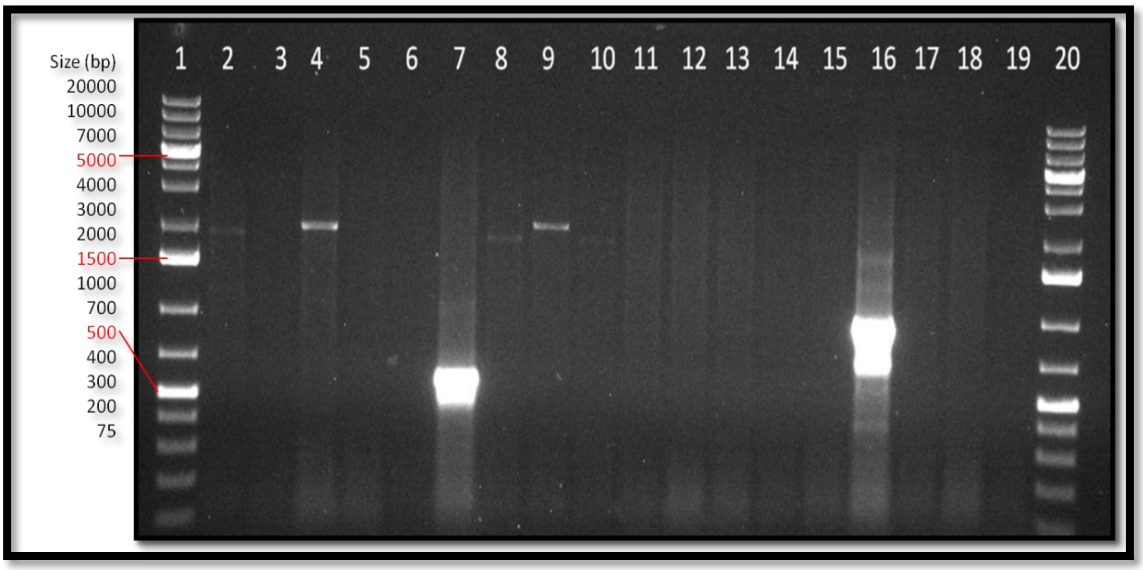

Figure 14: Colony PCR for colonies transformed with recombinant pET SUMO with Nitra-S gene [insert amplified by Nitr_exp1_F and Nitr_exp2_R], showing positive results with only one colony (lanes 7 & 16). lanes 1 & 20: GeneRuler™ 1Kb plus DNA ladder (Thermo Scientific). lanes 2-10: PCR reaction using Nitr6_F and T7 reverse (581bp expected size at lane 7). lanes 11-19: PCR reaction using Sumo_F and Nitr3_R (953bp expected size at lane 16).

Coumassie-stained 12% SDS-PAGE gels were used to visualize over-expressed Nitra-S, at 37°C for different time intervals, using IPTG concentrations of 0.1mM (figure 15), 0.2mM (figure 16), 0.4mM (figure 17) and 0.8mM (figure 18). Although the native Nitra-S size should be 37.12kDa, calculated using Expasy Compute pI/Mw tool (47)(48)(49), the over-expressed band had an apparent molecular mass of 51kDa. This is because of the attached SUMO fusion protein, which helps in the solubilisation of the recombinant protein, in addition to the His-tag sequence. From the results shown, it can be inferred that Nitra-S is highly over-expressed when incubated with IPTG for 2 hours, and that any increase in induction time, does not result in a considerable increase in the amount of the over-expressed protein. Preliminary observations revealed that the amount of induced protein increases with the increase in IPTG concentration figures 15-18. However, after cell lysis of the 2-hour-induced cell pellets, the increase in induction, accompanied with the increase in IPTG concentration, had shown that most of the over-expressed Nitra-S is preferentially precipitated with the cell debris (insoluble fraction) (figure 19) without an increase in its amount in the supernatant (soluble fraction) (figure 20). Figure 21 shows a 12% SDS-PAGE gel of both the supernatant and cell debris after 2 hours of induction with different IPTG concentrations. These results strongly suggest that any increase in the IPTG concentration, increases the expression of Nitra-S as inclusion bodies. Accordingly, it was concluded that, at 37°C, induction for 2 hours with 0.1mM IPTG is the best condition for the expression of the protein under study.

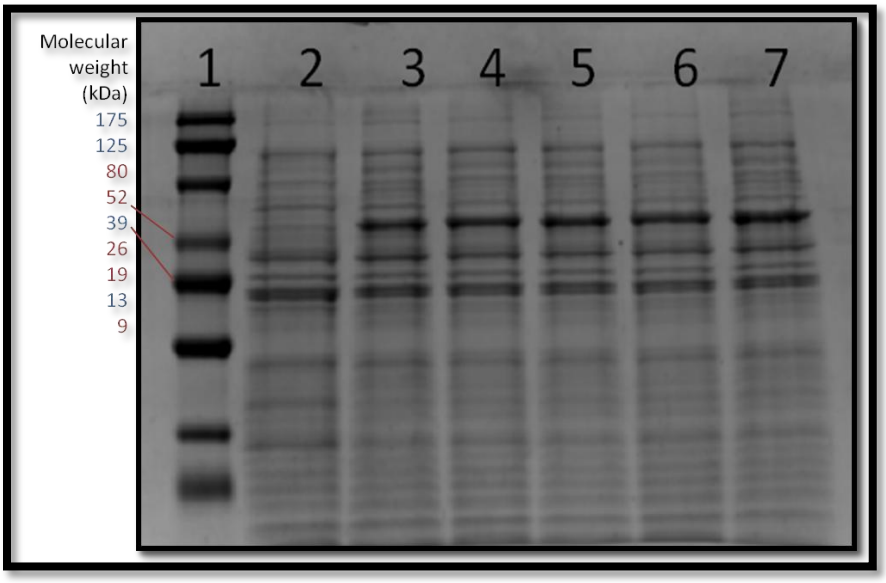

Figure 15: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S induced at 37°C with 0.1mM IPTG at different time intervals. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 1 hour induction, lane4: 2 hours induction, lane5: 3 hours induction, lane6: 4 hours induction, lane7: 5 hours induction.

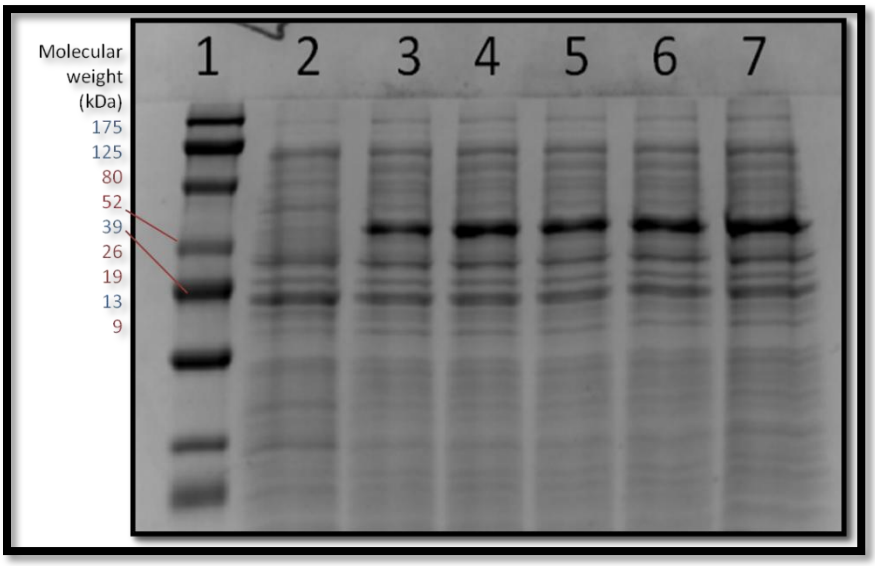

Figure 16: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S induced at 37°C with 0.2mM IPTG at different time intervals.12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 1 hour induction, lane4: 2 hours induction, lane5: 3 hours induction, lane6: 4 hours induction, lane7: 5 hours induction.

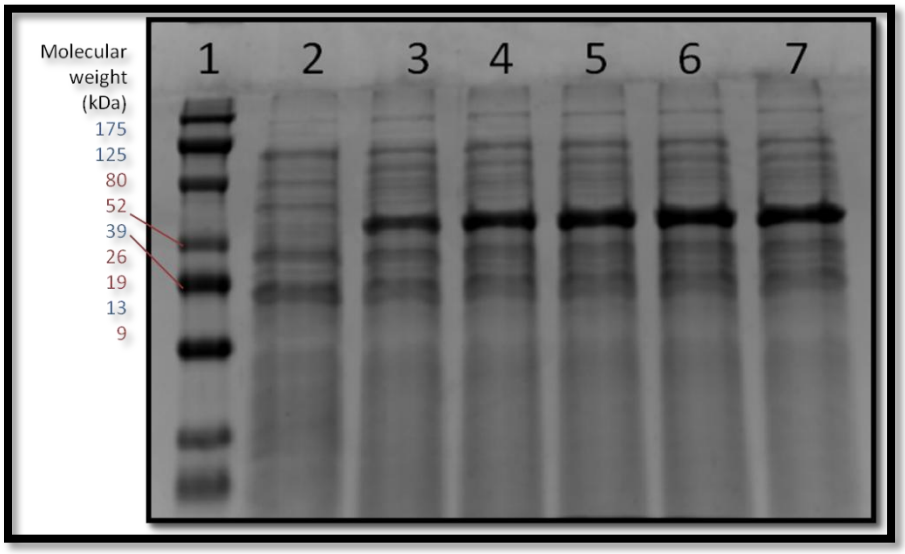

Figure 17: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S induced at 37°C with 0.4mM IPTG at different time intervals. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 1 hour induction, lane4: 2 hours induction, lane5: 3 hours induction, lane6: 4 hours induction, lane7: 5 hours induction.

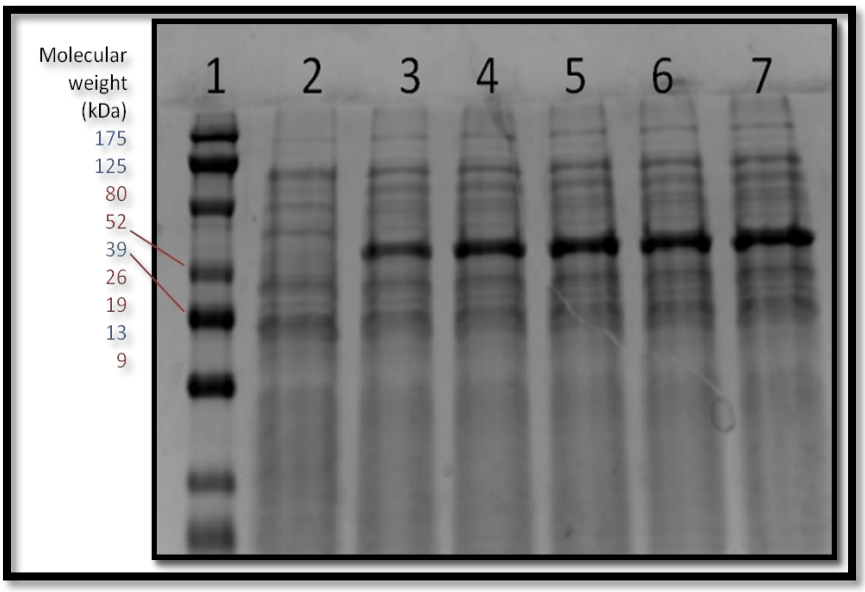

Figure 18: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S induced at 37°C with 0.8mM IPTG at different time intervals. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 1 hour induction, lane4: 2 hours induction, lane5: 3 hours induction, lane6: 4 hours induction, lane7: 5 hours induction.

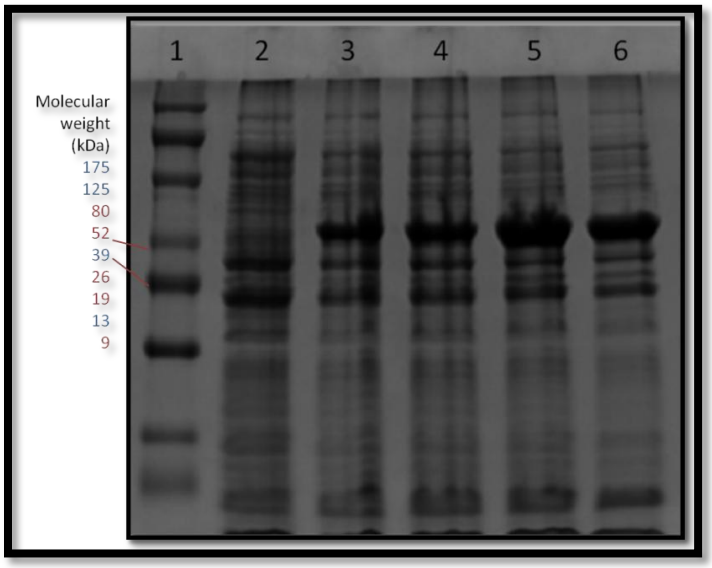

Figure 19: Analysis of the debris of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S after 2 hours induction. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cells' debris, lane3: 0.1mM IPTG induction, lane4: 0.2mM IPTG induction, lane5: 0.4mM IPTG induction, lane6: 0.8mM IPTG induction.

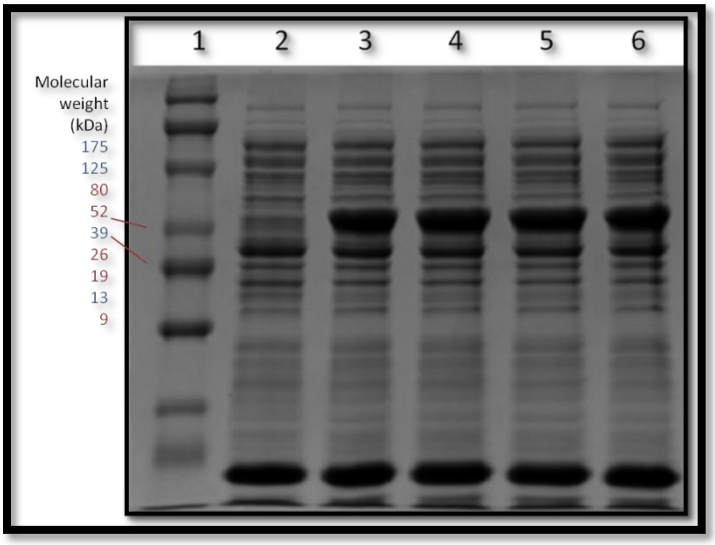

Figure 20: Analysis of the supernatant of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S after 2 hours induction. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cells' supernatent, lane3: 0.1mM IPTG, lane4: 0.2mM IPTG, lane5: 0.4mM IPTG, lane6: 0.8mM IPTG.

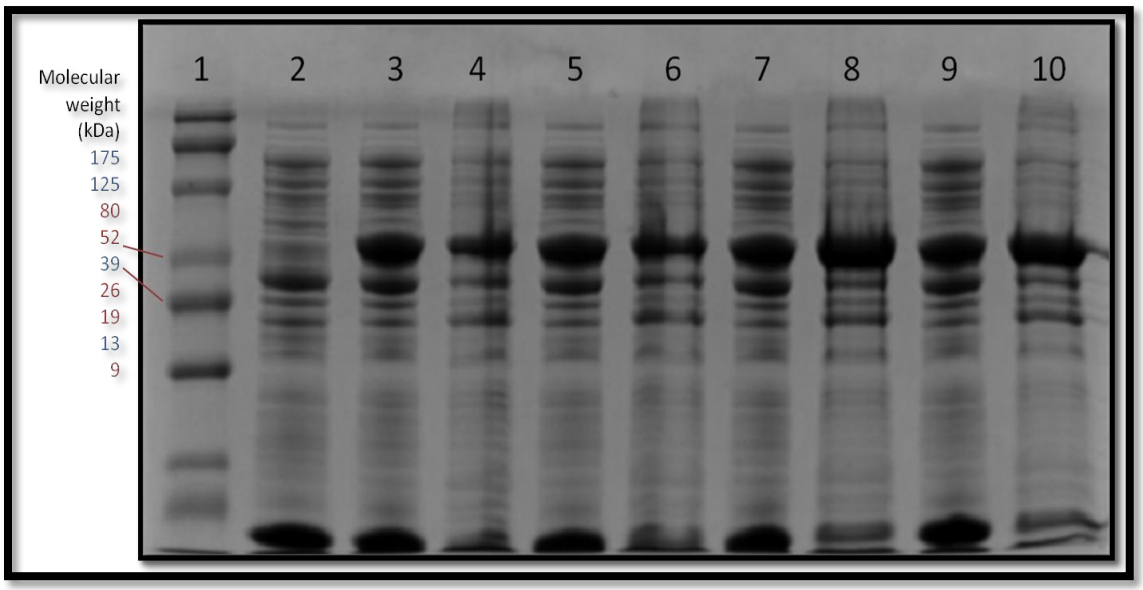

Figure 21: Analysis of the supernatant and debris of *E. coli* BL21 (DE3) transformed with pET SUMO-Nitra-S after 2 hours induction. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cells' supernatant, lane3: supernatant 0.1mM IPTG, lane4: debris 0.1mM IPTG, lane5: supernatant 0.2mM IPTG, lane6: debris 0.2mM IPTG, lane7: supernatant 0.4mM IPTG, lane8: debris 0.4mM IPTG, lane9: supernatant 0.8mM IPTG, lane10: debris 0.8mM IPTG.

4. Expression of Nitra-S from the recombinant pET-28b+ vector and optimization of the expression conditions:

A synthesized gene (based on the sequence amplified from the metagenomic DNA) with optimized codon usage was obtained from GenScript in pET-28b+ with a C- terminal His-tag. The His-tag was chosen to be in the C-terminus to lower the possibility of any effect on the activity, as the active site was found to be towards the N-terminus. The recombinant plasmid was transformed into *E. coli* BL21 (DE3) for expression. The apparent molecular mass of the induced protein is 38.83KDa, calculated using Expasy Compute pI/Mw tool (47)(48)(49). This small increase in size than the native protein is due to the extra histidine residues at the C- terminus and 2 extra methionine and glycine residues at the N-terminus that were added to put the gene sequence in frame in the expression vector between *Sac*I and *Hind*III restriction sites.

Expression was first done at 16°C with longer induction times (15, 18 and 21 hours) as shown in most studies using the pET system for expression. IPTG concentrations of 0.1mM, 0.2mM (figure 22), 0.5mM and 1mM (figure23) were used. From the results obtained, it was clear that an IPTG concentration of 0.1mM was adequate to yield a satisfactory level of Nitra-S over-expression. It was also observed that the amount of expressed protein has decreased after 21-hour-induction. Perhaps, this is due to degradation of part of the protein after long incubation times. However, most of the expressed protein was found in the cell debris (figure 24) and the solubilised percentage was very low (figures 25 & 26). Using 0.1mM IPTG at 37°C, different induction times were studied as done previously with transformed cells with pET SUMO-Nitra-S (figure 27). The highest amount of induced protein was observed after 2 hours of induction. In addition, a great improvement and a higher percentage of the soluble protein were observed when the induction was done at 37°C for few hours (figure 28). These conditions did not eliminate the presence of most of the expressed protein in the insoluble fraction; however, a considerable amount of soluble protein that could be purified under native conditions was found as well. Thus, induction for 2 hours at 37°C were considered the best expression conditions and were used to obtain the highest possible amount of protein to be purified and further characterized.

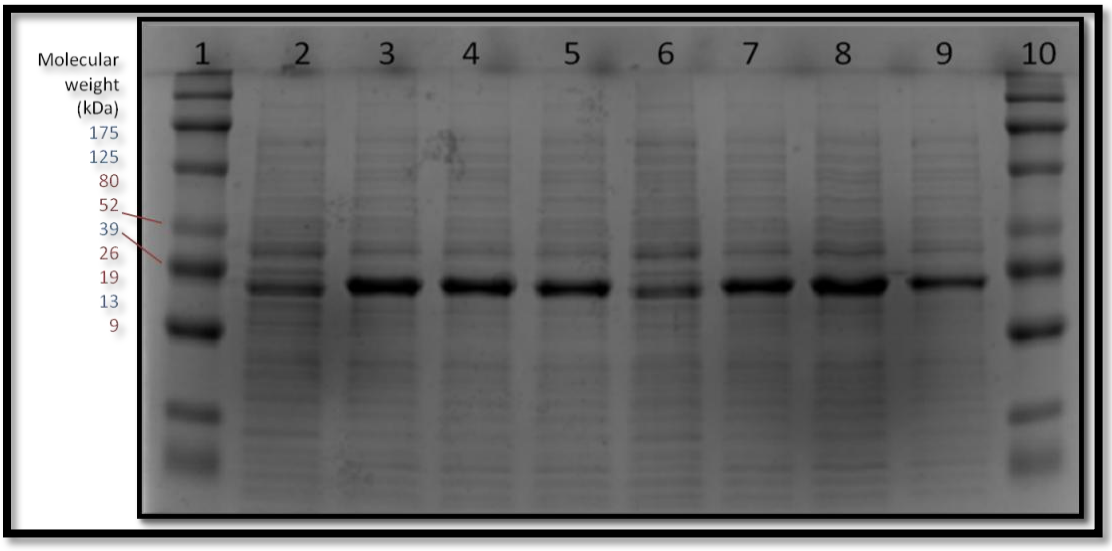

Figure 22: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S induced at 16°C with 0.1mM IPTG & 0.2mM IPTG at different time intervals. 12% SDS-PAGE of: lanes 1 & 10: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 15 hours 0.1mM IPTG, lane4: 18 hours 0.1mM IPTG, lane5: 21 hours 0.1mM IPTG, lane6: uninduced cell lysate, lane7: 15 hours 0.2mM IPTG, lane8: 18 hours 0.2mM IPTG, lane9: 21 hours 0.2mM IPTG.

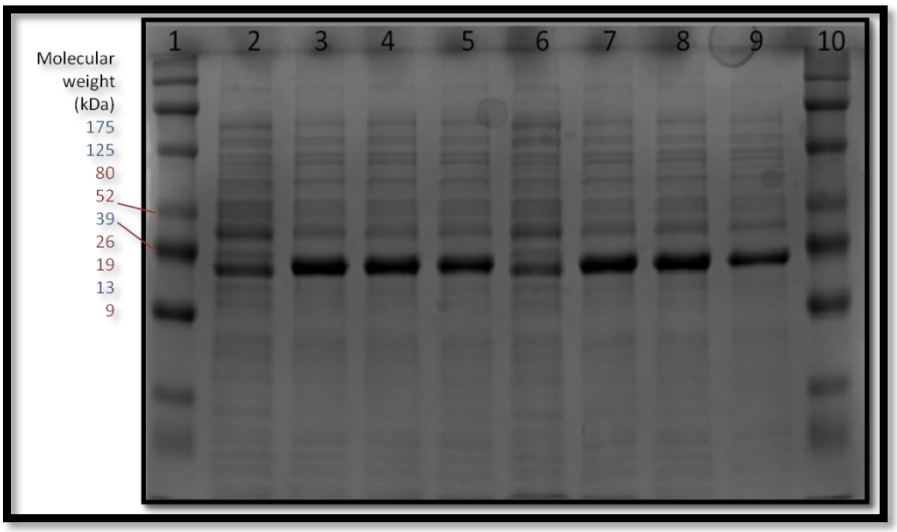

Figure 23: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S induced at 16°C with 0.5mM IPTG & 1mM IPTG at different time intervals. 12% SDS-PAGE of: lanes 1 & 10: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 15 hours 0.5mM IPTG, lane4: 18 hours 0.5mM IPTG, lane5: 21 hours 0.5mM IPTG, lane6: uninduced cell lysate, lane7: 15 hours 1mM IPTG, lane8: 18hours 1mM IPTG, lane9: 21 hours 1mM IPTG.

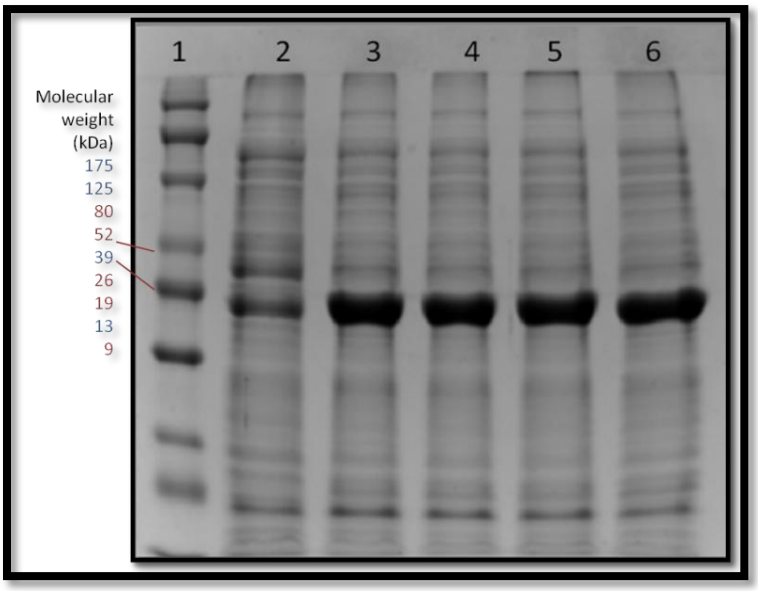

Figure 24: Analysis of the cell debris of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S after 15 hours induction at 16°C. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cells' debris, lane3: 0.1mM IPTG, lane4: 0.2mM IPTG, lane5: 0.5mM IPTG, lane6: 1mM IPTG.

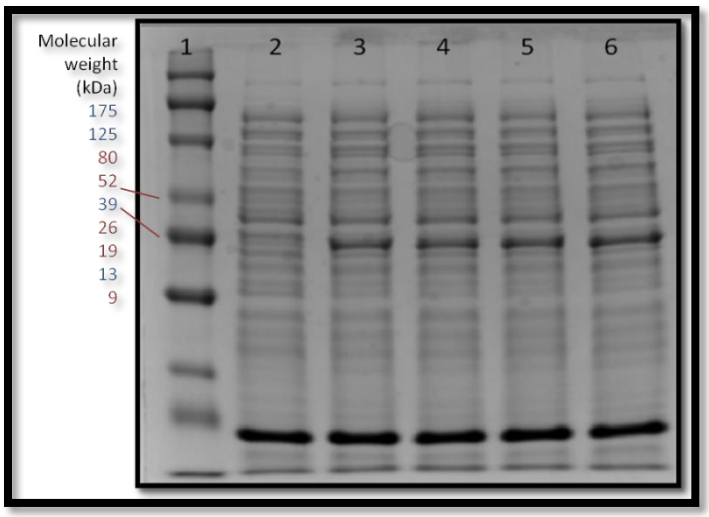

Figure 25: Analysis of the supernatant of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S after 15 hours induction at 16°C.12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cells' supernatent, lane3: 0.1mM IPTG, lane4: 0.2mM IPTG, lane5: 0.5mM IPTG, lane6: 1mM IPTG.

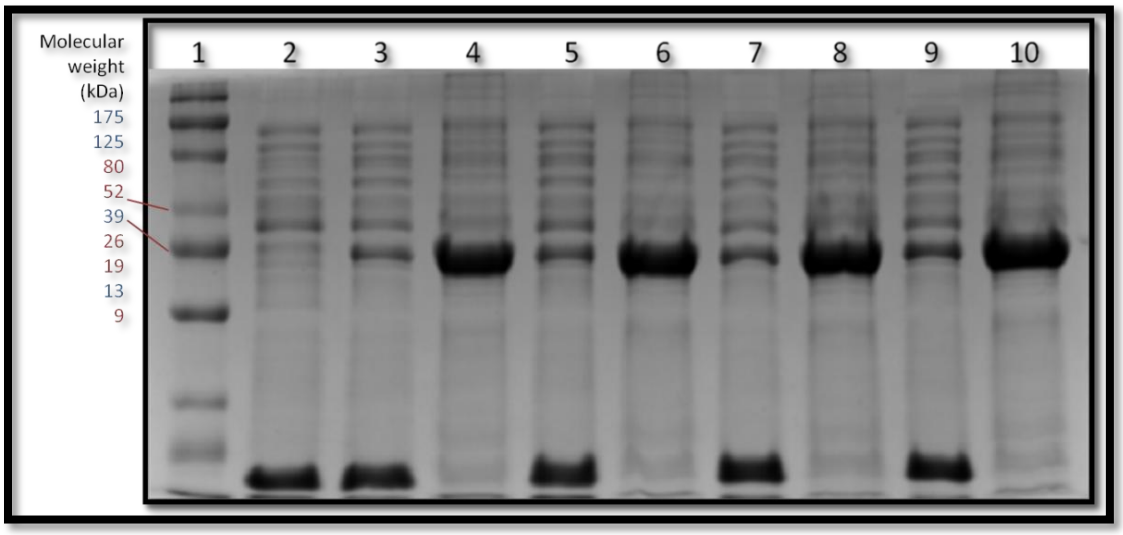

Figure 26: Analysis of the supernatant and debris of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S after 15 hours induction at 16°C. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced supernatant, lane3: supernatant 0.1mM IPTG, lane4: debris 0.1mM IPTG, lane5: supernatant 0.2mM IPTG, lane6: debris 0.2mM IPTG, lane7: supernatant 0.5mM IPTG, lane8: debris 0.5mM IPTG, lane9: supernatant 1mM IPTG, lane10: debris 1mM IPTG.

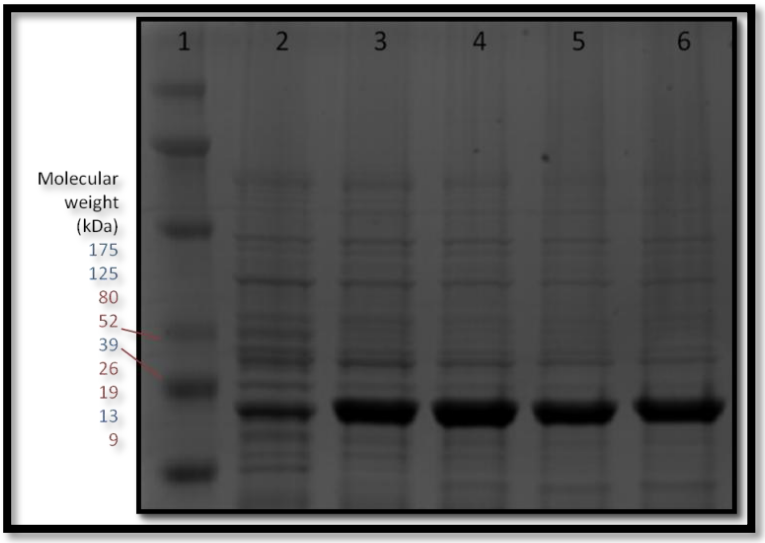

Figure 27: Analysis of the cell lysate of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S induced at 37°C with 0.1mM at different time intervals. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cell lysate, lane3: 1 hour induction, lane4: 2 hours induction, lane5: 3 hours induction, lane6: 4 hours induction.

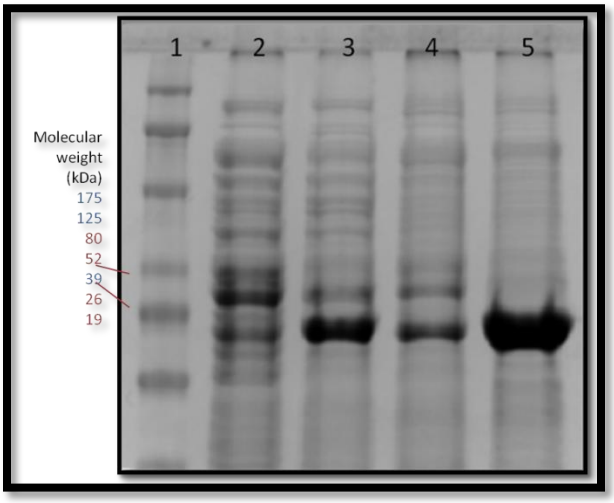

Figure 28: Analysis of the supernatant and debris of *E. coli* BL21 (DE3) transformed with pET-28b+ Nitra-S after 2 hours induction with 0.1mM IPTG at 37°C. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: uninduced cells' supernatant, lane3: induced cells' supernatant, lane4: uninduced cells' debris, lane5: induced cells' debris.

5. Purification of C-terminal His-tagged Nitra-S obtained from the expression of the synthesized gene:

Purification was done to the C-terminal His-tagged protein only. A stretch of 6 histidine residues is able to bind strongly to the Ni^{+2} ions immobilized on the column matrix with NTA. Elution was done with a high concentration of imidazole (500mM) which has higher affinity to Ni^{+2}, so it can displace the bound protein easily (50). Although the effect of the pH on the activity of Nitra-S was not examined, the pH of all used buffers was adjusted to 8, as the calculated isoelectric point (pI) of the protein was 6.41; thus, it was preferable to use a pH far away from its pI to prevent any possible precipitation of the protein.

First purification trials were done using binding buffer containing 20mM imidazole (figure 29). The concentration of imidazole was then raised to 40mM to decrease contamination with other proteins (figure 30). The protein (Nitra-S) was eluted in an elution buffer with 50% glycerol to be preserved at -20°C; and eluates with high protein concentration, from different purifications, were pooled together for further analysis of the protein. The elution buffer did not affect the activity assays; however, for determination of protein concentration, 10-fold diluted protein was used as the high concentrations of imidazole and glycerol were found to interfere with the BCA assay method. This assay depends on the reduction of Cu^{+2} to Cu^{+1} by the protein in alkaline medium. Cu^{+1} can be detected by the use of a reagent which contains bicinchoninic acid (BCA),

where 2 molecules of BCA chelate with one Cu^{+1} ion producing a purple color that can be measured spectrophotometrically (51). The highest protein concentration that I got was 2mg/ml.

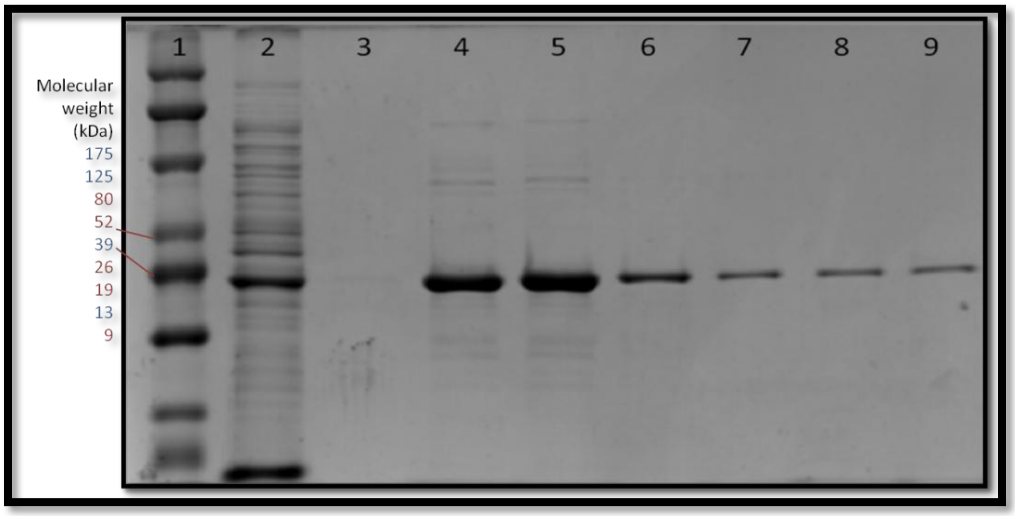

Figure 29: Analysis of eluates (500ul each) of Nitra-S protein from *E.coli* BL21 (DE3) transformed with pET-28b+ Nitra-S using Ni-NTA column under native conditions and 20mM imidazole in binding/ washing buffer and elution buffer with 500mM imidazole and 50% glycerol. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: supernatant before purification, lane3: eluate fraction 1, lane4: eluate fraction 2, lane5: eluate fraction 3, lane6: eluate fraction 4, lane7: eluate fraction 5, lane8: eluate fraction 6, lane9: eluate fraction 7.

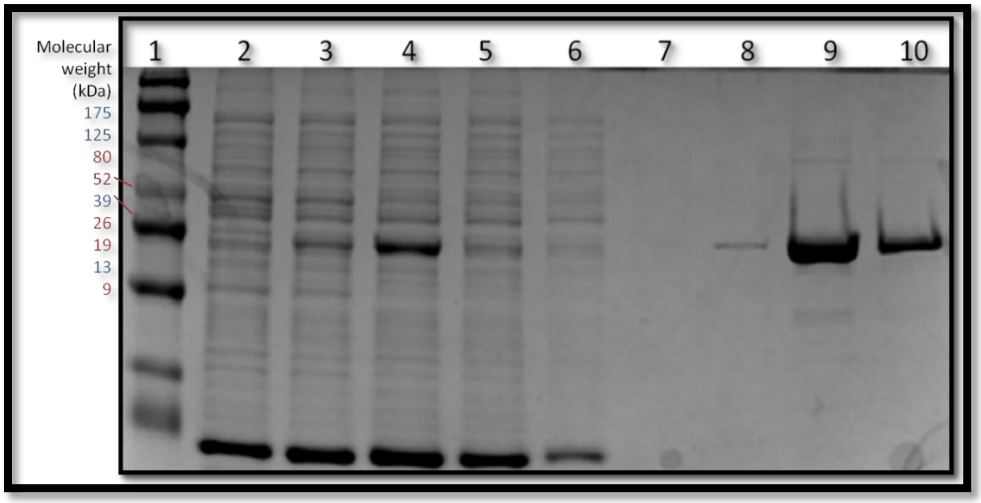

Figure 30: Purification of Nitra-S protein from *E.coli* BL21 (DE3) transformed with pET-28b+ Nitra-S using Ni-NTA column under native conditions and 40mM imidazole in binding/ washing buffer and elution buffer with 500mM imidazole and 50% glycerol. 12% SDS-PAGE of: lane1: ProSieve® Colour Protein Markers (Lonza), lane2: non-transformed cells' supernatant, lane3: Uninduced cell's supernatant, lane4: induced supernatant before purification, lane5: flow-through, lane6: wash 1, lane7: wash 2, lane8: eluate fraction 1, lane9: eluate fraction 2, lane10: eluate fraction 3.

6. Nitrilase activity and substrate specificity of Nitra-S:

Nitrilases are characterized by one-step reaction at which nitriles are hydrolysed into their corresponding carboxylic acid and ammonia. Different assays were developed for determination of the nitrilase activity; most are based on determination of the amount of released ammonia. Some assays depend on measuring the increased acidity in the medium upon the hydrolysis of nitriles. In a simple, qualitative, colorimetric assay, bromothymol blue was used as an acid base indicator where its color changes from green to yellow after incubating the induced transformed cells with the reaction mixture for 6 hours at 50°C (20). To ensure the validity of this assay, the nitrilase activity of induced *E. coli* BL21(DE3) cells, transformed with recombinant pET28a+ which carries a previously characterized nitrilase gene(6), was assayed using acetonitrile and mandelonitrile as substrates. A yellowish green color was observed with the induced cells rather than the uninduced ones. The difference in color between the positive and negative controls was observed with both substrates (figure 31).

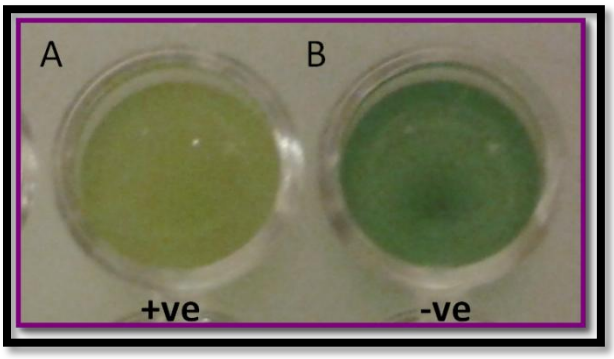

Figure 31. Qualitative detection of nitrilase activity on acetonitrile using cultures of *E.coli* BL21 (DE3) transformed with a recombinant pET28a+ with a nitrilase gene obtained from Dr. Oh D-K, Konkuk University, Korea. A: Yellow color obtained with induced culture. B: Green color obtained with uninduced culture.

The same assay was used as a preliminary test for nitrilase activity using non transformed *E. coli* BL21 (DE3) cultures, induced and uninduced cultures transformed with either recombinant pET SUMO or pET-28b+ (figure 32). Negative results were obtained with acetonitrile, mandelonitrile and benzonitrile, while positive results were only observed with glutaronitrile and succinonitrile. While the non-transformed cells yielded a dark green color, transformed cells – both induced and uninduced – showed yellowish discoloration indicating the presence of basal level of expression of the recombinant protein and a leakage in the T7*lac* promoter in *E. coli* BL21 (DE3). In case of cells transformed with recombinant pET-28b+, a drop in the pH was more prominent with the induced cells as observed from the color of the indicator.

However, the difference in color between transformed and non-transformed cells was clearer with succinonitrile than glutaronitrile. In case of cells transformed with recombinant pET SUMO, an equal drop in the pH was observed with induced and uninduced transformed cultures compared to the non-transformed cultures; this drop was more prominent with succinonitrile. The absence of a clear difference in color between the induced and the uninduced cells was surprising, since a considerable amount of the induced protein was found to be soluble as shown in the SDS-PAGE results (figure 21). This could be attributed to the presence of the SUMO peptide, which could result in improper folding of the expressed protein, thus lowering the activity of Nitra-S. These explanations can never be conclusive, being based on a qualitative assay. However, from the previous results, it could be suggested that the specificity of Nitra-S is towards dinitriles. Figure 33 shows the chemical structures of nitriles used in this study.

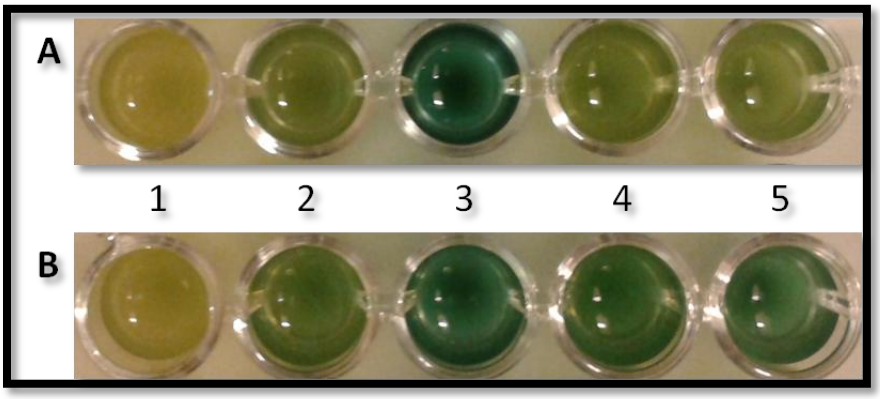

Figure 32. Qualitative detection of nitrilase activity on succinonitrile (A) & glutaronitrile (B) substrates using cultures of *E.coli* BL21 (DE3) transformed with recombinant pET-28b+ and recombinant pET SUMO. 1: induced pET-28b+ transformed cells, 2: uninduced pET-28b+ transformed cells, 3: non-transformed cells, 4: induced pET SUMO transformed cells, 5: uninduced pET SUMO transformed cells.

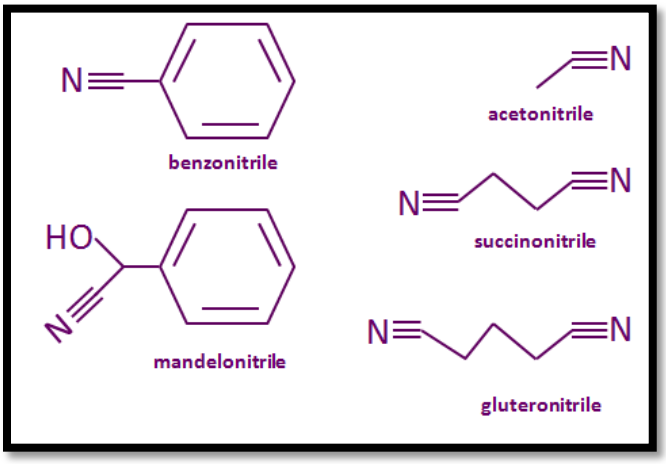

Figure 33: Chemical structures of nitriles used as substrates in studying substrate specificity.

For quantitative analysis of the nitrilase activity of Nitra-S, a spectrophotometric assay, which measures the amount of released ammonia, was utilized. The enzyme was incubated with different nitriles for 30 minutes before being inactivated by the addition of HCl. DTT (2mM) was added to the reaction mixture to ensure that all cysteine residues are in their reducing form. Although, the DiANNA 1.1 web server (52)(53)(54) has detected a possible disulfide bond in Nitra-S with a high score of 0.99659, one of the cysteine residues involved in this bond was found to be in the catalytic triad. The scores obtained by the DiANNA 1.1 web server range from 0 to 1; and the higher the score, the higher prediction reliability. Thus, based on the high possibility of the oxidation of the active cysteine and formation of a disulfide bond and from the previous knowledge that thiol complexing agents inactivate nitrilases (8), the necessity of DTT addition emerged. After the reaction termination, 10ul of the reaction mixture were added to a buffered alcoholic *o*-phthaldialdehyde/β-mercaptoethanol reagent. *O*-phthaldiladehyde reacts with ammonia in presence of β-mercaptoethanol to form an isoindole derivative (figure 34). The formed product could be assessed either colorimetrically or fluorometrically. In this study, the color intensity was measured at 405nm and the amount of released ammonia was calculated based on the standard curve, which was done using a standard NH_4Cl solution (figure 4). A more sensitive assay was to measure the fluorescence of the isoindole derivative at 467nm when exited at 412nm; however, the lack of suitable filters hindered the use of the fluorometric assay.

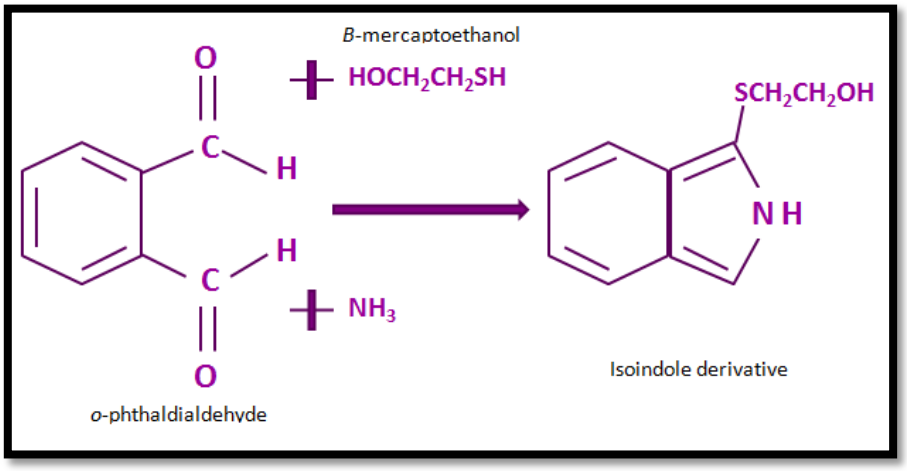

Figure 34. Reaction of ammonia with *o*-phthaldialdehyde and β-mercaptoethanol to form an isoindole derivative.

The activity was measured for whole cell suspensions (recombinant pET SUMO and pET-28b+ transformed cells), crude extracts (supernatants) and cell debris. Positive results were only obtained with glutaronitrile and succinonitrile when used as substrates, confirming the previous results of the dinitrile specificity of the nitrilase under study. The enzyme was more

active towards succinonitrile than glutaronitrile. The cell suspensions of both transformed cells showed that the activity towards glutaronitrile is about 13.24% of that towards succinonitrile. For the purified protein, nearly similar results were obtained, as the activity towards glutaronitrile was about 12.85% of that towards succinonitrile.

In contrast to the observation found in the qualitative colorimetric activity assay, the nitrilase activity of the induced recombinant pET SUMO transformed cells was greater than the uninduced cells. The later showed 42.6% of the activity shown by the induced cells. The discrepancy between the results of the two assays might be due to the qualitative nature of the first assay or it could be of lower sensitivity. However, this does not explain the difference observed between both induced and uninduced recombinant pET-28b+ transformed cells in both assays.

When the nitrilase activity was assayed for the cells' debris, after cell lysis of induced cultures of both recombinant pET SUMO and pET-28b+ transformed cells, low activity was observed with succinonitrile in both types of cells. This means that the used cell lysis technique was not sufficient for complete cell lysis, thus part of the soluble protein was still entrapped in undisturbed cells, or that part of the protein in the inclusion bodies is active. For proper explanation of these results, further analysis is to be done.

As some nitrilases have shown amidase activity beside their main function as nitrilases (1), acetamide and benzamide were used to test this possibility. No activity towards these amides was observed whatsoever; however, we cannot confirm the absence of any amidase activity without further analysis using other amides as substrates.

7. Enzyme characterization:

7.1. Enzyme kinetics:

Many attempts were done to determine the kinetic parameters of Nitra-S using succinonitrile as a substrate. However, the activity continued to increase proportionally with the substrate concentration, although very high concentrations of succinonitrile were used in the study (up to 600mM). Lowering the protein concentration to less than 100mg/ml or doing the reaction for less than 10 minutes resulted in non-measurable activity at low substrate concentrations. Perhaps, this protein does not obey the Michaelis-Menten equation rendering its kinetic study a problematic issue (figure 35) or perhaps, a more sensitive assay is needed such as measuring the fluorescence of the isoindole derivative instead of the absorbance. The specific

activity of the enzyme towards succinonitrile was calculated and was found to be 0.22 umol NH_4^+/min/mg using 400mM succinonitrile in the reaction (figure 36). The very low specific activities and the apparent high Km values observed indicate that succinonitrile is not the best substrate for studying the enzyme kinetics. A wide range of nitriles is to be used for better understanding of the protein and for a proper kinetic experiment; however, the use of succinonitrile in this study was because the enzyme activity towards it had been the highest among the nitriles available in the lab.

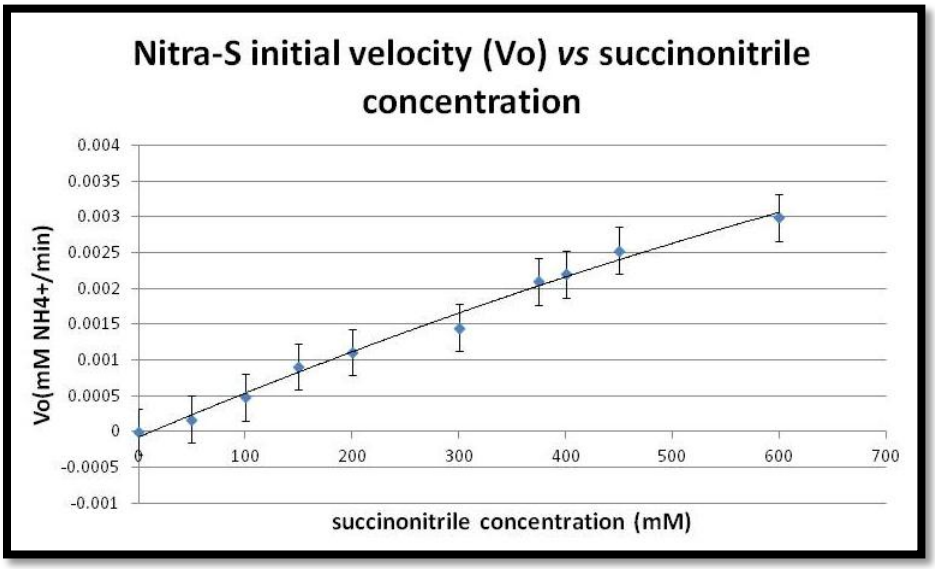

Figure 35. "Nitra-S" initial velocity (Vo) *vs* different substrate concentrations.

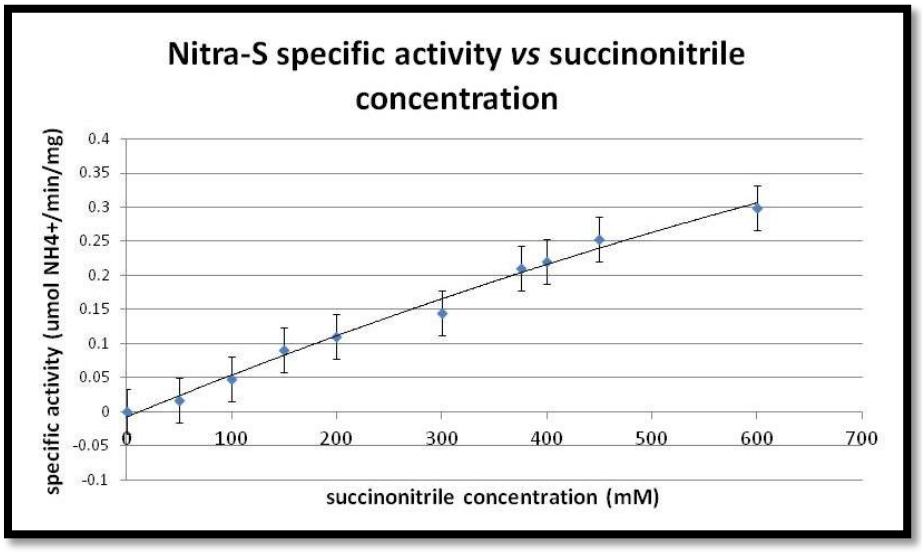

Figure 36. "Nitra-S" specific activity *vs* different substrate concentrations

7.2. Enzyme thermal stability and thermosensitivity:

Although it was highly expected that the isolated enzyme (Nitra-S) is to be thermostable, the activity of the enzyme decreased as the incubation temperature of the enzyme increased until it lost all remaining activity at 60°C as shown in figure 37. Upon examination of thermosensitivity, a sharp decrease in activity was observed as the reaction temperature increased (figure 38). Perhaps some components in the surrounding intracellular environment stabilize Nitra-S, thus it loses this gained thermal stability once purified or when it is no longer within the microorganism from which it was isolated. Another possibility is that native Nitra-S might be only expressed for a very short period to perform a specific function in a certain metabolic pathway.

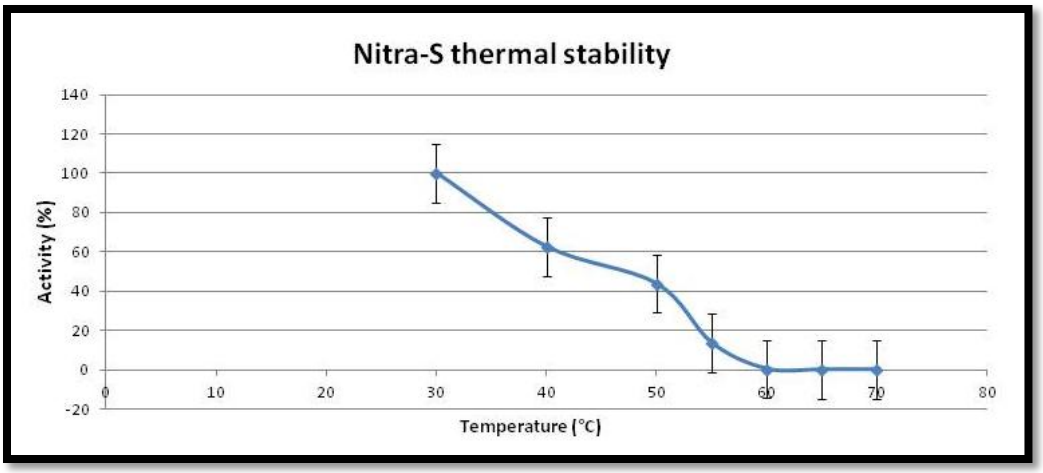

Figure 37. Evaluation of "Nitra-S" thermal stability. "Nitra-S" was incubated for one hour at different temperatures ranging from 30-70°C. The residual activity was determined and compared to the residual activity after 1-hour incubation at 30°C (assigned as 100%).

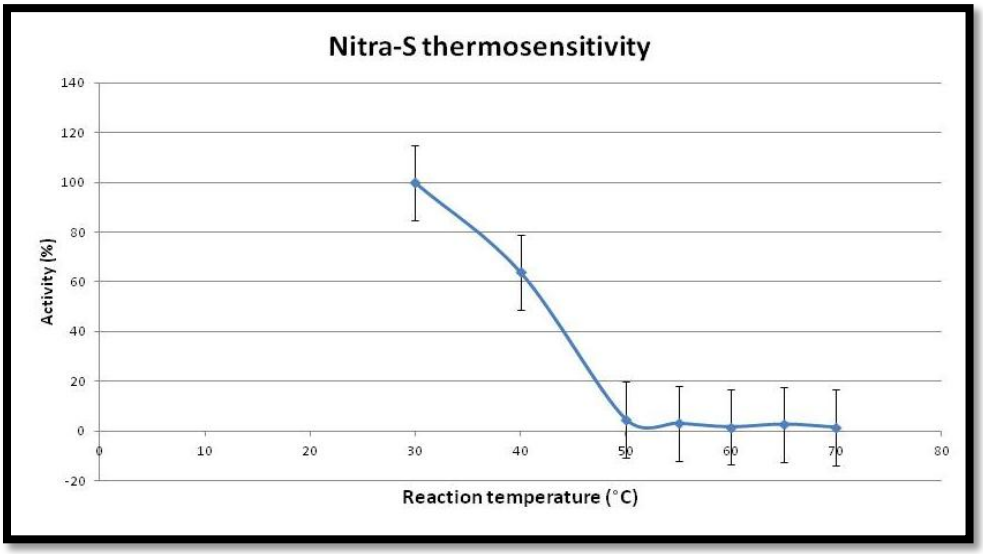

Figure 38. Evaluation of "Nitra-S" thermosensitivity. The activity assay was held at different temperatures ranging from 30-70°C. "Nitra-S" activity was assessed and compared to the activity at 30°C (assigned as 100%).

7.3. Enzyme halotolerance:

Examination of the stability of Nitra-S at different salt concentrations revealed that the activity of Nitra-S kept deceasing upon increasing the salt concentration as shown in figure 39. As the activity decreased with the increase in salt concentration, we can conclude that the identified protein is non-halophilic; however, the enzyme has shown a degree of halotolerance as its residual activity after 1-hour-incubation at 3.5M was 46.8% of that after incubation for the same time at 0.167M NaCl. Nevertheless, some studies reported different behaviour for some enzymes with different salts (55); thus, for thorough understanding of the halotolerance and halophilicity of the enzyme, different salts and higher molarities could be used in upcoming studies. Still, there is a possibility that Nitra-S might be derived from a halophilic microorganism, which can survive at high salt concentrations by accumulating organic osmolytes in the cytoplasm instead of the salting in mechanism. The later mechanism is a characteristic of extremely halophilic microorganisms, at which potassium ions accumulate inside the cell; and in this case, all the machinery within the microorganism should adapt to the high salt concentration in the internal environment. On the other hand, for halophilic microorganisms, which adapt to hypersaline environments by accumulating uncharged organic osmolytes, the enzymatic machinery could be as that observed in non-halophilic microorganisms (56).

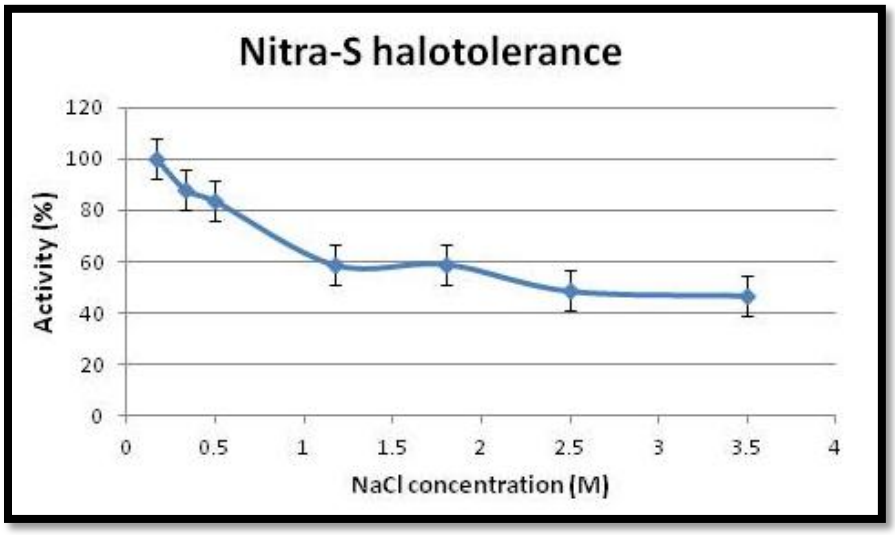

Figure 39. Evaluation of "Nitra-S"halotolerance. "Nitra-S" was incubated for one hour in different salt concentrations (0.167-3.5M). The residual activity was determined and compared to the residual activity after 1-hour incubation with 0.167M NaCl (assigned as 100%).

CHAPTER 4: CONCLUSIONS AND RECOMMENDATIONS

Identification of new nitrilases is of utmost importance due to the promises they hold in lots of applications. Mining the available metagenomic databases is an opportunity that has to be grasped by many researchers for the identification of such useful enzymes. Myriads of uncharacterized enzymes are pending on these databases, waiting for a chance to be studied and characterized. In this study, a nitrilase (Nitra-S) gene was identified, sequenced and cloned followed by expression and purification of the protein. Preliminary functional studies using a limited number of substrates revealed that the enzyme is active towards dinitriles. In addition, in contrast to our expectations, Nitra-S did not show tolerance towards high temperatures but it was halotolerant to some extent. Nevertheless, full characterization of Nitra-S is yet to be done. A larger number of substrates is needed to identify more specifically the substrate specificity of this enzyme. Besides, different nitrile isomers could be used to recognize the stereo- and regioselectivity of the identified nitrilase. Moreover, effect of the pH on Nitra-S activity is another important step that was not done in this study and needs to be analysed. Further investigation on salt tolerance using higher salt concentrations and other salts is an additional step that is to be studied. In this study, no attempts were done to identify the number of subunits of the active protein; size exclusion chromatography and native PAGE could help in that point. Generally, a lot of work and effort are needed to fully understand and characterize this newly identified nitrilase.

REFERENCES

1. Kaul P, Banerjee A, Banerjee UC. Nitrile Hydrolases. In: Polaina J, MacCabe AP, editors. Ind Enzym [Internet]. Springer Netherlands; 2007 [cited 2013 Feb 21]. p. 531–47. Available from: http://link.springer.com.library.aucegypt.edu:2048/chapter/10.1007/1-4020-5377-0_30

2. Gupta N, Balomajumder C, Agarwal VK. Enzymatic mechanism and biochemistry for cyanide degradation: A review. J Hazard Mater. 2010 Apr 15;176(1–3):1–13.

3. Robertson DE, Chaplin JA, DeSantis G, Podar M, Madden M, Chi E, et al. Exploring Nitrilase Sequence Space for Enantioselective Catalysis. Appl Environ Microbiol. 2004 Apr 1;70(4):2429–36.

4. Banerjee A, Sharma R, Banerjee UC. The nitrile-degrading enzymes: current status and future prospects. Appl Microbiol Biotechnol. 2002;60:33–44.

5. Kobayashi M, Yanaka N, Nagasawa T, Yamada H. Purification and characterization of a novel nitrilase of *Rhodococcus rhodochrous* K22 that acts on aliphatic nitriles. J Bacteriol. 1990 Sep;172(9):4807–15.

6. Yeom S-J, Kim H-J, Lee J-K, Kim D-E, Oh D-K. An amino acid at position 142 in nitrilase from *Rhodococcus rhodochrous* ATCC 33278 determines the substrate specificity for aliphatic and aromatic nitriles. Biochem J. 2008;415:401–7.

7. Raczynska JE, Vorgias CE, Antranikian G, Rypniewski W. Crystallographic analysis of a thermoactive nitrilase. J Struct Biol. 2011 Feb;173(2):294–302.

8. Layh N, Parratt J, Willetts A. Characterization and partial purification of an enantioselective arylacetonitrilase from *Pseudomonas fluorescens* DSM 7155. J Mol Catal B Enzym. 1998 Oct 15;5(5–6):467–74.

9. Cowan D, Cramp R, Pereira R, Graham D, Almatawah Q. Biochemistry and biotechnology of mesophilic and thermophilic nitrile metabolizing enzymes. Extermophiles. 1998;2:207–16.

10. Kim J-S, Tiwari MK, Moon H-J, Jeya M, Ramu T, Oh D-K, et al. Identification and characterization of a novel nitrilase from *Pseudomonas fluorescens* Pf-5. Appl Microbiol Biotechnol. 2009;83:273–83.

11. Bayer S, Birkemeyer C, Ballschmiter M. A nitrilase from a metagenomic library acts regioselectively on aliphatic dinitriles. Appl Microbiol Biotechnol. 2011;89:91–8.

12. Liu Z-Q, Zhou M, Zhang X-H, Xu J-M, Xue Y-P, Zheng Y-G. Biosynthesis of iminodiacetic acid from iminodiacetonitrile by immobilized recombinant *Escherichia coli* harboring nitrilase. J Mol Microbiol Biotechnol. 2012;22(1):35–47.

13. Stalker DM, McBride KE. Cloning and expression in Escherichia coli of a *Klebsiella ozaenae* plasmid-borne gene encoding a nitrilase specific for the herbicide bromoxynil. J Bacteriol. 1987 Mar;169(3):955–60.

14. Gong J-S, Lu Z-M, Li H, Shi J-S, Zhou Z-M, Xu Z-H. Nitrilases in nitrile biocatalysis: Recent

progress and forthcoming research. Microb Cell Factories. 2012;11(1):142.

15. Mueller P, Egorova K, Vorgias CE, Boutou E, Trauthwein H, Verseck S, et al. Cloning, overexpression, and characterization of a thermoactive nitrilase from the hyperthermophilic archaeon *Pyrococcus abyssi*. Protein Expr Purif. 2006 Jun;47(2):672–81.

16. Podar M, Eads J, Richardson T. Evolution of a microbial nitrilase gene family: a comparative and environmental genomics study. BMC Evol Biol. 2005 Aug 6;5(1):42.

17. Thuku RN. The structure of the nitrilase from *Rhodococcus rhodochrous* J1: homology modeling and three-dimentional reconstitution. [South Africa]: University of Western Cape; 2006.

18. Martinkova L, Kren V. Biotransformations with nitrilases. Current Opinion in Chemical Biology. 2010;14:130–7.

19. Gong J-S, Lu Z-M, Li H, Zhou Z-M, Shi J-S, Xu Z-H. Metagenomic technology and genome mining: emerging areas for exploring novel nitrilases. Appl Microbiol Biotechnol. 2013 Aug 1;97(15):6603–11.

20. Banerjee A, Kaul P, Sharma R, Banerjee UC. A High-Throughput Amenable Colorimetric Assay for Enantioselective Screening of Nitrilase-Producing Microorganisms Using pH Sensitive Indicators. J Biomol Screen. 2003;8:559–65.

21. Fawcett JK, Scott JE. A rapid and precise method for the determination of urea. J Clin Pathol. 1960;13:156–60.

22. Banerjee A, Sharma R, Banerjee UC. A rapid and sensitive fluorometric assay method for the determination of nitrilase activity. Biotechnol Appl Biochem. 2003 Jun 1;37(3):289.

23. Goyal SS, Rains DW, Huffaker RC. Determination of ammonium ion by fluorometry or spectrophotometry after on-line derivatization with o-phthalaldehyde. Anal Chem. 1988 Jan 15;60(2):175–9.

24. Yazbeck DR, Durao PJ, Xie Z, Tao J. A metal ion-based method for the screening of nitrilases. J Mol Catal B Enzym. 2006 May 2;39(1–4):156–9.

25. Seffernick JL, Samanta SK, Louie TM, Wackett LP, Subramanian M. Investigative mining of sequence data for novel enzymes: A case study with nitrilases. J Biotechnol. 2009;143:17–26.

26. Siam R, Mustafa GA, Sharaf H, Moustafa A, Ramadan AR, Antunes A, et al. Unique Prokaryotic Consortia in Geochemically Distinct Sediments from Red Sea Atlantis II and Discovery Deep Brine Pools. PLoS ONE. 2012 Aug 20;7(8):e42872.

27. Winckler G, Aeschbach-Hertig W, Kipfer R, Botz R, Rübel AP, Bayer R, et al. Constraints on origin and evolution of Red Sea brines from helium and argon isotopes. Earth Planet Sci Lett. 2001 Jan 30;184(3–4):671–83.

28. Antunes A, Ngugi DK, Stingl U. Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. Environ Microbiol Rep. 2011;3(4):416–33.

29. Atlantis II Deep (basin, Red Sea) [Internet]. Encycl. Br. [cited 2013 Oct 25]. Available from:

http://www.britannica.com/EBchecked/topic/41272/Atlantis-II-Deep

30. Schoell M, Hartmann M. Detailed temperature structure of the hot brines in the Atlantis II Deep area (Red sea). Mar Geol. 1973;14:1–14.

31. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. Nucleic Acids Res. 2011 Nov 29;40(D1):D290–D301.

32. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 1998 Mar;8(3):186–94.

33. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 1998 Mar;8(3):175–85.

34. Gordon D. Viewing and editing assembled sequences using Consed. Curr Protoc Bioinforma Ed Board Andreas Baxevanis Al. 2003 Aug;Chapter 11:Unit11.2.

35. Gordon D, Desmarais C, Green P. Automated finishing with autofinish. Genome Res. 2001 Apr;11(4):614–25.

36. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res. 1998 Mar;8(3):195–202.

37. Bprom tool [Internet]. Available from: http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb

38. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403–10.

39. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3--new capabilities and interfaces. Nucleic Acids Res. 2012 Aug;40(15):e115.

40. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999;41:95–8.

41. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970;227:680–5.

42. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000 Oct 1;16(10):944–5.

43. ClustalW2 [Internet]. Available from: http://www.ebi.ac.uk/Tools/msa/clustalw2/

44. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics. 2006;22:195–201.

45. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. Nucleic Acids Res. 2009;37:D387–D392.

46. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinforma Oxf Engl. 2011 Feb 1;27(3):343–50.

47. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein

Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. Proteomics Protoc Handb [Internet]. Humana Press; 2005 [cited 2013 Jul 1]. p. 571–607. Available from: http://link.springer.com.library.aucegypt.edu:2048/protocol/10.1385/1-59259-890-0%3A571

48. Bjellqvist B, Basse B, Olsen E, Celis JE. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. Electrophoresis. 1994 Apr;15(3-4):529–39.

49. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez JC, et al. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. Electrophoresis. 1993 Oct;14(10):1023–31.

50. Schmitt J, Hess H, Stunnenberg HG. Affinity purification of histidine-tagged proteins. Mol Biol Rep. 1993 Oct 1;18(3):223–30.

51. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, et al. Measurement of protein using bicinchoninic acid. Anal Biochem. 1985 Oct;150(1):76–85.

52. Ferrè F, Clote P. DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res. 2005 Jul 1;33(Web Server issue):W230–232.

53. Ferrè F, Clote P. DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. Nucleic Acids Res. 2006 Jul 1;34(suppl 2):W182–W185.

54. Ferrè F, Clote P. Disulfide connectivity prediction using secondary structure information and diresidue frequencies. Bioinforma Oxf Engl. 2005 May 15;21(10):2336–46.

55. Madern D, Ebel C, Zaccai G. Halophilic adaptation of enzymes. Extremophiles. 2000 Apr 1;4(2):91–8.

56. Kunte HJ, Truper HG, Stan-Lotter H. Halophilic microorganisms. In: Horneck G, Baumstark-Khan C, editors. Astrobiol Quest Cond Life. Koln, Germany: Springer; 2002. p. 185–200.