Identification, expression and characterization of genes encoding nitrilases from *Trichoderma reesei*

Rajeev Koundinya, MSc

This thesis is presented for the degree of
Doctor of Philosophy

Department of Chemistry and Biomolecular Sciences
Faculty of Science
Macquarie University, Australia
July 2013

Table of contents

Abstract

Nitrilases are attractive as industrial catalysts due to their ability to break down toxic nitriles into carboxylic acids and ammonia. Although nitrilases have been identified from several sources, identification of relevant nitrilases with attributes of enantio- and regioselectivity remains a challenge and research continues in this area. The current research explores the filamentous fungus *Trichoderma reesei* as a potential source of nitrilases.

Identification of genes encoding nitrilases in the *T. reesei* genome was performed *in silico* using bioinformatics and phylogenetic techniques. The *in silico* analysis enabled identification of four *T. reesei* genes that could potentially encode enzymes from branch 1 of the nitrilase superfamily; the gene *nit1* was judged as most likely to encode a nitrilase. To induce the expression of the nitrilase gene(s), *T. reesei* was grown in media supplemented with nitriles such as adiponitrile and benzonitrile. Due to the predicted intracellular nature of the enzymes, the mycelial protein was examined for the presence of nitrilases using a combination of approaches including nitrilase activity assays, high performance liquid chromatography (HPLC), 1 and 2-dimensional gel electrophoresis and mass spectrometry. Verification of the expression of the *nit1* gene by *T. reesei* grown in adiponitrile-supplemented minimal medium was achieved and the protein was found to be functionally active as a nitrilase, enabling the complete hydrolysis of adiponitrile to adipic acid.

Thereafter, to allow the recombinant expression of the *nit1* gene in *E. coli*, generation of the intron-less *nit1* complementary-DNA (cDNA) was attempted using reverse transcription polymerase chain reaction (RT-PCR) and overlap extension polymerase chain reaction (OE-PCR). The *nit1* cDNA was successfully obtained by RT-PCR using the total RNA extracted from *T. reesei* mycelia. The isolated *nit1* cDNA was cloned into a pET expression vector and transformed into two *E. coli* strains, Origami^{TM}B(DE3) and Tuner^{TM}DE3.

Expression of the *nit1* cDNA was successfully achieved in Origami™B(DE3) cells induced with 0.4 mM IPTG at 22 °C. Although the recombinant nitrilase appeared to aggregate in inclusion bodies in the cell, the protein could be solubilized and refolded to an enzymatically active form, as shown by nitrilase activity assays. The recombinant *T. reesei* nitrilase was purified using the Ni-NTA His-tagged protein purification system. Optimal activity of the nitrilase was achieved at 40 °C, pH 7.5, against the substrate adiponitrile. The enzyme could be of interest for future improvement and development for industrial processes involving the biocatalysis of adiponitrile such as in the manufacture of nylons and resins.

Declaration

I certify that the research presented in this thesis entitled “Identification, expression and characterization of genes encoding nitrilases from *Trichoderma reesei*” is an original work conducted by the author. Any form of assistance received for this work has been appropriately acknowledged. I also certify that this research material has not been submitted for any other degree to any other institution and to the best of my knowledge contains no information that has been previously written or published by any other person except where due reference is made in the text.

Rajeev Koundinya

July 2013

Abbreviations

bp	base pair
cDNA	complementary DNA
CDM	Czapek Doz medium
d	day
DEPC	diethylpyrocarbonate
DIG	didoxigenin
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
1D	one dimensional
2D	two dimensional
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
h	hour
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
kDa	kilodalton
LA	Luria Bertani agar
LB	Luria Bertani broth
MALDI	matrix assisted laser desorption ionization
MCS	multiple cloning site
min	minute
MM	minimal medium
MS	mass spectrometry
OE-PCR	overlap extension polymerase chain reaction
PCR	polymerase chain reaction
PDA	potato dextrose agar
RNA	ribonucleic acid
RT	room temperature
RT-PCR	reverse transcription polymerase chain reaction
sec	second
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
sp	species
TBE	tris borate EDTA
TE	tris EDTA
V	volt
v/v	volume per volume
w/v	weight per volume

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Professor Helena Nevalainen and Dr. Junior Te'o for granting me this opportunity, their invaluable time, support and guidance over the period of my PhD.

A super special thank you to two important people without whose support I would have been lost: firstly, Dr. Robyn Peterson, for all the excellent friendly help and mentorship that I have received during the course of my PhD especially during my thesis writing and secondly, Dr. Liisa Kautto, for her anytime technical assistance, brilliant laboratory management and never ending supply of delicious cakes which has kept me going during the course of my PhD. I would also like to express my acknowledgements to Dr. Moreland Gibbs, Professor Peter Bergquist and Dr. Anwar Sunna for their timely and expert inputs during the course of my PhD. Further thanks to Dylan Xavier from Australian Proteome Analysis Facility (APAF) for all the mass spectrometric analyses and Paul Worden from the Macquarie University Sequencing Facility for DNA sequencing.

I would also like to acknowledge Sridevi Muralidharan for all her timely advices and patient assistance for all my silly queries regarding the operation of mass spectrometry, IT support and EndNote and Dr. Shingo Miyauchi, my martial artist guru and good friend for helping me relax with hours of Kung Fu and Tai Chi. Many thanks to my colleague and friend Arun for all the discussions that we have had including hours of cricket both on field and on the computer over the past years and to Angela being inspirational and supportive whenever needed. I am also grateful to all the members of our laboratory including Vignesh for all the cricket and the good times we had, to Karthik for all the HPLC related assistance, to Suja, Clara, Andrew and Jashan for their friendship and providing a great work environment.

Above all and most importantly, I am deeply grateful to my parents to whom I dedicate this thesis, for all the encouragement, care and love which brought the best out of me. Thanks to my sister Joy for all her support and motivation throughout the period of my research. Completion of this thesis would have not been possible without them. And finally, I thank God for all his blessings.

Chapter 1: Introduction

Nitrile degrading enzymes are of great interest due to their considerable commercial value in the manufacture of pharmaceuticals and commodity chemicals. Nitrile degrading enzymes have also been widely used for bioremediation and biocatalysis of nitriles, due to their ability to break down harmful nitrile compounds. In addition, the enantioselective property of nitrile degrading enzymes makes them highly effective biocatalysts in the synthesis of enantiomerically pure compounds. Nitrile degrading enzymes have been broadly classified as nitrilases, nitrile hydratases and amidases. These enzymes have been identified and isolated from both plants and microorganisms with the majority of them isolated from bacterial species. Although nitrilases have been purified and characterized from certain species of filamentous fungi, there is considerable scope for further investigation into the production of nitrile degrading enzymes by known commercially exploited filamentous fungal species including exploration of the biochemical and biocatalytic properties of these nitrilases.

In the current work, *Trichoderma reesei* has been explored as a source of nitrile degrading enzymes. The purpose of this chapter is to provide information on nitrile degrading enzymes, their substrate nitriles, distribution of the nitrile degrading enzymes, *Trichoderma reesei* as a source of nitrilases and the heterologous expression of genes encoding nitrilases in *Escherichia coli*.

1.1 Introduction to nitriles

Nitrilase enzymes act by breaking down nitrile compounds to their respective carboxylic acids. Nitriles are organic compounds featuring the cyano functional group, "-C≡N" containing a carbon atom attached to a nitrogen atom by a triple bond. Hence nitrile compounds are usually denoted with the prefix "cyano-". Although nitrile compounds are known to be highly toxic due to the presence of the cyano group, they are found in plants, microorganisms and arthropods.

More than 3000 plants including several commercially important species are known to produce various forms of nitrile compounds such as cyanoglycosides, cyanolipids and phenylacetonitriles and are hence cyanogenic in nature (Piotrowski, 2008). Plants release toxic nitrile compounds during injury to their surface thus protecting them from herbivores. The nitrile compounds also play a pivotal role as biosynthetic intermediates in plant metabolism by helping the plants in their growth and development processes. For example, indole-3-acetonitrile, a nitrile compound identified in plants is the precursor for the plant growth hormone indole-3-acetic acid which helps in plant growth, ripening of fruits, root formation and plant cell elongation (Maor *et al*., 2004). In microorganisms, nitrile compounds are derived from the decarboxylation of amino-acids, which produces aldoximes. The aldoximes are converted to their corresponding nitriles through a complex dehydration pathway and used by the microorganisms for defense mechanisms and as a source of nitrogen (Fleming, 1999).

Although most nitrile compounds are known to be non-biodegradable, highly toxic and carcinogenic in nature, they have considerable commercial importance for the industrial generation of ketones, amines, amidines, carboxylic acids, esters and amides (Clarke and Read, 1941; Pollak *et al*., 2000). Nitrile compounds are widely used in the manufacture of several pharmaceuticals (as chiral synthons), polymers (nylon 66 polymers) (Banerjee *et al*., 2002), agro-chemicals (dichlobenil, ioxynil and Buctril) (Kobayashi *et al*., 1992) and are used as solvents in several rubber and plastic industries (Gupta *et al*., 2010). Owing to an ever-increasing demand, nitriles are commercially produced as chemical intermediates through reactions such as ammoxidation and hydrocyanation. Nitriles are also prepared by dehydration of oximes and primary amides, while aromatic nitriles are obtained from diazonium compounds by a chemical process called the Sandmeyer reaction (Clarke and Read, 1941).

One of the important applications of nitriles is in the production of carboxylic acids and amides by hydrolysis of nitrile substrates. Carboxylic acids obtained from the corresponding nitriles are used as precursors in the manufacture of several drugs and chemicals as discussed in Section 1.5. However, the chemical synthesis of carboxylic acids from nitriles has several drawbacks such as the high acidic or basic conditions required for the reaction, formation of large quantities of salts, difficulties in separation of the reaction products and high energy consumption (Nagasawa and Yamada, 1990; 1995). In recent years, enzymatic biocatalysis of nitriles to their corresponding acids has gained considerable importance. Enzyme-catalyzed hydrolysis is stereo-selective and nitrile re-racemization can yield optically pure conversions. Both the microbial nitrilase producing cells and their enzymes can be successfully immobilized thus enabling their reuse.

1.1.1 Biocatalysis of nitriles using nitrile degrading enzymes

Several industrial processes that traditionally have involved chemical catalysts can benefit from the use of enzymes, and biocatalysis has become a part of chemical biotechnology. Biocatalysis has several advantages over chemical catalysis such as improved product recovery, non-requirement of strong acidic or basic conditions for the reaction, and reduction in formation of toxic and unwanted by-products (Godtfredsen *et al*., 1985; Kim *et al*., 2009).

Chemical reactions involving nitriles attached to labile functional groups often require additional reaction steps for nitrile hydrolysis thus generating harmful waste. In contrast, nitrile hydrolysis reactions catalyzed by regio- and enantioselective enzymes are performed under controlled conditions and are therefore highly compatible with the labile functional groups attached to the nitriles. The reactions can be carried out at optimum temperatures and atmospheric pressures thus aiding isomerization and racemization without affecting the functional groups on the substrates.

Enzymatic biocatalysis helps in the production of single enantiomer intermediates which have considerable importance in the manufacture of drugs and chemicals (Patel *et al*., 2003) (Section 1.5.1). For example, derivatives of 1,4-benzodioxane-2-carboxylic acid are used extensively in several pharmaceutical companies. The failure to produce enantiopure 1,4-benzodioxane-2-carboxylic acid using chemical reactions has led to the production of considerable amount of wastes and by-products. The ability of nitrilase enzymes to act selectively on the chiral carbon at the 2-position without affecting its configuration has aided in the production of enantiopure 1,4-benzodioxane-2-carboxylic acid without significant wastage (Benz *et al*., 2007).

1.1.2 Nitrile metabolism pathways

Enzymes capable of metabolizing nitriles are found in several plants and microorganisms such as bacteria, fungi and yeasts (Section 1.4). The nitrile degrading enzymes are capable of breaking down nitriles into products such as formamide, formic acid, cyanohydrins, aldehydes and hydrocarbons (Figure 1.1). The different pathways depicted in Figure 1.1 are individually or collectively used by plants and microorganisms for their growth and metabolism. For example, the pathway involving oxygenases used exclusively by plants and insects to produce aldehydes are absent in most microorganisms (Johnson *et al*., 2000).

The nitrile metabolism pathway that is significant to this research is hydrolysis of organonitriles to carboxylic acids or amides (chiral α- hydroxycarboxylic acids or α-hydroxycarboxamides) either by nitrilases or by nitrile hydratases (NHase) and amidases (highlighted box in Figure 1.1) (Rustler *et al*., 2008).

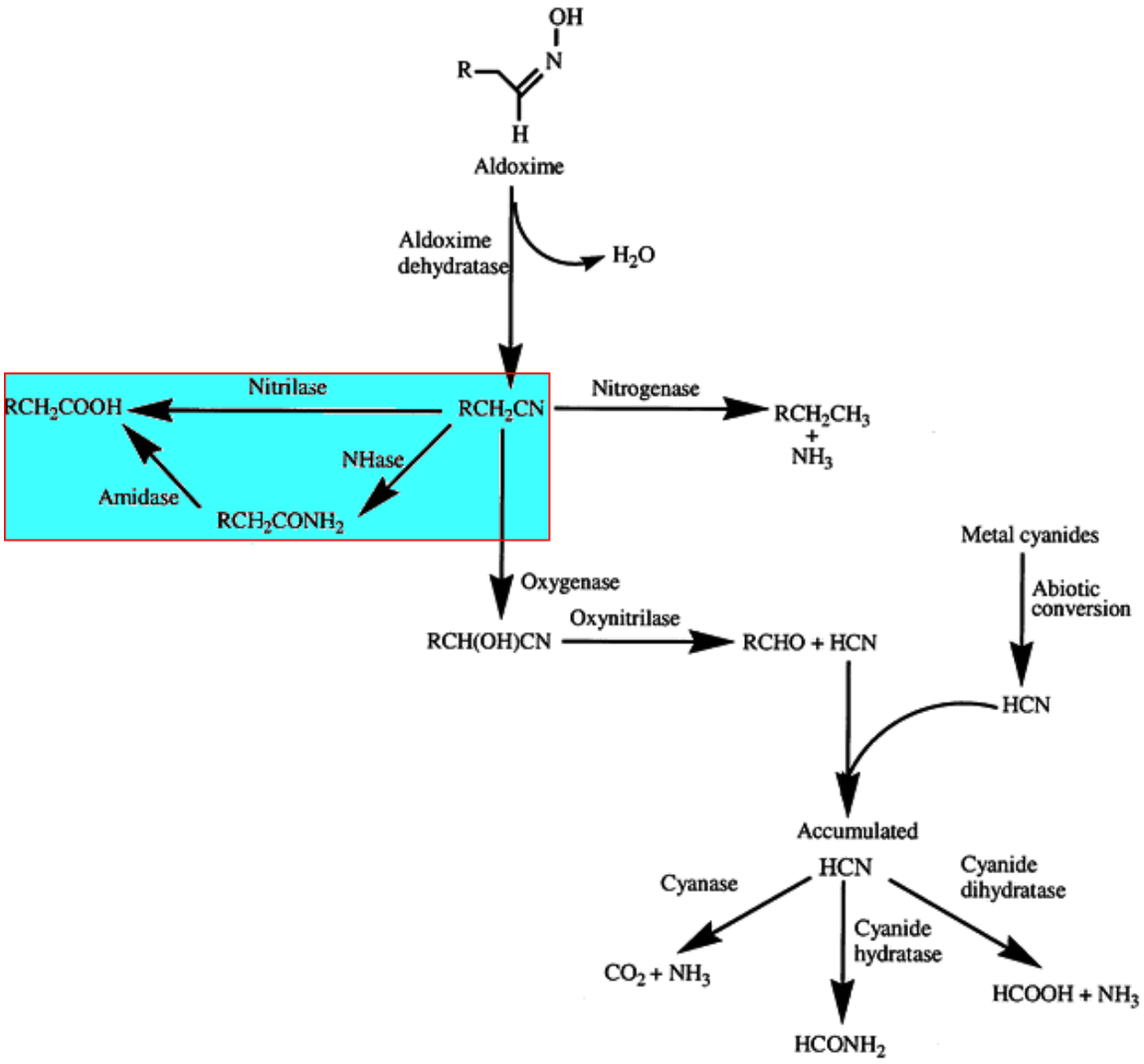


Figure 1.1 Nitrile metabolism pathways (modified from Banerjee *et al.*, 2002). The boxed area highlights the pathways significant to the current research involving enzymatic hydrolysis of nitriles (RCH_2CN) by nitrilases, nitrile hydratases (NHase) and amidases.

Nitrilases catalyze the conversion of nitriles to their respective carboxylic acids and ammonia while nitrile hydratases metabolize nitriles to amides which are converted to their corresponding carboxylic acids by amidases. Nitrilases and amidases are hydrolases with a relatively simple catalytic process, are independent of a coenzyme requirement and form a part of the nitrilase superfamily discussed below.

1.2 The nitrilase superfamily

The nitrilase superfamily contains several groups of thiol enzymes, collectively termed "C-N hydrolases", that are responsible for a variety of carbon-nitrogen hydrolysis reactions (Bork

and Koonin, 1994). Based on amino acid sequence analysis, the nitrilase superfamily has been divided into thirteen branches: 1) Nitrilases, 2) Aliphatic amidases, 3) Amino-terminal amidases, 4) Biotinidases, 5) β-ureidopropionases, 6) Carbamylases, 7) Prokaryote NAD synthetases, 8) Eukaryote NAD synthetases, 9) Apolipoprotein N-acyltransferases, 10) Nit; 11) Nitfhit; 12) NB11 and 13) NB12 Non-fused outliers (Pace and Brenner, 2001; Thuku *et al*., 2009). The first branch of the superfamily encompasses nitrilases (the focus of this thesis), cyanide hydratases and cyanide dihydratases. The remaining twelve branches predominantly contain different forms of amidases. Carbamylases decarbamylate D-amino acids and NAD-synthetases catalyze amide nitrogen transfer by utilizing either glutamine or ammonia.

Although members of the same branch of the nitrilase superfamily share considerable sequence similarity, the similarity drops to a maximum of 24% between branches. Most plants, animals and microorganisms have enzymes belonging to more than one branch of the superfamily. However, certain branches have been identified solely from prokaryotes; based on sequence and phylogenetic analyses, there is a strong possibility that some of the branches emerged prior to the evolution of plants, animals and fungi (Pace and Brenner, 2001).

All members of the nitrilase superfamily contain a highly conserved active site known as the catalytic triad containing cysteine, glutamate and lysine residues. Cysteine acts as a nucleophile, lysine stabilizes the tetrahedral intermediate and the glutamate acts as a general base catalyst (Wang *et al*., 1999; Nakai *et al*. 2000; Pace and Brenner, 2001; Hung *et al*., 2007). Each branch of the superfamily is further characterized by a specific highly conserved motif containing two amino acid residues present at the carboxy-terminus of the active-site cysteine residue. Members of the nitrilase branch of the superfamily, have a motif containing the active-site cysteine, tryptophan and glutamic acid residues (Andrade *et al*., 2007).

Members of the nitrilase superfamily share close sequence and structural similarities with certain metalloenzymes known as nitrile hydratases. Although nitrile hydratases breakdown

nitriles, they lack the catalytic triad conserved in the nitrilase superfamily. Instead they have a conserved metal binding domain which plays a pivotal role in the catalysis process. This metal binding domain is absent in all members of the nitrilase superfamily (Marron *et al*., 2012).

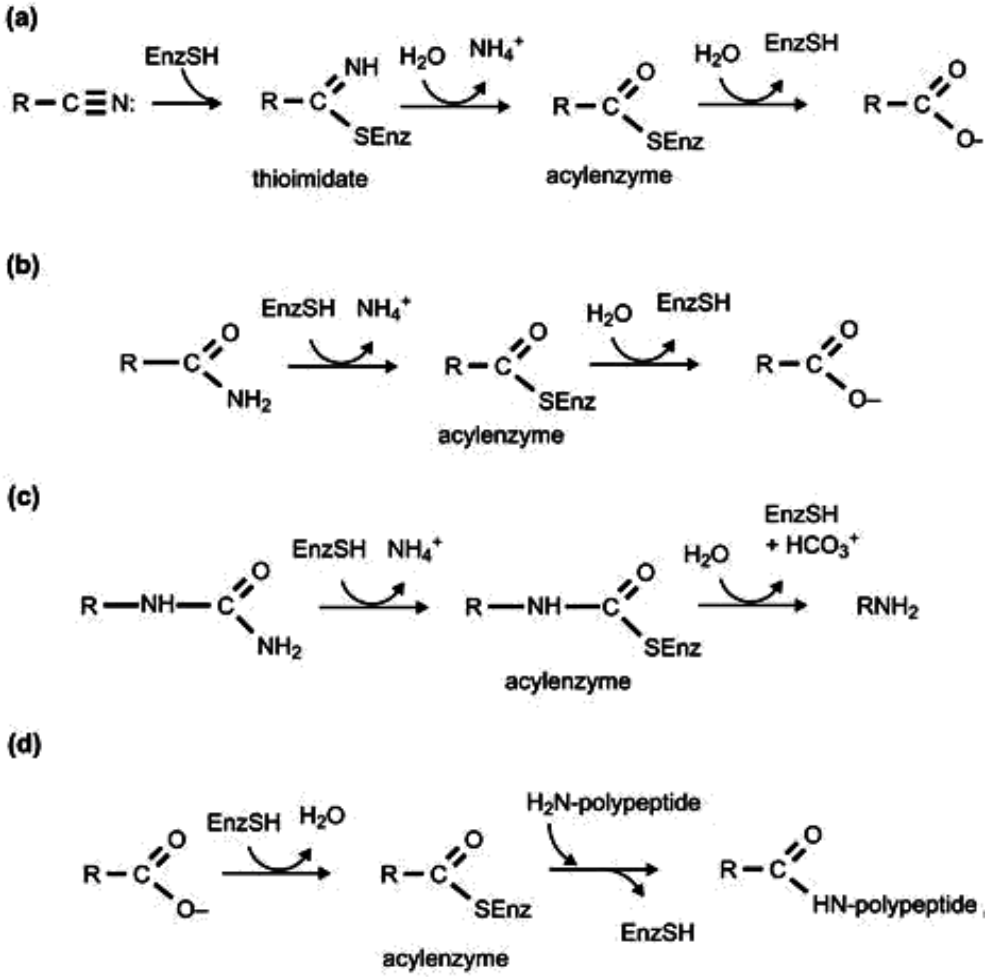


Figure 1.2 Four major types of reactions carried out by nitrilase superfamily members. (a) reactions performed by branch 1 nitrilases; (b) reactions performed by enzymes in branches 2-4 and branches 7-8 which include all amidases and NAD synthetases; (c) carbamylase reactions carried out by branch 5 and 6 enzymes; (d) reactions performed by branch 9-13 N-acyltransferases (Pace and Brenner, 2001).

Four types of reactions are performed by the members of the nitrilase superfamily (Figure 1.2; Pace and Brenner, 2001). The reaction that is most relevant to this work is the hydrolysis of nitriles to carboxylic acids (Figure 1.2a, explained further in Section 1.2.5); this reaction is carried out only by the nitrilases of the nitrilase branch (branch 1) of the nitrilase superfamily, hereafter simply referred to as "nitrilases" in this thesis.

1.3 Nitrilases

The first enzyme from the nitrilase superfamily to be discovered was a nitrilase from barley leaves (Thimann and Mahadevan, 1964). By the year 2004, another twenty nitrilases had been discovered (Robertson *et al*., 2004). However, in spite of their commercial potential, nitrilases have largely been under-utilized in industry due to lack of sufficient knowledge about their structure, function and stability. A recent rise in demand for stable, enantioselective nitrilases appropriate for environmentally sound manufacture of various chemicals and pharmaceuticals and bioremediation (detailed in Section 1.5) has increased interest in these enzymes. In the following sections the current state of knowledge about the structure, reaction mechanisms and distribution of nitrilases is reviewed and assessed.

1.3.1 Structural features of nitrilases

Although nitrilases have provoked industrial interest and research activity, very little information on the structure of nitrilases is available today. The structure of only one nitilase has been studied to date, namely that of the thermophilic bacterium *Pyrococcus abyssi* (Raczynska *et al*., 2011). However, more information is available about the structure of other members of the nitrilase superfamily with which nitilases are thought to share high structural homology (Sewell *et al*., 2005; Thuku *et al*., 2009). Hence, based on the solved structures of these other superfamily members, inferences have been made on the general protein structure of nitrilases such as the protein fold and conserved domains.

All nitrilases are thought to share the characteristic monomeric fold and catalytic active site containing conserved cysteine, glutamate and lysine residues that are typical of every member of the nitrilase superfamily (Barglow *et al*., 2008). The monomers are known to associate to form homodimeric building blocks having a *αββα*- *αββα* sandwich fold. The catalytic triad, α helices and the *β* sheets of the prokaryotic Nit protein from the bacterium *Xanthomonas campestri* (Chin *et al*., 2007) are shown in Figure 1.3.

a) **b)**

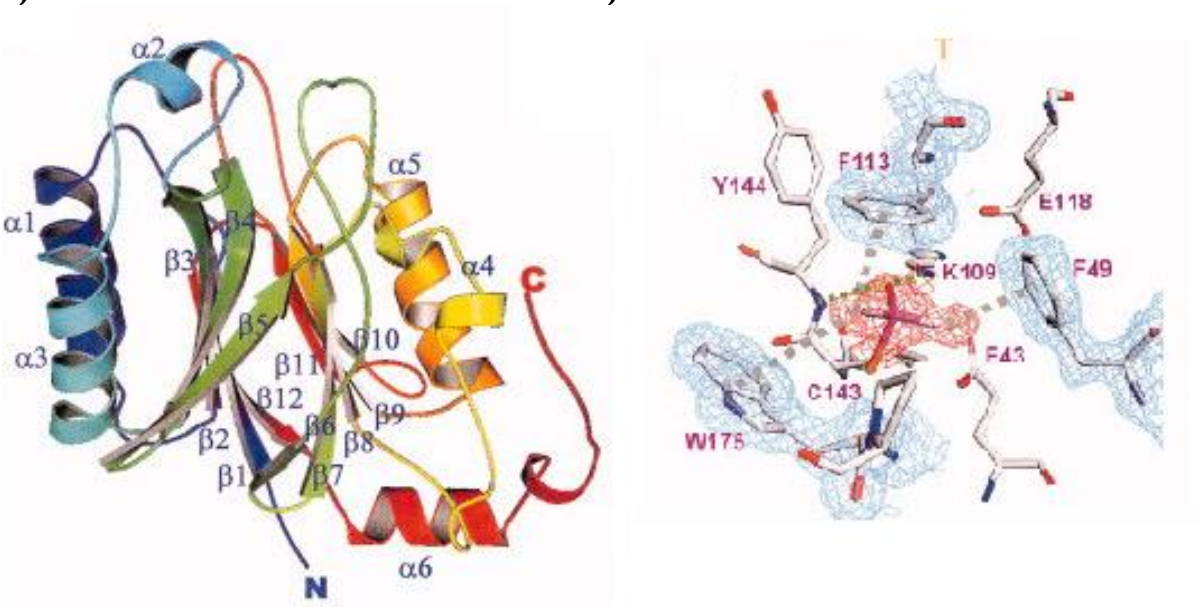


Figure 1.3 Structure of nitrilase from *Xanthomonas campestri* (Chin *et al*., 2007). (a) Tertiary structure of the *X. campestri* nitrilase monomer. The blue color denotes the N-terminus and the red denotes the C-terminus; (b) The active site of the nitrilase. C143, K109 and E43 denote the catalytic triad (cysteine, lysine and glutamic acid).

Nitrilases are generally inactive in their monomeric form, with the only known exceptions being nitrilases from *Arthrobacter sp*. strain J1 (Bandyopadhyay *et al*., 1986) and *Rhodococcus rhodochrous* PA-34 (Bhalla *et al*., 1992) which are active as monomers. Most nitrilases, along with the other branches of the superfamily, exist as oligomeric spirals with variable length helices. The formation of oligomers leads to conformational changes in the enzyme allowing the substrate molecule to attach to the enzyme's active site. Although the average molecular weight of a single nitrilase polypeptide ranges between 32 – 45 kDa, aggregates ranging between 6 – 26 subunits have been recorded (O'Reilly and Turner, 2003) (Table 1.1). Oligomerization is known to occur in response to heat treatment, the availability of particular substrates, and in the presence of salts such as ammonium sulphate (Harper, 1977; Harper, 1985; Stevenson *et al*., 1992; Nagasawa *et al*., 2000; Thuku *et al*., 2009).

Table 1.1 Molecular masses of nitrilases in monomeric and active oligomeric form.

Organism	Monomer (kDa)	Oligomer (kDa)	Reference
Alcaligenes faecalis ATCC 8750	32	460	Yamamoto *et al*., 1991
Alcaligenes faecalis JM3	44	260	Nagasawa and Yamada, 1990
Fusarium solani IMI196840	76	620	Harper, 1977
Nocardia NCIB11216	45	560	Hoyle *et al*., 1998
Rhodococcus ATCC39484	40	560	Stevenson *et al*., 1992
Fusarium oxysporum	37	170 – 880	Goldlust and Bohak, 1989
Rhodococcus rhodococcus K22	41	650	Kobayashi *et al*., 1989

Oligomerization gives rise to several isoforms of the enzyme, often resulting in increased functionalities and a broader range of substrate activity. For example, ZmNit1 and ZmNit2 are the two isoforms of the nitrilase enzyme in maize. They form a complex which is able to synthesize indole-3-acetic acid as well as hydrolyze β-cyanoalanine (Kriechbaumer *et al*., 2007).The nitrilase enzymes are distinct from other members of the nitrilase superfamily due in part to an additional 35 – 100 amino acids in their C-terminus (Sewell *et al*., 2005; Thuku *et al*., 2009). The extended C-terminal region has been shown to play an important role in the process of oligomerization (Sewell *et al*., 2005; Lundgren, 2008). The extended C-terminal region has also been suggested to be critical to the properties of the enzyme. A deletion of around 47 and 75 amino acids of *Pseudomonas fluorescens* nitrilase decreased the overall enzyme activity and stability. The deletion also increased the affinity towards a particular enantiomer of mandelonitrile (Kiziak *et al*., 2007). More details about the properties of nitrilases are provided in the following section.

1.3.2 Properties of nitrilases

Nitrilases are classified as aromatic, aliphatic or arylacetonitrilases on the basis of substrate specificity (Table 1.2). Most aromatic nitrilases are capable of hydrolyzing only aromatic nitriles such as benzonitrile, with little or no activity detected on aliphatic nitrile substrates. The enzyme undergoes a conformational change whenever an aromatic substrate binds to the active site thus boosting enzyme activity. Maximum activity is observed when the nitrile substrates have a methyl or halogen group present at the *meta*- and *para*- position. However, the activity of the enzymes drops in the case of substitutions at the *ortho*- position due to steric hinderance (Stevenson *et al*., 1992).

Table 1.2 Classification of nitrilases on the basis of substrate specificity.

Substrate Class	Organism	Optimal pH	Optimal Temp (°C)	Reference
Aromatic	*Fusarium solani* IMI 196840	7.8-9.1	40-45	Harper 1977a
	Fusarium oxysporum	6-11	40	Goldlust and Bohak, 1989
	Aspergillus niger K10	8	45	Kaplan *et al*., 2006
	Arthrobacter strain J1	8.5	40	Bandyopadhyay *et al*., 1986
Aliphatic	*Acidovorax facilis* 72W	8-9	65	Gavagan *et al*., 1999
	Comamonas testosterone	7	30	Levy-Schil *et al*., 1995
	Bradyrhizobium japonicum	7-8	45	Zhu *et al*., 2008
	Synechocystis PCC6803	7	50	Heinemann *et al*., 2003
	Pyrococcus abyssi	7.4	100	Mueller *et al*., 2006
Arylacetonitrilase	*Alcaligenes faecalis* ATCC 8750	7.5	45	Yamamoto *et al*., 1991
	Pseudomonas fluorescens DSM 7155	9	55	Layh *et al*., 1998
	Pseudomonas putida	7	40	Banerjee *et al*., 2006

A few aromatic nitirlases are known to be active on non-aromatic nitriles (Heinemann *et al*., 2003). For example, an aromatic nitrilase purified from *R. rhodochrous* NCIMB11216 was active on aliphatic nitriles such as propionitrile and a nitrilase from *R. rhodochrous* J1 could hydrolyze acrylonitriles (Hoyle *et al*., 1998; Nagasawa *et al*., 2000).

Aliphatic nitrilases hydrolyze aliphatic nitriles to their respective carboxylic acids. Aliphatic nitrilases in general have broad substrate specificities. For example, aliphatic nitrilase from *Synechosystis* sp. strain PCC6803 was reported to be active on several aromatic, heterocyclic and aliphatic nitriles. Addition of certain organic solvents such as methanol at various concentrations increased the activities of aliphatic nitrilases on long chain hydrophobic aliphatic nitriles. The activities of the aliphatic nitrilases depend upon the nature and size of the substrates (Heinemann *et al*., 2003).

Arylacetonitrilases are the third class of nitrilases. This class of nitrilases is highly enantioselective on account of the additional functional groups attached to the arylacetonitriles and is used to obtain optically pure products such as R-(-) mandelic acid (Yamamoto *et al*., 1991). Most known arylacetonitrilases form little or no amide by-products and have no activity towards either aliphatic or aromatic nitriles. For example, a nitrilase from *P. fluorescens* sp. was found to be significantly enantioselective towards arylacetonitriles with no activity towards benzonitrile (Layh *et al*., 1998).

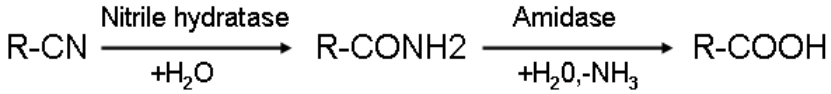
Most nitrilases are active over a wide pH range, from 5 – 10 (Table 1.2). Nitrilases from both bacterial and fungal sources are known to be moderately thermostable with an activity range of 30 – 50 ºC (Martinkova *et al*., 2009). In addition, certain mesophilic and thermophilic nitrilases with activity over a wide temperature range have been identified (Cowan *et al*., 1998; Almatawah and Cowan, 1999; Kabaivanova *et al*., 2005, Mueller *et al*., 2006). For example, a thermostable nitrilase from *Bacillus pallidus* Dac52 hydrolyzes 3-cyanopyridine at 60 ºC and pH 8 with no trace of nicotinamide being formed. The enzyme was reported to

retain full activity at temperatures ranging from 50 ºC – 60 ºC (Cowan *et al*., 1998). The mechanism of the nitrilases to act on a variety of nitrile substrates under different reaction conditions to yield optically active α- hydroxyl and α- fluoro acids is discussed in the following section (Mylerova and Martinkova, 2003).

1.3.3 Nitrilase reaction mechanisms

Biocatalysis of nitriles to carboxylic acids proceeds either via a bi-enzymatic or mono-enzymatic pathway. In the bi-enzymatic pathway, a nitrile hydratase first catalyses the conversion of the nitrile to the corresponding amide (R-C=O(NH_2)), then an amidase hydrolyzes the amide to the corresponding carboxylic acid (-COOH) with the release of ammonia (Figure 1.4a). In the monoenzymatic pathway, a nitrilase directly hydrolyzes the nitrile substrate to carboxylic acid and ammonia by the successive addition of two molecules of water (Figure 1.4b).

a)



b)

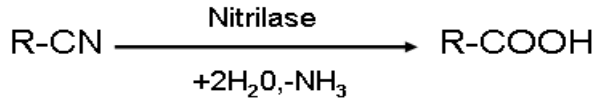


Figure 1.4 (a) Biocatalysis of nitriles using the bi-enzymatic pathway involving nitrile hydratase and amidase; (b) Biocatalysis of nitriles using the mono-enzymatic pathway involving nitrilase.

Most nitrilases prefer bulky substrates with additional functional groups as the enzymes bind by placing the 2 carbon closer to 120º than to 180º from the cyano nitrogen (Harper, 1985; Schmidt *et al*., 1996). The reaction is initiated by a nucleophilic attack on the nitrile substrate by cysteine in the catalytic triad to form a tetrahedral structure (Pace and Brenner, 2001; Stevenson *et al*., 1992). The glutamate from the catalytic triad aids in the transfer of a proton to produce ammonia. The tetrahedral structure is stabilized by lysine, the final member of the

catalytic triad. A second tetrahedral structure is formed after the nucleophilic attack by water on the thioester intermediate which is formed by the release of ammonia. This tetrahedral structure then breaks down releasing the corresponding carboxylic acid. A thioimidate formed after the first nucleophilic attack is hydrolyzed to give ammonia and an acyl intermediate. This acyl intermediate prevents further hydrolysis of the substrate to carboxylic acid. It has been suggested that a second glutamate residue acts as a general base to help catalyze hydrolysis of the acyl intermediate (Kimani *et al*., 2007; Thuku *et al*., 2009). In reactions involving substrates with weak scissile bonds, the thiomidiate is converted to formamide instead of the acyl intermediate. Therefore, traces of amides have been observed in several nitrilase catalyzed reactions. In reactions involving nitrilases from plants, rate of formation of amide by-products is considerably higher (Goldlust and Bohak, 1989; Kaplan *et al*., 2006; Kiziak *et al*., 2005; Layh *et al*., 1998; Thuku *et al*., 2009; Wajant and Effenberger, 2002).

1.4 Distribution of nitrilases

Genes encoding nitrilases have been identified from a variety of plants and microorganisms. After the first discovery of a nitrilase in barley leaves (Section 1.3), the first microbial nitrilase was isolated from a soil bacterium (*Pseudomonas* species) (Hook and Robinson, 1964). Since then, nitrilases have been identified and purified from several plants and microbial species as described in the following sections (Harper, 1976; O'Reilly and Turner, 2003).

1.4.1 Nitrilases from plants

Nitrilases have been successfully purified from several plant species including *Arabidopsis thaliana* (Piotrowski *et al*., 2001), *Brassica rapa* (Chinese cabbage; Rausch and Hilgenberg, 1980), and *Hordeum vulgare* (barley; Thimann and Mahadevan, 1964). As mentioned earlier, the nitrilase enzymes in plants convert indole-acetonitrile to indole-3-acetic acid, better

known as auxins. The auxins play a pivotal role in plant growth by producing growth promoting hormones, assisting in cell elongation and root formation. Besides assisting in growth and metabolism, nitrilases aid in the process of cyanide detoxification by converting β-cyanoalanine to useful metabolites such as asparagine, aspartic acid and ammonia. For example, nitrilase from *S. bicolor* metabolizes cyanide glycoside to carboxylic acids without releasing toxic cyanide compounds (Piotrowski, 2008).

1.4.2 Microbial nitrilases

A wide range of microorganisms maintaining an ecological relationship with the flora and fauna of their habitat are known to produce nitrilases. These nitrilases play an important role in the detoxification of harmful cyanide compounds released by plants as well as in synthesis of nitrogen by hydrolysis of nitriles (Pace and Brenner, 2001).

Since purification of a nitrilase from *Pseudomonas* sp. in 1964, nitrilases have been successfully identified and purified from several bacterial species such as *Rhodococcus*, *Alcaligenes*, *Bacillus* and *Pseudomonas* (Harper, 1976; O'Reilly and Turner, 2003) (Table 1.3). A large scale screening for genes encoding nitrilases was undertaken by a commercial group, Diversa Corporation, U.S.A (Mathur *et al*., 2005). Nucleic acids were extracted from microorganisms present in environmental samples such as soil, sediments and leaf litter and a metagenome library was created. Out of approximately 600 environmental DNA libraries, around 137 genes encoding nitrilase enzymes were identified (DeSantis *et al*., 2002; Robertson *et al*., 2004; Mathur *et al*., 2005). The gene library has been utilized to identify, produce and purify nitrilases from organisms such as *B. xenovorans* and *B. japonicum* (Seffernick *et al*., 2009).

Today nitrilase genes are known to be spread across various bacterial, fungal and yeast species. Examples of bacteria, plants, fungi and yeasts that produce nitrilases are shown in Table 1.3.

Table 1.3 Organisms with well characterized nitrilase activity.

Source	Organism	Reference
Plants	*Arabidopsis thaliana*	Piotrowski *et al*., 2001
	Hordeum vulgare	Thimann and Mahadevan, 1964
	Brassica rapa	Rausch and Hilgenberg, 1980
Fungi	*Fusarium oxysporum*	Goldlust and Bohak, 1989
	Fusarium solani IMI196840	Harper, 1977a
	Penicillium multicolor	Kaplan *et al*., 2006
	Aspergillus niger K10	Kaplan *et al*., 2006
Bacteria	*Bacillus* sp. strain OxB-1	Kato *et al*., 2000
	Pseudomonas sp. (SI)	Dhillon *et al*., 1999
	Acido vorax facilis 72W	Gavagan *et al*., 1999
	Bacillus pallidus Dac521	Almatawah *et al*., 1999
	Alcaligenes faecalis JM3	Nagasawa and Yamada, 1990
Yeast	*Candida famata*	Linardi *et al*., 1996
	Candida guilliermondi	Dias *et al*., 2000

Considerably more work has been carried out on nitrilases of bacterial and plant origins than on nitrilases from fungi. Although the first fungal nitrilase from *Fusarium solani* was described in 1977 (Harper, 1977b), little attention has been given to fungal nitrilases until recently, thus there is scope for further research in the field of nitrilases from fungi (Martinkova *et al*., 2009).

1.4.3 Nitrilases from filamentous fungi

Although most of the nitrilases identified until recently originate from bacteria, screening approaches such as the use of nitrilase specific selection media (Asano, 2002), metagenomic

analysis of DNA libraries to screen for genes encoding nitrilases (Robertson *et al*., 2004) and database mining (Seffernick *et al*., 2009) have shown that filamentous fungi are also a rich source of nitrilase enzymes (Martinkova *et al*., 2009). Fungi are now known to contain genes that code for nitrilases, cyanide hydratases and amidases. Fungal nitrilases are used to breakdown nitriles which are utilized by fungi as nitrogen source. Cyanide hydratases on the other hand are used by fungi to detoxify cyanide compounds released by plants after injury.

Although the first fungal nitrilase enzyme was identified from *Fusarium solani* by Harper (1977), the structure and amino acid sequence of the enzyme were not studied until recently (Petrickova *et al*., 2012). Moreover, in spite of considerable research being carried out on nitrilases from several fungal species such as *Fusarium*, *Aspergillus* and *Penicillium*, the catalytic properties and functions of several nitrilases from filamentous fungi are still unknown (Martinkova *et al*., 2009). The first fungal nitrilase that was purified and characterized both as its native form (Kaplan *et al*., 2006) and as a recombinant enzyme produced in *E. coli* was from *Aspergillus niger* (Kaplan *et al*., 2011) . The properties of the identified fungal nitrilases are discussed below.

1.4.3.1 Properties of fungal nitrilases

All the known fungal nitrilases are inducible enzymes, as are most nitrilases from bacteria. Although the transcriptional mechanism of induction for genes encoding nitrilases from bacteria such as *R. rhodochrous* has been studied (Komeda *et al*., 1996), nitrilase induction is yet to be fully understood in filamentous fungi. However, research has revealed that the properties and stabilities of fungal nitrilases depend upon the length and position of the functional groups attached to nitriles used for nitrilase induction. For example, the specific activity of a nitrilase from *A. niger* was increased by about 400-fold when 2-cyanopyridine was used as an inducer as compared to 3-cyanopyridine (Kaplan *et al*., 2006).

Fungal nitrilases are active on a wide range of substrates with their maximum activity being reported on aromatic nitrile substrates such as benzonitrile and mandelonitrile. For example, an aromatic nitrilase from *F. solani* O1 was active on both branched and substituted aliphatic nitriles such as isobutyronitrile and 2-phenylpropionitrile (Vejvoda *et al*., 2008). Furthermore, fungal nitrilases are reported to be diastereoselective with the ability to differentiate between *cis*- and *trans*- isomers of nitriles. For example, the *cis*- isomer of γ- substituted cyclopentane nitrile was hydrolyzed by aromatic fungal nitrilases from *F. solani* and *A. niger* whereas the *trans*- isomer was not hydrolyzed by either enzyme (Winkler *et al*., 2009). This ability of the fungal nitrilases to selectively hydrolyze one isomeric form of a substrate can be utilized commercially to separate isomeric mixtures.

Although arylaliphatic nitrilases are known to be enantioselective towards their substrates, the enantioselectivity of aromatic nitrilases is yet to be fully examined. Aromatic nitrilases from both *F. solani* and *A. niger* are reported to be moderately enantioselective (Kaplan *et al*., 2006). In comparison to bacterial nitrilases available commercially, enantioselectivities of fungal nitrilases are relatively low for industrial applications. Further research on the effect of different nitrile substrates on fungal nitrilase enantioselectivity might assist in understanding the mechanisms involved, and could potentially lead to increased enantioselective biotransformations.

Fungal nitrilases are comparatively more thermostable than bacterial nitrilases with an active temperature range of 30 – 50 °C. They appear more stable at higher temperatures when compared to bacterial nitrilases, with the exception of the thermostable nitrilase from *P. abyssi* (Mueller *et al*., 2006). For example, the half life of nitrilase from *B. subtilis* dropped from 92.6 h to 24 h with an increase in temperature from 30 to 35 °C. In contrast, nitrilase from *F. solani* was comparatively stable at this temperature. Fungal nitrilases like most bacterial nitrilases are active in a pH range from 5 to 8. The pH and temperature optimum of the enzymes vary on the basis of the substrates used in their production. Several techniques

such as immobilization of the nitrilases are used to boost nitrilase stabilities. The nitrilase enzymes immobilized in gel matrices (polyvinyl alchohol-polyethylene glycol copolymer) were reported to be stable even at pH 11. However, no significant effect of immobilization was observed on the enzyme thermostability (Vejvoda *et al*., 2006).

Nitrilases obtained from *A. niger*, *F. solani* and *Penicillum* have shown great promise, indicating the potential of fungal nitrilases as commercially relevant enzymes in terms of stability and substrate specificity. In addition, there is added scope for research into new fungal sources of nitrilases with attributes of enantioselectivity and chemoselectivity.

1.4.3.2 *Trichoderma reesei* as a source for nitrilases

Trichoderma reesei is an industrially important mesophilic filamentous fungus that is used as a source as well as an expression host for several gene products. *T. reesei* has been certified as a safe organism for commercial usage (GRAS classified- Generally Regarded As Safe) by the Organization of Economic and Commercial Development (OECD). *T. reesei* is one of the most well studied of the filamentous fungi with considerable amount of research carried out on its distribution, classification, defense mechanisms, growth, production of enzymes and secretory pathways (Nevalainen *et al*., 1994; Kuhls *et al*., 1996; Gams and Bissett, 1998; Druzhinina *et al*., 2006; Druzhinina *et al*., 2012; Li *et al*., 2012). The genomes of *Trichoderma reesei*, *T. virens* and *T. atroviride* have been sequenced by the Joint Genome Institute (http://genome.jgi.doe.gov; Grigoriev *et al*., 2012).

T. reesei Rut-C30 is a high protein secreting, low protease hypercellulolytic mutant. With an ability to secrete proteins in the range of 40 – 100 grams per litre of supernatant, *T. reesei* Rut-C30 is used to produce several industrially significant enzymes such as cellulases and xylanases (Durand *et al*., 1988; Cherry and Fidantsef, 2003). Recombinant expression of both endogenous and heterologous proteins is carried out in *T. reesei* Rut-C30, typically under the *cbh1* (cellobiohydrolase I) gene promoter which is considered to be one of the strongest

inducible promoters in *T. reesei* (Harkki *et al*., 1991). In the current research, *T. reesei* Rut-C30 has been used as a source for genes encoding nitrilases with a view of recombinant expression of the isolated genes in *T. reesei* at a later stage.

1.5 Industrial applications of nitrilases

Industrial use of enzymes has gained momentum in the past few decades. The use of enzymes in commercial chemical reactions has several advantages such as reduction in process costs, increased time and process efficiency and reduction in harmful by-products. The high enantio-, regio- and chemo- selectivity of enzymes have made biocatalysis an important and commercially significant process. According to a report generated by Freedonia Group (Freedonia, 2009), the global demand for industrial enzymes is expected to exceed $7 billion by the end of 2013.

Nitrilases are generating considerable interest due to their role in the biocatalysis of nitriles. As of 2007, the global annual production of nitrilase enzymes from both plant and microbial sources was around 30,000 tons (Demain, 2007). Figure 1.5 illustrates the increase in the number of research articles regarding enzymatic hydrolysis of nitriles hosted in Scopus (http://www.scopus.com/home.url).

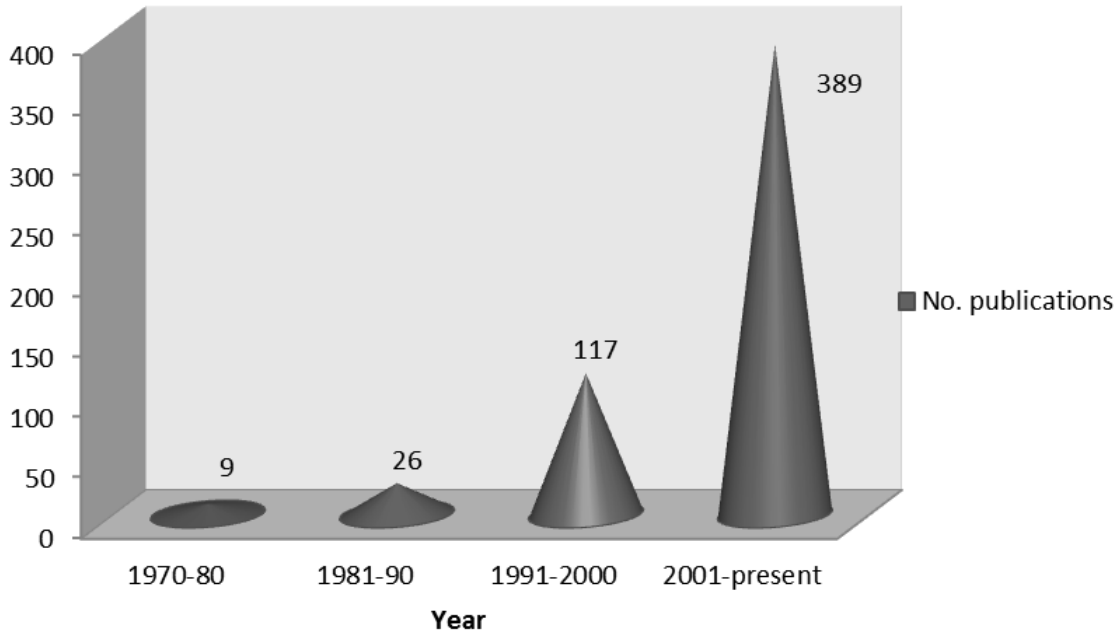


Figure 1.5 Number of research publications on nitrile degrading enzymes hosted in the Scopus database (http://www.scopus.com/home.url).

The increased interest in nitrilases has widened the scope of the possible applications of the enzyme. With versatile biocatalytic properties and their capability to breakdown toxic nitriles, nitrilases have been utilized in the manufacture of chemicals and drugs and in the environmental biodegradation of nitriles. More specific applications of nitrilases are described below.

1.5.1 Drug and chemical manufacture

The economic value of nitrilases lies primarily in their ability to achieve high substrate to product conversion rates with high specificity and purity of the products. In addition, the biocatalysis of nitriles performed using nitrilases generates less waste when compared to the chemical hydrolysis reaction. For example, isonicotinic acid (INA), a major intermediate in the synthesis of antituberculostatic drugs and anti-malarial drugs such as "Terefenadine", is chemically synthesized from 4-methylpyridine using high amounts of aqueous sulphuric acid generating a significant amount of waste. Recently, INA was synthesized using a nitrilase from *Nocardia globerula* without the generation of any amide by-product using considerably

milder reaction conditions (Sharma *et al*., 2012). Also as discussed earlier, the carboxylic acids produced using nitrilases have added attributes of enantio- and regio-selectivities. Optically pure enantiomers such as (R)-mandelic acid, (S)-ibuprofen and (R)-4-cyano-3-hydroxybutyric acid obtained using nitrilases are used as chemical building blocks in the manufacture of semisynthetic cephalosporins, skin creams and agro-chemicals. They are also used in the manufacture of commercial drugs such as Pfizer's "Lipitor" and (S)-naproxen (anti-inflammatory drug) (Effenberger and Bohme, 1994; DeSantis *et al*., 2002; Breuer *et al*., 2004).

Immobilized nitrilase enzymes have been used to produce carboxylic acids of high purity. An example is the regioselective hydrolysis of (E)-2-methylbut-2-enoic acid from (E/Z)-2-methylbut-2-enenitrile using an immobilized nitrilase enzyme from *Acidovorax facilis* 72W. The conversion was of high purity as no trace of (Z)-2-methylbut-2-enoic acid was observed (Singh *et al*., 2006). Furthermore these immobilized enzymes can be retrieved and reused.

1.5.2 Environmental biodegradation of toxic nitriles

Nitrilases have also been utilized in the biodegradation process of nitriles which are highly toxic and carcinogenic in nature. Hazardous nitrile compounds such as bromoxynil and ioxynil are present in herbicides. These compounds contaminate soil and water and can become a health hazard causing symptoms such as fever, vomiting and urinary problems if not prevented from entering the food chain (Freyssinet *et al*., 1996). Nitrilases purified from several organisms such as *Nocardia,* and *Arthrobacter* have been used to degrade these harmful compounds in the soil into less toxic compounds (Stalker *et al*., 1988a; Stalker *et al*., 1988b). In addition, bromoxynil resistant transgenic plants have been developed by the introduction of recombinant DNA encoding a specific nitrilase from the soil bacterium *Klebsiella ozaenae* that converts bromoxynil to a harmless primary metabolite (Stalker *et al*., 1988b). This technique has led to the development of bromoxynil tolerant cotton (trade name

BXN cotton), which is a commercially profitable transgenic crop (Duke, 2005; Dyer *et al*., 1993). Live nitrilase producing bacteria have also been applied as mixed cultures in the treatment of industrial effluents containing the toxic nitriles acrylonitrile and fumaronitrile (Wyatt and Knowles, 1995). Thus nitrilase enzymes have been successfully used in different fields ranging from production of chemicals to biodegradation of toxic nitriles.

1.6 Assessment of microorganisms as sources of novel nitrilases

The search for new nitrilases can involve broadscale screening of environmental microorganisms for nitrilase activity. Alternatively, well characterized microorganisms that are already widely used in research and/or industry can be assessed as potential sources of novel nitrilases as undertaken in this project; previous knowledge of the organism can be a valuable resource. Assessment of microorganisms as potential sources of novel nitrilase(s) involves several approaches. Techniques involving enrichment cultures and activity assays are discussed in the sections below.

1.6.1 Selection of nitrilase-producing microorganisms using enrichment cultures

Identification of nitrile-degrading microorganisms can be performed using enrichment cultures containing nitriles as the sole source of nitrogen and carbon (Ferrer-Miralles *et al*., 2009). Survival depends on the microorganism's ability to hydrolyze the nitriles into ammonia and carboxylic acid, compounds that can then be used for nutrition and growth; microorganism without this ability will die. Repeated sub-cultivations on either solid or liquid cultivation media containing nitriles ensure that the most efficient nitrile-hydrolyzing strains are selected due to their ability to grow faster and dominate over other strains. For example, *Arthrobacter* sp. strain J-1 was isolated from a soil sample using an enrichment medium consisting of acetonitrile as the sole source of carbon and nitrogen (Yamada *et al*., 1979). The temperature and/or pH of the enrichment cultures can be modified to select for

microorganisms capable of producing nitrilases that are active under certain conditions. For example, an acidotolerant strain of *Exophiala mesophila* species capable of hydrolyzing nitriles was isolated using an acidic enrichment medium containing 2-phenylacetonitrile as the source of nitrogen (Rustler and Stolz, 2007).

One limitation of the enrichment culture technique is that it fails to distinguish those microorganisms that produce nitrilases from those that only produce nitrile hydratases and amidases. As nitriles can be hydrolyzed either monoenzymatically (via nitrilases) or bienzymatically (via nitrile hydratases and amidases; Section 1.3.3), microorganisms capable of hydrolyzing nitriles via either pathway will be selected. Further characterization of the enzyme activity of these organisms then needs to take place.

1.6.2 Nitrilase activity assays

Nitrilases are typically intracellular enzymes (Gong *et al*., 2012) so to enable assessment of nitrilase activity the cells first need to be ruptured e.g. by sonication or lysozyme treatment. Following centrifugation to remove cell biomass, nitrilase activity is normally detected from the soluble protein fraction. Nitrilase activity assays are broadly classified as either photometric, chromatographic assays or ammonia based activity assays as discussed below.

1.6.2.1 Photometric and chromatographic activity assays

Traditional methods to determine nitrilase activity and enantioselectivity include spectrophotometric and chromatographic analyses of the rate of nitrile hydrolysis as well as the products obtained after the biocatalysis. Spectrophotometric analyses using UV-spectroscopy, H NMR (Hydrogen Nuclear Magnetic Resonance) spectroscopy (Jallageas *et al*., 1979) and mid-infrared spectroscopy (Dadd *et al*., 2000) have allowed continuous monitoring of substrate consumption and product formation. These approaches enable enzymes to be classified as nitrilases, nitrile hydratases or amidases by comparison of the spectral signals obtained from the products of the nitrile hydrolysis (Hook and Robinson,

1964; Brennan *et al*., 1996; Murakami *et al*., 2000). Chromatographic techniques such as reverse phase high performance liquid chromatography (HPLC) coupled with UV detection, thin layer chromatography and gas chromatography have gained considerable significance due to their high sensitivity and accuracy in measuring enzyme activity (Martinkova *et al*., 2008).

1.6.2.2 Ammonia-based nitrilase activity assays

As their name implies, ammonia-based nitrilase activity assays are based on the detection of ammonia released as a consequence of nitrile hydrolysis. Although unable to distinguish between ammonia produced as a consequence of the monoenzymatic nitrilase pathway or bienzymatic nitrile hydratase/ amidase pathway, ammonia-based assays are considerably faster and cheaper than the photometric and chromatographic methods described above. Consequently, ammonia-based assays are particularly useful in the characterization of purified enzymes that have already been classified as nitrilases, or in the assessment of crude samples in which a nitrilase enzyme is known to be present.

Several ammonia based techniques have been developed, including Nessler's, Roth's, indophenol and indothymol, chemiluminiscent and ion selective electrode methods (Molins-Legua *et al*., 2006). Assays that enable high throughput screening of samples have been devised by Banerjee *et al*. (2003b) and Yazbeck *et al*. (2006). Banerjee *et al*. (2003b) described a fluorometric plate assay which allowed simultaneous screening for nitrilase activity on a large scale. The technique was based on the fluorometric detection of an isoindole derivative formed by a reaction between ammonia and o-phthaldialdehyde. The advantage of this method is its high sensitivity and specificity towards ammonia. Yazbeck *et al*. (2006) described a colorimetric plate assay based on detection of a color change due to formation of an ammonia-cobalt complex. Although this method is considerably faster than

the fluorometric activity assay described by Banerjee *et al*. (2003b) in terms of reaction time, it is not as sensitive to ammonia as the fluorometric approach.

For the research presented in this thesis, nitrilase activity of crude cell lysate and purified enzyme fractions were assessed by fluorometric (Banerjee *et al*., 2003b) and colorimetric (Yazbeck *et al*., 2006) ammonia-based assays and UV-coupled HPLC analysis using a range of nitriles as substrates.

1.7 Heterologous expression of microbial nitrilases

In the past few decades there has been an increased interest in the use and exploitation of microorganisms and other cell-based systems for the expression of biological products (Arnold, 1995; Yoon *et al*., 2001). With a variety of hosts available for the purpose, the number of recombinantly produced proteins has risen considerably. The bacterium *Escherichia coli* is widely used for amplification of the cloned DNA and initial expression and characterization of the recombinant gene products. Filamentous fungi such as *T. reesei* and *Aspergillus niger* are frequently applied for industrial scale production of recombinant proteins, and mammalian cell lines including Chinese hamster ovary (CHO) cells and murine cells are used for production of complex recombinant proteins of mammalian origin (Merten *et al*., 2001; Walsh, 2002; Steinborn *et al*., 2006; Ferrer-Miralles *et al*., 2009). A few cell free or *in vitro* protein expression systems have been developed but due to their high costs and low yields they have remained as a research tool (Kigawa *et al*., 2004).

Not all hosts are suitable for the production of all types of recombinant proteins. A thorough understanding about host-cell metabolism, potential co-factors required for recombinant protein production and the necessary post translational modifications of the recombinant protein is needed. For example, an important factor to be considered during heterologous expression is whether the recombinant enzyme retains the properties of the native enzyme.

For example, the stability and optimal temperature and pH of the recombinant nitrilase from *A. niger* expressed in *E. coli* were considerably lower than those of the native *A. niger* enzyme (Kaplan *et al*., 2011).

Table 1.4 Examples of nitrilases expressed in *E. coli.*

Original source	***E. coli* strain**	**Reference**
Arabidopsis thaliana	BL21	Vorwerk *et al*., 2001
Pseudomonas fluorescens	JM 109	Kiziak *et al*., 2005
Rhodococcus rhodochrous	BL21	Luo *et al*., 2010
Aspergillus niger	BL21	Kaplan *et al*., 2011
Alcaligenes faecalis	JM 109	Jain *et al*., 2012

Considerable research has been carried out to boost heterologous protein production by using various growth media and a range of host organisms and expression vectors (Chae *et al*., 2000; Mahadevan and Doyle, 2003; Hewitt and McDonnell, 2004; Braud *et al*. 2005). Nitrilases from plants and bacteria have been expressed recombinantly in *E. coli* (Table 1.4). However, to date, the only published account of a recombinant fungal nitrilase is from *A. niger*, expressed in *E. coli* by Kaplan *et al*. (2011).

1.7.1 *Escherichia coli (E. coli)* as an expression host

In the current research, recombinant expression of a eukaryotic fungal nitrilase from *T. reesei* was attempted in *E. coli*. *E. coli* is highly popular for gene cloning, plasmid propagation and studies to investigate recombinant expression of prokaryotic and eukaryotic proteins due to its well-studied genetics, cost effective growth and availability of well developed molecular tools (Yin *et al*., 2007; Ferrer-Miralles *et al*., 2009).

Plasmid based expression systems ensures a relatively simple transformation process without the need for chromosomal integration of the recombinant gene into the host cell. Thus

introduction of multiple copies of a gene into a single cell is possible with *E. coli*. Several *E. coli* expression plasmids are available with the most popular being the pET expression system using the T7 promoter (Yin *et al*., 2007). The pET system has been developed since its discovery for several expression applications (Studier *et al*., 1990; Dubendorf and Studier, 1991).

The main drawbacks of expressing eukaryotic proteins in *E. coli* are the inability of the bacterial cells to splice introns, perform certain post-translational modifications such as phosphorylation and glycosylation that play an important role in protein folding and stability, and the formation of protein aggregates or insoluble inclusion bodies due to improper protein folding. Post-translational modifications are not a known feature of prokaryotic nitrilases and hence the lack of post-translational modification machinery in *E. coli* has not been an issue for their recombinant expression. However, fungi are eukaryotic organisms and thus post-translational modifications are more likely to be involved in the production of functional fungal proteins, implying that expression in *E. coli* poses more of a challenge.

When eukaryotic proteins are expressed in *E. coli* they sometimes form insoluble aggregates or inclusion bodies. There are several factors that are thought to be associated with the formation of these aggregates including incorrect protein folding, the absence of appropriate molecular chaperones and necessary post-translational modifications of the protein such as glycosylation, and the instability of disulphide bonds caused by the reducing environment of the *E. coli* cytoplasm (Mitraki and King, 1989; Cardamone *et al*., 1995; Doyle and Smith 1996, Guise *et al*., 1996). Association of several partially folded proteins results in the formation of inclusion bodies. Besides misfolded proteins, inclusion bodies may also contain polypeptides and phospholipids (Middelberg, 1996). Modification of growth parameters (temperature, pH), the concentration of induction agents, and *in vitro* protein refolding using approaches such as dialysis are some of the approaches that have been used to successfully

overcome the problem of the aggregation of recombinant proteins in *E. coli* (Kopetzki *et al*., 1989; Forciniti, 1994).

1.7.2 Protein purification using affinity tags

The ability to purify a heterologous protein is important for its biochemical characterization. Also, pharmaceutically relevant proteins often need to undergo stringent purification procedures before their administration to patients (Cho *et al*., 2001). Thus considerable work has been dedicated to develop robust techniques to purify target proteins from crude cell extracts.

Affinity tags used as highly efficient protein purification tools are developed on the basis of the affinity of amino-acid sequences, peptides or small proteins for a particular ligand immobilized to a solid support. Another common class of tags features an antibody as the protein binding partner which recognizes specific peptide epitopes. For example, the FLAG peptides are used to bind to the anti- FLAG antibody resins. Some of the common protein purification tags used today are His-tag, FLAG, Steptag II, HA-tag, GST, S-tag and MBP-tag (Lichty *et al*., 2005).

The main advantage of affinity tags is that they allow single step purification of the tagged protein based on affinity chromatography, thereby ensuring high recovery rates. For example, a recombinant urokinase-type plasminogen activator expressed in *E. coli* was purified using a His-tag fused to the coding sequence. The tag also facilitated increased protein refolding efficiency and high protein yield (Tang *et al*., 1997). Affinity tags that themselves are proteins such as maltose binding protein (MBP) have been known to potentially increase solubility of the recombinant protein to which they are fused (Dyson *et al*., 2004).

The His-tag is one of the most widely used protein purification tags containing an amino acid motif with six to nine histidine residues and was used for the purification of the nitrilase

enzyme in the current research. Histidine has a great affinity for metal ions such as copper and nickel which are immobilized with chelating agents such as nitrilotriacetic acid and sepharose matrices on a chromatography column (immobilized metal-ion affinity chromatography, IMAC) thus allowing the affinity based separation of the proteins fused to the His-tag. The detailed procedure for the purification of a His-tagged nitrilase from the crude cell lysate of *E. coli* is detailed in Chapter 7.

1.8 Aims of this project

The broad aims of the study were to identify nitrilase-encoding genes from *Trichoderma reesei* with a view of isolating the genes and producing the protein using recombinant technology.

The more specific aims of the project were:

- Identification of genes encoding nitrilases from *T. reesei* using bioinformatics and phylogenetic methods *in silico*;
- Exploration of of production of nitrilase enzymes in *T. reesei* using specific induction media;
- Cloning and expression of a gene encoding a true nitrilase from *T. reesei* in *E. coli*;
- Purification and characterization of the recombinant nitrilase enzyme produced in *E. coli*.

In the current research, a gene encoding a nitrilase enzyme from *Trichoderma reesei* (named *nit1*) was identified, isolated and expressed in *E.coli*. This research acts as a prelude to the large scale production of nitrilases in a fungal host.

Chapter 2: Materials and methods

General materials and methods used in this research are outlined in this chapter. Modifications to the following methods and specific methods only relevant to particular phases of the work are outlined in the corresponding chapters.

2.1 Common solutions and buffers

The chemicals used for this study were of analytical grade and were purchased from Sigma-Aldrich (Australia) unless otherwise specified. The water used for all reagents, buffers and media was filtered through the Millipore Milli-Q Academic filtration system. The commonly used solutions/buffers are listed in Table 2.1. Percentage concentrations are expressed as weight per volume (w/v) unless otherwise stated.

Table 2.1 Commonly used solutions.

Solution	**Components**
TBE (10×)	90 mM Tris-HCl, 1.25 mM EDTA, 90 mM boric acid, pH 8.4
TE (10×)	100 mM Tris-HCl, 10 mM EDTA, pH 8.0
Revco stock	42 mM K_2HPO_4, 22 mM KH_2PO_4, 1.7 mM sodium citrate, 0.4 mM $MgSO_4$, 30% (v/v) glycerol
DNA gel loading buffer	0.005% bromophenol blue, 0.005% xylene cyanol, 30% (v/v) glycerol
Protein gel loading buffer	0.5 M Tris-HCl, pH 6.8, 50% (v/v) glycerol, 10% SDS, 5% (v/v) β-mercaptoethanol, 0.005% bromophenol blue
SDS-PAGE running buffer	25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3

2.2 Molecular biology techniques

2.2.1 Digestion of DNA with restriction enzymes

All restriction endonucleases used in this research were obtained from Fermentas (USA) unless otherwise stated. Digestions of DNA were carried out using buffers and protocols provided by the manufacturer. Typical quantities used for the restriction enzyme digestion of DNA are shown in Table 2.2.

Table 2.2 Components used for restriction enzyme digestion.

Components	**Amount**
DNA	0.4 µg
10× Digestion buffer	5 µl
Restriction enzyme	0.5-1 µl
H_2O	Add up to final volume of 50 µl

2.2.2 Ligation of DNA fragments

Ligation reactions were performed with T4 DNA ligase (Fermentas, USA). All the reactions were performed as per manufacturer's instructions. The ligation reaction mix was incubated overnight at 22 °C. Before the transformation process, the enzyme was inactivated by incubation at 65 °C for 10 min.

2.2.3 Sequencing of DNA

Sequencing reactions of the recombinant plasmid DNA was carried out using the ABI prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, USA). The reaction mixture consisted of 350 ng plasmid DNA, 3.2 pmol of primers, 2 µl of Big Dye and 4 µl buffer. The DNA was precipitated as per the manufacturer's protocol. The precipitated DNA was analyzed by the Macquarie University Sequencing Facility.

2.2.4 Polymerase Chain Reaction

PCR Extender System (5 Prime, USA) and Amplitaq Gold® Polymerase (Applied Biosystems, USA) were used for all the DNA amplification reactions. The reactions were carried out in 50 µl volumes on a Thermocycler (Eppendorf, USA).

A standard 50 µl PCR performed using PCR Extender Polymerase System contained 0.2 U PCR Extender Polymerase mix, 1 x High Fidelity Buffer with Mg^{2+} and 0.2 mM dNTPs. A standard 50 µl PCR reaction performed using Amplitaq Gold® Polymerase contained 1.25U Amplitaq Gold® Polymerase, 10 x PCR buffer II, 3mM $MgCl_2$ and 0.2 mM dNTPs. The standard PCR protocol is shown in Table 2.3.

Table 2.3 Standard PCR protocol.

Steps	**Temp (°C)**	**Time**
Preheating	94	15 min
Template Denaturation	94	30 sec
Primer annealing	50-55	15 sec
Primer elongation	72	1 min

All the primers used in this study were manufactured by Integrated DNA Technologies (Australia). Detailed descriptions of the primers and the reaction conditions used are presented in the separate chapters.

2.2.5 Colony PCR

Colony PCR was carried out using a single colony picked using a sterile inoculating loop and resuspended in the standard PCR mixture (Section 2.2.4). All colony PCR was carried out using Amplitaq Gold® Polymerase under the conditions described in Table 2.3.

2.2.6 Purification of PCR products

PCR products were separated by electrophoresis on 1% (w/v) agarose gels using 1x TBE buffer. PCR products were purified from the agarose gel using Qiagen QIAquick PCR purification kit (QIAGEN, USA) as per manufacturer's instructions.

2.3 *Trichoderma reesei*

2.3.1 Fungal strain

T. reesei Rut-C30, a third generation, high cellulase secreting mutant of *Trichoderma reesei* (Montenecourt and Eveleigh, 1979) was used throughout this work.

2.3.2 Fungal growth media and culture conditions

All culture media and materials used for the cultivation and storage of microorganisms in this project were sterilized by autoclaving at 121 °C for 20 min. Fungi were inoculated onto Potato Dextrose Agar (PDA) plates and incubated at 28 °C for 5-7 d for conidiation. Conidia were collected by flooding the plates with 5 ml of 0.9% (w/v) NaCl/0.01% (v/v) Tween 80 solution and by gently abrading the surface of the PDA plates with a glass rod spreader. The conidial suspension was filtered through cotton wool packed in micropipette tips to remove the hyphae. The number of conidia in the solution was counted using a haemocytometer. The conidia were resuspended in 12.4% (v/v) glycerol/0.04% (v/v) Tween 80 for storage at -20 and -80 °C.

Various liquid media and their components used for the cultivation of *T. reesei* throughout this project are shown in Table 2.4. Freshly harvested conidia (2×10^6) were used to inoculate 50 ml of liquid culture medium in flat bottom 250 ml conical flasks. Cultures were grown for seven days at 28 °C with constant shaking at 250 rpm. The specialized culture conditions used for nitrilase induction for *T. reesei* are detailed in the following chapters.

Table 2.4 Liquid growth media for *T. reesei* Rut-C30.

Medium	**Components (per liter)**
Minimal medium (Penttilä *et al*., 1987)	20 g glucose, 5 g $(NH_4)_2SO_4$, 15 g KH_2PO_4, 0.6 g $MgSO_4$, 0.6 g $CaCl_2$, 50 mg $FeSO_4 \cdot 7\ H_2O$, 14 mg $ZnSO_4 \cdot 7H_2O$, 20 mg $CoCl_2$, pH 6.5
Modified Czapek-Dox (Kaplan *et al*.. 2006)	30 g sucrose, 1 g K_2HPO_4, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.5 g KCl, 0.01 g $FeSO_4.7H_2O$, 0.001 g $CoCl_2 \cdot 6H_2O$, 0.0067 g $ZnSO_4 \cdot 7H_2O$
Cellobiose/ Lactose/ Soy hydrolysate (CLS) (Lim *et al*., 2001)	Minimal medium, 1 g cellobiose, 0.1 g lactose, 0.73 g soy hydrolysate

2.3.3 Extraction of genomic DNA

T. reesei Rut-C30 was grown for 4 d over cellophane discs placed on PDA plates. The hyphae was harvested, lyophilized and ground to a fine powder using mortar and pestle. The genomic DNA was isolated from the lyophilized mycelia using GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, Australia). The quality of the extracted genomic DNA was visualized on a 1% (w/v) agarose gel.

2.3.4 Extraction of total RNA

Total RNA was extracted from *T. reesei* Rut-C30 in RNase-free environment using Trizol Reagent (Invitrogen, USA) as per manufacturer's instructions. Mycelium from five day old liquid cultures was collected by centrifugation. The pellets were thoroughly washed in cold diethylpyrocarbonate (DEPC)-treated H_2O and flash frozen using liquid nitrogen. The frozen pellet was ground to a fine powder. A 1 ml volume of Trizol Reagent was added to 10 mg of mycelial powder. The insoluble material was removed by centrifugation at 10,000 × *g* for 15 min at 4 °C. Chloroform was added to the supernatant for phase separation followed by incubation at RT for 2 min. The colorless supernatant containing the RNA was collected by

centrifugation at 10,000 × *g* for 15 min at 4 °C. The RNA was precipitated from the supernatant using isopropyl alcohol and was incubated at RT for 10 min. The RNA pellet was obtained by centrifugation at 10,000 × *g* for 15 min at 4 °C. The RNA pellet was washed with 75% ethanol and was air dried at RT for 20 min. The RNA pellet was dissolved completely in DEPC-treated water. The RNA quality was determined by electrophoresis on a 1% (w/v) agarose gel and visualization under UV light.

2.4. Expression of recombinant proteins in *Escherichia coli*

2.4.1 *E. coli* strains

E. coli strain DH5α, a low endonuclease strain (New England Biolab, USA) was used for all cloning work (Inoue et al., 1990). *E. coli* Origami^{TM}B(DE3) and Tuner^{TM}DE3 competent cells (Novagen, USA) were used for gene expression. Both the Origami^{TM}B(DE3) and the Tuner^{TM}(DE3) cells were selected due to their reported high performance in enabling correct protein folding and solubility of recombinantly expressed proteins.

2.4.2 Plasmid

The plasmid pETDuet-1 was used for construction of all gene vectors. The pETDuet plasmid contains a region encoding a 6 x His-tag at the 5 ′ end of the multiple cloning site-1 (MCS-1) that can be used in the purification of recombinant proteins (Figure 2.1).

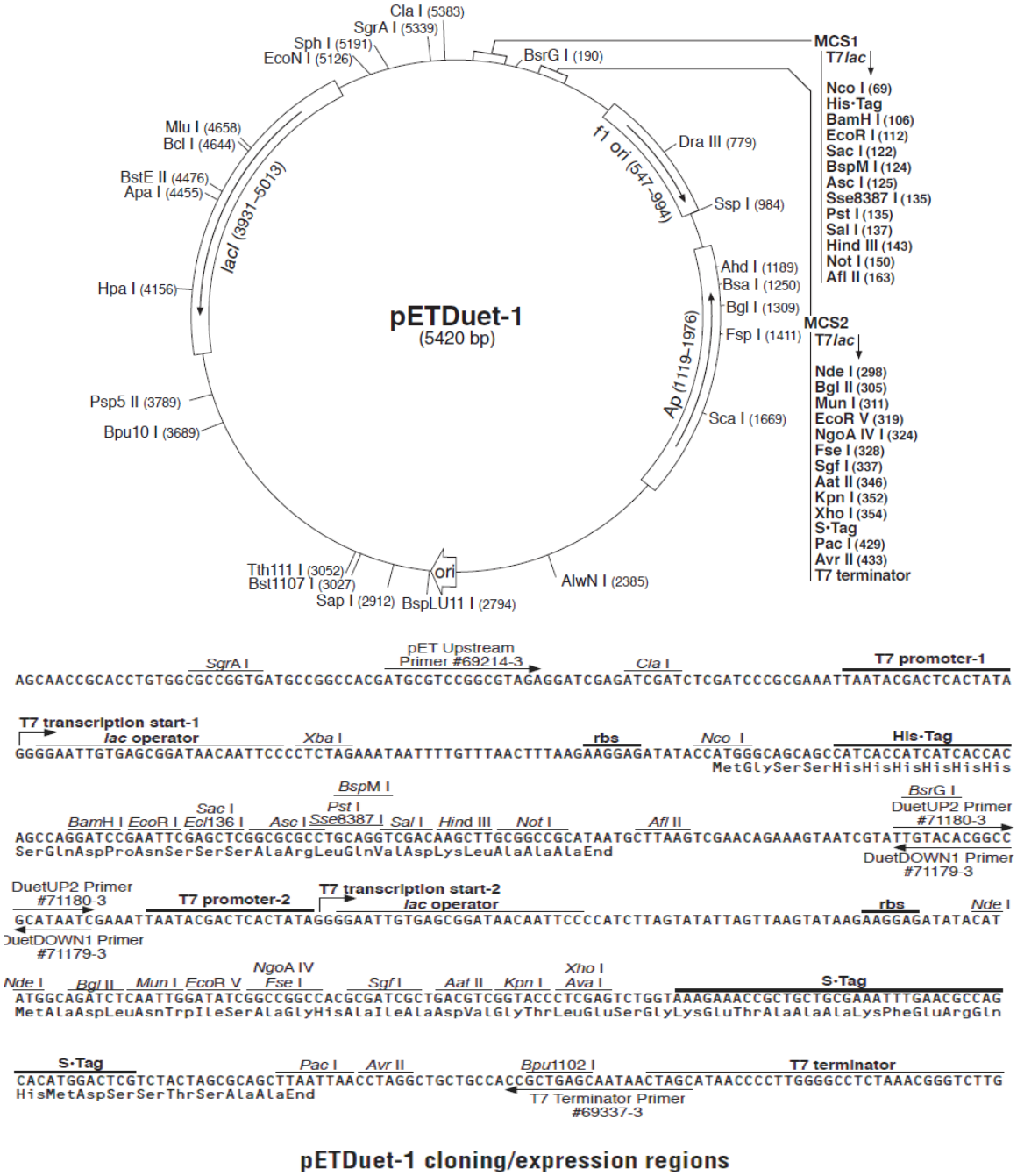


Figure 2.1 Map of the pETDuet-1 vector containing two multiple cloning sites (MCS) proceeded by a T7 promoter/*lac* operator and a ribosome binding site (rbs). The vector also carries the ampicillin resistance gene (Ap). Vector map and sequence of cloning regions have been adapted from the Merck-Millipore website (www.merckmillipore.com.au).

2.4.3 Bacterial growth media and culture conditions

Growth media for the cultivation of *E. coli* strains are shown in Table 2.5. Ampicillin to a final concentration of 100 µg/ml was added to the Luria agar (L-agar; Luria and Burrous, 1957) plates for the selection of recombinant *E. coli* containing the expression vector with the

ampicillin resistant gene (Figure 2.1). *E. coli* strain DH5α was cultivated at 37 °C. The *E. coli* Origami™B(DE3) and Tuner™DE3 competent cells were cultivated at 22 °C and 37 °C.

Table 2.5 Growth media for *Escherichia coli.*

Medium	**Components**
Luria broth	1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0
Luria agar	Luria broth with 2% agar
SOB	25% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$
SOC	SOB with 20 mM glucose
Terrific broth	1.2% Tryptone, 2.4% yeast extract, 0.5% (v/v) glycerol, 0.017 M KH_2PO_4, 0.072 M K_2HPO_4

2.4.4 Transformation of *E. coli* DH5α cells

A total of 0.5 µl of DNA ligation mixture was added to 50 µl aliquots of *E. coli* DH5α cells. The mixture was incubated on ice for 30 min. The tubes were heat pulsed at 42 °C for 30 sec and transferred back onto ice. About 0.75 ml of SOC medium was added to the *E. coli* DH5α cell aliquots containing the DNA ligation mixture followed by incubation at 37 °C for 1 h. Thereafter the sample mixture was plated onto Luria agar plates containing ampicillin to a final concentration of 100 µg/ml.

2.4.5 Extraction of total RNA from *E. coli*

Total RNA was extracted from *E. coli* Origami™B(DE3) and Tuner™DE3 competent cells cultivated at 22 °C and 37 °C using TRI Reagent® LS (Sigma, USA) as per manufacturer's instructions. All the experiments were performed under RNase free conditions.

2.5 Detection of RNA transcripts by Northern blotting

2.5.1 Preparation of RNase –free solutions

Glassware used for this technique was washed thoroughly with DEPC-treated H_2O, autoclaved and baked overnight at 180 °C. The work benches, pipettes and the electrophoresis tank were washed with Pyroneg, 70% (v/v) ethanol and DEPC-treated H_2O. Buffers and solutions were treated overnight with 0.1% (v/v) DEPC, then autoclaved at 121 °C for 30 min. The solutions used for Northern blotting are listed in Table 2.6.

Table 2.6 Solutions used for Northern blotting.

Solution	Composition
10x saline-sodium citrate (SSC) buffer	3 M NaCl, 0.3 M sodium citrate, pH 7.0
Maleic acid buffer	100 mM maleic acid, 150 mM NaCl, pH 7.5
Washing buffer	100 mM maleic acid, 150 mM NaCl, pH 7.5, 0.3% (v/v) Tween 20
Blocking solution	1% blocking reagent (Roche, Germany) in maleic acid buffer
Detection buffer	100 mM Tris-HCl, 100 mM NaCl, pH 9.5

2.5.2 Preparation of RNA for Northern blotting

Total RNA (20 µg) was mixed with three volumes of RNA loading buffer (250 µl deionised formamide, 83 µl of 37% (v/v) formaldehyde, 50 µl 10 × MOPS buffer and 10 µl of 2.5% (v/v) DEPC- treated bromophenol blue). The RNA was denatured at 65 °C for 10 min. The samples were separated on a formaldehyde-agarose (1% w/v) gel as described by Sambrook and Russell (2001). After electrophoresis the gel was stained with 0.5 µg/ml acridine orange for 10 min. The gel was de-stained in DEPC-treated H_2O and examined under UV illumination (Sambrook and Russell, 2001).

2.5.3 Blotting and detection protocol

The RNA in the gel (Section 2.5.2) was transferred onto a positively-charged nylon membrane (Roche, Germany) with a vacuum blotter (BioRad, USA). The transferred RNA was fixed onto the membrane by baking at 120 °C for 15 – 20 min. DIG Luminescent Detection Kit for Nucleic Acids (Roche, Germany) was used for the hybridization and detection of the mRNAs. The membrane was incubated in 20 ml of hybridization solution at the relevant hybridization temperature for 1 h. The membrane was transferred to fresh hybridization buffer with DIG- labeled probes (20 ng/ml) that had been denatured at 100 °C for 10 min. The probe was then allowed to hybridize to the membrane by incubating overnight at the relevant temperature. After overnight hybridization, the blot was washed twice with 2 × saline-sodium citrate (SSC), 0.1% SDS for 5 min. The blot was then washed twice with 0.1× SSC, 0.1% SDS for 15 min. All the wash steps mentioned above were carried out at room temperature. The membrane was placed in blocking solution for 1 h then incubated in Anti-Digoxigenin-AP antibody (Roche, Germany) diluted 1:10000 in blocking buffer for 30 min. The membrane was washed twice with washing buffer at RT to remove excess antibody and was transferred into a plastic bag with approximately 150 µl of "ready-to-use" CDP-star (Roche, Germany). Excess CDP- star solution was squeezed out from the bag. The membrane was exposed to CL-Xposure film (Thermo Fisher Scientific Inc, USA) and developed using Kodak GBX developing solutions.

2.6 Protein analysis

2.6.1 Protein extraction from *E. coli*

The soluble protein fractions from the *E. coli* transformants were extracted using B-PER protein extraction reagent (Thermo, USA). Extraction of total protein, including insoluble protein contained in the *E. coli* inclusion bodies, was carried out using two different solubilization buffers, a 4 × SDS buffer made in 50 mM Tris-HCl (pH 6.8) and an 8 M

guanidine hydrochloride (Gu-HCl) buffer (pH 8). A detailed description of the protein extraction procedures is contained in Chapter 6. Protein concentration was determined by the method of Bradford (1976).

2.6.2 Separation of proteins using SDS-PAGE

A 0.25 volume of 5 × protein loading buffer was added to the protein samples. The proteins were denatured by boiling the solution for 5 – 10 min. NuPAGE® 4 – 12% Bis-Tris gels (Invitrogen, USA) was used for protein separation. The protein samples were run at 180 V for 50 min. A molecular weight marker (Invitrogen, USA) was used to estimate the size of the target protein. The gel was fixed in 7% (v/v) acetic acid/ 10% (v/v) methanol for 30 min and was stained overnight in Coomassie Blue 250. The gel was destained in 1% acetic acid.

2.7 Nitrilase activity assays

2.7.1 Fluorometric assay of nitrilase activity

The fluorometric nitrilase activity assay was performed as described by Banerjee *et al*. (2003). The solutions and buffers are listed in Table 2.7. Adiponitrile was used as substrate for all assays in this work unless otherwise specified.

Table 2.7 Solutions and buffers used for fluorometric nitrilase assay.

Solution/Buffer	**Concentration**
Phosphate buffer	100 mM, pH 7
HCl	0.1 M
o-phthaldialdehyde/2-mercaptoethanol	o-phthaldialdehyde- 75 mM 2-mercaptoethanol- 72 mM

In a typical reaction, 25 – 40 μg of protein (soluble intracellular protein or purified nitrilase) was suspended in 175 μl of 100 mM phosphate buffer (pH 7) and equilibrated at 37 ºC prior

to the addition of 5 µl of the nitrile substrate. The mixture was incubated at 37 ºC for 20 min before 200 µl of 0.1 M HCl was added to stop the reaction. Samples were centrifuged at 5000 × *g* for 15 min and 10 µl of the supernatant was added to 300 µl of 75 mM o-phtaldialdehyde/ 72 mM 2-mercaptoethanol in a 96 well plate. Following incubation at room temperature for 20 min the fluorescence was measured using a FLUOstar microplate reader (BMG Labtech, Germany) at excitation and emission wavelengths of 405 and 460 nm respectively. Typically a range of dilutions of the enzyme sample was tested in triplicate and assessed relative to a commercial nitrilase (Codexis) with a known enzyme activity of 43 U/mg at 37 °C. One unit of nitrilase activity was defined as the amount able to release 1 µmol of ammonia per mg of protein under the assay conditions.

2.7.2 Colorimetric assay of nitrilase activity

The colorimetric nitrilase activity assay was performed as described by Yazbeck *et al*. (2006). The solutions and buffers used for the colorimetric nitrilase activity assay are listed in Table 2.8. Adiponitrile was used as substrate unless otherwise specified.

Table 2.8. Solutions and buffers used for colorimetric nitrilase assay.

Solution/Buffer	**Concentration**
Tris buffer	10 mM , pH 7
$CoCl_2$/Tris	$CoCl_2$- 10 mM, Tris- 10 mM, pH 7

Approximately 25 – 40 µg of the nitrilase enzyme was added to the nitrile substrate. The mixture was incubated at 37 ºC for 2 h before 50 µl of this reaction mixture was added to 50 µl of $CoCl_2$/ Tris solution in a 96 well plate and incubated for 5 min at RT. The absorbance was measured at 375 nm using a FLUOstar microplate reader (BMG Labtech, Germany).

Chapter 3: Identification and analysis of putative nitrilase genes from *Trichoderma reesei*

3.1 Introduction

The past few decades have seen the development of increasingly rapid genome sequencing techniques, making possible the availability of extensive genomic data of many organisms. *In-silico* screening of genome databases using tools such as NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), ClustalW (Thompson *et al*., 1994) and phylogenetic analysis software have helped in the identification of several novel sources of nitrilases and their genes (Seffernick *et al*., 2009). Many organisms contain genes encoding enzymes from multiple branches of the nitrilase superfamily (Section 1.2). Moreover, several organisms such as *Pseudomonas*, *Klebsiella* and *Burkholderia* have been found to contain more than one nitrilase gene, a phenomenon attributed to gene duplications and horizontal gene transfer (Podar *et al*., 2005).

The amino acid sequences of enzymes serve as important predictors of their function (Minshull *et al*., 2005). The properties of novel nitrilases can thus be suggested on the basis of the closest matched and characterized nitrilase homolog in a database. The structural features of nitrilases are also critical determinants in their identification. The common *αββα* fold with a catalytic triad consisting of lysine, cysteine and glutamic acid is a defining characteristic of all members of the nitrilase superfamily and is absent in non-family members such as nitrile hydratases (Section 1.2). The availability of protein structure prediction tools such as PROSITE (http://prosite.expasy.org/) can thus assist in determining whether a predicted protein derived from a gene sequence could be a nitrilase.

Nitrilases have been successfully identified and purified from filamentous fungi such as *Aspergillus niger*, *Fusarium solani and Penicillium* as reviewed by Martinkova *et al*. (2009). However, nitrilases from other fungal sources are yet to be fully explored and exploited

despite a considerable body of work carried out on nitrilases from bacteria (Hoyle *et al*., 1998; Kabaivanova *et al*., 2005; Kiziak *et al*., 2007). Although fungal nitrilases have been reported to have significant substrate specificity towards aromatic, aliphatic as well as aryl-aliphatic nitriles (Section 1.4.3.1) and are therefore of potential industrial interest, comparatively few fungal nitrilases have been fully characterized and large scale recombinant expression of fungal nitrilases has not been reported to date. Recombinant expression of an *A. niger* nitrilase in *E. coli* enabled characterization of the enzyme but production levels were low and the stability of the recombinant enzyme was significantly lower than that of the native version (Kaplan *et al*., 2006; 2011). In this work, the filamentous fungus *Trichoderma reesei* presented a likely source of nitrilases worthy of exploration, and well-established high-secreting mutant strains of *T. reesei* would be ready and appropriate hosts for large scale recombinant expression of homologous nitrilases in the future.

The *in silico* identification of genes encoding nitrilases from *T. reesei* using bioinformatic and phylogenetic analyses is described in this chapter. At the time this work was carried out the partially annotated genome sequence of the wild-type *T. reesei* QM6a (Martinez *et al*., 2008) had just been made available by the Joint Genome Institute (JGI; http://genome.jgi.doe.gov/) and was used for all data mining studies described here. The work performed by Seffernick *et al*. (2009) in the identification of genes encoding nitrilases from *B. xenovorans* and *B. rhizobium* along with the work investigating nitrilase genes from the filamentous fungi *A. niger* (Kaplan *et al*., 2006) and *F. solani* (Harper, 1977) were used to guide the *in silico* analysis of the *T. reesei* genome.

3.2 Materials and methods

3.2.1 Sequence analysis

The genome sequence of the *T. reesei* wild-type QM6a has been translated into its corresponding predicted protein sequences by JGI (http://genome.jgi-psf.org/Trire2). To

identify putative nitrilase genes, the predicted *T. reesei* protein sequences were searched for sequence similarities to known nitrilases; any of the predicted *T. reesei* protein sequences that had high sequence similarity to known nitrilases could then be traced back to their corresponding gene sequences in the *T. reesei* genome.

The entire set of predicted *T. reesei* protein sequences were compared to the sequences of known nitrilases (Section 3.2.2) using the Basic Local Alignment Search Tool (BLAST; http://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple sequence alignment of the proteins was carried out with ClustalW (www.ebi.ac.uk/ Tools/ services/web_clustalw; Thompson *et al*., 1994). A phylogenetic tree was constructed with the beta test PAUP4* software using default parameters (http://paup.csit.fsu.edu/; Swofford, 1991). A maximum likelihood tree was created using the MEGA 5 phylogenetic software (Tamura *et al*., 2011). Structure based sequence analyses were performed using ExPASy PROSITE tool (http://prosite.expasy.org/) and PRALINE software (http://www.ibi.vu.nl/programs/pralinewww/; Simossis and Heringa, 2003).

3.2.2 Sequence library of nitrilases and other closely related enzymes

Protein sequences of identified nitrilases from bacteria, plants and filamentous fungi were obtained from the NCBI database. Sequences were selected for inclusion in the library if they encoded a nitrilase that has been characterized either totally or partially. Predicted or hypothetical protein sequences of nitrilases which had not been characterized were not considered for the sequence library. Nitrilase sequences of considerable diversity were selected and inclusion of two or more similar sequences (more than 90% similarity) in the library was avoided. Since this research work focused on nitrilases from *T. reesei*, special attention was given to known nitrilase sequences from filamentous fungi; however, only three sequences from characterized fungal nitrilases were available. A final set of eleven sequences encoding nitrilase enzymes was created with sequence similarities ranging from 15% to 90%.

Three nitrilase sequences from filamentous fungi (*Aspergillus niger*, *Fusarium solani* and *Penicillum* sp.), five from bacteria (*Pseudomonas fluorescens*, *Rhodococcus rhodochrous*, *Bradrhizobium japonicum*, *Geobacillus pallidus* and *Acidovorax facilis*), and three from plants (*Arabidopsis thaliana*, *Hordeum vulgare* and *Brassica rapa*) were included in the sequence library.

In addition, sequences of enzymes from other branches of the nitrilase superfamily and of nitrile hydratases were also selected for clustering and phylogenetic analyses: three amidases (*Rhodococcus erythropolis*, *Helicobacter bizzozeronii* and *A. niger)*, three Nitfhit proteins (*Drosophila melanogaster*, *Caenorhabditis elegans* and *Schistosoma japonicum*) and four nitrile hydratases (*Burkholderia gladioli*, *Pseudonocardia dioxanivorans*, *A. niger* and *Pseudomonas putida*). Although nitrile hydratases are not a part of the nitrilase superfamily and lack the catalytic triad (Section 1.2), they were used for phylogenetic analyses as they share considerable sequence similarity with the nitrilases.

3.3 Results

Potential nitrilases were identified amongst the predicted proteins derived from the *T. reesei* genome sequence (http://genome.jgi-psf.org/Trire2) based on sequence similarities of the predicted proteins with nitrilases and closely related enzymes from other filamentous fungi and bacteria.

3.3.1 Identification of putative nitrilase genes by sequence alignment

Each of the 11 protein sequences of the known fungal, bacterial and plant nitrilases (Section 3.2.2) were used as the query sequence in BLASTp (protein BLAST) searches against the predicted proteins derived from the *T. reesei* genome sequence. Twelve *T. reesei* predicted proteins were identified that had high sequence homology to the known nitrilases. The *T. reesei* predicted protein sequences of interest were designated Nit 1 – 12 and were traced

back to their encoding genes in the *T. reesei* genome sequence (http://genome.jgi-psf.org/Trire2). The location (scaffold) of the corresponding gene in the genome was determined, and the expected molecular weight and function of the protein as predicted by JGI were collated (Table 3.1).

Table 3.1 Predicted proteins (Nit 1 to 12) derived from the genome sequence of *T. reesei* that shared sequence similarity to known nitrilases from other organisms. The predicted length, molecular weight (MW) and function of the proteins and the location of the corresponding gene in the *T. reesei* genome are as determined by JGI (http://genome.jgi-psf.org/Trire2).

Predicted protein	Predicted length (amino acids)	Predicted MW (kDa)	Predicted protein family	Gene location
Nit 1	360	40.1	Carbon-nitrogen hydrolase	Scaffold 15
Nit 2	329	35.7	Carbon-nitrogen hydrolase	Scaffold 15
Nit 3	335	36.5	Carbon-nitrogen hydrolase	Scaffold 15
Nit 4	369	39.7	Carbon-nitrogen hydrolase	Scaffold 32
Nit 5	284	31.1	Carbon-nitrogen hydrolase	Scaffold 19
Nit 6	683	77.2	Carbon-nitrogen hydrolase	Scaffold 1
Nit 7	311	33.6	Carbon-nitrogen hydrolase	Scaffold 19
Nit 8	316	33.8	Carbon-nitrogen hydrolase	Scaffold 16
Nit 9	306	33.1	Carbon-nitrogen hydrolase	Scaffold 4
Nit 10	303	32.6	Carbon-nitrogen hydrolase	Scaffold 2
Nit 11	298	32.8	Carbon-nitrogen hydrolase	Scaffold 1
Nit 12	1087	118.7	Carbon-nitrogen hydrolase	Scaffold 6

All of the twelve *T. reesei* predicted proteins sharing high homology with the known nitrilases in the NCBI database were designated by JGI as carbon-nitrogen hydrolases (Table 3.1). Most had a predicted molecular weight of 30 – 40 kDa, although Nit 6 and Nit 12 were predicted to

be over 75 kDa. Out of all the identified nitrilases to date only the nitrilase from *F. solani* has a molecular weight of over 75 kDa (76 kDa; Harper, 1977).

To gain further information about the possible catalytic nature of the predicted *T. reesei* carbon-hydrolases another series of BLASTp searches was performed using the protein sequences of the twelve putative carbon-nitrogen hydrolases (Table 3.1) as query sequences against the entire NCBI protein database. Closest sequence homologs are shown in Table 3.2.

Table 3.2 Proteins in the NCBI database with highest sequence similarity to the predicted *T. reesei* carbon-nitrogen hydrolases (Table 3.1). Protein function is as defined by the NCBI database.

Predicted protein	Closest sequence homolog in the NCBI protein database	Sequence similarity	Protein function
Nit 1	*A. niger* nitrilase	60%	Nitrilase
Nit 2	*Fusarium* cyanide hydratase	59%	Nitrilase
Nit 3	*Rhodococcus rhodochrous* K22 aliphatic nitrilase	64%	Nitrilase
Nit 4	Plant indole-3-acetonitrilase	60%	Nitrilase
Nit 5	Plant N-carbamoylputrescine amidohydrolase	57%	Amidase
Nit 6	*Beauveria bassiana* ARSEF 2860 NAD-synthase	48%	NAD-synthase
Nit 7	*C. elegans* Nit-Fhit	55%	NitFhit Protein
Nit 8	Plant N-carbamoylputrescine amidohydrolase	59%	Amidase
Nit 9	*C. elegans* Nit-Fhit	62%	NitFhit Protein
Nit 10	*C. elegans* Nit-Fhit	44%	NitFhit Protein
Nit 11	*Colletotrichum higginsianum* N-terminal amidase	41%	N-acyltransferase
Nit 12	*Verticillium dahliae* VdLs.17 N-terminal amidase,	57%	NAD-synthase/ N-acyltransferase

From the NCBI BLAST searches, the protein sequences of four of the twelve predicted carbon-hydrolases (Nit 1 – 4) had highest sequence similarity to enzymes designated in the NCBI database as "nitrilases" *i.e.* from the "nitrilase" branch (branch 1) of the nitrilase superfamily: an *A. niger* nitrilase, *Fusarium* cyanide hydratase, *R. rhodochrous* (bacterial) aliphatic nitrilase and a plant indole-3-acetonitrilase, respectively (Table 3.2). The remaining eight putative carbon-hydrolase sequences (Nit 5 – 12) had highest sequence similarity to fungal, plant or nematode amidases, NitFhit (nitrilase and fragile histidine triad fusion proteins), NAD-synthases or N-terminal acyltransferases.

3.3.2 Phylogenetic analyses of potential *T. reesei* nitrilase sequences

The phylogenetic relationships between nine of the predicted carbon-nitrogen hydrolases of *T. reesei* (Table 3.1) and known nitrilases from other sources was carried out using PAUP4 (parsimony analyses) software. The three Nit sequences found to have high similarity to NAD-synthase or N-terminal acyltransferase (Nit 6, Nit 11 and Nit 12; Table 3.2) were excluded from the phylogenetic analyses. The remaining nine protein sequences were aligned with protein sequences of nitrilases identified by Diversa Corporation (Mathur *et al*., 2005) along with nitrilases, NitFhit and amino terminal amidases from the nitrilase sequence library (Section 3.2.2). The bootstrap values for the branches were maintained low and all the gaps were excluded.

Out of the nine predicted *T. reesei* protein sequences used for the phylogenetic analysis, only four (Nit 1 – Nit 4) clustered with known nitrilases. Three of the *T. reesei* protein sequences (Nit 7, 9 and 10) clustered with Nitfhit proteins and the remaining two (Nit 5 and 8) with amino terminal amidases (Figure 3.1). These results were consistent with the closest alignments of the *T. reesei* predicted Nit proteins with the NCBI nitrilases from the BLASTp searches (Table 3.2).

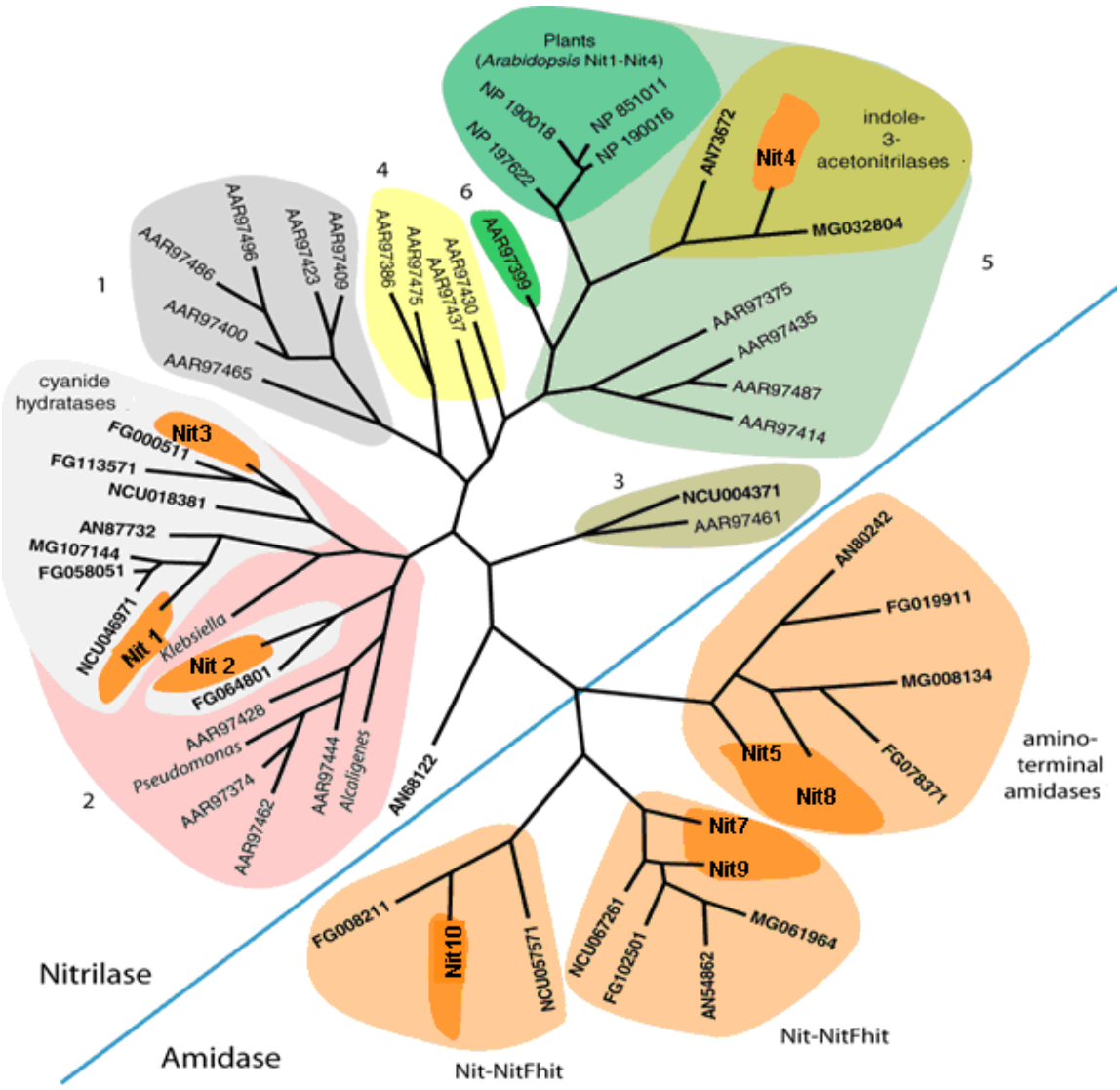


Figure 3.1 Phylogenetic analysis of the protein sequences of predicted *T. reesei* carbon-nitrogen hydrolases (Nit 1 – 10) in relation to carbon-nitrogen hydrolases of known function from the Diversa bacterial nitrilase library (AAR; DeSantis *et al.*, 2002), *N. crassa* (NCU), *A. nidulans* (AN), *F. graminearum* (FG), and *M. grisea* (MG). Analysis was carried out using PAUP4 software.

Out of the four *T. reesei* predicted proteins clustered with known nitrilases, three (Nit 1 – 3) were clustered with nitrilases from fungal (*Fusarium* sp.) and bacterial (*Pseudomonas* and *Alcaligenes* sp.) origins. One of the *T. reesei* predicted proteins (Nit 4) clustered with plant nitrilases characterized as indole-3-acetonitrilases; the similarity of this sequence with indole-3-acetonitrilases was also detected in the BLASTp search of the NCBI database (Table 3.2).

3.3.3 Phylogenetic analysis of predicted *T. reesei* proteins Nit 1 – 4 using a maximum likelihood tree

A new database was created containing the four predicted protein sequences (Nit 1 – 4) derived from the *T. reesei* genome that had the closest similarities to protein sequences of enzymes of the nitrilase branch of the nitrilase superfamily, along with protein sequences of known amidases, nitrile hydratases and Nitfhit enzymes from other organisms. The new database was used to create a maximum likelihood tree (Figure 3.2) using MEGA 5 phylogenetic software (Tamura *et al*., 2011).

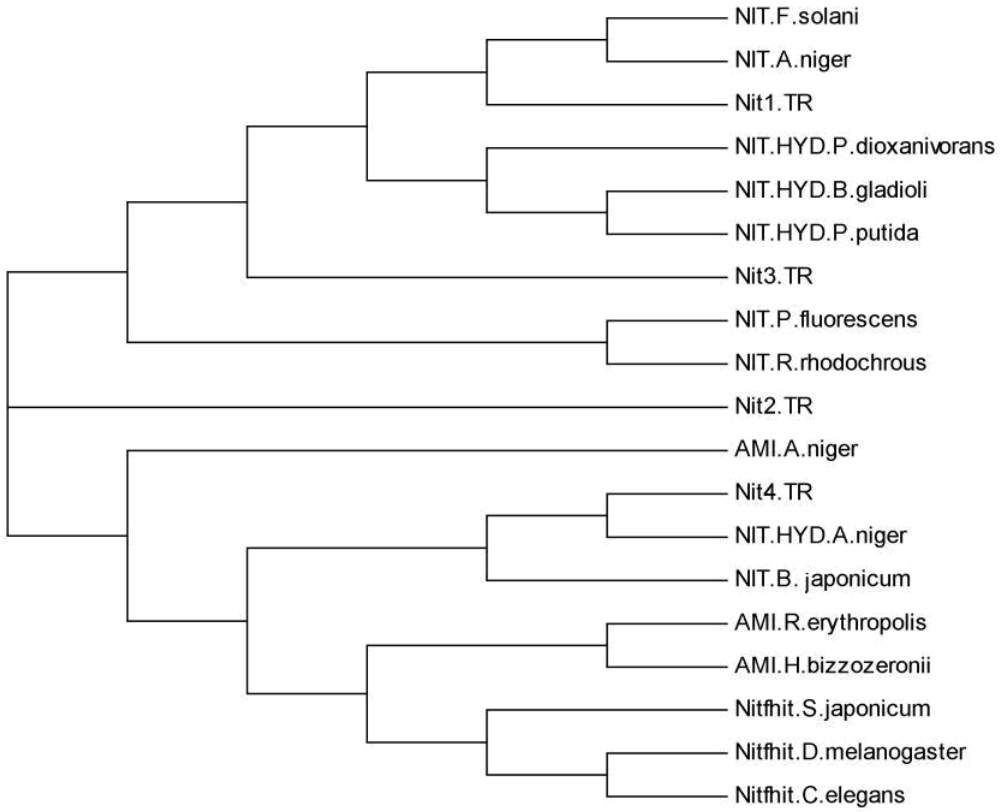


Figure 3.2 Phylogenetic analysis of predicted *T. reesei* Nit proteins (Table 3.1) by Maximum likelihood tree. The evolutionary differences in the sequences of the major branches of the nitrilase superfamily were inferred using the Maximum likelihood method. Nit1.TR – Nit4.TR represent the four predicted proteins from *T. reesei* sharing high sequence homology to enzymes from the nitrilase branch of the nitrilase superfamily; NIT.F.solani, NIT.A.niger, NIT.P.fluorescens, NIT.R.rhodochrous, NIT.B. japonicum denote nitrilase enzymes from *F. solani, A. niger, P. fluorescens, R. rhodochrous* and *B. japonicum* respectively; NIT.HYD.P.dioxanivorans, NIT.HYD.P.putida, NIT.HYD.B.gladioli, NIT.HYD.A.niger represent nitrile hydratases from *P. dioxanivorans, P. putida, B. gladioli* and *A. niger* respectively; AMI.R.erythropolis, AMI.H.bizzozeronii, AMI.A.niger represent amidases from *R. erythropolis, H. bizzozeronii* and *A. niger* respectively; NitFhit.S.japonicum, NitFhit.D.melanogaster and Nitfhit.C.elegans represents NitFhit proteins from *S. japonicum, D. melanogaster* and *C. elegans* respectively.

The evolutionary data was inferred by using the Maximum likelihood method which is based on the JTT matrix-based model (Jones *et al*., 1992). The tree with the highest log likelihood (-7372, 1917) is shown in Figure 3.2. Initial trees for the heuristic search were obtained by using maximum parsimony method in cases where the common sites were less than 100 or less than one fourth of the total number of sites. For all other cases, the BIONJ method with maximum composite likelihood distance matrix was used. All the positions containing gaps and missing data were eliminated. The four *T. reesei* Nit sequences were designated as Nit1.TR, Nit2.TR, Nit3.TR and Nit4.TR.

The Maximum likelihood tree (Figure 3.2) was broadly separated into three main branches. The first branch consisted of nitrilases and nitrile hydratases, the second branch consisted only of the predicted *T. reesei* protein Nit2.TR. The third branch consisted of the NitFhit proteins, amidases, two nitrilases and a nitrile hydratase. In the third branch all the NitFhit proteins clearly branched away separately from the other enzymes. This denotes a significant sequence difference between the NitFhit proteins and the other branches of the nitrilase superfamily. The amidases on the other hand were split into a prokaryotic amidase branch and a separate, distant branch containing the only eukaryotic amidase from *A. niger*.

The Nit2.TR branched separately from the other known nitrilases, amidases and NitFhit proteins. The BLASTp search (Section 3.3.1) revealed that the closest characterized homolog of the Nit2.TR protein sequence in the NCBI database was a cyanide hydratase from *Fusarium* sp. (Table 3.2). No other cyanide hydratases were used to construct the maximum likelihood tree so the isolation of this sequence was not unexpected. Furthermore the separation of the predicted Nit 2 protein from nitrilases in the first branch of the tree reduces the likelihood that Nit 2 represents a nitrilase. Similarly, Nit4.TR was assigned to the third main branch of the tree alongside amidases and the Nitfhit proteins and thus less likely to represent a nitrilase. The Nit3.TR was included in the first branch near nitrilases but was on a separate sub-branch between nitrilases and nitrile hydratases. The Nit1.TR branched closely

with two well-characterized fungal nitrilases from *A. niger* and *F. solani* (Martinkova and Kaplan, 2010) and hence the predicted Nit 1 protein seemed the most likely candidate as a nitrilase enzyme amongst the *T. reesei* predicted proteins based on the Maximum likelihood tree.

3.3.4 Sequence comparison using ClustalW

Multiple sequence alignment of the predicted *T. reesei* proteins Nit 1 – 4 (Table 3.1) and the sequence library of known fungal and bacterial nitrilases (Section 3.2.2) was performed using ClustalW software to identify conserved residues. The four predicted *T. reesei* proteins were aligned against the sequence library to determine the conserved regions in each of the predicted nitrilase sequences. The ClustalW result using Nit 1 protein as the query sequence is shown in Figure 3.3. The number and color-based scoring denotes "0" and "blue" for the least conserved alignments and "10" and "red" for the most conserved alignments. Multiple sequence alignments were also carried out using the predicted Nit 2, Nit 3, Nit 4 proteins. A summary of the results from the multiple alignments of each of the four Nit proteins with the known nitrilases is provided in Table 3.3.

Table 3.3 Summary of multiple alignment analyses of the predicted *T. reesei* Nit 1 – 4 proteins with known nitrilases using ClustalW.

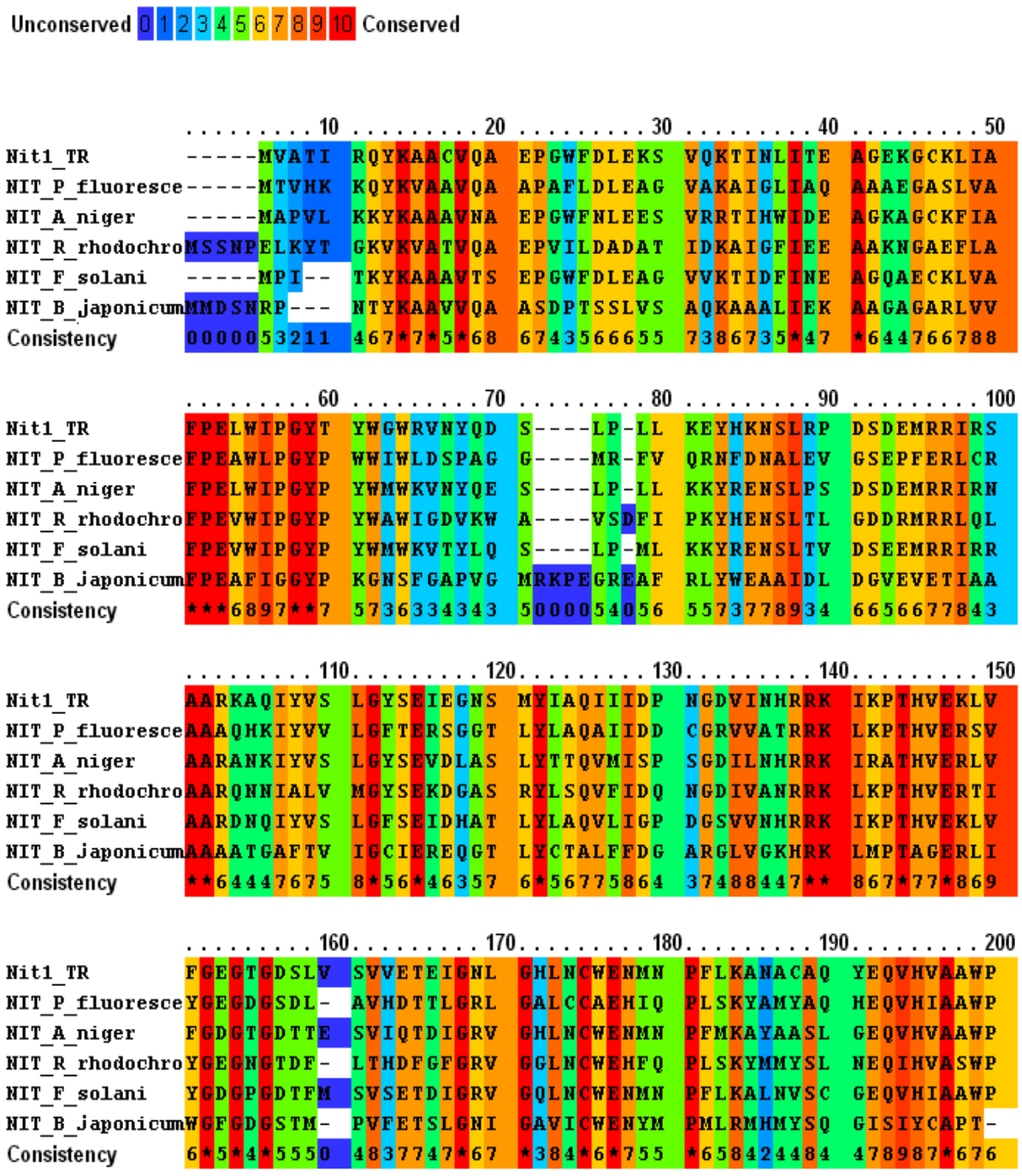
Protein	Sequence identity (%) [a]	Consensus scores > 7 [b]	Total alignment score [c]
Nit 1	39	77 residues	58276
Nit 2	37	74 residues	55706
Nit 3	36	75 residues	55842
Nit 4	34	64 residues	52904

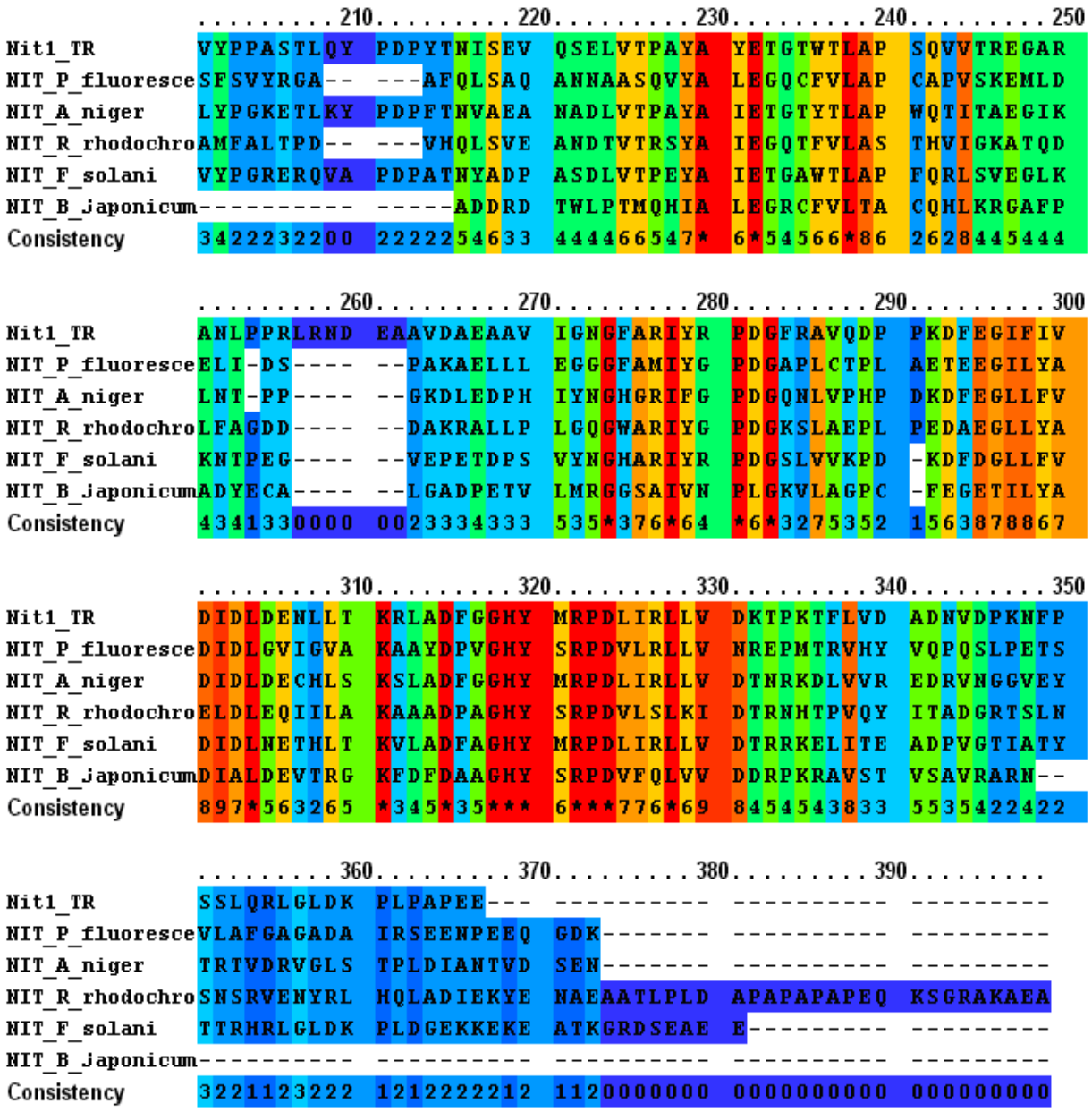
[a] Sequence identity: the percentage of amino acid residues in the query sequence (Nit protein) that aligned to identical residues in all the known nitrilases included in the analysis.

[b] Consensus scores > 7: the number of residues in the sequence that were scored 7 or higher based on their consensus with the corresponding residues in the known nitrilase sequences.

[c] Total alignment scores: the compilation of the residue-specific multiple alignment indices. The scores are calculated on iterative modes and the alignments are evaluated using BAliBASE benchmark database.

Figure 3.3 Multiple sequence alignment of the predicted *T. reesei* Nit 1 protein against known microbial nitrilases from *P. fluorescens, A. niger, R. rhodochrous, F. solani* and *Bradrhizobium japonicum*, carried out using ClustalW.





The ClustalW sequence alignment provided further evidence that amongst the four predicted *T. reesei* proteins Nit 1 – 4, the Nit 1 protein sequence had the greatest similarity to known nitrilases (Table 3.3). In addition, specific amino acids conserved amongst known nitrilases and the predicted Nit 1 protein were identified (Figure 3.3).

3.3.5 Sequence-based structural analysis of the predicted *T. reesei* Nit 1 – 4 proteins

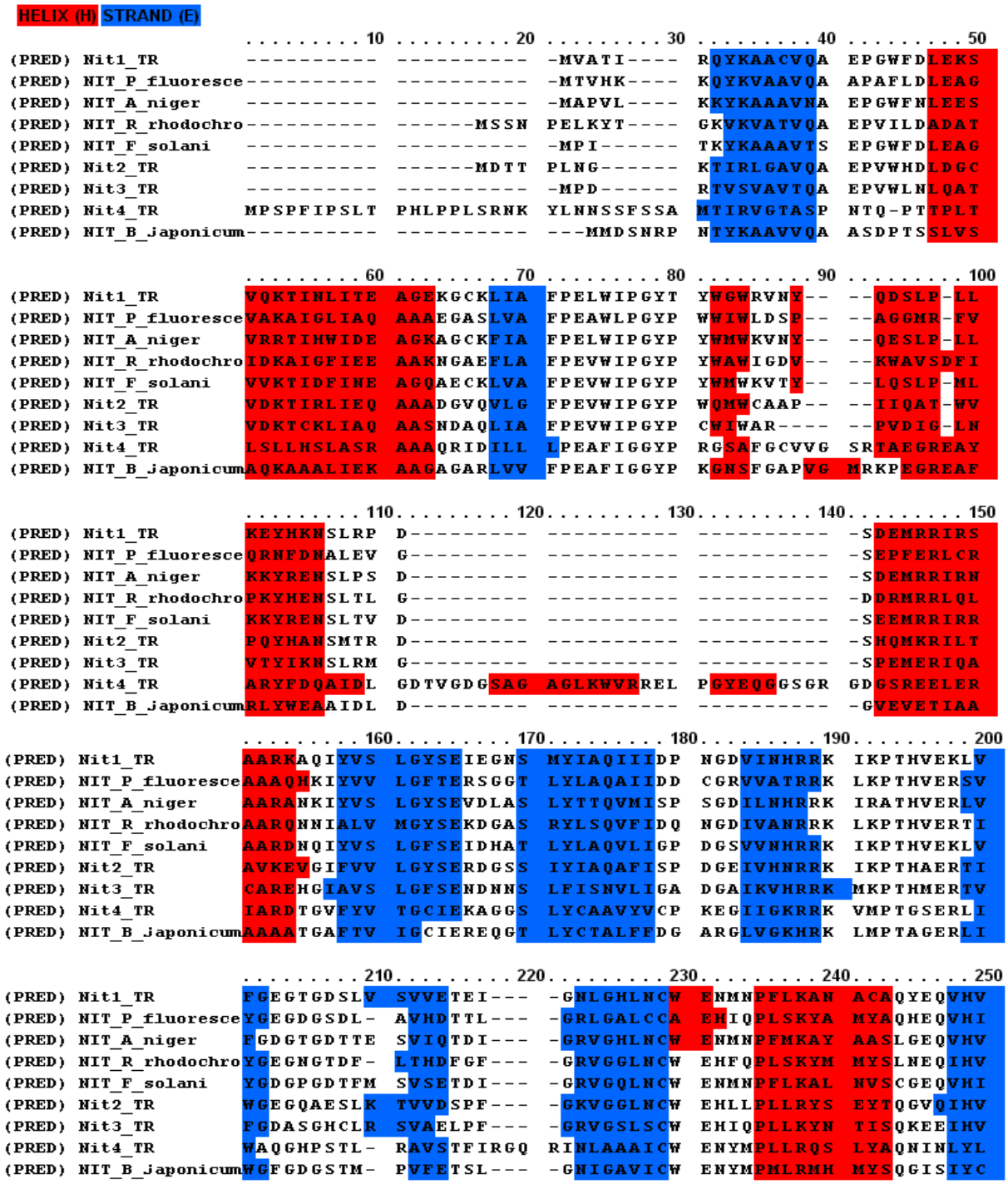
Nitrilases have been shown to share a characteristic monomeric fold and a catalytic active site containing three conserved active site residues: a highly conserved cysteine nucleophile, a glutamate base and a conserved lysine residue (Section 1.3.1). To identify and locate the catalytic triad in the predicted *T. reesei* proteins Nit 1 – 4, the sequences were analyzed using the ExPASy PROSITE tool (http://prosite.expasy.org/) and PRALINE software (http://www.ibi.vu.nl/programs/pralinewww/). All four of the Nit sequences were predicted to contain the conserved catalytic triad typical of nitrilases (Table 3.4).

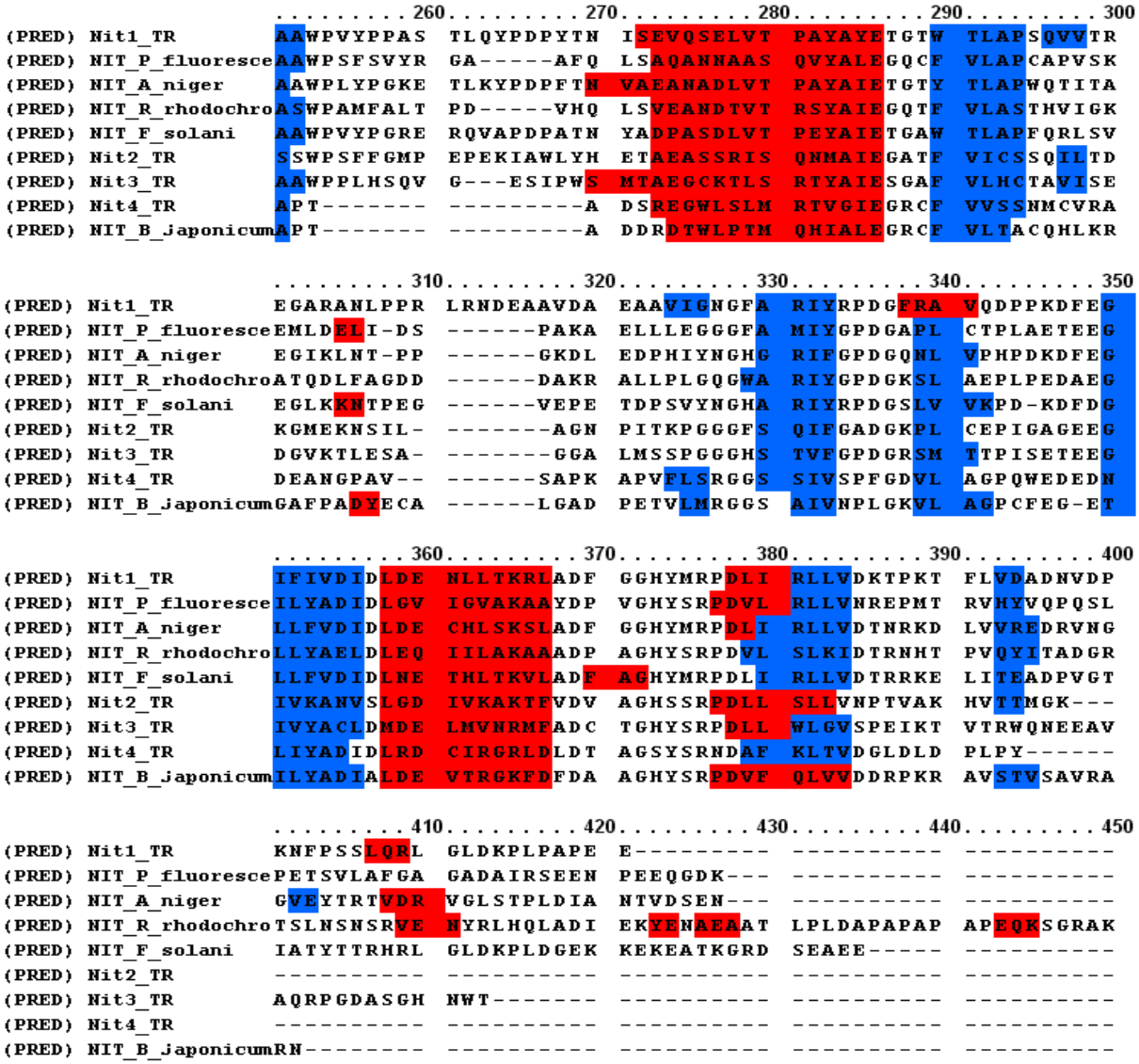
Table 3.4 Positions of amino acids that form the characteristic nitrilase catalytic triad in the *T. reesei* predicted proteins Nit 1 – 4, as determined by PROSITE (http://prosite.expasy.org/).

Protein	Proton acceptor (lysine) position	Base catalyst (glutamine) position	Nucleophile (cysteine) position
Nit 1	48	130	165
Nit 2	51	133	168
Nit 3	46	126	161
Nit 4	72	189	227

Features of the predicted secondary structure of Nit 1 – 4 from *T. reesei* and several known nitrilases based on their amino acid sequences are shown in Figure 3.4. Predicted helical regions (highlighted in red) and beta strands (highlighted in blue) were shown to occur in similar amino acid positions across all the known nitrilases and in the *T. reesei* Nit proteins. In addition, Nit 4 had two additional predicted helices between amino acid positions 118 and 132 which were not shared by any of the other nitrilases used in the alignment. *T. reesei* predicted proteins Nit 1, Nit 2 and Nit 3 shared an exclusive predicted beta strand at 296 – 298.

Figure 3.4 Alignment of the amino acid sequences of the *T. reesei* predicted Nit 1 – 4 proteins (Nit1_TR, Nit2_TR, Nit3_TR, Nit4_TR) and known nitrilases from other organisms (*P. fluorescens*, *A. niger*, *R. rhodochrous*, *F. solani*, *B. japonicum*), indicating the location of the predicted secondary structural features of helices (red) and beta strands (blue) as determined by PRALINE (http://www.ibi.vu.nl/programs/pralinewww/).





3.4 Discussion

The identification of putative nitrilase genes from *Trichoderma reesei* was carried out using bioinformatics and phylogenetic analyses of predicted proteins derived from the *T. reesei* wild-type QM6a genome sequence made available by JGI. Known nitrilases from fungi, bacteria and plants were used as models for all the analyses resulting in the identification of twelve predicted carbon-nitrogen hydrolases from *T. reesei*, four of which showed high homology to enzymes from the nitrilase branch (branch 1) of the nitrilase superfamily. Of these proteins, Nit 1 was predicted to have the highest similarity to known nitrilases.

3.4.1 Sequence similarities between predicted *T. reesei* proteins and known nitrilases

Twelve of the predicted proteins derived from the *T. reesei* genome sequence were carbon-nitrogen hydrolases showing significant homology to known nitrilases from other organisms. The genes encoding the predicted carbon-nitrogen hydrolases were from different locations in the *T. reesi* genome including scaffold 15 (Nit 1 – Nit 3), scaffold 19 (Nit 5 and 7), scaffold 1 (Nit 6 and Nit 11) and scaffold 32, 16, 4, 2 and 6 (Nit 4, 8, 9, 10 and 12 respectively). The expected sizes of most of the predicted Nit proteins were in the range of 32 – 45 kDa which is consistent with the molecular weights of several known nitrilases (Table 1, Chapter 1), with the exception of Nit 6 and Nit 12, which had higher predicted molecular weights (77.2 and 118.7 kDa respectively).

The predicted *T. reesei* carbon-nitrogen hydrolases were in turn used as query sequences for BLASTp search against the NCBI database which yielded their closest characterized homologs. Certain properties and characteristics of enzymes are reported to be sequence dependent (Minshull *et al*., 2005). Thus the sequence homologs of the predicted Nit proteins (Table 3.2) were used to assist in understanding likely functions of the *T. reesei* proteins. Based on the sequence alignment analyses performed using BLASTp, only four of the predicted *T. reesei* carbon-nitrogen hydrolases (Nit 1 – 4) were expected to have functions typical of enzymes from branch 1 of the nitrilase superfamily. Nit 1 had the greatest similarity to a nitrilase from a filamentous fungus (*A. niger*), thus provided the best indication of the existence of a nitrilase gene in the genome of the filamentous fungus *T. reesei*. Nit 2 and Nit 4 had the highest sequence similarity to a cyanide hydratase and indole-3-acetonitrilase respectively, branch 1 enzymes but not those that were the target of this research (*i.e.* nitrilases). The closest sequence homolog to the predicted Nit 3 protein was an aliphatic nitrilase of bacterial origin. Therefore Nit 3 continued to be of interest for possible representation of a putative nitrilase gene in the *T. reesei* genome at this stage. Based on the

BLASTp analysis, the remaining eight predicted *T. reesei* Nit proteins had closest sequence similarity to either amidases or acyltransferases (Table 3.2).

3.4.2 Phylogenetic analysis of the predicted *T. reesei* proteins

The possible functions of the predicted proteins were further investigated using phylogenetic analyses. In the initial analyses using PAUP 4 software, the tree of known enzymes from the nitrilase superfamily was broadly divided into two sections, one section containing enzymes from branch 1 of the nitrilase superfamily and the other containing enzymes with amidase activity (Figure 3.1). The four predicted *T. reesei* proteins (Nit 1 – 4) identified to share close sequence similarity to nitrilases in the BLASTp searches were the only proteins to be on the branch 1 "nitrilase" section of the tree. Out of the four enzymes, three of them (Nit1 – 3) clustered together with known fungal and bacterial nitrilases, whereas Nit 4 clustered in the section dominated by plant based nitrilases having indole-3-acetonitrilase properties (Figure 3.1). Although the role of auxins in fungi has not been well studied (Maor *et al*., 2004), it has been suggested that fungi may produce auxins to increase plant-fungal interaction (Robinette and Matthysse, 1990; Jouanneau *et al*., 1991).

3.4.3 Maximum likelihood tree

The Maximum likelihood tree (Figure 3.2) broadly separated the proteins into three main branches; the first branch contained nitrilases and nitrile hydratases, the second contained only the predicted *T. reesei* Nit 2 protein (Nit2.TR) and the third branch contained the amidases and Nitfhit proteins. The separation of the amidases and the Nitfhit proteins from the nitrilases based on their protein sequences indicated distinct evolution of the different members of the nitrilase superfamily. Although nitrile hydratases are not part of the nitrilase superfamily since they lack the catalytic triad specific to the superfamily, they still clearly clustered with the nitrilases based on sequence similarity thus justifying the combined

approach involving both sequence and structure based analyses (Section 3.4.4) to identify the nitrilases used in this study.

The predicted *T. reesei* Nit 1 protein clustered with the fungal nitrilases from *A. niger* and *F. solani*. The predicted Nit 3 enzyme was assigned to the first branch of the tree but was placed between nitrilases and nitrile hydratases. Although, both the enzyme families share some similar sequences, nitrile hydratases do not have nitrilase activity (ability to break down nitriles to carboxylic acid and ammonia in a single step; Section 1.2.5). Therefore the placement of the Nit 3 enzyme in the tree placed some doubts as to its function as a nitrilase despite the sequence similarity to a bacterial nitrilase in the BLASTp analysis (Table 3.2).

The predicted *T. reesei* Nit 2 protein branched separately from both the nitrilase/ nitrile hydratases and the amidase/ Nitfhit proteins in the Maximum likelihood tree (Figure 3.2). In the BLASTp analysis Nit 2 was most similar to a cyanide hydratase. No other cyanide hydratases were included amongst the proteins used in the Maximun likelihood so the isolation of Nit 2 was not surprising. Yet the placement of the Nit 2 protein on the tree served to further decrease the likelihood that the predicted protein actually represented a nitrilase. Likewise the placement of the Nit 4 protein in a branch dominated by NitFhit proteins, amidases and nitrile hydratases reduced the confidence in this predicted protein representing a nitrilase.

3.4.4 ClustalW and structure based sequence analysis

Based on the results of the multiple alignment analyses performed using ClustalW, the four predicted *T. reesei* proteins (Nit 1 – 4) were found to share sequence similarity with the other known nitrilases. The predicted Nit 1 protein had the highest total alignment score with the known nitrilases and a greater number of residues having consensus scores greater than 7 (Table 3.3), thus again suggesting that the Nit 1 sequence was the most likely candidate as a nitrilase. The structure prediction software PROSITE enabled location of the catalytic triad in

Nit 1 – 4, a defining feature of members of the nitrilase superfamily (Section 1.2). Furthermore, sequence alignment analyses and structure prediction using PRALINE software revealed similarity in the location of predicted structural features of the four *T. reesei* predicted Nit proteins with other known nitrilases.

3.5 Chapter summary

It was established in the work described in this chapter that *T. reesei* contained at least four genes which could encode enzymes from branch 1 of the nitrilase superfamily based on bioinformatics and phylogenetic analysis of predicted proteins derived from the *T. reesei* wild-type genome sequence hosted by JGI. A gene (now to be named *nit1*) on scaffold 15 of the genome was predicted to encode a protein (Nit 1) which had the highest similarity amongst the predicted *T. reesei* proteins to known nitrilases from other organisms based on BLASTp searches, multiple sequence analysis, phylogenetic and maximum likelihood tree investigations. Three other genes (now to be named *nit2 – 4*) from scaffolds 15, 15 and 32 respectively were predicted to encode enzymes of branch 1 of the nitrilase superfamily bearing some similarity to either nitrilases, indole-3-acetonitrilases or cyanide hydratases; however, the predicted proteins (Nit 2 – 4) were not considered as likely to represent nitrilases as Nit 1 based on the work described here. The *in silico* analysis for nitrilase gene identification paved the way for experimental work to investigate expression of the gene(s) by *T. reesei* when subjected to nitriles (Chapter 4).

Chapter 4: Expression of nitrilase gene(s) by *T. reesei* Rut-C30

4.1 Introduction

In silico analyses to identify new sources of nitrilases have shown that filamentous fungi are indeed a rich source of putative nitrilase genes (Martinkova *et al*., 2009). However there are only a few works that report on the production of nitrilases by filamentous fungi. Nitrilases have been successfully purified from fungal species such as *Aspergillus* and *Fusarium* (Winkler *et al*., 2009; Petrickova *et al*., 2012) and have displayed several advantages over their bacterial counterparts such as activity on a wide range of substrates, better substrate to product conversion rates, and added attributes of chemo- and enantioselectivity (Martinkova *et al*., 2009). Most known fungal nitrilases are reported to be inducible intracellular enzymes (Gong *et al*., 2012) and hence their production and characterization has involved finding appropriate culture media for induction, and techniques to identify the production and activity of the expressed nitrilases in the fungal mycelia.

In the work described in Chapter 3, *in silico* analysis of the *T. reesei* wild-type genome revealed four genes (*nit1 – 4*) predicted to encode proteins (Nit 1 – 4) with high sequence similarity to previously characterized enzymes from branch 1 of the nitrilase superfamily. The predicted Nit 1 protein sequence had particularly close homology to a branch 1 nitrilase from a filamentous fungus, *A. niger*, and therefore the corresponding *nit1* gene sequence presented the most likely candidate as a putative nitrilase gene for potential nitrilase expression. The work discussed in this chapter involved attempted induction of the putative nitrilase gene(s) in *T. reesei* using various media enriched with a selection of nitriles. The expression of nitrilase(s) was assessed by enzyme activity assays and proteomic approaches such as SDS-PAGE, 2D gel electrophoresis and mass spectrometric analyses.

The *T. reesei* strain Rut-C30 was used in the work described in this chapter. Fungal strain improvement programs have enabled the development of high secreting *T. reesei* strains from

the wild-type QM6a, particularly those with high capacity for cellulase production. The high-secreting hypercellulolytic mutant *T. reesei* Rut-C30 (Montenecourt and Eveleigh, 1979) is well characterized and used for large scale recombinant expression of homologous and heterologous proteins in research and industry. The strain would be an excellent host for large scale recombinant expression of a homologous nitrilase in the future. The genome sequence of *T. reesei* Rut-C30 was not available when this work was carried out so presence of the putative nitrilase genes in the Rut-C30 genome could not be verified at the time by the *in silico* analysis alone. Since then the *T. reesei* Rut-C30 genome has been sequenced (http://genome.jgi.doe.gov/ TrireRUTC30_1/home.html) and the presence of the putative genes *nit1 – 3* (Chapter 3) was confirmed.

4.2 Materials and methods

4.2.1 Cultivation media and culture conditions

Freshly harvested *T. reesei* Rut-C30 conidia were obtained as detailed in Section 2.3.2. Induction of the putative nitrilase genes in *T. reesei* was attempted using two different approaches. In the first approach, Czapek Dox medium (CDM; Table 2.4, Section 2.3.2) was used as the basis for five different culture media, each supplemented with a different nitrile (20 mM 3-cyanopyridine, acetonitrile, benzonitrile or adiponitrile) or ammonium sulphate (control) as the nitrogen source (Martinkova *et al.*, 2006). *T. reesei* conidia (2×10^6) were used to inoculate 50 ml cultures in 250 ml conical flasks. Incubation was at 28 ºC for 1 – 5 d at 250 rpm. Five parallel cultures were prepared for each nitrogen source and one of these cultures was harvested each day to assess mycelial mass. Mycelia were collected from the cultures by centrifugation at 5000 × *g* for 15 min.

In the second approach, pre-cultures of *T. reesei* were prepared in 250 ml conical flasks containing 50 ml of minimal medium (MM; Table 2.4, Section 2.3.2) inoculated with 2×10^6 conidia. After 60 h of incubation at 28 ºC, 250 rpm, mycelia were harvested by centrifugation

at 2000 × *g* for 15 min. The mycelia were washed three times in MilliQ water and added into five different culture media containing minimal medium supplemented with different nitriles (20 mM 3-cyanopyridine, acetonitrile, benzonitrile or adiponitrile) or ammonium sulphate (control) as the nitrogen source. Cultures were incubated at 250 rpm, 28 ºC for up to 48 h. The choice of the nitrile inducers used in this research was based on results from preliminary studies performed in our laboratory. Similar to the Czapek Dox-based cultures, five parallel cultures were prepared for each nitrogen source and one of these cultures was harvested each day to assess mycelial mass.

4.2.2 Assessment of the growth of *T. reesei* in the various media

The general health of *T. reesei* following growth in the various media (Section 4.2.1) was monitored by examination of the hyphae under a light microscope. To quantify growth rate, the mycelia were harvested from a series of parallel cultures of *T. reesei* grown for 1, 2, 3, 4 and 5 days in the various growth media. Mycelia were pelleted by centrifugation, washed with MQ water, collected on pre-weighed filter papers (Whatman, England) and freeze-dried. The mycelial dry weight was calculated by subtracting the weight of the filter paper.

4.2.3 Signal-P analysis to determine expected cellular location of predicted Nit proteins

The amino acid sequences of the predicted proteins Nit 1 – 4 (Table 3.1) were searched for signal peptide cleavage sites using the SignalP 4.1 software to determine the expected location of the proteins (intracellular or extracellular/secreted).

4.2.4 Extraction of protein from the *T. reesei* mycelia

Mycelial pellets from liquid cultures of *T. reesei* (Section 4.2.1) were used for the extraction of protein. The mycelial pellets were snap-frozen in liquid nitrogen and ground to a fine powder using a pre-chilled mortar and pestle. The ground mycelia were added to cold 0.1 M

phosphate buffer (pH 7.4) at a ratio of 1 g mycelia per 3 ml of buffer. The mycelial solution was then placed in an ice-water bath and subjected to sonication in short pulses of twelve seconds following which the tubes containing the mycelia were centrifuged at 5000 × *g* for 45 min at 4 ºC. Thereafter, the supernatant containing the total cellular proteins was collected and stored at 4 ºC until further use. The total protein concentration was determined using the protocol described by Bradford (1976) using bovine serum albumin (BSA) as protein standard. The mycelial proteins were subjected to fluorometric and colorimetric nitrilase activity assays (Section 2.7.1/ Section 4.2.5), HPLC analysis (Section 4.2.6), 1D SDS-PAGE (Section 2.6.2), 2D SDS-PAGE (Section 4.2.7) and mass spectrometric analysis (Section 4.2.8).

4.2.5 Enzyme activity assays

Nitrile-hydrolyzing activity by protein extracted from the *T. reesei* mycelia (Section 4.2.4) was investigated using fluorometric and colorimetric activity assays (Section 2.7). All the protein extracts were desalted by dialysis with regenerated cellulose dialysis tubing (Thermo Scientific, USA) before they were used for the assays. The enzyme activities were determined relative to a commercial nitrilase provided by Codexis Pty Ltd, USA that had a claimed activity of 43 U/mg when reactions were performed at 37 ºC. Approximately 25 µg of protein was used for all the assay reactions. Enzyme activity was investigated using 20 mM of adiponitrile, benzonitrile, acetonitrile or 3-cyanopyridine as the nitrile substrate.

Because the *T. reesei* mycelial protein could have contained nitrile hydratases and amidases (which together can convert nitriles to carboxylic acid and ammonia via a bi-enzymatic pathway), as well as nitrilases (which directly convert nitriles to carboxylic acid and ammonia via a mono-enzymatic pathway; Section 1.3.3), detection of ammonia could not be a definitive indication of the presence of a nitrilase. Therefore the assays were used to give an

indication of the possible presence of a nitrilase for further investigation using HPLC (Section 4.2.6) and proteomic analysis (Section 4.2.8).

4.2.6 Analytical High Pressure Liquid Chromatography to determine nitrilase activity

Analytical HPLC was performed using a Shimadzu model 10AD VP HPLC using a Shimadzu SPD-M10A diode array detector. The conversions were detected using a Phenomenex Luna c18 (2) column (250 x 4.6 mm, 5 µm) at ambient temperature. Solutions of 0.1% (v/v) phosphoric acid and acetonitrile were used as phase A and phase B, respectively. The standard reaction mixture contained 150 µl of the total protein suspended in 100 mM phosphate buffer (pH 7.5). Benzonitrile (4 mM) and adiponitrile (5 mM) were used as nitrile substrates for reactions involving total protein from cultures induced with benzonitrile and adiponitrile respectively. The reaction mixtures were incubated at 37 ºC for 35 min.

The products of nitrile hydrolysis (nitrile, amide and carboxylic acid) were separated using a gradient starting with 20% phase B maintained for 5 min, increased linearly to 70% phase B over 15 min and held at 70% phase B for further 2 min. The flow rate was maintained at 1 ml/min.

4.2.7 Two-Dimensional (2D) gel electrophoresis of mycelial proteins

4.2.7.1 Extraction of protein from *T. reesei* mycelia for 2D gel electrophoresis

Extraction of protein from the mycelia of *T. reesei* grown in media with and without nitriles was obtained using LS-TRI reagent following manufacturer's protocol. The protein pellets were resuspended in 2D buffer containing 6 M urea, 2 M thiourea, 4% CHAPS, 5 mM tributylphosphine, 10 mM acryamide and protease inhibitor cocktail tablets (Roche) in 100 mM Tris-HCl buffer, pH 8.5.

4.2.7.2 Isoelectric focusing and 2D electrophoresis

To approximately 250 µg of mycelial protein, 6 µL of 1 M dithiothretol (DTT), 0.5 µL of 4 – 7 IPG buffer (ampholytes), 0.5 µL of 3 – 10 IPG buffer, 3 µL of 0.001% bromophenol blue were added and the volume was adjusted to 215 µL using iso-electric focusing (IEF) rehydration buffer consisting of 6 M urea, 2 M thiourea and 4% (w/v) CHAPS in 100 mM Tris-HCl (pH 8.5). After centrifugation at 10,000 × g for 5 min, the solution was used to rehydrate the IPG strips. IPG strips, 11 cm in length with a pI range of 4 – 7 (Amersham Bioscience, Sweden) were used for the 2D- gel electrophoresis analysis. The IPG strips were rehydrated for approximately 4 h to ensure complete absorption of the sample solution. The IPG strips were later transferred onto focusing trays with the gel side facing down for 10 h at 50v. Moist absorbent paper wicks were placed at each of the electrodes to prevent collection of salts at the electrodes which affects the flow of current across the gradient. IPGs were focused using Biorad PROTEAN IEF Cell unit to a total of 78,000 Vh, using a four step focusing program starting at 200 volts (v) for 1 h, 1000 v for 1 h, 4000 v for 3 h and 8000 v for 8 h. The proteinswere allowed to stabilize at 500 volts for a further 2 h before removing the strips.

The IPG strips were equilibrated for 20 min with gentle rocking in a buffer containing 6 M urea, 20% (v/v) glycerol, 2% (w/v) SDS, 130 mM DTT in 50 mM Tris-HCl (pH 8.8). The strips were further re-equilibrated to SDS-PAGE running conditions using 1 x Tris-glycine buffer containing 10% (w/v) SDS. The IPG strips were placed on top of Precast 11 cm Criterion SDS-PAGE gels (BioRad, Australia) and were run at a constant 50 mA. Special care was taken to prevent the formation of air bubbles underneath the strips. Following the run, the gels were stained with Coomassie Blue G-250.

4.2.8 Protein identification using mass spectrometry

4.2.8.1 MALDI TOF/TOF MS/MS

Lanes from 1D SDS-PAGE gels of *T. reesei* mycelial protein (Section 4.2.4) were cut into small pieces, destained and dried. The proteins were digested using molecular grade trypsin for 16 h at 37 °C. The peptides were re-acidified, desalted and concentrated by zip-tip (C18, Millipore, U.S.A). The samples were eluted (Xcise, Proteome Systems) onto the sample plate with 1 µl of matrix (α-cyano-4-hydroxy cinnamic acid, 1 mg/ml in 90% (v/v) acetonitrile, 0.1% (v/v) TFA) and were air dried.

MALDI (Matrix Assisted Laser Desorption Ionization) was performed using an Applied Biosystems 4800 Proteomics Analyzer. A Nd: YAG laser (355nm) was used to irradiate the sample. Spectra were acquired in reflectron mode in the mass range of 700 to 3500 Da. The instrument was switched to TOF/TOF (tandem time of flight) mode where the strongest peptides from the MS scan were isolated, fragmented and re-accelerated to measure their masses and intensities. A near point calibration was applied to give a typical mass accuracy of 50 ppm or less. The peak lists of the proteins were exported to be searched against the suitable protein database.

4.2.8.2 In gel trypsin digestion and nanoLC MS/MS mass spectrometry

Each of the SDS-PAGE gel lanes (Section 4.2.2) and 2D-PAGE protein spots (Section 4.2.7) containing *T. reesei* mycelial proteins were cut into small pieces using a scalpel. The gel pieces were destained, rehydrated, reduced and alkylated before digestion using molecular grade trypsin. The digested peptides were then subjected to mass spectrometric analyses using nanoLC MS/MS (nanoscale liquid chromatography tandem mass spectrometry). Automated peak recognition, dynamic exclusion and MS/MS of the top six most intense precursor ions at 35% normalization collision-induced dissociation energy were performed using Xcalibur software (Version 2.0, Thermo).

4.2.8.3 Protein identification from mass spectra

The peptide mass data of the protein samples produced by the mass spectrometry was analyzed using the search engine software called MASCOT (Perkins *et al*., 1999). An in-house protein database was created using predicted *T. reesei* proteins derived from the *T. reesei* genome sequence (http://genome.jgi-psf.org/Trire2/). The peptide fragments and the fragmented ions obtained from the mass spectra were compared to the predicted fragmentation patterns of the translated proteins from *T. reesei*. The proteins were identified by MASCOT software on the basis of their probability based Mowse scores (Perkins *et al*., 1999).

4.3 Results

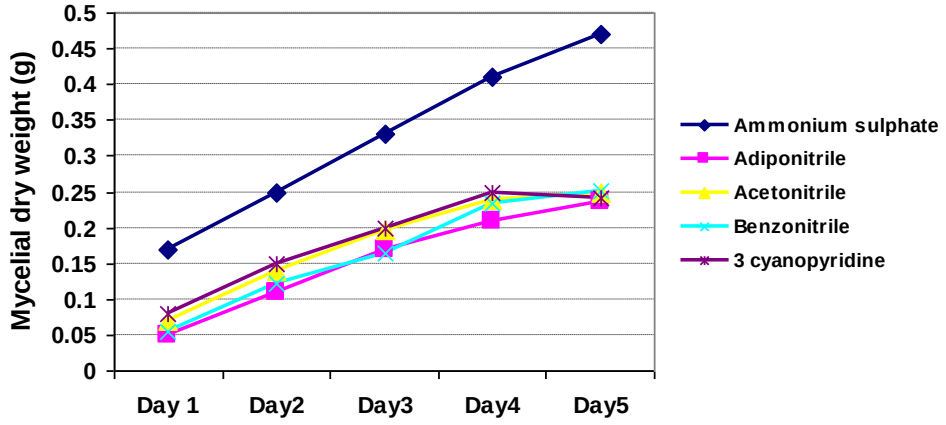
4.3.1 Growth and protein production of *T. reesei* in the various growth media

Hyphae from 5 d cultures grown in liquid minimal medium (MM) and Czapek Doz medium (CDM) supplemented with different nitriles were viewed under a light microscope and compared to hyphae from the "control" culture containing MM and ammonium sulphate as nitrogen source. Considerable morphological differences such as a decrease in hyphal diameter and increased vacuolation were seen in the cultures grown in media enriched with nitriles and those containing ammonium sulphate as the nitrogen source (control culture). Substantial differences were also observed in the amount and rate of mycelial growth (Figure 4.1).

For *T. reesei* cultures cultivated in CDM- and MM-based media, maximum culture biomass was obtained when ammonium sulphate was used as nitrogen source (control). In Figure 4.1 (b) showing the mycelial growth of *T. reesei* grown in CDM, cultures in nitrile-supplemented medium exhibited much slower growth rates when compared to the ammonium sulphate-supplemented control. Furthermore, in cultures cultivated in CDM enriched with 3

cyanopyridine, the mycelila mass was seen to fall after day four. Slow growth rate and decrease in mycelial biomass can be attributed to stressful growth conditions (Riley *et al*., 2000). It appeared that the *T. reesei* cultures were under considerable physiological stress when forced to use nitriles as their primary nutrient source.

(a) Czapek Doz medium base



(b) Minimal medium base

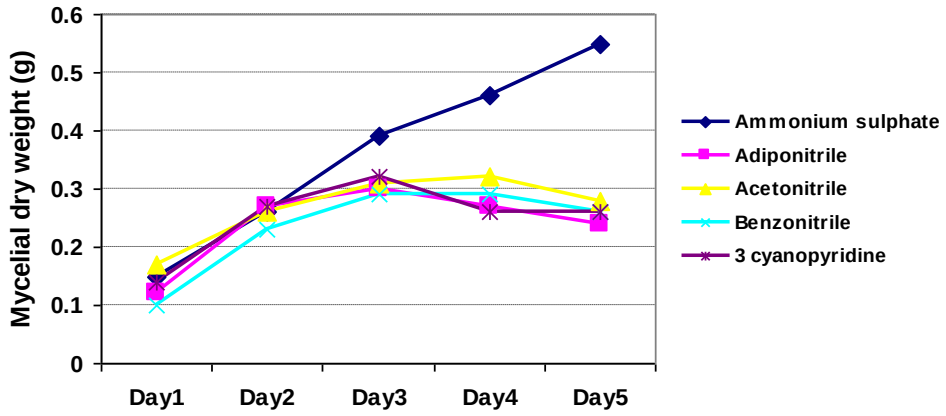


Figure 4.1 Mycelial growth during cultivation of *T. reesei* in (a) Czapek Doz medium and (b) minimal medium, supplemented with ammonium sulphate or different nitriles as nitrogen source. The Czapek Doz-based cultures contained nitriles prior to conidial inoculation. However, the first 60 h of growth shown for the minimal medium-based cultures occurred in pre-cultures containing ammonium sulphate as nitrogen source (data not shown); mycelia from these cultures were then used for inoculation of the nitrile-supplemented or ammonium sulphate-supplemented minimal media. Mycelial growth is represented by the dry weight of the mycelia harvested each day from a series of parallel cultures (Section 4.2.1).

In Figure 4.1 (b) showing the growth of *T. reesei* in MM-based cultures, the mycelial biomass of pre-cultures increased over the first 60 h. However after the mycelia were transferred to nitrile-supplemented MM the growth not only ceased but the mycelial biomass decreased. Therefore, the formation of healthy mycelia prior to the introduction of nitriles did not increase the ability of *T. reesei* to tolerate and use the nitriles as a nitrogen source.

4.3.2 Assessment of nitrilase activity using fluorometric and colorimetric assays

The predicted *T. reesei* proteins Nit 1 – 4 (Chapter 3) were expected to be intracellular proteins by SignalP analysis (http://www.cbs.dtu.dk/services/SignalP/). Therefore, fluorometric (Section 2.7.1) and colorimetric (Section 2.7.2) nitrilase activity assays were carried out primarily using protein extracted from the *T. reesei* mycelia (Section 4.2.4). Some exploratory nitrilase activity assays were carried out on the *T. reesei* culture supernatants but no nitrilase activity was detected.

Total mycelial protein extracted from *T. reesei* cultivated in Czapek Doz (CDM) and minimal medium (MM) enriched with adiponitrile, benzonitrile, acetonitrile or 3 cyanopyridine were used as the potential enzyme source for the assays. The activities were determined using 25 µg of protein, and adiponitrile, benzonitrile and 3 cyanopyridine as substrates. Despite repeated assays using a variety of reaction conditions, the production of ammonia (the expected end product of nitrile hydrolysis) was not detected using the mycelial protein of any of the *T. reesei* cultures as potential enzyme source. The inability to detect ammonia could have been due to the absence of nitrile-hydrolyzing enzymes (nitrilases and/or nitrile hydratase and amidases) or a lack of sensitivity of the assays for detecting low levels of enzyme activity in crude cell samples. Therefore a more sensitive technique was used as described below.

4.3.3 Determination of nitrilase activity by analytical HPLC

High Performance Liquid Chromatography (HPLC) provided a more sensitive and refined technique for the detection of nitrilase activity in the *T. reesei* mycelia. As well as detecting the final products of nitrile degradation, HPLC also allowed the detection of amides, intermediary products only produced during the bi-enzymatic nitrile hydratase/ amidase pathway (Section 1.3.3); hence, nitrilase activity could be distinguished from nitrile hydratase/ amidase activity.

The HPLC analysis revealed that the mycelial protein of adiponitrile-supplemented minimal media cultures exhibited nitrilase activity. The chromatogram of the products of the adiponitrile hydrolysis reaction catalyzed by the mycelial protein is shown in Figure 4.2. The retention times for the standards adiponitrile, adipamide and adipic acid were 20.6, 13.3 and 5.69 min respectively.

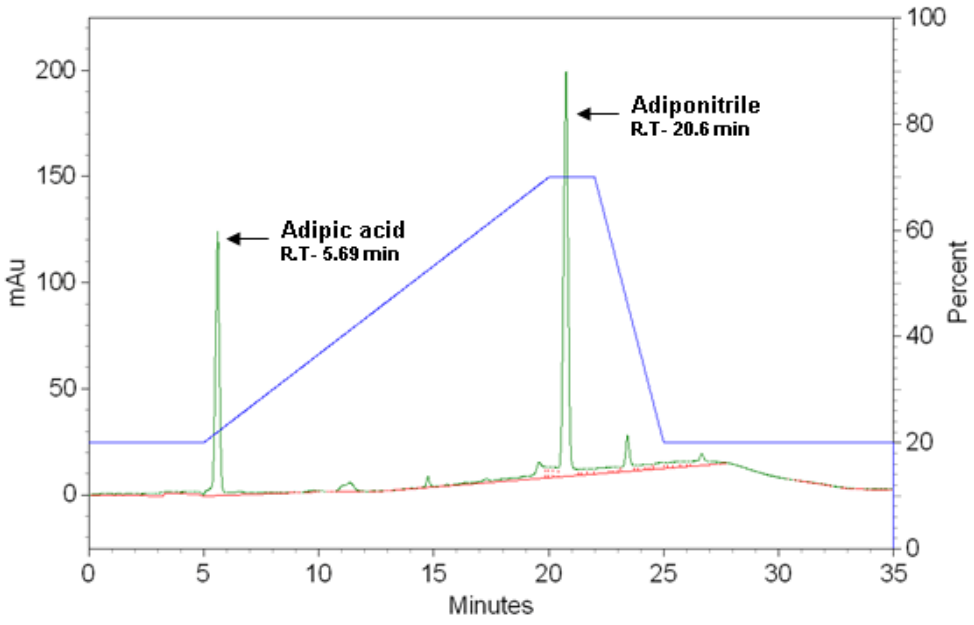


Figure 4.2 HPLC chromatogram of adiponitrile degradation by mycelial protein from *T. reesei* grown in adiponitrile-supplemented minimal media. Peak identification: at retention time 5.69 min- adipic acid; at retention time 20.6 min- adiponitrile.

The peak indicating the formation of adipic acid which was the expected product of the hydrolysis of adiponitrile clearly suggested the presence of a nitrilase enzyme in the total

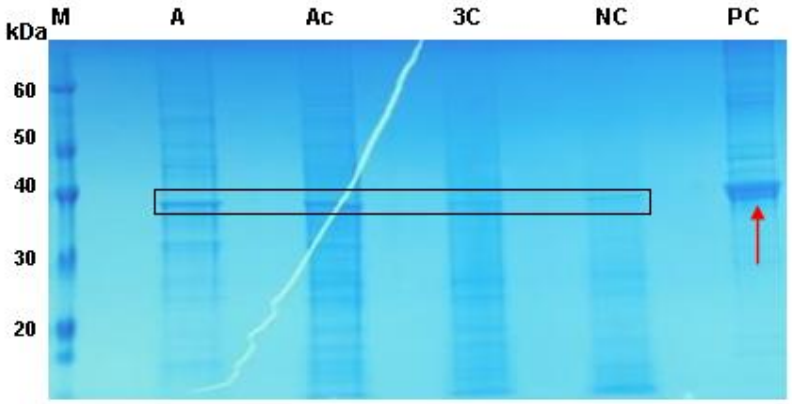
mycelial protein fraction of adiponitrile-supplemented minimal media cultures of *T. reesei*. No trace of adipamide as intermediary product of a bi-enzymatic nitrile hydratase/ amidase pathway was detected by the HPLC, strongly indicating that a *T. reesei* nitrilase was the key player in the enzyme activity observed.

However, no trace of benzoic acid or benzamide were observed after the HPLC analysis of potential benzonitrile degradation by mycelial protein from benzonitrile-supplemented cultures, nor were the substrates acetonitrile or 3-cyanopyridine degraded by the mycelial protein from benzonitrile-supplemented cultures. Furthermore, no trace of nitrile hydrolysis was evident towards any of the nitrile substrates using the total mycelial protein from *T. reesei* cultures supplemented with either 3 cyanopyridine or acetonitrile.

4.3.4 SDS-PAGE of total mycelial protein from *T. reesei* cultures supplemented with different nitriles

The mycelial protein from *T. reesei* grown in nitrile-supplemented media (CDM and MM) was subjected to SDS-PAGE to enable separation and visualization of the proteins according to molecular weight (Figure 4.3).

(a) Czapek Doz medium



(b) Minimal medium

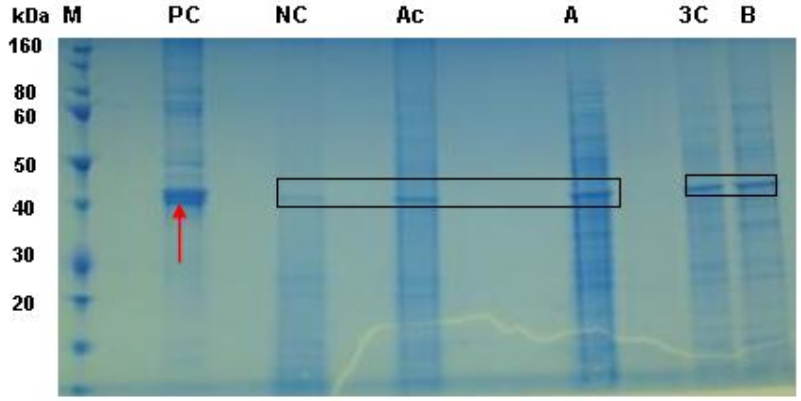


Figure 4.3 SDS-PAGE of mycelial protein from *T. reesei* grown in (a) Czapek Doz and (b) minimal medium, supplemented with adiponitrile (A), acetonitrile (Ac), 3-Cyanopyridine (3C) or benzonitrile (B), or without nitrile supplementation (NC). Lane M: molecular weight ladder; Lane PC: Commercial nitrilase from Codexis, USA (red arrow). Black boxes surround the most prominent proteins in the *T. reesei* mycelia, which were of the approximate molecular weight of the predicted Nit 1 protein (40 kDa).

The most prominent protein in the *T. reesei* mycelia, regardless of the nitrogen source in the culture media, was of a molecular weight of approximately 40 kDa, similar to the molecular weight of the commercial nitrilase (PC; Figure 4.3). In relation to the molecular weight marker, the prominent bands appeared slightly lower on the gel containing mycelial protein from the Czapek Doz cultures (Fig. 4.3a) than in the gel containing mycelial protein from the minimal medium cultures (Fig. 4.3b) but this was considered to be an artifact due to variations in the migration of the marker and/or sample protein rather than a function of the culture media because the commercial nitrilase also dropped to a similar degree.

The expected molecular weight of the predicted *T. reesei* Nit 1 protein that had the greatest sequence similarity to known nitrilases was 40.1 kDa (Table 3.1, Section 3.3.1). Therefore it was considered possible that the prominent protein in the *T. reesei* mycelia (Figure 4.5) was the predicted Nit 1 nitrilase. The *T. reesei* predicted protein Nit 4 was expected to have a similar molecular weight (39.7kDa, Table 3.1) but as Nit 4 had a greater sequence similarity to plant indole-3-acetonitrilases than fungal nitrilases, and the nitrile sources included in this experiment did not include indole-3-acetonitrile it was considered less likely that the prominent protein represented the predicted Nit 4.

The prominent protein of approximately 40 kDa appeared most abundant in the mycelia from minimal media-based cultures supplemented with adipontirile (Fig. 4.5b, lane A); correspondingly, it was only the mycelia from these cultures that nitrilase activity was detected by HPLC (Section 4.3.3). However, the ~40 kDa protein was also present, although at lower abundance, in mycelia from *T. reesei* cultures supplemented with other nitriles and could even be observed in the mycelia from cultures that did not contain nitriles (NC, Fig. 4.5). This suggested that the protein(s) of ~40 kDa in the *T. reesei* mycelia could have included constitutive intracellular proteins; it was also possible that the predicted Nit1 protein was produced constitutively at basal levels. Further investigation was carried out by mass spectrometry (Section 4.3.5).

4.3.5 Identification of proteins from the 1D SDS-PAGE gels

To identify the mycelial proteins present in the gels (Figure 4.3), 1 cm wide portions including the prominent 40 kDa protein bands were cut for MALDI TOF/TOF analysis. Identified proteins are shown in Table 4.1.

Table 4.1 Proteins from the *T. reesei* predicted protein database that were identified by MALDI TOF/TOF MS/MS of 35 – 45 kDa protein on SDS-PAGE gels (Fig. 4.3) of mycelial protein of *T. reesei* grown in Czapek Doz (C) or minimal medium (M) supplemented with adiponitrile (Adipo), acetonitrile (Aceto), 3-Cyanopyridine (3Cyan) or benzonitrile (Benzo), or in minimal medium without nitriles (Control).

Sample	Protein	Score [a]	Confidence [a]
3Cyan.C	Glyceraldehyde-3-phosphate dehydrogenase	154	High
3Cyan.M	Glyceraldehyde-3-phosphate dehydrogenase	362	High
Adipo.C	Glyceraldehyde-3-phosphate dehydrogenase	102	High
Adipo.M	Malate dehydrogenase activity	54	High
Aceto.C	Glyceraldehyde-3-phosphate dehydrogenase	263	High
Aceto.M	Glyceraldehyde-3-phosphate dehydrogenase	84	High
Control	Glyceraldehyde-3-phosphate dehydrogenase	32	Low

[a] The Mascot score (Score) and confidence level of identification was produced by the search engine software Mascot and indicate the significance of the match with the predicted proteins.

MALDI TOF/TOF MS/MS (Table 4.1) was used to identify the protein in the smaller gel pieces since the technique can identify only the top 10% of the abundant proteins present in a sample. NanoLC MS/MS (Table 4.2) is a much more powerful tool with capacity to generate a higher number of mass spectra with much better resolution and so was used to identify proteins present in entire gel lanes. The entire gel lanes were cut into 16 small fragments, trypsin digested and subjected to nanoLC/MS/MS analyses (Section 4.2.4). The resulting protein identifications are shown in Table 4.2.

Table 4.2 Proteins from the *T. reesei* predicted protein database that were identified by nanoLC MS/MS of SDS-PAGE gel lanes (Fig. 4.5) of total mycelial protein of *T. reesei* grown in Czapek Doz (C) or minimal medium (M) supplemented with adiponitrile (Adipo), acetonitrile (Aceto), 3-Cyanopyridine (3Cyan) or benzonitrile (Benzo), or in minimal medium without nitriles (Control).

Sample	Protein	Score [a]	Confidence [a]
3 Cyan.C	Glyceraldehyde-3-phosphate dehydrogenase	251	High
	Glycoside hydrolase	146	High
	Unknown	98	High
3 Cyan.M	Unknown (similar to NmrA- like family protein)	133	High
	Unknown	112	High
	Glycoside hydrolase	74	High
Adipo.C	Unknown	243	High
	Beta-Ig-H3/Fasciclin	188	High
	Glyceraldehyde-3-phosphate dehydrogenase	138	High
Adipo.M	Malate dehydrogenase	129	High
	FAD linked oxidase	102	High
	Extracellular matrix protein precursor	79	High
Aceto.C	Glyceraldehyde-3-phosphate dehydrogenase	272	High
	Glycoside hydrolase	161	High
	Unknown	132	High
Aceto.M	ATP synthase	134	High
	Glycoside hydrolase	109	High
	Glyceraldehyde-3-phosphate dehydrogenase	91	High
Control	Glyceraldehyde-3-phosphate dehydrogenase	366	High
	Zn-finger, C2H2 type	236	High
	Folylpolyglutamate synthetase	207	High

[a] The Mascot score (Score) and confidence level of identification was produced by the search engine software Mascot and indicate the significance of the match with the predicted proteins.

Most of the *T. reesei* mycelial proteins that were identified from the 1D SDS-PAGE gels using MALDI TOF/TOF MS/MS were also identified by nanoLC MS/MS (Table 4.1, Table 4.2). However, none of the predicted *T. reesei* carbon-nitrogen hydrolases (Nit 1 – 12) generated by *in silico* translation of the *T. reesei* genome (Table 3.1, Section 3.3.1) were identified from the 1D SDS-PAGE gels of *T. reesei* mycelial protein. In the MALDI TOF/TOF MS/MS analyses, the mass spectra were dominated mainly by peptides assigned to Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). GAPDH is an abundant intracellular protein that assists in the break down of glucose for energy besides playing an important role in several other cellular "housekeeping" functions (Riley *et al*., 2000). The predicted *T. reesei* GADPH has an expected molecular weight of 36.2 kDa, which was within the range of the small gel pieces cut for MALDI TOF/TOF MS/MS analysis.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was also identified by the nanoLC MS/MS analyses of gels containing mycelial protein from *T. reesei* grown in nitrile-supplemented Czapek Dox medium and minimal medium. Other intracellular proteins were also identified (*e.g.* glycoside hydrolase, ATP synthase, beta-Ig-H3/fasciclin, folylpolyglutamate synthetase) but no carbon-nitrogen hydrolases or nitrilases were detected. It was possible that the nitrilase, if present, could have been at very low levels in comparison to the abundance of the other intracellular proteins detected, and if so, the peptides of the more abundant proteins in the gel slices could have dominated the mass spectra and masked the presence of the nitrilase. To overcome this potential problem, further separation of the mycelial proteins prior to MS analysis was carried out using 2D electrophoresis.

4.3.6 2D gel electrophoresis of mycelial protein

Mycelial proteins from *T. reesei* grown in minimal medium supplemented with adiponitrile, and minimal medium supplemented with benzonitrile, were subjected to 2D electrophoresis. As the expected isoelectric points (pI) of the predicted *T. reesei* carbon nitrogen hydrolases of

interest (Table 3.1) were between 4 and 7, the 2D gel analysis was performed using pH 4 to 7 IPG strips.

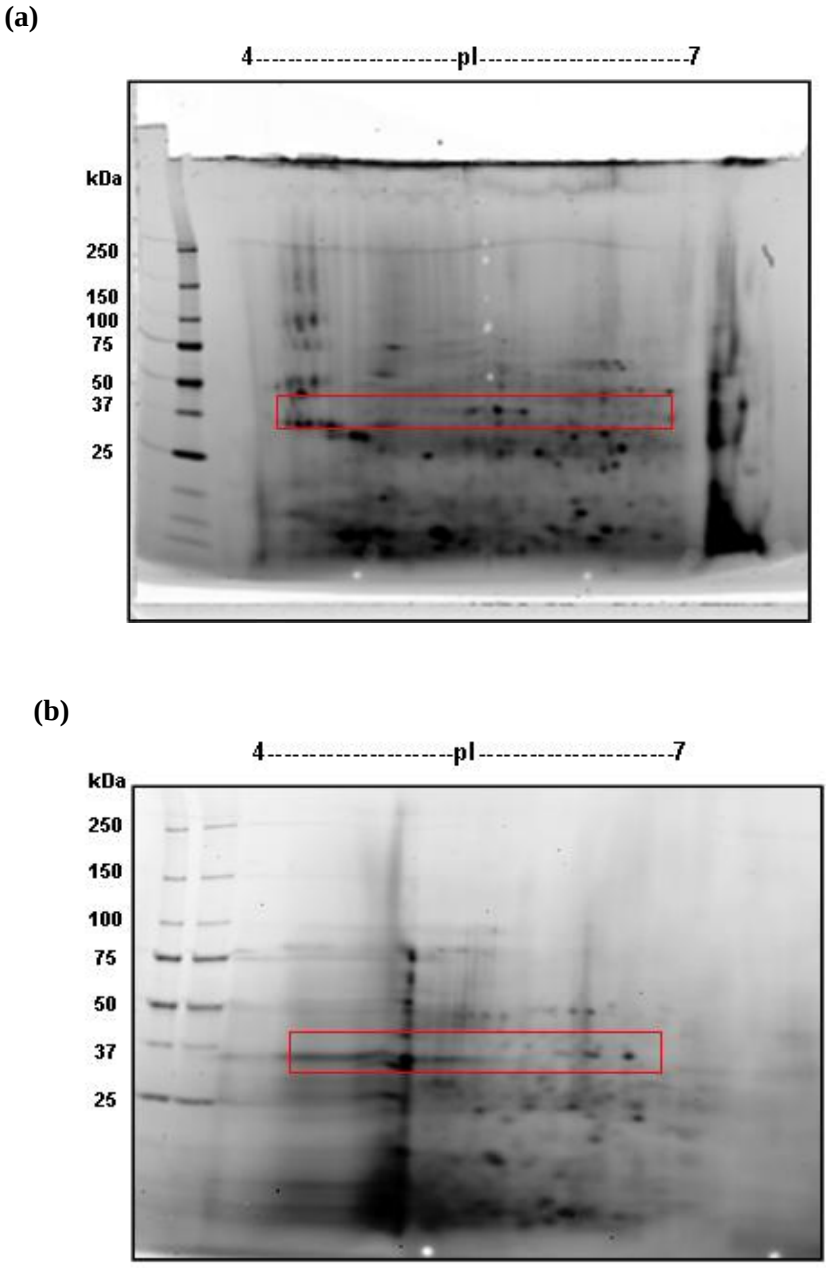


Figure 4.4 2D gels produced by the electrophoresis of mycelial protein from *T. reesei* grown in minimal medium supplemented with (a) adiponitrile and (b) benzonitrile. Approximately 250 µg of total protein was used per gel. Proteins were stained with Coomassie Blue G. The red rectangle indicates the portion of the gels subjected to nanoLC MS/MS for protein identification.

Portions of each gel containing proteins of approximate molecular weight 35 – 45 kDa (red rectangles, Figure 4.4) were sliced and diced into small pieces and each piece was subjected to nanoLC/MS/MS as described above. The top five identifications assigned to the proteins

from the 2D gels by Mascot search of the database of *T. reesei* predicted proteins are listed in Table 4.3.

Table 4.3 Proteins from the *T. reesei* predicted protein database that were identified by nano LC MS/MS of portions of 2D gels (Fig. 4.4) of total mycelial protein of *T. reesei* grown in minimal medium supplemented with (a) adiponitrile and (b) benzonitrile.

(a)

Protein name	**Predicted MW (kDa) [a]**	**Predicted pI [a]**	**Score [b]**
Glyceraldehyde-3-phosphate dehydrogenase	36.3	6.61	805
Unknown (similar to NmrA- like family protein)	35	5.87	415
Nit 1 (Nitrilase)	40.1	5.24	267
Peptidase	42.4	4.8	140
Subtilase	37.9	6.40	42

(b)

Protein name	**Predicted MW (kDa) [a]**	**Predicted pI [a]**	**Score [b]**
Glyceraldehyde-3-phosphate dehydrogenase	36.3	6.61	680
Zn-finger	38.3	6.8	362
Peptidase	42.4	4.8	140
Hypothetical protein	46.9	6.26	98
Unknown (similar to NmrA- like family protein)	35	5.87	33

[a] The molecular weight (MW) and isoelectric point (pI) of the proteins were derived using the ExPASy pI/MW tool (http://web.expasy.org/compute_pi/).

[b] The Mascot score (Score) was produced by the search engine software Mascot and indicates the significance of the match with the predicted proteins.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which was identified earlier to be the most dominant protein in the mycelia of *T. reesei* (Table 4.1, Table 4.2) was again identified from both the 2D gels (Figure 4.4) with Mascot scores of 680 and 805 respectively (Table

4.3). Other intracellular proteins were also identified, but most notably, the predicted Nit 1 nitrilase was identified from the mycelia of *T. reesei* that had been grown in minimal medium supplemented with adiponitrile (Table 4.3a).

The predicted Nit 1 nitrilase was not identified from the 2D gel of mycelial protein of *T. reesei* grown in benzonitrile-supplemented minimal medium (Figure 4.6b; Table 4.3) or 2D gels produced from mycelial protein from 3 cyanopyridine- and acetonitrile-supplemented cultures. However, 2D gel electrophoresis and subsequent mass spectrometric analysis of mycelial protein from another adiponitrile-supplemented culture of *T. reesei* again resulted in the identification of the predicted Nit 1 protein.

4.4 Discussion

Out of the twelve predicted *T. reesei* carbon-nitrogen hydrolases identified earlier in this work (Chapter 3), only the Nit 1 protein could be identified as an expressed protein when *T. reesei* Rut-C30 was subjected to nitriles. Expression of Nit 1 was achieved in the presence of the aliphatic nitrile adiponitrile but not in the presence of the aliphatic acetonitrile or the aromatics benzonitrile and 3 cyanopyridine. Based on the *in silico* analysis, the predicted Nit 1 protein had the greatest sequence similarity to known nitrilases compared to the other predicted carbon-nitrogen hydrolases in the *in silico* analysis (Chapter 3). None of the predicted proteins with less sequence similarity to nitrilases (Nit 2 – 12), including those with greater similarity to cyanide hydratases, indole-3-acetonitrilases, amidases, NAD-synthases and Nit-Fhit proteins (Table 3.1) were detected in the mycelia *T. reesei* Rut-C30 when grown in the presence of any of the nitriles tested. The effect of the different media used for nitrilase induction on the growth and nitrilase activity of *T. reesei* Rut-C30 is discussed below, along with the advantages and limitations of the analysis techniques.

4.4.1 Growth of *T. reesei* in the presence of nitriles

Induction of the putative nitrilase gene(s) of *T. reesei* was attempted using culture media supplemented with various nitriles, including 3-cyanopyridine and benzonitrile which were already known inducers of nitrilases in fungi (Harper 1977; Kaplan *et al*., 2006). Post-induction, to gauge the physiological effect of nitrile enrichment on *T. reesei*, the dry weight of *T. reesei* mycelia grown both in the presence and absence of the nitriles was determined (Figure 4.1). The biomass production (mycelial growth rate) and the culture cell density have been reported to act as indicators of cellular metabolism and stress affecting cell growth (Riley *et al*., 2000). However, slow growth can also be associated with higher protein production. For example, in the work performed by Arvas *et al*. (2011) on *T. reesei*, it was reported that the rate of protein production was inversely proportional to the total biomass of the cultures wherein high biomass concentrations of the cultures affected the amount of protein produced by directing more energy towards cell growth rather than protein production (Arvas *et al*., 2011). This phenomenon may help explain why the SDS-PAGE gel (Figure 4.3b) revelaed a higher protein content in the *T. reesei* mycelia from the adiponitrile-supplemented minimal media compared to that grown in the presence of other nitriles despite similar or reduced mycelial weight in adiponitrile-supplemented minimal medium (Figure 4.1).

4.4.2 Detection of nitrilase production by *T. reesei*

Detection of nitrilase activity in the *T. reesei* mycelia was first attempted using fluorometric and colorimetric activity assays, which depend on the production of ammonia as an indicator of nitrile hydrolysis (Section 1.6.2.2). The production of ammonia from nitriles can result from a bi-enzymatic pathway involving nitrile hydratase and amidase, or a mono-enzymatic pathway involving nitrilase alone (Section 1.3.3). It was quite possible that the mycelia might have contained nitrile hydratase(s) and amidase(s); therefore the assays were used only as an

indicator of possible nitrilase activity prior to further analysis. However the presence of ammonia in the mycelia or supernatants of any of the *T. reesei* cultures was not detected using these assays despite repeated attempts using a variety of assay conditions (Section 4.3.2). The inability to detect ammonia production could have been due to the absence of nitrile-hydrolyzing enzymes. Alternatively, the enzymes may have been present at levels below that detected by the sensitivity of the assays, and/or the presence of other native proteins may have interfered with the assay readings, for example by blocking the formation of the fluorophore complex.

High performance liquid chromatography provided a more sensitive method of analysis to detect nitrilase activity in the *T. reesei* mycelia. The technique not only allowed the detection of the end products of nitrile hydrolysis, even when produced at low levels, it also allowed the detection of intermediary products of the bi-enzymatic pathway of nitrile hydrolysis to distinguish nitrilase activity from nitrile hydratase/ amidase activity (Section 1.3.3). Consequently, nitrilase activity could be identified in the mycelia of *T. reesei* grown in adiponitrile-supplemented minimal media (Section 4.3.3). The absence of any trace of amide by-product indicated that the breakdown of the nitrile substrate was carried out by a nitrilase rather than via the nitrile hydratase/amidase pathway.

Production of nitrilase by *T. reesei* was then investigated using proteomic approaches. The 1D SDS-PAGE gels of *T. reesei* mycelial proteins revealed proteins of the estimated size of the predicted Nit 1 nitrilase (Figure 4.3) but mass spectrometric analysis of the protein in the gel did not result in the detection of peptide fragments matching those of Nit 1 (Section 4.3.5). Similar to the experience with the fluorometric and colorimetric assays, this could have been due to the enzyme being expressed at low levels. More abundant proteins of a similar molecular weight in the fungal mycelia, such as glyceraldehyde-3-phosphate dehydrogenase, could have dominated the mass spectra and masked the presence of the nitrilase. Glyceraldehyde-3-phosphate dehydrogenase is responsible for the oxidation of the D-

glyceraldehyde-3-phosphate to high energy compounds such as 1,3-bisphosphoglycerate and NADH (Punt *et al*. 1990; Ma *et al*. 2001) and thus would be expected to be expressed constitutively at high levels. However the separation of mycelial proteins according to both pI and molecular weight, as achieved by 2D electrophoresis, enabled subsequent identification of the Nit 1 nitrilase by mass spectrometry from the *T. reesei* mycelial protein from the adiponitrile-supplemented minimal media. Separation of the more abundant proteins from the less abundant Nit 1 protein prevented their domination of the mass spectra allowing the Nit 1 peptides to be detected.

4.5 Chapter summary

To investigate the induction and expression of putative nitrilase gene(s) identified in the genome of the wild-type *T. reesei* by *in silico* analysis (Chapter 3), *T. reesei* Rut-C30 was grown in liquid media containing aliphatic nitriles (adiponitrile and acetonitrile) or aromatic nitriles (benzonitrile and 3 cyanopyridine) as the primary nitrogen source. Fluorometric and colorimetric assays did not reveal nitrilase activity in the mycelia or supernatant of any of the *T. reesei* cultures, probably due to lack of sensitivity. However a more sensitive technique, HPLC, enabled the identification of nitrilase activity in the mycelial protein of *T. reesei* Rut-C30 cultivated in adiponitrile-supplemented minimal medium; the enzyme activity seemed to be specific to the aliphatic nitrile substrate adiponitrile and not effective against the aromatic nitrile substrates. Subsequently, identification of the Nit 1 protein was achieved from a 2D gel of mycelial protein from the adiponitrile-supplemented culture, verifying the expression of the *nit1* gene in *T. reesei* Rut-C30. Establishing the appropriate cultivation conditions for the expression of *nit1* laid the foundation for the isolation of RNA from *T. reesei* for cDNA production (Chapter 5) and subsequent recombinant expression of *nit1* in *E. coli* (Chapter 6).

Chapter 5: Isolation of a nitrilase gene from *Trichoderma reesei* using gene specific primers

5.1 Introduction

Different types of gene products are produced by recombinant means. Examples include proteins of which the yields are low in the original host and proteins that are not naturally secreted outside the cell. The Nit 1 nitrilase of *T. reesei* described in this work displayed both of these characteristics. A list of recombinantly expressed nitrilases of bacterial and fungal origins was presented in Table 1.4, Chapter 1. The only nitrilase of fungal origin that has been recombinantly expressed to date is the nitrilase from *A. niger*, which was expressed in *E. coli* by generation of intron-less nitrilase gene transcripts using approaches such as reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE) (Asano, 2002; Kaplan *et al*., 2011).

In this work, a putative nitrilase gene (*nit1*) was identified in the *T. reesei* genome by *in silico* analysis (Chapter 3) and expression of *nit1* by in the high-secreting hypercellulolytic strain *T. reesei* Rut-C30 was verified using proteomic analysis of the cell lysate (Chapter 4). In the work presented in this chapter, generation of the intron-less transcripts of the *nit1* gene was attempted using RT-PCR and overlap extension polymerase chain reaction (OE-PCR) (Lundberg *et al*., 1991; Orgogozo *et al*., 2011) to allow the recombinant expression of the *T. reesei* nitrilase gene in *E. coli* (Chapter 6).

Reverse transcription polymerase chain reaction (RT-PCR) performed using the reverse transcriptase enzyme is a widely used method for the generation of the intron-less complementary DNA (cDNA) from the RNA of eukaryotic organisms. Currently, several commercial RT-PCR systems manufactured by companies such as Invitrogen (USA) and Agilent (USA) are available today. Another approach to obtain intron-less cDNA gene copies is by overlap-extension PCR (OE-PCR). With the ability to combine multiple DNA fragments

of the size less than 20 kb, OE-PCR has been successfully used for the construction of cDNA by isolation and fusion of only the coding regions (exons) of the gene (Orgogozo *et al*., 2011). A schematic representation of the isolation of exons and their subsequent fusion by OE-PCR is shown in Figure 5.1.

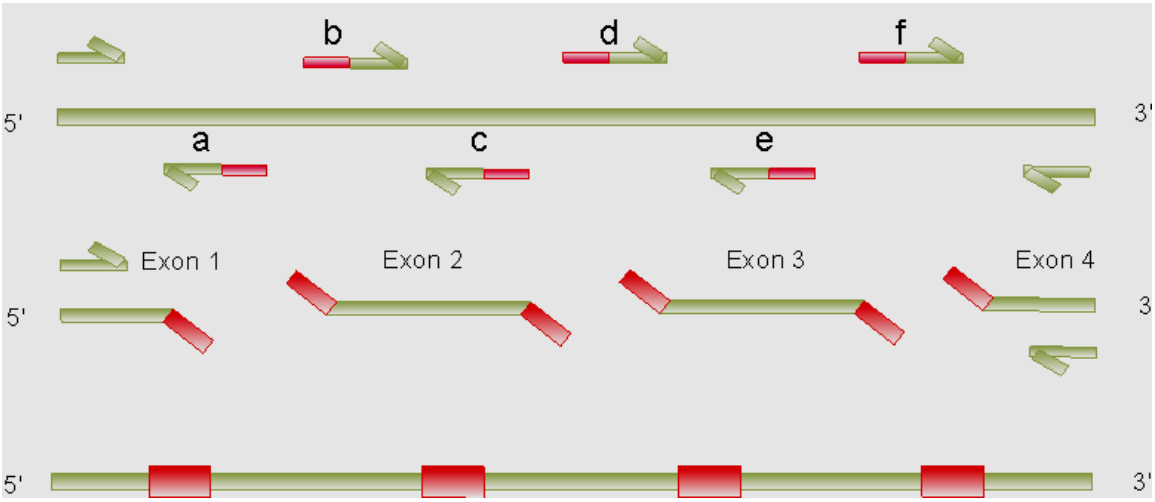


Figure 5.1 Schematic of the construction of intron-less cDNA using OE-PCR. The red extensions attached to the primers a, c and e have nucleotide sequences complementary to the extensions attached to the primers b, d and f respectively. The fusion of the subsequent complementary extensions during a standard PCR reaction generates a full length cDNA containing only the coding regions (exon) of the gene.

Culture conditions appropriate for the expression of *nit1* by *T. reesei* Rut-C30 were established in the work described in Chapter 4. In the work described in this chapter, *T. reesei* Rut-C30 was grown using the same culture medium and RNA was extracted from the mycelia for the isolation and amplification of the intron-less cDNA of *nit1* by RT-PCR. Amplification of *nit1* cDNA from RNA from *T. reesei* Rut-C30 grown in an alternative culture medium was also attempted. Furthermore, construction of a synthetic *nit1* gene was carried out using OE-PCR.

5.2 Materials and methods

5.2.1 Primers for the isolation of the *nit1* gene from *T. reesei*

Gene specific primers were designed to isolate and amplify the full length *T. reesei nit1* gene using the standard PCR reaction (Section 2.2.4). The gene primers along with their restriction sites are shown in Table 5.1.

Table 5.1 Primers used for the amplification of the *T. reesei nit1* gene using standard PCR. *nit1*.Fwd and *nit1*.Rev represent the forward and reverse primers respectively. The restriction sites are in red and the corresponding restriction enzyme is indicated.

Gene	Primer	Sequence (5′ – 3′)	Enzyme
nit1	*nit1*.Fwd	AAAA**CCATGG**TTGCCACTATTCGT	*NcoI*
	nit1.Rev	AAAA**GGATCC**TTACTCCTCAGGTGCAGG	*BamHI*

The forward and the reverse primers were designed with suitable restriction sites on the basis of the *nit1* gene sequences in the JGI *T. reesei* genome database (http://genome.jgi-psf.org/Trire2/).

5.2.2 Isolation and amplification of the full length chromosomal *nit1* gene by PCR

T. reesei Rut-C30 was grown in adiponitrile-supplemented minimal medium as described in Section 4.2.1. Genomic DNA was extracted from the mycelia using GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, Australia) as described in Section 2.3.3. The *nit1* gene (1.3 kb) was amplified from the *T. reesei* genomic DNA using gene-specific primers (Table 5.1). The PCR reactions were carried out using Amplitaq Gold® Polymerase under standard PCR conditions (Section 2.2.4).

5.2.3 Amplification of the intron-less *T. reesei nit1* cDNAs by RT-PCR

Amplification of the complementary DNA transcripts of the *nit1* gene (*nit1* cDNA) was attempted by RT-PCR using total RNA extracted from *T. reesei* Rut-C30 grown in adiponitrile-supplemented minimal medium (Section 4.2.1). In addition, generation of *nit1*

cDNA was attempted using RNA extracted from *T. reesei* Rut-C30 cultures grown in a medium containing cellobiose, lactose and soy hydrolase ("CLS" medium, Lim *et al*., 2001, Table 2.4, Section 2.3.2). The CLS medium did not contain nitriles and was known to promote healthy growth and protein production by *T. reesei*.

5.2.3.1 Extraction of total RNA from *T. reesei* mycelia

Mycelia from liquid cultures of *T. reesei* cultivated in adiponitrile-supplemented minimal medium or CLS medium were collected as described in Section 4.2.1. The mycelia were snap-frozen in liquid nitrogen and ground to a fine powder using a pre-chilled mortar and pestle. Thereafter, total RNA was extracted from the mycelia using Trizol reagent (Invitrogen, USA) as described in Section 2.3.4. The extracted RNA pellet was completely dissolved in DEPC-treated water. The quality of the extracted RNA was determined on a 1% (w/v) agarose gel. The concentration of the RNA was determined by measuring the absorbance at 260 nm using a Biophotometer (Eppendorf, Germany). The RNA samples were stored at -80º C until further use.

5.2.3.2 Reverse-transcription Polymerase Chain Reaction (RT-PCR)

Amplification of the cDNA of the *nit1* gene from *T. reesei* total RNA was attempted by RT-PCR. The first strand synthesis of the *nit1* cDNA was performed using SuperScript® III RT-PCR system (Invitrogen, USA) following the manufacturer's protocol. One µl of 50 µM oligo$(dT)_{20}$ primer and 1 µl of 10 mM dNTP were added to 5 µg total RNA extracted from *T. reesei* mycelia (Section 5.2.3.1). The reaction mixture was then incubated at 65 ºC for 5 min. The mixture was kept on ice for 1 min followed by the addition of 10 µl of cDNA Synthesis Mix containing 2 µl of 10 × RT buffer, 4 µl of 25 mM $MgCl_2$, 2 µl of 0.1 mM DTT, 1 µl of RNaseOUT (40 U/µl) and 1 µl of SuperScript® III RT (200 U/µl). Thereafter, the mixture was gently vortexed and subjected to brief centrifugation at 3000 × *g* for 2 min. The mixture

was incubated at 50 ºC for 1 h and the reaction subsequently terminated by incubation at 85 ºC for 5 min. The mixture was then put on ice for 3 min followed by brief centrifugation at 3000 × *g* for 2 min. To prevent the formation of cDNA-RNA hybrid molecules, 1 µl of RNase H was added to the mixture, followed by incubation at 37 ºC for 20 min. RNase H removes any traces of RNA present in the cDNA mixture. An aliquot of 1 – 4 µl of the first strand cDNA product was used to perform a standard PCR reaction (Section 2.4.4) using gene specific primers (Section 5.2.1). The PCR reaction was carried out in an Eppendorf Master Cycler using PCR Extender System (5 PRIME, USA) (Section 2.2.4). The PCR products were subjected to electrophoresis on a 0.8% (w/v) agarose gel and visualized under UV light.

Molecular grade DNase was used to degrade traces of genomic DNA present in the *T. reesei* total RNA and the treated RNA was used in further RT-PCR experiments. RT-PCR experiments were also carried out using an alternative RT-PCR system from Agilent Technologies (USA) using the manufacturer's protocol performed under similar reaction conditions as mentioned above. Moreover, to assist in the denaturation of the covalent bonds present in "GC" rich *nit* mRNAs, the RT-PCR experiments were also performed using 3 µl of the SuperScript® III Reverse Transcriptase enzyme and at a cDNA synthesis temperature of 60 ºC.

5.2.4 Construction of synthetic *nit1* cDNA by overlap extension PCR

A synthetic *nit1* cDNA was constructed by fusing the coding regions (exons) of the *nit1* gene using the overlap extension PCR approach. The *nit1* gene sequence corresponding to the proposed Nit 1 nitrilase as presented by JGI (http://genome.jgi.doe.gov/) contained four exons and three introns. Amplification of the individual coding regions (exons) from the *nit1* gene was performed using the primers shown in Table 5.2.

Table 5.2 Primers used for amplification of the four coding regions (exons) of the *nit1* gene. The highlighted region of the primers indicates the primer extensions. The "Fwd" and "Rev" denote the forward and reverse primers respectively.

Exon	Primer	Sequence (5′ – 3′)
1	*nit1*.Fwd	AAAACCATGGTTGCCACTATTCGT
	nit1.Rev.4	GTGCCCACCAAAGTCAGCCAGACGCT
2	*nit1*.Fwd.2	CAGAGCTTTCTGACTGTGACGATAGATA
	nit1.Rev.3	GGGTTCATGTTCTCCCAGCAGTTGA
3	*nit1*.Fwd.3	TGGGAGAACATGAACCCCTTCCTCAAG
	nit1.Rev.2	CACAGTCAGAAAGCTCTGGGAATGCAAT
4	*nit1*.Fwd.4	CTGACTTTGGTGGGCACTACATGC
	nit1.Rev	AAAAGGATCCTTACTCCTCAGGTGCAGG

Each of the exon specific reverse primers except *nit1*.Rev were designed to have extensions that were complementary to the next immediate forward primer (Figure 5.1).

Genomic DNA was extracted from lyophilized mycelia of *T. reesei* Rut-C30 grown over cellophane discs placed on PDA plates for four days using GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, Australia). The overlap extension was performed using approximately 200 ng of genomic DNA as template. Amplification of all the four exons of the *nit1* gene was carried out in a total reaction volume of 50 µl using the PCR Extender Polymerase system (5 PRIME, USA). The amplification of the four exons by PCR was attempted using multiple time frames for preheating, denaturation of the template and primer annealing. The optimum PCR conditions for the isolation and amplification of exons 1 and 2 were: 1× (94 °C, 2 min), 28× (96 °C, 10 sec; 52.5 °C, 15 sec; 72 °C, 30 sec), 1× (72 °C, 2 min). Exons 3 and 4 were amplified using the following reaction conditions: 1× (94 °C, 2 min), 28× (96 °C, 10 sec; 55 °C, 15 sec; 72 °C, 30 sec), 1× (72 °C, 2 min).

The PCR products were separated on a 0.8% (w/v) agarose gel and the gel bands were purified using QIAgen gel purification kit (QIAGEN, USA) using the manufacturer's instructions. The concentrations of all four purified exon fragments of the *nit1* gene were determined by measuring the absorbance at 260 nm using a Biophotometer (Eppendorf, Germany). The individual exon fragments of the *nit1* gene were then fused together by overlap extension PCR. The purified exon fragments were mixed in equal concentrations and were used as template for a standard PCR carried out under the following conditions: 1× (94 °C, 2 min), 28× (96 °C, 10 sec; 55 °C, 15 sec; 72 °C, 30 sec), 1× (72 °C, 2 min).

5.2.5 Sequencing of DNA

Sequencing of the *T. reesei nit1* gene isolated from the genomic DNA of *T. reesei* Rut-C30 was carried out using the ABI prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, USA). The reaction mixture consists of 350 ng DNA, 3.2 pmol of primer, 2 µl of Big Dye and 4 µl buffer. The DNA was precipitated as per the manufacturer's protocol and was sequenced by Macquarie University Sequencing Facility using an Applied Biosystems 3130xl Genetic Analyzer.

5.2.6 Bio-informatic analysis

All the multiple alignment analyses involving the predicted Nit 1 protein sequence and other known nitrilases was carried out with ClustalW software (Thompson *et al*., 1994; www.ebi.ac.uk/ Tools/ services/ web_clustalw). Translation of the *nit1* cDNA sequences into their corresponding protein sequences were performed using ExPASy software (http://web.expasy.org/translate/).

5.3 Results

5.3.1 PCR amplification of the *T. reesei nit1* gene using gene specific primers

The expected size of the *nit1* gene was 1.3 kDa based on the wild-type *T. reesei* genome sequence in the JGI database. To experimentally determine the size and the nucleotide sequence of *nit1* from the mycelia of *T. reesei* Rut-C30, isolation and amplification of the gene was attempted from the *T. reesei* genomic DNA by standard PCR (Section 2.2.4) using gene specific primers (Table 5.1). The amplified PCR product is shown in Figure 5.2.

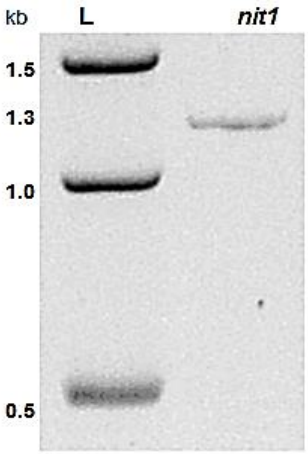


Figure 5.2 Agarose gel electrophoresis of the *nit1* gene isolated from genomic DNA of *T. reesei* Rut-C30 using gene specific primers in a standard PCR reaction (Section 2.2.4). Lane L: DNA Ladder (Fermentas, USA). 8 µl of PCR product was loaded onto the gel. The expected size of *nit1*was 1.3 kDa (http://genome.jgi-psf.org/Trire2).

A DNA band of approximately 1.3 kb was observed from the PCR product when subjected to agarose gel electrophoresos (Figure 5.2) which was in line with the expected size of the *nit1* gene (JGI database). The amplified *nit1* gene from *T. reesei* Rut-C30 was sequenced in both directions as described in Section 2.2.3 and had 100% sequence similarity to the *nit1* gene in of the wild-type *T. reesei* in the JGI database.

5.3.2 Amplification of *nit1* cDNA by RT-PCR

Amplification of *nit1* cDNA by RT-PCR was attempted from the total RNA of adiponitrile-supplemented cultures of *T. reesei* Rut-C30 as described in Section 5.2.3.2. Based on the JGI

database, the *nit1* gene was expected to contain three introns and the expected size of the intron-less transcript of the *nit1* gene was 1.08 kb. The gene specific primers used to amplify the DNA of the *nit1* gene (Section 5.2.1) were used to isolate the intron-less cDNA from the *T. reesei* RNA by RT-PCR. As a negative control, the RT-PCR experiment was carried out with all components except the reverse transcriptase enzyme thus inhibiting reverse transcription of the RNA template. An RT-PCR product was obtained (Figure 5.3); however, the RT-PCR product was of the same size (1.3 kb) as the *nit1* gene isolated from genomic DNA by standard PCR (Figure 5.2, Section 5.3.1).

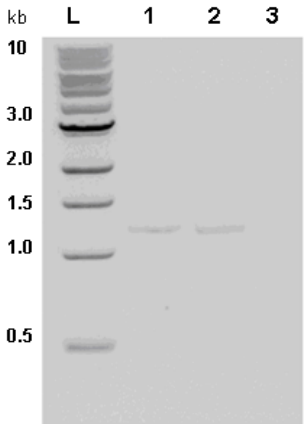


Figure 5.3 Agarose gel electrophoresis of the RT-PCR product obtained using *nit1* gene specific primers and total RNA extracted from *T. reesei* Rut-C30 mycelia grown in adiponitrile-supplemented medium (lane 1). Expected size of the intron-less *nit1* cDNA was approximately 1.08 kb but a band of 1.3 kb was obtained. The full length *nit1* gene (containing introns) obtained by standard PCR using the same primers and genomic DNA from the same *T. reesei* mycelia is shown in lane 2. Expected size of the *nit1* gene was approximately 1.3 kb as shown. A negative control containing the product of the RT-PCR reaction performed without the reverse transcriptase enzyme is shown in lane 3. Lane L: DNA ladder (Fermentas, USA). 3 µl of the PCR products were loaded per lane.

Based on the nucleotide sequence of *nit1* deposited in the JGI database, a size difference of approximately 230 bp was expected between the cDNA (1.080 kb) and the genomic DNA (1.310 kb). Sequencing of the RT-PCR product revealed a nucleotide sequence that had retained all the expected introns, suggesting that reverse transcription of the RNA had not occurred. This could have been due to contamination of the total RNA with genomic DNA.

Hence, to remove any trace of genomic DNA contamination in the RNA template, the total RNA was treated with molecular grade DNase (Invitrogen, USA) and the entire RT-PCR procedure was repeated (Figure 5.4).

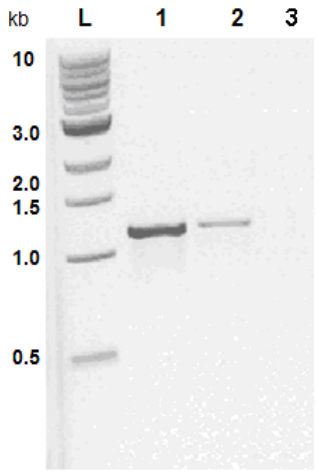


Figure 5.4 Agarose gel electrophoresis of the RT-PCR product obtained using DNase treated *T. reesei* RNA as template (lane 1). The full length *nit1* gene (containing introns) obtained by standard PCR is shown in lane 2. A negative control containing the product of the RT-PCR reaction performed without the reverse transcriptase enzyme is shown in lane 3. Lane L: DNA ladder (Fermentas, USA). 5 µl of the PCR products were loaded per lane.

Again, very little size difference was observed between the RT-PCR product obtained using DNase treated RNA as template and the full length *nit1* DNA containing introns (Figure 5.4) and sequencing again revealed that all introns had been retained. However, in all the RT-PCR experiments no bands were observed in the lane containing the negative control (product of the RT-PCR reaction performed with the same RNA sample and all reaction components except the reverse transcriptase enzyme). If contamination of the template RNA with genomic DNA had occured, it would be expected that the negative control would also contain amplified DNA. In order to overcome any technical errors associated with the protocol or the SuperScript® III RT-PCR system itself, the entire RT-PCR experiment was repeated several times with a fresh batch of RNA and by using Agilent Reverse Transcriptase kit (Agilent Technologies, USA). However, no differences were observed in the size or the nucleotide sequences of the RT-PCR products and the full length *nit1* DNA containing introns.

5.3.3 Synthetic *nit1* cDNA by overlap extension PCR

Due to the inability to obtain a full length cDNA transcript of the *T. reesei nit1* gene using the RT-PCR approach, construction of synthetic *nit1* cDNA was attempted using overlap extension PCR. Fragments of DNA that were predicted by the JGI database to represent the exons of the *T. reesei nit1* gene were amplified using PCR with modified reaction conditions (Section 5.2.4) using exon specific primers (Table 5.2). The amplified exon fragments were then run on a 0.8% (w/v) agarose gel and viewed under UV light (Figure 5.5).

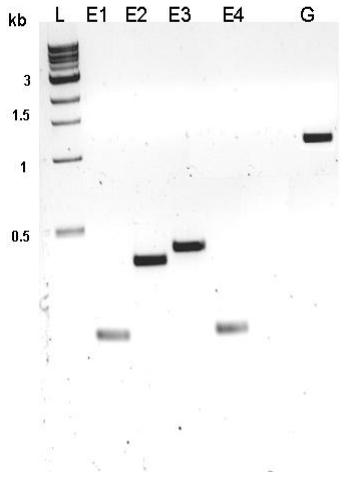


Figure 5.5 Agarose gel electrophoresis of the predicted *nit1* exons isolated from *T. reesei* genomic DNA by PCR using exon specific primers. Lane L: DNA ladder; Lanes E1, E2, E3 and E4: *nit1* exons. The expected sizes of the exons are shown in Table 5.4. Lane G: full length *nit1* gene. 6 µl of the PCR products were loaded on the agarose gel.

The expected sizes of the individual exon fragments according to the JGI *T. reesei* genome database are shown in Table 5.3.

Table 5.3 Expected sizes of the four *nit1* exons (http://genome.jgi-psf.org/Trire2/).

Exon	Size (bp)
1	148
2	362
3	415
4	155

The sizes of the amplified exons generated by PCR (Figure 5.5) were consistent with their expected sizes based on the *T. reesei* genome sequence as annotated by JGI (Table 5.3). The exon DNA fragments were gel-purified (Section 2.2.6) and the purified fragments were fused together by standard PCR carried out under the conditions described in Section 5.2.4. The fusion of the exon fragments was attempted using equal concentrations of the individual fragments (~ 5 ng/ exon) at different annealing temperatures (50 ºC – 55 ºC) in order to identify the optimum annealing temperature. Significant non-specific binding was observed at annealing temperatures ranging from 50 ºC to 54 ºC. For example, two prominent gel bands with approximate sizes of 800 bp and 500 bp were generated when the PCR was performed at an annealing temperature of 53 ºC (data not shown). A full length synthetic *nit1* cDNA with no non-specific binding was obtained of the expected size (1080 bp) at an annealing temperature of 55 ºC (Figure 5.6).

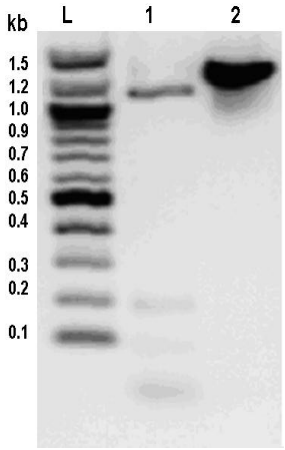


Figure 5.6 Agarose gel electrophpresis of the synthetic *nit1* cDNA obtained using overlap extension PCR. Lane L: DNA ladder; Lane 1: synthetic *nit1* cDNA (expected size 1.08 kb); Lane 2: *nit1* gene containing introns (size 1.3 kb). 5 µl of the PCR products were loaded on the agarose gel.

Sequencing of the synthetic *nit1* cDNA revealed that the nucleotide sequence of the synthetic *nit1* cDNA was exactly the same as the proposed *nit1* cDNA sequence based on the annotated *T. reesei* gene in the JGI database.

5.3.4 Amplification of *nit1* cDNA from RNA of *T. reesei* grown in a nitrile-free nutrient-rich medium (CLS medium)

In addition to the construction of synthetic *nit1* cDNA, another attempt was made to amplify *nit1* cDNA from *T. reesei* RNA. However, instead of using RNA from *T. reesei* grown in nitrile-supplemented culture media, RNA from *T. reesei* grown in cellobiose, lactose and soy hydrolase (CLS) medium (Table 2.4, Section 2.3.2) was used as a template for RT-PCR (Section 5.2.3.2). CLS is a rich medium known to support production of a wide range of proteins in *T. reesei* (Lim *et al*., 2001). The RT-PCR was performed using SuperScript® III RT-PCR system (Invitrogen, USA) as described in Section 5.2.3.2. The PCR products were then run on a 0.8% (w/v) agarose gel (Figure 5.7).

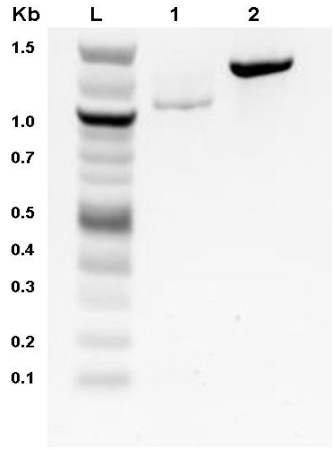


Figure 5.7 Agarose gel electrophoresis of the RT-PCR product obtained using total RNA extracted from *T. reesei* Rut-C30 mycelia grown in CLS medium (lane 1). A band of the approximate expected size of the intron-less *nit1* cDNA (1.08 kb) is shown. The full length *nit1* gene (containing introns) is shown in lane 2 (1.3 kb). Lane L: DNA ladder (Fermentas, USA). 5 µl of the PCR products were loaded per lane.

A size difference of approximately 220 – 230 bp was evident between the RT-PCR product obtained using RNA obtained from the CLS culture of *T. reesei* and the full length *nit1* DNA which contained introns. In contrast to previous RT-PCR experiments carried out using RNA from *T. reesei* cultivated in nitrile supplemented media, this time the size of the RT-PCR product was similar to the expected size of the *nit1* cDNA. To verify that the nucleotide sequence was consistent with that of the expected *nit1* cDNA based on the JGI database, the

RT-PCR product was gel purified and sequenced using the ABI prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, USA) as described in Section 5.2.5. The analysis of the nucleotide sequence is described below (Section 5.3.5).

5.3.5 Analysis of the *nit1* cDNA obtained from *T. reesei* grown in CLS medium

The predicted *nit1* cDNA sequence in the JGI database (http://genome.jgi-psf.org/Trire2/) has a total length of 1080 nucleotides and is derived from four predicted exons in the *nit1* gene. The nucleotide sequence of the *nit1* cDNA obtained in this research from RT-PCR of RNA extracted from *T. reesei* Rut-C30 grown in CLS medium (Section 5.3.4) was not identical to the *nit1* cDNA sequence proposed by JGI; sequence similarity was ~92% and the length of the *nit1* cDNA sequence was 10 bp longer than the predicted *nit1* cDNA sequence in the JGI database.

Comparison of the nucleotide sequences revealed that the additional 10 bp in the *nit1* cDNA corresponded to nucleotides predicted by JGI to form part of an intron (intron 1) in the *nit1* gene (Figure 5.8). Based on the results obtained in this research, these nucleotides are transcribed and thus represent an additional exon in the *nit1* gene not previously described. Hence, the 148 nucleotides originally predicted by JGI to form intron 1 in the *nit1* gene actually appeared to represent two introns separated by a 10 bp exon (Figure 5.8). The inclusion of the additional exon thus increased the size of the experimentally obtained *nit1* cDNA to 1.09 kb, 10 bp above that predicted by JGI.

(a)

```
1.........  .........20 ..........  .........40 ..........
GTAAGTTTCA  AAGGCGGAAT  TGGAAGCCGA  GGGAACATGG  CTGATGATTT

.........60 ..........  .........80 ..........  .........100
GTTCTAGGGA  TTCCTGGGTA  GGAGACGCAG  AAACTGACGG  TCATGGCAAT

..........  ...
GACTGATTCG  GAG
```

(b)

```
1.........  .........20 ..........  .........40 ..........
GTAAGTTTCA  AAGGCGGAAT  TGGAAGCCGA  GGGAACATGG  CTGATGATTT

.........60 ..........  .........80 ..........  .........100
GTTCTAGGGA  TTCCTGGGTA  GGAGACGCAG  AAACTGACGG  TCATGGCAAT

..........  ...
GACTGATTCG  GAG
```

Figure 5.8 (a) The nucleotide sequence of intron 1 of the *nit1* gene as predicted by JGI. (b) The location of a 10 bp exon (highlighted in light blue) within the predicted intron 1 sequence that was discovered in this work by sequencing of *nit1* cDNA generated from *T. reesei* Rut-C30 RNA. The newly exposed exon defines two separate introns within the predicted intron 1. The first intron (57 bp) is shown in red font and the second intron (46 bp) is in dark blue font.

From the nucleotide sequence analysis described above, it appeared that the *T. reesei nit1* gene consisted of five exons and four introns as opposed to four exons and three introns as predicted in the JGI database (Table 5.4).

Table 5.4 Number and sizes of the introns and exons in the *nit1* gene as predicted by JGI, and that were evident from *nit1* cDNA isolated in this work. The additional 10 bp exon identified from the *nit1* cDNA in this work is in red font (exon 2). Intron 1 of the predicted *nit1* gene (113 bp) corresponds to two introns of size 57 bp and 46 bp respectively in the *nit1* gene according to this work (blue font).

***nit1* gene**	**Number of exons**	**Number of introns**	**Size of exons (bp)**					**Size of introns (bp)**			
			1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**
Predicted (JGI)	4	3	148	362	415	155	-	**113**	54	63	-
This work	5	4	148	**10**	362	415	155	**57**	**46**	54	63

The addition of the ten newly discovered exon nucleotides (5′ GGATTCCTGG 3′) to the translated DNA sequence was expected to result in a frame shift that could also change the corresponding amino acid sequence in the predicted Nit 1 protein. To determine the effect of the additional 10 bp exon in the *nit1* gene, a complete amino acid sequence of the Nit 1 protein was predicted by the translation of the experimentally derived *nit1* cDNA nucleotide sequence using ExPASy protein translation software (Section 5.2.6). The predicted amino acid sequence of Nit 1 translated from the *nit1* cDNA sequence differed from the predicted Nit 1 amino acid sequence in the JGI database primarily by the addition of amino acids encoded by the nucleotides present in the newly discovered exon. The additional amino acids present in the predicted Nit 1 from the experimentally derived *nit1* cDNA were tryptophan (W), isoleucine (I), proline (P) and glycine (G) respectively. Analysis of how the amino acid additions affected the sequence similarity of the Nit1 protein with known nitrilases was carried out using ClustalW.

5.3.6 *In silico* analysis of the predicted Nit 1 protein based on experimentally-obtained *nit1* cDNA

Multiple alignment analyses of the amino acid sequences of the Nit 1 protein as predicted by JGI (from the *nit1* gene predicted to contain 3 introns and 4 exons) and the Nit 1 protein as predicted from the *nit1* cDNA isolated in this work (implicating a *nit1* gene containing 4 introns and 5 exons) with other known nitrilases was performed using ClustalW software (Section 5.2.6). The two predicted Nit 1 sequences differed primarily in the region shown in Figure 5.9.

(a)

```
Nit1-T. reesei.JGI      AFPELSDCDDRYTYW
Nit- A. niger           AFPELWIPGYPYWMW
Nit- F. solani          AFPEVWIPGYPYWMW
Nit- R. rhodochrous     AFPEVWIPGYPYWAW
Nit- Nocardia sp.       AFPEVWIPGYPYWAW
Nit- B. japonicum       AFPEVWIPGYPWWLW
```

(b)

```
Nit1-T. reesei.iden     AFPELWIPGYTYWGW
Nit- A. niger           AFPELWIPGYPYWMW
Nit- F. solani          AFPEVWIPGYPYWMW
Nit- R. rhodochrous     AFPEVWIPGYPYWAW
Nit- Nocardia sp.       AFPEVWIPGYPYWAW
Nit- B. japonicum       AFPEVWIPGYPWWLW
```

Figure 5.9 A selected area from the multiple alignment analysis by ClustalW of the amino acid sequences of known nitrilases and (a) the predicted Nit 1 protein sequence from the JGI database (Nit1-*T. reesei*.JGI); and (b) the sequence of the Nit 1 protein translated from the *nit1* cDNA obtained in this work (Nit1- *T. reesei*.iden). The red font indicates a region containing the amino acids W, I, P and G identified in most of the known nitrilases. The green font exposes the presence of the "WIPG" region in the Nit 1 protein translated from the *nit1* cDNA sequence (b) and the absence of "WIPG" in the predicted Nit 1 protein in the JGI database (a). Nit- *A. niger, F. solani, R. rhodochrous, Nocardia* sp. and *B. japonicum* indicates nitrilases from *Aspergillus niger, Fusarium solani, Rodococcus rhodochrous, Nocardia* sp. and *Bradyrhizobium japonicum* respectively.

The multiple alignment analyses (Figure 5.9) revealed that the amino acid sequence "W, I, P, G", which was encoded by the DNA from the additional exon in the *nit1* gene discovered in this work (Section 5.3.5) was present in all the known nitrilases in the analysis, including nitrilases of fungal origin such as those from *A. niger* and *F. solani*. Hence, it appeared that the predicted intron/ exon assignment of the *nit1* gene sequence in the JGI database may have been incorrect, or the transcription of the DNA segment was the result of an alternatively spliced transcript as discussed in Section 5.4. The translation of the additional exon slightly increased the sequence similarity of the predicted Nit 1 protein with the other known nitrilases. For example, the sequence similarity between the predicted *T. reesei* Nit 1 and the

closest homolog, a nitrilase from *A. niger* (Table 3.1, Chapter 3), increased from approximately 60% to approximately 62%.

The synthetic *nit1* cDNA (Section 5.3.3) was constructed based on the predicted *nit1* cDNA sequence in the JGI database, which was based on a *nit1* gene consisting of four exons instead of five, and thus lacking the additional 10 bp that were shown in the work just described to encode conserved amino acid sequences in nitrilase genes. Therefore, the synthetic *nit1* cDNA was not used for any further experiments. All the experiments performed from this point forward were carried out using the *nit1* cDNA generated from the RNA of *T. reesei* Rut-C30 grown in CLS medium.

5.4 Discussion

In silico identification of the putative *nit1* nitrilase gene and the design of all the primers used in this work were based on the genome sequence of *T. reesei* wild-type QM6a in the JGI database. In the work described in this chapter, the *nit1* gene was amplified from the genomic DNA of *T. reesi* Rut-C30 by standard PCR and both the size and the nucleotide sequence of the isolated *nit1* gene were in line with the proposed sequence of the *nit1* gene in the JGI database. Amplification of the *nit1* cDNA was attempted using different PCR based approaches, firstly to allow the identification of the introns in the *nit1* gene and secondly, to allow the recombinant expression of the *nit1* cDNA in a bacterial host. The isolation of *nit1* cDNA was ultimately achieved by RT-PCR using total RNA extracted from *T. reesei* Rut-C30 grown in a nutrient rich culture medium (CLS medium) that did not contain nitriles. The difficulties encountered in isolating *nit1* cDNA from nitrile-supplemented cultures of *T. reesei* are discussed in Section 5.4.1. Features of the *nit1* cDNA isolated from *T. reesei* grown in the CLS medium are discussed in Section 5.4.2.

5.4.1 RNA from *T. reesei* grown in adiponitrile-supplemented minimal medium as template for RT-PCR

The isolation and amplification of the *nit1* cDNA was first attempted by RT-PCR using total RNA extracted from *T. reesei* cultures induced with adiponitrile. Adiponitrile was chosen for induction of the *nit1* gene based on the successful identification of the expressed Nit 1 protein in *T. reesei* cultures supplemented with adiponitrile (Chapter 4). The RT-PCR approach was also used by Kaplan *et al*. (2011) in conjunction with the Rapid Amplification of cDNA Ends (RACE) system for isolation and amplification of the cDNA of the gene encoding a nitrilase from *A. niger.*

The expected size of the *nit1* cDNA was around 1080 bp which is approximately 230 bp smaller than the full length *nit1* gene (1310 bp). Although, the RT-PCR using RNA isolated from the mycelia of *T. reesei* grown in adiponitrile-supplemented media yielded a product, no size difference was observed between the RT-PCR product and the *nit1* genomic DNA. Moreover, the nucleotide sequence of the RT-PCR product contained all the three predicted non-coding regions (introns) of the *nit1* gene, as defined by JGI. The most probable explanations for the RT-PCR products retaining all the introns were, firstly that the RNA template used for the reaction contained traces of genomic DNA which was amplified by the standard PCR reaction, secondly, RNA with a higher number of guanine (G) and cytosine (C) bases inhibited the activity of the reverse transcriptase due to their highly stable secondary structures, and thirdly, possible handling and technical errors including contaminated reverse transcriptase enzyme and buffers.

To ensure there was no contamination with genomic DNA, the *T. reesei* total RNA was treated with molecular grade DNase and the RT-PCR experiment was repeated using the treated RNA as template. However, no difference in the sizes of the full length *nit1* genomic DNA and the RT-PCR product was observed. Moreover, no PCR product was observed for the negative control in which the RT-PCR reaction was performed without the reverse

transcriptase enzyme. In case of genomic DNA contamination of the RNA template, it would be expected that the *nit1* genomic DNA would have been amplified during the standard PCR reaction performed using the RT-PCR products as template. Therefore, it appeared that the absence of reverse transcription of the *T. reesei* RNA was not due to genomic DNA contamination.

The predicted intron-less *nit1* gene transcript (1.08 kb) contained 277 guanine bases (25.6%) and 316 cytosine (29%) nucleotide bases with a total "GC" content of 55%. Hence, to assist in the denaturing of the covalent bonds formed due to the presence of considerably high "GC" content in the proposed *nit1* mRNA, the RT-PCR was repeated using a higher amount of reverse transcriptase (3 µl) and the cDNA synthesis temperature was raised to 60 ºC. Yet, the RT-PCR products appeared to retain all of their respective introns. With neither the genomic DNA contamination nor the secondary structures identified to prevent the reverse transcription process, it appeared that a possible reason for the lack of appropriate RT-PCR products was either faults in the RT-PCR reagents or physical handling errors.

The SuperScript® III Reverse Transcriptase enzyme was initially chosen for all the RT-PCR experiments due to the reported stability of the enzyme, including high thermo-stability for RNA with high "GC" content. Hence the RT-PCR analyses were performed at higher temperatures with modified reaction conditions. Furthermore, to overcome possible faults with the SuperScript® III First- Strand Synthesis System including the protocol and contamination of the reverse transcriptase and buffers, RT-PCR was attempted using an RT-PCR kit from Agilent Technologies (USA). Yet, no intron-less *nit1* gene transcript was obtained.

The identification of the Nit 1 enzyme in the mycelia of *T. reesei* Rut-C30 in the adiponitrile-supplemented culture medium as detailed in Chapter 4 provided evidence of successful transcription, translation and expression of the *nit1* gene under these defined culture

conditions. Therefore, it appeared that the inability to obtain the intron-less *nit1* cDNA was due to some unexplained technical difficulty. The inability to obtain the *nit1* cDNA using the conventional RT-PCR approach led to the construction of a synthetic *nit1* cDNA using overlap extension PCR. Hence, a full length synthetic *nit1* cDNA was successfully constructed based on the nucleotide sequence of the proposed *nit1* cDNA (JGI database). Concurrently, further attempts were made to isolate *nit1* cDNA from *T. reesei* RNA, this time using RNA extracted from *T. reesei* grown in a more nutrient rich culture medium that did not contain nitriles, as discussed below.

5.4.2 Generation of *nit1* cDNA from RNA extracted from *T. reesei* grown in CLS medium

The growth of *T. reesei* Rut-C30 in the adiponitrile-supplemented minimal medium was slow and the fungus appeared to be under stress in the presence of the nitrile; mycelia were thin and sparse and total protein production was low (Section 4.3.1). In contrast, *T. reesei* Rut-C30 was known to exhibit good growth and protein production when grown in a frequently used culture medium in the laboratory containing cellobiose, lactose and soy hydrolase (CLS medium). *T. reesei* cDNA had successfully been obtained from RNA extracted from *T. reesei* grown in CLS medium in the past so it was considered possible that the *nit1* cDNA could also be isolated from such cultures despite the fact that the culture medium lacked nitriles as a potential inducer. Exploration of this possibility was rewarded by successful isolation of the *nit1* cDNA.

The isolation of the *nit1* cDNA from *T. reesei* grown in a medium that had not been supplemented with nitriles was unexpected. Most known nitrilases are inducible enzymes (Gong *et al*., 2012) and the CLS medium did not contain any known nitrilase inducers. However, genes encoding nitrilases from sources such as *B. subtilis* ZJB-063 and *Alcaligenes* sp. ECU0401 have been reported to be expressed to varying degrees under both induced and

non-induced conditions (Zheng *et al*., 2008; He *et al*., 2010). The *T. reesei nit1* gene appeared to display at least a basal level of constitutive gene expression. Possibly due to the overall better health and protein production of *T. reesei* when grown in the CLS medium, constitutive expression of the *nit1* gene was sufficient for isolation of the *nit1* cDNA. Although *nit1* was also expressed in the presence of adiponitrile (Chapter 4), the resulting retarded growth of *T. reesei* in the presence of the nitrile (when no other nutrient source was available) may have affected the quantity and quality of the total RNA that could be extracted from it, and hence the subsequent outcome of the RT-PCR to obtain *nit1* cDNA. A closer investigation of the complexities involved in the regulation and expression of *nit1* in *T. reesei* was beyond the scope of this work, which was concerned primarily with the recombinant expression of the *nit1* gene.

An intron-less *nit1* cDNA transcript was successfully isolated and amplified by RT-PCR from the RNA extracted from *T. reesei* grown in CLS medium using the SuperScript® III First-Strand Synthesis System. The isolated *nit1* cDNA sequence contained an additional 10 nucleotides that were absent in the proposed *nit1* cDNA sequence in the JGI database. This appeared to be due to the presence of an additional previously undescribed exon within the first predicted intron in the *nit1* gene as proposed by JGI. Moreover, *in silico* translation of the nucleotide sequence of the *nit1* cDNA isolated in this work (containing the additional 10 nucleotides) resulted in a predicted Nit 1 protein sequence that contained an additional tryptophan (W), isoleucine (I), proline (P) and glycine (G) due to the newly discovered exon. Multiple alignment analyses performed using ClustalW revealed that the "WIPG" sequence was conserved amongst many known nitrilases including fungal nitrilases from *A. niger* and *F. solani*.

The assignment of intron and exon boundaries in the *nit1* gene by JGI did not appear to be entirely correct. However, it was also possible that the predicted and observed *nit1* cDNA sequence could represent alternatively spliced transcripts. Alternative splicing of *T. reesei*

genes has been reported previously *e.g.* a *hex1* transcript resulted from alternative splicing of the intron as reported by Curach *et al*. (2004). However, in this work, no other mRNA/ cDNA transcripts indicating alternate splicing were obtained.

5.5 Chapter summary

The work described in this chapter involved attempts to isolate *nit1* cDNA from *T. reesei* Rut-C30 to enable recombinant expression of the nitrilase in *E. coli* (Chapter 6) for further characterization of the enzyme (Chapter 7). Although evidence of the expression of the Nit 1 nitrilase in the mycelia of *T. reesei* grown in adiponitrile-supplemented culture medium was obtained using proteomic techniques (Chapter 4), generation of *nit1* cDNA from RNA extracted from the same mycelia was not achieved by RT-PCR. However, *nit1* cDNA was successfully obtained from RNA extracted from the mycelia of *T. reesei* grown in a nutrient rich culture medium that did not contain nitriles. This suggested a basal level of constitutive expression of the *nit1* gene in *T. reesei*. Sequencing of the *nit1* cDNA obtained in this work revealed the presence of an additional exon in the *nit1* gene which had not been predicted in the gene sequence by JGI. The additional exon encoded amino acids that were conserved in known nitrilases, thus slightly increasing the sequence similarity of the predicted Nit 1 protein with other known fungal nitrilases. Establishment of the correct sequence of the *nit1* cDNA was thus achieved by this work, a critical step prior to expression of the *T. reesei* nitrilase in *E. coli*.

Chapter 6: Cloning and expression of the *nit1* gene from *T. reesei* in *E. coli*

6.1 Introduction

A variety of fungal and bacterial hosts have been utilized for the recombinant expression of proteins. *Escherichia coli* is highly popular for initial expression studies, gene cloning and plasmid propagation due to well-studied genetics, cost effective growth and the availability of well-developed molecular tools (Yin *et al*. 2007; Ferrer-Miralles *et al*., 2009). In the work described in this chapter, the *nit1* gene from *T. reesei* was expressed in *E. coli* to enable subsequent characterization of the recombinant nitrilase (Chapter 7), an important procedure before recombinant expression of the nitrilase in a more complicated fungal system might be considered in the future.

The expression of eukaryotic proteins in *E. coli* can sometimes pose quite a challenge due to the comparative simplicity of protein production mechanisms in the prokaryotic host. For example, proteins requiring post-translational modifications such as glycosylation may not be successfully produced in a functional form in *E. coli*. Hence, one of the first steps prior to expression of the *nit1* nitrilase in *E. coli* in this project was bioinformatic analysis of the predicted Nit 1 protein to establish whether any probable glycosylation sites were present. Eukaryotic proteins are also known to form protein aggregates (inclusion bodies) due to improper folding when expressed in *E. coli*. Therefore, the expression of *nit1* was carried out using two types of *E. coli* competent cells, Origami^{TM}B(DE3) and Tuner^{TM}DE3, that have been specifically modified to increase the production, stability and activity of recombinant proteins. Origami ^{TM}B(DE3) cells are reported to enhance disulphide bridge formation and correct protein folding due to mutations in thioredoxin reductase and glutathione reductase genes that alter the reducing conditions in the *E. coli* cytoplasm (Prinz *et al*., 1997; Stewart *et al*., 1998). Tuner^{TM}DE3 cells are also known to increase the solubility of recombinant

proteins and reduce protein aggregation. In addition, a lacZY mutation in the Tuner^{TM}DE3 cells ensures uniform entry of the inducer IPTG into the cells (Maurer, 2012).

In the work described below, the *nit1* cDNA obtained from RNA extracted from *T. reesei* grown in CLS medium (Chapter 5) was used to enable the expression of the *T. reesei* Nit 1 nitrilase in Origami^{TM}B(DE3) and Tuner^{TM}DE3 cells. The soluble and the insoluble fractions of the recombinant *E. coli* cells were analyzed for the presence of the Nit 1 enzyme using SDS-PAGE, mass spectrometry and nitrilase activity assays.

6.2 Materials and methods

6.2.1 Bacterial strains and plasmids

E. coli strains Origami^{TM}B(DE3) and Tuner^{TM}DE3 purchased from Novagen (USA) were used for *nit1* expression. *E. coli* DH5α (New England Biolab, USA), a low endonuclease strain was used for plasmid propagation. The *E. coli* gene expression vector pETDuet-1 (Novagen, USA) was used for all the transformation experiments. The pETDuet-1 plasmid (5420 bp) contains a "6 × His-tag" encoding region at the 5′ end of the multiple cloning site-1 (MCS1) which enables purification of the recombinant protein (Chapter 7).

6.2.2 Construction of the pET-*nit1*cDNA vector

The *nit1* cDNA obtained using total RNA extracted from *T. reesei* cultures grown in cellobiose, lactose and soy hydrolase (CLS) medium (Section 5.3.4), was digested with 5 U of *NcoI* enzyme (Fermentas, USA). The reaction mixture was incubated at 37 ºC for 1 h and the digested products were purified using QIAgen PCR purification kit (Section 2.2.6). The purified products were further digested with 5 U of *BamHI* enzyme (Fermentas, USA) for 1 h at 37º C. The digested *nit1* cDNA was subjected to electrophoresis on a 0.8% (w/v) agarose gel and purified using the QIAgen gel purification kit (Section 2.2.6). The recombinant pET-

nit1 vector was generated by inserting the purified *nit1* cDNA fragment into the pETDuet-1 vector using T4 DNA ligase (Fermentas, USA) under conditions described in Section 2.2.2.

6.2.3 Amplification of the pET-*nit1* plasmid

The pET-*nit1* plasmid was transformed into *E. coli* DH5α cells and the cells were grown in selective medium containing ampicillin at a final concentration of 100 µg/ml (Section 2.4.3). The presence of the *nit1* cDNA fragment in the vector was determined by colony PCR (Section 2.2.5) using gene specific primers (Table 5.1, Chapter 5). The expected size of the *nit1* cDNA was approximately 1.09 kb (Section 5.3.5). The pET-*nit1* plasmids were isolated from the transformed DH5α cells using the QIA Prep Spin Mini Prep kit (QIAGEN, USA) following manufacturer's directions.

6.2.4 Transformation and selection of transformants

The pET-*nit1* vector was transformed into *E. coli* Origami^{TM}B(DE3) and Tuner^{TM}DE3 cells. Ampicillin at a concentration of 100 µg/ml in Luria Bertani agar plates was used to select the transformants. Single transformant colonies were selected and grown overnight in selective medium at 37 ºC (Section 2.4.3). The presence of the *nit1* cDNA fragment in the multiple cloning site of the pETDuet-1 vector was determined by analyzing the recombinants by colony PCR (Section 2.2.5) using gene specific primers (Table 5.1, Chapter 5). The expected size of the *nit1* cDNA was approximately 1.09 kb (Section 5.3.5). Cells were also transformed with vectors without *nit1* cDNA to be used as a negative control in future experiments.

6.2.5 Culture conditions for recombinant expression of *nit1*

The transformed Origami^{TM}B(DE3) and Tuner^{TM}DE3 cells were grown overnight in Terrific broth (Table 2.5, Chapter 2) containing ampicillin at a concentration of 100 µg/ml. The cultures were maintained at 37 ºC with constant shaking at 250 rpm. The overnight cultures

were diluted 1:20 into 200 ml of Terrific broth. Two 200 ml cultures were prepared from each overnight culture, one for continued growth at 22 ºC and the other for continued growth at 37 ºC, with shaking at 250 rpm. Once the cultures had reached an optical density at 600 nm of 0.5, they were induced with 0.1 mM isopropylthiogalactoside (IPTG) and the growth temperatures for the cultures were maintained. After 5 h of induction, the bacterial cells were collected and harvested by centrifugation at 5000 × *g* for 15 min. The pellet was stored at -20 ºC until further use. Bacterial cell pellets were also obtained from parallel cultures that were not induced with IPTG but maintained at 37 ºC for 3 h after the optical density of the cultures at 600 nm was 0.5. In addition, cultures of Origami™B(DE3) recombinant cells were prepared using different concentrations of IPTG (0.1 mM – 0.6 mM) for induction.

6.2.6 Preparation of protein extracts

6.2.6.1 Extraction of soluble protein

The bacterial cell pellets from transformed Origami™B(DE3) and Tuner™DE3 cells were lysed using B-PER (Bacterial Protein Extraction Reagent) reagent (Pierce, USA). The soluble protein fractions were extracted from the cell lysates using manufacturer's directions. One-dimensional (1D) SDS-PAGE analyses were performed using NuPAGE® 4 – 12% (w/v) Bis-Tris gels (Invitrogen, USA) (Section 2.6.2) with 15 µl of protein samples extracted using B-PER reagent.

6.2.6.2 Extraction of protein from inclusion bodies (insoluble protein)

Proteins that form insoluble aggregates (inclusion bodies) in the cell also tend to be insoluble in the detergent-free aqueous B-PER reagent described above. These proteins (hereafter called "insoluble proteins") require further treatment before their solubilization can be achieved. Therefore to obtain all the protein (including soluble and insoluble proteins) from the bacterial cell pellets (Section 6.2.5) two different protein solubilization buffers were used: 4 × SDS

buffer made in 50 mM Tris-HCl (pH 6.8) and 8 M Gu-HCl buffer containing 2 M urea (pH 8). The cell pellets were gently resuspended in 5 ml of solubilization buffer per gram wet weight of the cells. The cell suspensions were lysed on a platform rocker by gentle vortexing for 120 min. The cell lysates containing the proteins were collected by centrifugation at 10,000 × *g* for 45 min and stored at -20º C until further use.

6.2.6.3 Refolding of solubilized protein

The solubilized protein fractions were refolded by dialysis against a linear Gu-HCl gradient (7 M – 0 M) made in 50 mM Tris.HCl, pH 8. The dialysis was performed using regenerated cellulose dialysis tubing (Thermo Scientific, USA) and all the experiments involving protein refolding were carried out at 4 ºC with the buffers changed every 12 h.

6.2.7 Protein analysis

Fifteen microlitres of the refolded protein samples (Section 6.2.6.3) were subjected to 1D SDS-PAGE on NuPAGE® 4 – 12% (w/v) Bis-Tris gels (Invitrogen, USA) (Section 2.6.2). Gel pieces potentially containing the expressed Nit 1 protein were excised for nanoLC MS/MS analysis as previously described (Section 4.2.8.2). The spectral data obtained from the LC MS/MS analysis were searched against the in-house created *T. reesei* protein database using the Mascot search engine (Matrix Science, London, United Kingdom).

6.2.8 Northern hybridization

Northern blotting was performed in order to investigate *nit1* expression using total RNA extracted with TRI Reagent® LS (Sigma, USA) from induced *E. coli* Origami^{TM}B(DE3) and Tuner^{TM}DE3 transformants (Section 2.4.5). All the solutions and buffers used for the Northern analyses were prepared in DEPC-treated water to ensure an RNase free environment. The primer pair *nit1*.Fwd.3 and *nit1*.Rev.2 (Table 5.1, Chapter 5) was used to

amplify the hybridization probe (415 bp) for the Northern blot analysis. Total RNA extracted from Origami^{TM}B(DE3) and Tuner^{TM}DE3 transformants containing the pETDuet-1 plasmid without the *nit1* gene insert and cultivated at both 37 ºC and 22 ºC for 3 h post-induction were used as negative controls. The quality of the extracted total RNA used for the Northern analyses was determined on a 1% (w/v) agarose gel and by measuring the absorbance at 230, 260, 280 and 320 nm.

6.2.9 Dot blot analysis

A dot blot analysis was performed using the total RNA extracted from induced Origami^{TM}B(DE3) recombinant cells grown at 22 ºC and 37 ºC for 1 and 3 h to investigate *nit1* expression. The analysis was carried out using the Bio-Dot^{TM} apparatus (Bio-Rad, USA) following the manufacturer's protocol.

6.2.10 Nitrilase activity assays

The nitrilase activity was measured from the soluble and the insoluble fractions of the recombinant *E. coli* strains using fluorometric (Banerjee *et al*., 2003b) and colorimetric (Yazbeck *et al*., 2006) nitrilase activity assays (Section 2.7). The nitrilase activity was determined using both nitrilase activity assays to enable more confident assessment of the results. All the assays were performed in triplicate and enzyme activities are represented as percentages relative to the positive control (Codexis, USA). The activity of the positive control was claimed to be approximately 43 U/mg by Codexis Pty Ltd, USA when the reactions were performed at 37 ºC. One unit of nitrilase activity was defined as the amount able to release 1 µmol of ammonia per mg of protein under the assay conditions. Approximately 25 µg of total protein extracted from the *E. coli* transformants was used for all the assays. Adiponitrile was used as the substrate for all assays.

6.2.11 Tools for the detection of potential glycosylation sites of Nit1

Analyses of the Nit 1 amino-acid sequence for the detection of potential *N*- and *O*-glycosylation sites was carried out using the NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) and NetOGlyc 3.1(http://www.cbs.dtu.dk/ services/ NetOGlyc/) prediction softwares.

6.3 Results

6.3.1 The pET- *nit1* vector

The pET-*nit1* vector was generated by cloning the *nit1* cDNA into the pETDuet1 vector using the *NcoI* and *BamHI* restriction sites. The successful insertion of the *nit1* cDNA into the vector (Figure 6.1) was determined by sequencing the pET-*nit1* plasmid.

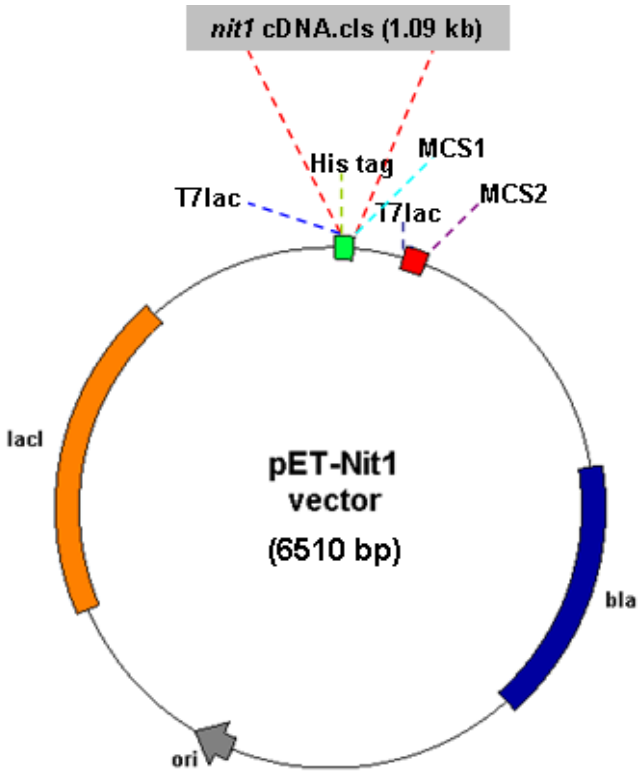


Figure 6.1 Representation of the pET-*nit1* vector. The *nit1* cDNA (1090 bp) was inserted into the pETDuet-1 vector (5420 bp) at the Multiple Cloning Site 1 (MCS1) using *SacI* and *PstI* restriction sites. The final size of the pET-*nit1* vector was 6510 bp. The "bla" represents the ampicillin resistance gene and the "lacI" represents the *lac* gene encoding the lactose repressor; T7 lac represents the T7 promoter region and "ori" represents the origin of replication. A region encoding for a 6×His-tag is present at the 5′ end of the MCS1 thus enabling the purification of the recombinant Nit 1 protein with an N-terminal His-tag.

The successful transformation of the Origami™B(DE3) and Tuner™DE3 cells with the pET-*nit1* vector was confirmed by colony PCR (Section 2.2.5) with bands seen at approximately 1.09 kb using single colonies picked from the *E. coli* transformants (data not shown).

6.3.2 Analysis of potential glycosylation sites of the Nit 1 protein

Glycosylation is a post-translational modification involving the attachment of sugar moieties (glycans) to proteins. Incorrect glycosylation affects the characteristics and properties of the gene products especially during heterologous expression in non-native hosts (Elliott *et al.*, 2004). Since many fungal gene products are reported to be glycosylated it was necessary to look for potential glycosylation sites in the predicted *T. reesei* Nit 1 before its recombinant production in *E. coli*. Hence, the Nit1 amino acid sequence was analyzed for potential *N*-glycosylation and *O*-glycosylation sites as described in Section 6.2.11. According to the results of the analyses, recombinant production of the Nit1 enzyme in *E. coli* was considered appropriate as the protein did not contain any predicted *N*-linked or *O*-linked glycosylation sites.

6.3.3 Cloning of the *nit1* gene into *E. coli* strains

The *nit1* cDNA from *T. reesei* grown in CLS medium was digested with *NcoI* and *BamHI* and the formation of DNA fragments of appropriate size was verified by agarose gel electrophoresis (Section 6.2.2). The pET-*nit1* vector was formed by inserting the purified *nit1* cDNA fragment into the pETDuet-1 vector using DNA ligase. Successful transformation of *E. coli* DH5α cells with the pET-*nit1* vector was confirmed by colony PCR using primers specific for the *nit1* cDNA (Section 6.2.3) and a DNA fragment of the expected size (1.09 kb) was obtained. Following isolation of the amplified pET-*nit1* plasmids from the transformed DH5α cells, the pET-*nit1* vector was transformed into *E. coli* Origami™B(DE3) and Tuner™DE3 cells. Successful transformation was again confirmed by colony PCR (Section

6.2.4) and DNA fragments of the expected size of the *nit1* cDNA (1.09 kb) were obtained. The transformed Origami™B(DE3) and Tuner™DE3 cells were cultured in Terrific broth and *nit1* expression was induced with 0.1 mM IPTG (Section 6.2.5).

6.3.4 Proteomic analysis of the soluble protein of *E. coli* transformants

The soluble proteins from Origami™B(DE3) and Tuner™DE3 transformants induced by 0.1 mM IPTG were subjected to 1D SDS-PAGE analyses to seek evidence of the expressed Nit 1 protein. As no glycosylation sites were identified for the Nit 1 protein, the expected size of the enzyme was about 40 kDa based on the Nit 1 sequence obtained using the *nit1* cDNA (Section 5.3.6).

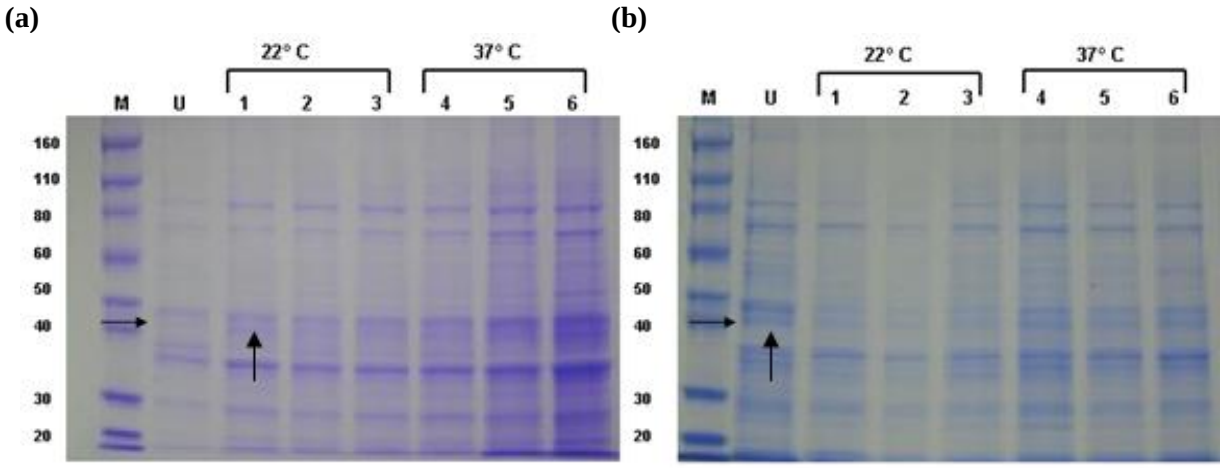


Figure 6.2 SDS-PAGE of soluble protein from (a) Origami™B(DE3) and (b) Tuner™DE3 cells transformed with the pET-*nit1* vector. Lane M: Molecular weight marker; Lane U: soluble protein from cells that were not induced with IPTG; Lanes 1, 2 and 3: soluble protein from cells grown at 22 °C for 1, 2 and 3 h post induction with 0.1 mM IPTG, respectively; Lanes 4, 5 and 6: soluble protein from cells grown at 37 °C for 1, 2 and 3 h post induction with 0.1 mM IPTG, respectively. The expected molecular weight of the Nit 1 was approximately 40 kDa as indicated by the arrows.

Prominent protein bands were seen on the SDS-PAGE gels at approximately 40 kDa as indicated by the arrows in Figures 6.2 (a) and 6.2 (b). In the gel containing the soluble protein fractions from the Origami™B(DE3) recombinants, a gradual increase in the intensity of the 40 kDa protein band was seen as the time since induction by IPTG increased from 1 h to 3 h. The intensity of the protein band at 40 kDa was the highest in the gel lane containing the

soluble protein fraction extracted from the Origami™B(DE3) transformants grown at 37 °C for 3 h post induction (Figure 6.2a).

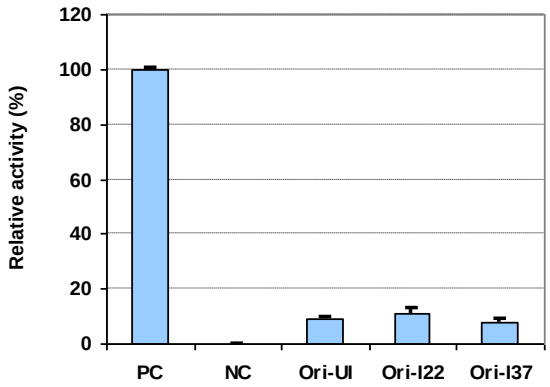
Since the T7 promoter system present in the pET-Duet-1 vector was used in this work, basal level expression of the Nit 1 protein was expected even under un-induced culture conditions (Mertens *et al*., 1995). Still, the band intensities of protein fractions from the induced cultures were expected to be higher than from the un-induced cultures. However, in the gel containing the soluble protein fractions from Tuner™DE3 recombinants (Figure 6.2b), the intensity of the protein band at 40 kDa was observed to be stronger for the soluble protein extracts from the un-induced Tuner cells than from the induced transformants. Hence, it appeared that the protein bands observed at 40 kDa in Figure 6.2b possibly contained either a native *E. coli* protein or a group of proteins including the Nit 1 enzyme sharing a similar molecular weight.

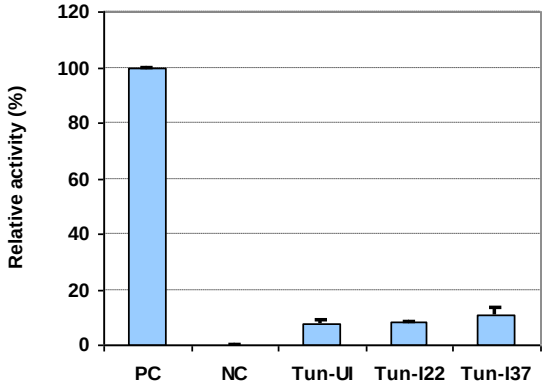
To find evidence of the Nit 1 enzyme in the soluble protein from the *E. coli* transformants by proteomic analyses, a portion of the gels containing proteins of approximately 38 – 42 kDa was excised for in-gel LC MS/MS analysis (Section 6.2.7). However, when the mass spectra were searched using the Mascot software against the in-house *T. reesei* protein database, no peptides from the Nit 1 enzyme were identified from the soluble protein of the Origami™B(DE3) and Tuner™DE3 cells.

6.3.5 Nitrilase activity of the soluble protein of *E. coli* transformants

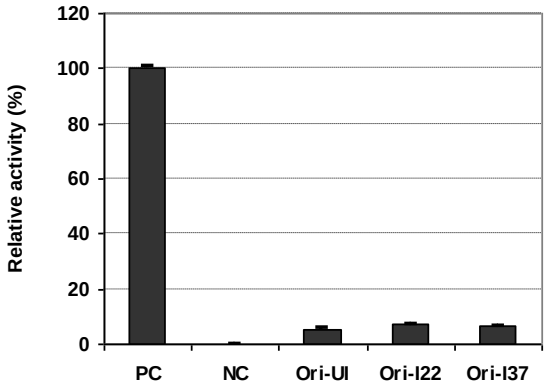
Nitrilase activity was determined from the soluble protein fractions of Origami™B(DE3) and Tuner™DE3 transformants 3 h post-induction with 0.1 mM IPTG using both the fluorometric and colorimetric nitrilase assays (Section 6.2.10). The assays were carried out in triplicate and the average activity of the soluble proteins from the *E. coli* transformants relative to a commercial nitrilase is presented in Figure 6.3. Soluble protein from cells transformed with vectors without *nit1* cDNA were used as a negative control, and the background florescence/ absorbance exhibited by the negative control was subtracted from all sample data.

(a) Origami™B(DE3) **(b) Tuner™DE3**





(c) Origami™B(DE3) **(d) Tuner™DE3**



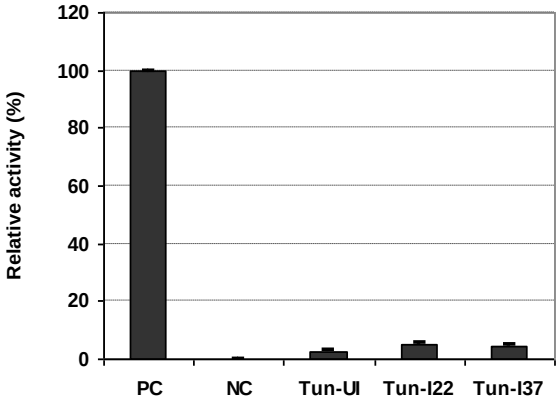


Figure 6.3 Nitrilase activity of soluble protein of recombinant Origami™B(DE3) and Tuner™DE3 cells as measured by fluorometric (a and b) and colorimetric (c and d) assays. Nitrilase activity is represented as a percentage relative to a positive control. PC: positive control, a commercial nitrilase from Codexis (USA) with an activity of approximately 43 U/mg; NC: cells transformed with vectors without *nit1* cDNA insert; UI: cells not induced with IPTG; I22 and I37: cells induced with 0.1 mM IPTG for 3 h at 22 and 37 °C, respectively. The background fluorescence and absorbance of the negative control (NC) was subtracted from all sample data. The fluorescence/ absorbance exhibited by the chemicals and buffers used in the assays was subtracted from the positive control data. Adiponitrile was used as a substrate for all the assay experiments. Approximately 25 µg of total protein was used for all the assays.

The activities of the soluble proteins from the all the *E. coli* transformants were less than 12% of the activity of the positive control (43 U/ mg). Slightly higher activities were recorded using the more sensitive fluorometric assay (Section 1.6.2.2) than the colorimetric assay but in either assay no significant differences were seen between the relative activities of soluble protein fractions obtained from the induced and un-induced Origami™B(DE3) and Tuner™DE3 transformants (Figure 6.3). Furthermore, as the assay readings were so low from

any of the samples it was difficult to be confident that accurate measurement of nitrilase activity was being achieved.

The failure to detect peptides unique to the Nit 1 enzyme (Section 6.3.4) and the difficulties experienced in detecting nitrilase activity in the soluble protein of the *E. coli* transformants could have been due to lack of *nit1* gene transcription, degradation of the Nit 1 protein due to proteolysis (Maurizi *et al*., 1992) or the possible aggregation of the Nit 1 protein in insoluble inclusion bodies in either the periplasm or cytoplasm of the *E. coli* strains (Section 1.7.1). Hence, the *nit1* expression at the mRNA level was investigated using Northern blot and Dot blot analyses.

6.3.6 Northern analysis of *nit1* gene expression

Total RNA (Figure 6.4) was isolated from both induced and un-induced Origami^{TM}B(DE3) and Tuner^{TM}DE3 recombinants grown at 22° and 37 °C and at two time points (1 h and 3 h) post induction for Northern blot analyses.

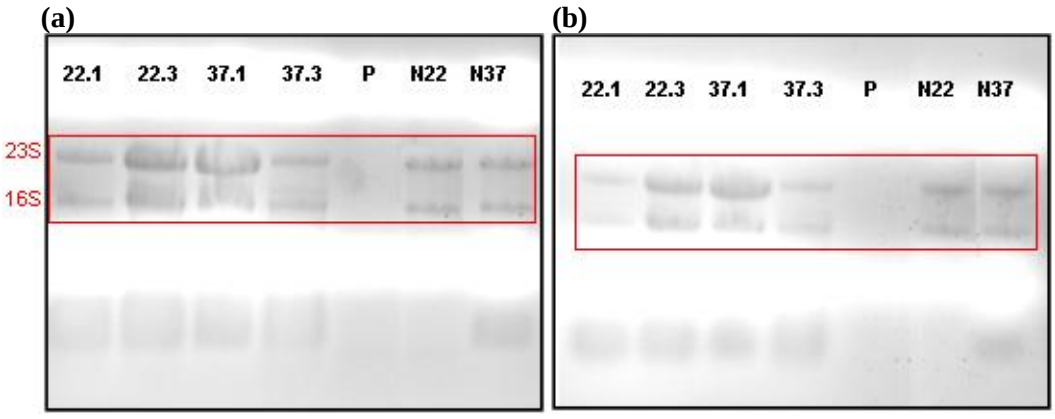


Figure 6.4 Agarose gel electrophoresis of total RNA extracted from (a) Origami^{TM}B(DE3) and (b) Tuner^{TM}DE3 recombinant cells for Northern analysis. Lanes 22.1 and 22.3: RNA from cells harboring the *nit1* gene and grown at 22 °C for 1 and 3 h after induction with 0.1 mM IPTG, respectively; Lanes 37.1 and 37.3: RNA from cells harboring the *nit1* gene and grown at 37 °C for 1 and 3 h after induction with 0.1 mM IPTG, respectively; Lane P: the hybridization probe (415 bp) which was amplified from the *nit1* gene using the primer pair *nit1*.Fwd.3 and *nit1*.Rev.2 (Table 5.2); Lane N22 and N37: RNA from cells containing the pETDuet-1 vector without the *nit1* cDNA insert grown for 3 h after induction at 22 °C and 37 °C, respectively. The 23S and the 16S ribosomal RNA are highlighted by the red box.

The Northern blot results for the detection of *nit1* transcripts at the mRNA level using total RNA extracted from both the Origami and Tuner cells are shown in Figure 6.5. Following hybridization with the *nit1* specific probe (0.415 kb) that targeted the *nit1* mRNA, a single prominent band (~ 1.1 kb) of high intensity was observed from the total RNA sample isolated from the Origami recombinants harboring the *nit1* gene and grown at 22 °C for 1 h after IPTG induction (Figure 6.5). A faint band was also observed at the same size for the Origami recombinants with the *nit1* gene grown at 22 °C for 3 h. However, no Northern blot signals denoting *nit1* transcription at mRNA levels was obtained from Origami recombinants grown at 37 °C (Figure 6.5).

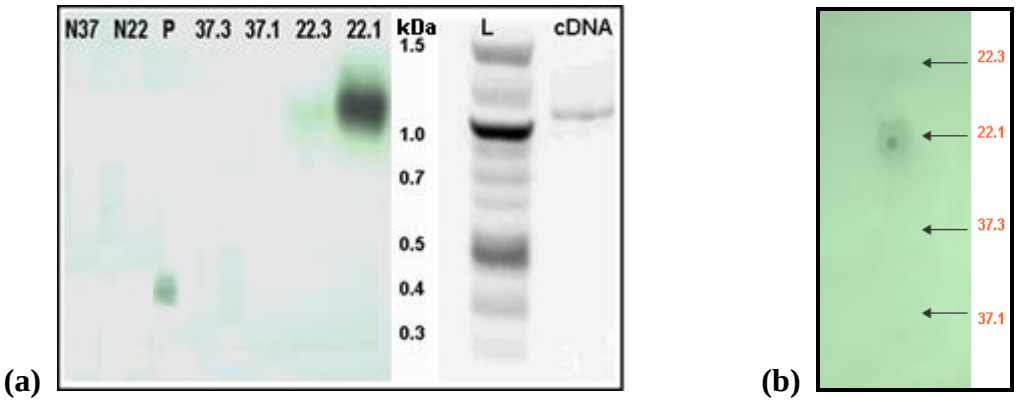


Figure 6.5 (a) Northern blot of total RNA from Origami™B(DE3) recombinant cells 10 µg of total RNA was loaded on each lane. Lanes N22 and N37: negative controls, RNA from cells containing the pETDuet-1 vector without the *nit1* cDNA insert grown for 3 h at 22 °C and 37 °C, respectively; Lane P: hybridization probe; Lanes 37.1 and 37.3: RNA from cells harboring the *nit1* gene and grown at 37 °C for 1 and 3 h after induction with 0.1 mM IPTG, respectively; Lanes 22.1 and 22.3: RNA from cells harboring the *nit1* gene and grown at 22 °C for 1 and 3 h after induction with 0.1 mM IPTG, respectively. The sizes of the Northern blot signals were determined by direct comparison with an agarose gel containing a DNA ladder. Lane L: DNA ladder; Lane cDNA: *nit1* cDNA. Expected size of the *nit1* transcript was approximately 1.1 kb. (b) Dot-blot analyses of total RNA from induced Origami™B(DE3) cells harboring the *nit1* gene.

Even though the quality of the total RNA extracted from the Tuner™DE3 cells was good, based on the clear separation of the 23S and 16S ribosomal RNA bands on an agarose gel (Figure 6.4b), no signals indicating *nit1* expression were obtained from Northern blot analysis. The lack of hybridization bands for all the Tuner samples could have been either due to low transformation efficiency or the formation of non-detectable levels of *nit1* transcripts.

Therefore, a RNA dot blot hybridization was performed to affirm the results of Northern analyses with Bio-Dot^{TM} apparatus (Bio-Rad, USA) using the total RNA extracted from both the Origami and Tuner cells grown at 22 °C and 37 °C for 1 h and 3 h post induction with IPTG. RNA dot-blot hybridization is a robust technique that involves the quantitation of the gene expression by transfer of total RNA onto the membrane without prior size separation on an agarose gel.

A prominent spot indicating *nit1* expression was obtained only for total RNA extracted from Origami transformants harboring the *nit1* gene grown at 22 °C for 1 h post induction with IPTG (Figure 6.5b). This was in line with the results obtained from the Northern blot experiment (Figure 6.5a). However, unlike in the Northern blot, no spot indicating *nit1* expression was observed for the samples from Origami cultures grown at 22 °C for 3 h (Figure 6.6). No signals indicating *nit1* transcription were obtained from RNA from any of the Tuner cells using Dot-blot analyses (data not shown).

The Northern and dot blot analyses provided evidence that *nit1* transcripts were formed in the Origami recombinants grown at 22 °C for 1 h post induction. Therefore, the inability to detect the Nit 1 protein in the soluble protein of these cells (Section 6.3.4) and the very low levels of nitrilase activity (Section 6.3.5) suggested that the Nit 1 protein was probably aggregating as inclusion bodies. Therefore, the insoluble protein of Origami transformants (Section 6.2.6.2) was used for further investigation of Nit 1 expression (Section 6.3.7).

6.3.7 Analysis of insoluble protein from transformed Origami cells

Total protein (including soluble and insoluble protein) was extracted from Origami transformants maintained at 22 °C for 1 and 3 h post-induction using 4 × SDS solubilization buffer (Section 6.2.6.2). The solubilized protein was analyzed by SDS-PAGE. A protein band was visualized at the expected molecular weight range of approximately 40 kDa in the gel

lanes containing protein fractions from recombinant Origami cells grown at 22 °C for 1 and 3 h after induction with 0.1 mM IPTG (Figure 6.6).

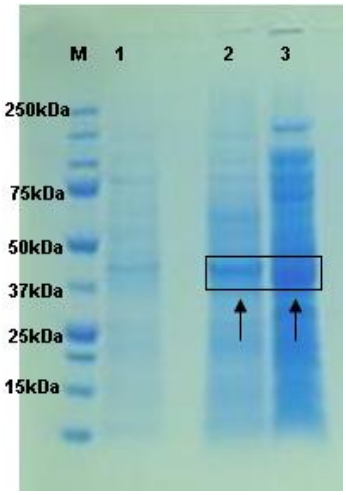


Figure 6.6 Total protein (soluble and insoluble) from Origami™B(DE3) recombinant cells containing the *nit1* gene, without IPTG induction (lane 1), and 1 h and 3 h post-induction with 0.1 mM IPTG at 22 °C (lanes 2 and 3, respectively). Lane M: molecular weight marker. The protein bands that may contain the Nit 1 protein are boxed and indicated by black arrows. Expected size of the Nit 1 is approximately 40 kDa. 15 µl of the solubilized protein were loaded per lane.

Modifications to cultivation parameters including IPTG concentration and growth temperatures has been reported to boost recombinant nitrilase production in *E. coli* (Kaplan *et al*., 2011). Hence to test the effect of IPTG concentration on Nit 1 production, recombinant Origami cells were induced for 1 h and 3 h at 22 °C with different concentrations of IPTG ranging from 0.1 mM to 0.6 mM (Section 6.2.5). The cell biomasses were weighed and resuspended in 5 ml of solubilization buffer. The protein concentrations of the samples were then compared using Bradford's assay (Bradford, 1976). The total culture biomass generally decreased with an increase in the IPTG concentration, and for each of the IPTG concentrations, the cells grown for 3 h post-induction consistently contained higher concentrations of protein per gram wet weight of the cells compared to cells grown for 1 h post-induction. The total protein produced per milligram of cell biomass for Origami recombinant cells maintained for 3 h at 22 °C after induction with different concentrations of IPTG is shown in Table 6.1.

Table 6.1 Protein production (mg/mg of cell biomass) of Origami™B(DE3) transformants induced for 3 h at 22 °C with various concentrations of IPTG.

IPTG concentration (mM)	Total protein per mg of cell biomass (mg/mg)
0.1	0.44
0.2	0.16
0.3	0.67
0.4	2.46
0.5	1.03
0.6	1.25

The highest protein production per milligram of cell biomass was achieved using 0.4 mM of IPTG (Table 6.1). To investigate the effect of the IPTG concentrations on Nit 1 protein production specifically, 15 µl of the resolubilized total protein from Origami cells induced with different concentrations of IPTG were used for SDS-PAGE analysis (Figure 6.7).

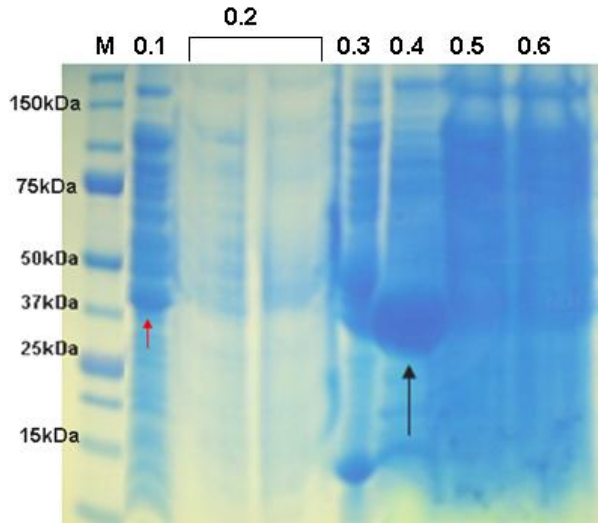


Figure 6.7 Total protein (soluble and insoluble) from Origami recombinants induced at 22 °C for 3 h with 0.1, 0.2 (two lanes), 0.3, 0.4, 0.5 and 0.6 mM IPTG. 15 µl of protein samples were loaded per lane. Lane M, molecular weight marker. The red arrow indicates a protein band that may represent Nit 1 protein expression using 0.1 mM IPTG for induction. The black arrow indicates an abundance of protein that may include Nit 1 produced by cells induced with 0.4 mM IPTG. Expected molecular weight of Nit 1 is 40 kDa.

The SDS-PAGE gels of protein (soluble and insoluble) from the recombinant Origami cells revealed very high total protein production using IPTG concentrations of 0.4 mM IPTG and

over; the protein content was so high in the samples that massive overloading of the gel had clearly occurred (Figure 6.7). However, whereas high production of proteins across a broad molecular weight range was evident for cells induced with 0.5 and 0.6 mM IPTG, protein of the approximate expected molecular weight of the Nit 1 protein (40 kDa) was predominant in cells induced with 0.4 mM IPTG. The fact that the large band of proteins in the sample from cells induced with 0.4 mM IPTG appeared a little lower than the indicator of 37 kDa in the molecular weight marker may have been due to overloading of protein in the gel. The total protein production (including proteins of approximately 40 kDa) was significantly lower in the cells induced with 0.2 mM IPTG (Figure 6.7), a phenomenon compatible with the low protein production of these cells per mg of cell biomass (Table 6.1). The possible explanations for the variable production of the Nit 1 enzyme at the different concentrations of IPTG are detailed in Section 6.4.

The prominent bands of 35 – 40 kDa from gel lanes containing protein from cells induced with 0.1 and 0.4 IPTG (Figure 6.7) were sliced, digested and subjected to in-gel 1D nano-LC MS/MS analysis to search for peptide sequences from the predicted Nit 1 protein. The LC MS/MS data were searched against the in-house created *T. reesei* database and confident matches were made to the predicted Nit 1 nitrilase from both the 40 kDa band from the cells induced with 0.1 mM IPTG and the large 35 – 40 kDa band from cells induced with 0.4 mM IPTG. In fact, the latter produced a total of 271 spectral matches from 18 peptides unique to the predicted Nit 1 protein, achieving a Mascot score of 3878 indicating high significance of the match of the sample protein to the predicted *T. reesei* Nit 1 nitrilase in the database.

The above analysis was carried out on total protein from the Origami transformants and identification of the Nit 1 protein was clearly achieved. The total protein that had been extracted from the cells for these experiments using the 4 × SDS buffer (Section 6.2.6.2) included both insoluble protein (from protein aggregates/ inclusion bodies) and soluble proteins. Identification of the Nit 1 protein in the soluble proteins of the Origami

transformants had not been achieved (Section 6.3.4), nor was there clear evidence of nitrilase activity (Section 6.3.5). Therefore it could be concluded that the Nit 1 protein was probably forming inclusion bodies in the Origami cells.

Although the 4 × SDS buffer had enabled solubilization and proteomic analysis of the insoluble proteins in the Origami cells, it also caused denaturation of the proteins which typically can affect enzyme activity. Therefore to measure the enzyme activity of the insoluble proteins another solubilization buffer was used to enable protein refolding as described below.

6.3.7.1 Refolding of the Nit 1 protein

Solubilization of total protein (soluble and insoluble protein) from Origami recombinant cells induced with 0.4 mM IPTG was repeated using an alternative protein solubilization buffer containing 8 M guanidine HCl (Gu-HCl) and 2 M urea (Section 6.2.6.2). The Gu-HCl buffer was inappropriate for SDS-PAGE analysis of the protein described above because Gu-HCl solubilized proteins tend to precipitate when treated with SDS for electrophoresis. However, the Gu-HCl solubilization buffer had advantages over the 4 × SDS buffer when planning further downstream protein renaturation for enzyme activity analysis; the urea in the Gu-HCl buffer helps to stabilize the protein upon refolding. After resolubilization in the Gu-HCl buffer the total protein from the Origami transformants (including insoluble protein from inclusion bodies) was refolded as described by Laymon *et al*. (1996) by dialysis against a linear 7 – 0 M Gu-HCl gradient (Section 6.2.6.3). The refolded protein was used for the activity assays.

6.3.7.2 Nitrilase activity of refolded protein from Origami transformants

Nitrilase activity of the refolded total protein (Section 6.3.4.2) from Origami recombinant cells induced with 0.4 mM IPTG was measured by both the fluorometric and colorimetric activity assays (Figure 6.8).

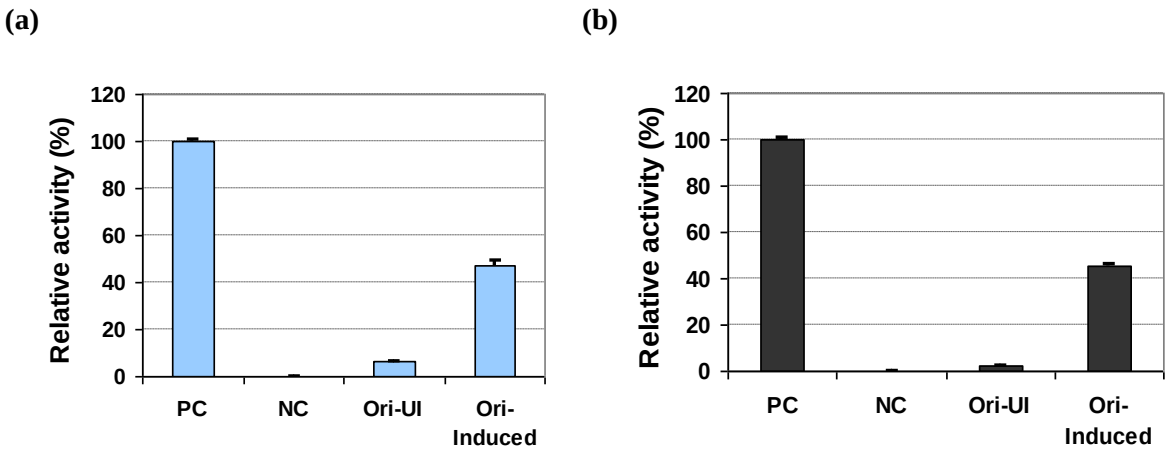


Figure 6.8 Nitrilase activity of refolded total protein (soluble and insoluble) of recombinant Origami™B(DE3) cells as measured by (a) fluorometric and (b) colorimetric assays. Nitrilase activity is represented as a percentage relative to a positive control. PC: positive control, a commercial nitrilase from Codexis (USA) with an activity of approximately 43 U/mg. NC: cells transformed with vectors without *nit1* cDNA insert; UI: cells not induced with IPTG; Ori-Induced: cells induced with 0.4 mM IPTG for 3 h at 22 °C. The background fluorescence and absorbance of the negative control (NC) was subtracted from all sample data. The fluorescence/ absorbance exhibited by the chemicals and buffers used in the assays was subtracted from the positive control data. Adiponitrile was used as substrate. Approximately 25 µg of total protein was used for the assays.

The activity of the resolubilized total protein from the Origami recombinants induced with 0.4 mM IPTG (Figure 6.8) was 45% of that displayed by the positive control (43 U/mg), much higher than the nitrilase activity of the soluble protein from the Origami recombinants induced with 0.1 mM IPTG which was barely detectable in the assay described in Section 6.3.5. Thus a nitrilase activity of approximately 19 U/mg could be attributed to total protein of the recombinant cells, which would have included the Nit 1 protein, probably in inclusion bodies.

6.4 Discussion

In the work described in this chapter, the *T. reesei nit1* cDNA isolated in this research (Chapter 5) was recombinantly expressed in Origami^{TM}B(DE3) cells and the Nit 1 protein was identified from the insoluble protein of the transformants. The work performed by Kaplan *et al*. (2011) on the recombinant expression of the gene encoding nitrilase from *A. niger* in *E. coli* was used to guide the production of recombinant Nit 1 enzyme in *E. coli*.

6.4.1 Expression of *T. reesei nit1* cDNA in *E. coli*

In the work described by Kaplan *et al*. (2011), the cDNA of a nitrilase encoding gene from *A. niger* was ligated into the pET expression vector (pET-30(+)) before transformation into *E. coli*. Similarly in this work, expression of the *T. reesei nit1* cDNA was attempted using a pET expression vector (pETDuet-1). One advantage of the pET vector system is the expression of genes under the control of the T7 promoter. Since the T7 promoter is not recognized by the native *E. coli* RNA polymerase, a specific T7 RNA polymerase can be used to exclusively facilitate the expression of the target gene expressed under the T7 promoter thereby boosting transcription efficiency. Several genes that have failed to express using *E. coli* promoters such as *tac* and *trc* have been successfully cloned and expressed using pET expression vectors (Studier *et al*., 1986). Also, bacterial nitrilases have been produced in *E. coli* using pET expression systems *e.g.* a regioselective aliphatic nitrilase from *Acidovorax facilis* (Chauhan *et al*., 2003).

The expression of the *A. niger* nitrilase cDNA was attempted by Kaplan *et al.* (2011) in nine different *E. coli* strains, including five BL21 variants deficient in the OmpT and Lon proteases which are known to interfere with the isolation of recombinant proteins, a Rosetta-gami strain derived from an Origami^{TM} strain, and three Artic strains. The production of the enzyme was achieved in only seven strains with the maximum nitrilase activity recorded from a BL21-Gold(DE3) strain. The absence of nitrilase activity in two BL21 strains was attributed

to low transformation efficiency or endonuclease activity. In all the seven strains, the recombinant nitrilase was reported to constitute almost half of the soluble proteins.

In this work, Origami™B(DE3) and Tuner™DE3 cells were chosen for the production of the recombinant *T. reesei* Nit 1 protein. Although the Origami™B(DE3) and the Tuner™DE3 cells contain the OmpT and Lon proteases which are absent in the BL21 strains, they are known to boost recombinant protein solubility. Origami ™B(DE3) cells have been reported to yield ten-fold more functionally active proteins than any other hosts (Prinz *et al*., 1997). Furthermore, the formation of disulphide bridges in recombinant proteins is enhanced in Origami™B(DE3) cells due to mutations to thioredoxin reductase and glutathione reductase genes which alter the reducing conditions in the *E. coli* cytoplasm and reportedly preventing protein misfolding (Stewart *et al*., 1998). Tuner™DE3 cells are also favored for recombinant protein production due to a lacZY mutation that ensures a uniform distribution of the inducer IPTG in the cells thereby influencing protein expression. Tuner™DE3 cells are also known to boost protein solubility by reducing protein aggregation (Maurer, 2012). It was hoped that utilization of the Origami™B(DE3) and Tuner™DE3 cells for the expression of the *T. reesei nit1* gene would increase the likelihood of obtaining soluble Nit 1 protein. In addition, protein expression was attempted at 22 °C, as well as the more standard temperature of 37 °C, since lower temperatures are known to reduce protein aggregation by slowing the rate of transcription, translation and protein folding, thus decreasing the incidence of protein misfolding. Moreover, hydrophobic interactions that are known to promote protein aggregation are also considerably reduced at lower temperatures. Lower temperatures are also reported to reduce the incidence of heat shock proteases that can cause degradation of the recombinant protein (Sorensen and Mortensen, 2005).

Due to the above explained reasons, the Nit 1 protein in the induced recombinant cells harboring the pET-*nit1* vector was expected to be located primarily amongst the soluble protein in the cell. However this was not the case since no peptides from the predicted Nit 1

protein were detected from the soluble protein by mass spectrometric analysis (Section 6.3.4) and enzyme assays could not definitively confirm nitrilase activity (Section 6.3.5). The difficulty experienced in detecting Nit 1 in the soluble protein of the *E. coli* strains could have been due to a number of reasons including poor transcription of the *nit1* gene, degradation of the Nit 1 protein by proteolysis or low sensitivity of the assays. Also, more abundant soluble proteins native to *E. coli* could have masked the presence of Nit 1 by dominating the mass spectra in the LC MS/MS analysis, or by interfering with the assay readings by blocking either the formation of the fluorophore in the case of the fluorometric assay or the ammonia-cobalt complex in the case of the colorimetric assay. However, it was also quite conceivable that the Nit 1 protein could be aggregating into inclusion bodies in either the periplasm or cytoplasm of the *E. coli* strains.

To investigation whether transcription of the *nit1* gene had occured, Northern blot and Dot blot experiments were carried out (Section 6.3.6). Results indicated that the *nit1* mRNA transcripts were being formed only in Origami™B(DE3) transformants grown at 22 °C. Transcription of *nit1* was not evident from Northern blot or dot blot analysis of any of the Tuner™DE3 transformants, although the presence of the *nit1* gene had been confirmed by colony PCR (Section 6.3.3) and the quality of the RNA extracted from the Tuner™DE3 cells appeared similar to that of the Origami™B(DE3) cells (Figure 6.4). A possible reason for poor transcription of *nit1* in the Tuner™DE3 transformants was the addition of glucose in the growth medium which in certain cases has been known to suppress protein expression in Tuner cells. For example, Tuner™DE3 cells have failed to express green fluorescent protein when grown in a medium containing glucose before induction whereas expression was achieved when the cells were grown without glucose supplementation (Novy and Morris, 2001). The decrease in protein expression was attributed to low cell density of the cells at stationary phase due to acidic culture conditions caused by the products of glucose metabolism (Novy and Morris, 2001). The reason for the lack of *nit1* mRNA transcripts in the

Tuner cells in this work was not fully investigated because transcription had been achieved in the Origami cells, enabling the project to proceed towards characterization of the recombinant protein.

6.4.2 The recombinant Nit 1 enzyme

Detection of *nit1* mRNA transcripts in the Origami^{TM}B(DE3) cells, despite difficulty in detecting Nit 1 amongst the soluble proteins, lead to the conclusion that most of the recombinant protein was probably aggregating into inclusion bodies. Therefore, further work to investigate the recombinant Nit 1 was carried out using a protein solubilization buffer that enabled analysis of both the soluble and insoluble (inclusion body) proteins in the cell (Section 6.3.7). Furthermore, Nit 1 production was explored using different concentrations of the IPTG inducer since in the work of Kaplan *et al*. (2011) a several fold increase in the expression of the recombinant *A. niger* nitrilase was achieved by manipulation of IPTG concentration.

The highest total protein production per weight of cell biomass was achieved when the Origami recombinants harboring the *nit1* gene were induced with 0.4 mM IPTG (Table 6.1, Section 6.3.4). The highest production of protein of the expected molecular weight of Nit 1 was also achieved using 0.4 mM IPTG, as demonstrated by SDS-PAGE (Figure 6.7); mass spectrometric analysis confirmed that the most abundant protein seen on the gel was Nit 1. However, the total culture biomass generally decreased with the increase in IPTG concentration (Section 6.3.4). It appeared that the cell growth was inhibited either by higher concentrations of the inducer IPTG or by the Nit 1 enzyme itself since the culture with the lowest biomass also appeared to be the highest producer of the Nit 1 enzyme when induced with 0.4 mM IPTG.

Confirmation of the presence of the recombinant Nit 1 protein in the insoluble proteins (inclusion bodies) of the Origami cells as described above was achieved using a 4 × SDS

solubilization buffer compatible with downstream SDS-PAGE and mass spectrometric analysis (Section 6.3.4). However this buffer caused the irreversible denaturation of the Nit 1 protein, which would also affect the enzyme activity. Therefore, analysis of the enzyme activity of Nit 1 required the use of a different solubilization buffer containing guanidine HCl and urea, which enabled stabilization of the protein during refolding by dialysis against a linear gradient of Gu-HCl (Laymon *et al*., 1996). Both the fluorometric and colorimetric assays detected hydrolysis of the substrate adiponitrile by the refolded protein from the Origami transformants carrying the *nit1* gene. As native nitrilases, nitrile hydratases or amidases are not known to exist in *E. coli* (http://genome.jgi-psf.org/Escherichi1), the hydrolysis of adiponitrile could be entirely attributed to monoenzymatic nitrilase activity of the recombinant *T. reesei* Nit 1 protein. Moreover, the nitrilase activity of the total protein of the Origami transformants, containing the recombinant Nit 1 nitrilase was 45% of that of a purified commercial nitrilase (Codexis, USA) with a reported activity of 43 Units/ mg.

6.5 Chapter summary

Recombinant expression of the *T. reesei nit1* gene in *E. coli* was attempted by inserting the cDNA isolated from *T. reesei* Rut-C30 grown in CLS medium (Chapter 5) into the pETDuet1 expression vector, plasmid amplification in DH5α cells, and subsequent transformation into Origami^{TM}B(DE3) and Tuner^{TM}DE3 cells. Although successful transformation of the Origami and Tuner cells was evident by colony PCR, expression of the Nit 1 protein in the soluble protein of the cells could not be detected by LC MS/MS or enzyme assays. Northern blot analysis of RNA confirmed transcription of the *nit1* gene in Origami transformants and analysis of total cellular protein revealed that the recombinant Nit 1 protein was aggregating in inclusion bodies in the Origami cells. Production of the Nit 1 protein was found to be highest when the Origami^{TM}B(DE3) transformants were cultivated at 22 °C and induced for 3 h with 0.4 mM IPTG. Although Nit 1 still appeared to aggregate into inclusion bodies, the

recombinant protein could be successfully solubilized and refolded into a functionally active form. Nitrilase activity of the total solubilized and refolded protein (containing the Nit 1 nitirilase) against the aliphatic nitrile substrate adiponitrile was demonstrated to be 45% of that of a commercial purified nitrilase. The next step was the purification of the recombinant *T. reesei* nitrilase using the attached His-tag for further biochemical characterization.

Chapter 7: Purification and characterization of the recombinant His-tagged Nit 1 protein

7.1 Introduction

Large scale purification of recombinant proteins has become theoretically relatively straight forward since the development of expression vectors that attach affinity tags to the protein of interest. Affinity tags facilitate rapid protein purification with most tags designed to allow single step purification through their affinity to an immobilized ligand. Polyhistidine-tag (His-tag), glutathione S-transferase (GST) tag and maltose binding protein (MBP) tag are widely used affinity tags for protein purification. Once purified, the tags can be cleaved from the target protein by endoproteases thus allowing their application in areas such as production of pharmaceuticals which require proteins of high purity (Waugh, 2011).

In this work, the 6× His-tag was used to purify the recombinant Nit 1 enzyme from *E. coli*. The 6× His-tag consists of six histidine residues which are attached to the protein of interest by inclusion of the corresponding coding sequence in the gene expression vector (Figure 6.1). Several single step protein purification systems using His-tags have been developed by companies such as QIAGEN and Roche. In this research, the purification of the His-tagged recombinant Nit 1 enzyme was carried out using the QIAexpress® System (QIAGEN, USA) which is based on the affinity of His-tagged proteins for nickel-nitrilotriacetic acid (Ni-NTA). The affinity and the interaction of the His-tagged proteins with the Ni-NTA is reported to be independent of the conformation and the tertiary structure of the proteins thus enabling protein purification under both native and denaturing conditions. Moreover, many nitrilases are known to form multimers in their active forms (Section 1.2.3) and the QIAexpress® System allows the elution of both monomeric and multimeric proteins with minor modifications to the elution buffers.

Purification of an enzyme enables more accurate characterization of its activity. Establishing substrate specificities and temperature and pH profiles can help to determine if the enzyme is appropriate for development for particular industrial, pharmaceutical or environmental applications. For example, a thermostable nitrilase from *B. pallidus* Dac521 was appropriate for the manufacture of nicotinic acid, a form of vitamin B3 (Almatawah and Cowan, 1999); nitrilases with broad substrate specificity are particularly useful for bioremediation (Gong *et al*., 2012). In the work described in this chapter, the pH and temperature profiles and substrate specificity of the recombinant Nit 1 enzyme were determined following purification of the His-tagged Nit 1 enzyme using the QIAexpress purification system.

7.2 Materials and methods

7.2.1 Culture conditions for recombinant expression of *nit1*

Culture conditions that were found to increase the production of Nit 1 in *E. coli* in work described in Chapter 6 were again used to produce the Nit 1 protein for the purification procedure described in this chapter. Origami^{TM}B (DE3) transformants harboring the pET-nit1 vector were grown overnight at 37 °C (250 rpm) in 50 ml Terrific broth supplemented with 100 μg/ml ampicillin. The culture was then diluted 1:20 into 200 ml of fresh Terrific broth and grown at 22 °C until the optical density at 600 nm was 0.5. Expression of *nit1* was induced with 0.4 mM IPTG for 3 h at 22 °C. The bacterial cells were harvested by centrifugation at 5000 × *g* for 15 min and the pellet was stored at -20 °C until further use.

7.2.2 Solubilization of total cellular protein

As the Nit 1 protein had been found to form insoluble aggregates in inclusion bodies of the cells (Section 6.3.7), the bacterial pellet (Section 7.2.1) was gently thawed on ice and resuspended in a protein solubilization buffer appropriate for the QIAexpress® protein purification system. The solubilization buffer contained 100 mM NaH_2PO_4, 10 mM Tris-HCl

and 8 M urea, pH 8 and 5 ml of the buffer was used per gram wet weight of the cells. The cell suspension was lysed by gentle vortexing for 60 min. The cellular lysate was collected by centrifugation at 10,000 × *g* for 35 min at room temperature. The cell supernatant containing all proteins (soluble and insoluble) was collected and stored at 4 ºC until the purification process. Approximately 20 µl of the protein was used for SDS-PAGE analysis to determine the presence of the Nit1 enzyme.

7.2.3 Batch purification of His-tagged Nit 1 protein

Purification of the His-tagged Nit 1 protein was performed using the QIAexpress® System (QIAGEN, USA) using the manufacturer's protocol with minor modifications. The purification of His-tagged proteins was performed under denaturing conditions which assists purification efficiency by exposing the His-Tag to the Ni-NTA matrix thus ensuring maximum binding.

QIAexpress Ni-NTA resin (1 ml) was added to 4 ml of cell lysate and mixed gently at room temperature on a rotary shaker at 200 rpm for 60 min. The resin-lysate mixture was carefully loaded into an empty polypropylene column (QIAGEN, USA) and the flow-through was collected. The column with the resin-lysate mixture was then washed twice with 4 ml of wash buffer (100 mM NaH_2PO_4, 10 mM Tris.HCl, 8 M urea, pH 6.3). The wash fractions were collected and stored for further analysis.

The recombinant Nit 1 protein was eluted using 0.7 ml of elution buffer 1 (100 mM NaH_2PO_4, 10 mM Tris.HCl, 8 M urea, pH 5.9) and elution buffer 2 (100 mM NaH_2PO_4, 10 mM Tris.HCl, 8 M urea, pH 4.5). Elution buffer 1 was designed for the elution of monomeric proteins whereas elution buffer 2 was designed for proteins that formed multimers. The proteins were eluted at room temperature and the elution fractions were collected for SDS-PAGE (Section 2.6.2) and activity assays (Section 7.2.5). The fractions were collected and stored at 4 ºC until further analysis.

To reduce the non-specific binding of non-target proteins to the Ni-NTA matrix, 20 mM imidazole was added to the lysis and wash buffer. Also, to prevent co-purification of the *E. coli* native proteins 20 mM β-mercaptoethanol was added to the lysis buffer. Furthermore, to boost the interaction of the His-tagged proteins with the nickel ions, the amount of the Ni-NTA resin was increased to 2 ml and the protein-Ni-NTA interaction time was increased from 1 h to 2 h.

7.2.4 Refolding of purified Nit 1 protein

Purified protein extracts were refolded following the method described by Laymon *et al*. (1996) with minor modifications. In brief, the protein extracts were transferred into dialysis tubes and the samples were dialyzed against a linear 7 M – 0 M urea gradient in 500 mM NaCl, 20 mM Tris-HCl buffer supplemented with 20% (v/v) glycerol at pH 7.4. The buffers were changed every 12 h and the dialysis was performed at 4 ºC with gentle stirring. The total protein concentration of the refolded protein fractions was determined using Bradford's protein assay (Bradford 1976).

7.2.5 Nitrilase activity assays

Out of the two nitrilase activity assays (Section 2.7), the fluorometric activity assay had been shown to be the more sensitive (Section 6.3.7.2) so was used to characterize the nitrilase activity of the purified Nit 1 protein. Temperature and pH profiles were determined with 20 mM adiponitrile as substrate. Enzyme activity assays at 20 to 60 °C were carried out in potassium phosphate buffer, pH 7. The pH profile was established at 40 °C in 50 mM potassium buffer (pH 6 to 7.5) or Tris-HCl buffer (pH 7.5 to 9.0). Assays to investigate substrate specificity were carried out at 40 °C, pH 7.5, using 20mM of the substrates adiponitrile, benzonitrile, 3-cyanopyridine, acetonitrile and mandelonitrile. Approximately 25 µg of the purified Nit 1 protein was used for all the assays and enzyme activity was measured

relative to the activity of the purified commercial nitrilase from Codexis (USA) with a reported activity of 43 U/mg at 37 °C and pH 7.5. One unit of nitrilase activity was defined as the amount able to release 1 µmol of ammonia per mg of protein under the assay conditions.

7.2.6 Protein identification

The purified protein fractions were subjected to SDS-PAGE and nanoLC MS/MS analysis as described in Section 4.2.8.2. The spectral data obtained from the mass spectrometric analysis were searched against the in-house created *T. reesei* protein database using the Mascot search engine (Matrix Science, London) as described in Section 4.2.8.3.

7.3 Results

7.3.1 Purification and analysis of the recombinant Nit 1 protein

Origami^{TM}B (DE3) transformants harboring the *nit1* gene were induced for 3 h with 0.4 mM IPTG at 22 °C and the total soluble and insoluble (inclusion body) proteins were extracted from the cells using the solubilization buffer described in Section 7.2.2. The protein was subjected to 1-D SDS-PAGE and the visualization of a band of approximately 40 kDa corresponding to the size of Nit 1 enzyme from *T. reesei* was achieved (data not shown). The His-tagged recombinant Nit 1 enzyme was purified from the total cellular protein using the QIAexpress® System (Section 7.2.3). Analysis of the protein fractions obtained during the purification process was carried out by SDS-PAGE (Figure 7.1).

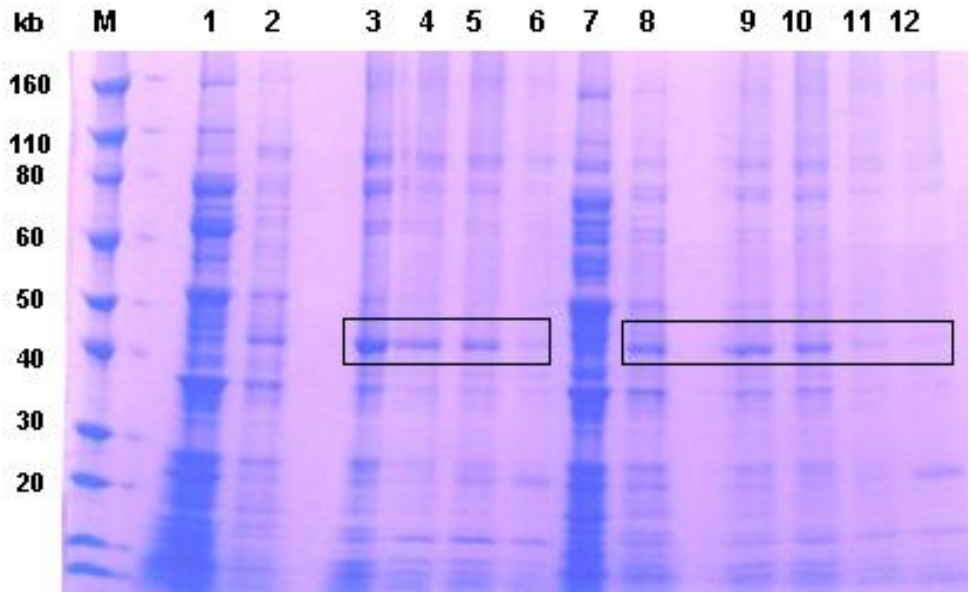


Figure 7.1 SDS-PAGE of protein fractions obtained during the purification of the recombinant Nit 1 protein from the total proteins of Origami™B(DE3) transformants harboring the *nit1* gene, induced with 0.4 mM IPTG for 3 h at 22 °C. Lane M: molecular weight marker; Lanes 1 and 7: total cellular protein prior to purification; Lane 2: flow through; Lanes 3 and 4: protein fractions eluted using elution buffer 1; Lanes 5 and 6: wash fractions; Lanes 8 and 9: protein fractions eluted using elution buffer 2; Lanes 10 to 12: wash fractions. Approximately 20 µl of the protein fractions were loaded per lane. The box indicates proteins of the expected molecular weight of Nit 1.

The SDS-PAGE gel (Figure 7.1) revealed a similar banding pattern of proteins eluted from the column with elution buffer 1 (lanes 3 and 4) and those eluted with elution buffer 2 (lanes 8 and 9), with the most abundant protein in each eluted fraction being of a molecular weight of 40 kDa, presumably representing the Nit 1 nitrilase. The elution buffer 2 is specifically designed to assist the elution of multimeric proteins (Section 7.2.3) and was used in this work as most known nitrilases form multimers (Section 1.3.3). However, slightly more proteins appeared to elute with elution buffer 1 (lanes 3 and 4) than 2 (lanes 8 and 9), and the predominance of the 40 kDa band in both fractions suggested that the Nit 1 protein had maintained a monomeric form at this stage of the analysis. However, as the purification and SDS-PAGE (Section 7.2.3) were carried out under denaturing and reducing conditions, it was possible that the Nit 1 protein had not yet started to form multimers, yet may still do so prior to acquiring an active form following refolding (Section 7.3.3). Therefore, no conclusions

could be made about the formation of a functionally active Nit 1 nitrilase as monomeric or multimeric from this analysis. Research that could enable future structural characterization of the functional Nit 1 protein is discussed in Chapter 8.

The fractions eluted from the column using buffer 1 or 2 (Figure 7.1) contained numerous proteins of various molecular weights other than the prominent 40 kDa protein, expected to represent the Nit1 protein, indicating considerable non-specific binding of non-target proteins to the nickel ions immobilized on the Ni-NTA matrix (Section 7.2.3). Furthermore, a prominent band was also observed at approximately 40 kDa in the gels lanes containing the wash fractions obtained during the Nit 1 purification (lanes 5, 6, 10 and 11, Figure 7.1). This suggested that a considerable amount of Nit 1 recombinant protein had not bound to the Ni-NTA resins, which could have been due to the blocking of the nickel ions on the Ni-NTA matrix by contaminants and native *E. coli* proteins. To reduce the non-specific binding and to increase the protein-Ni-NTA binding, the purification procedure was repeated with minor modifications which included the addition of imidazole at 20 mM concentration to the lysis and wash buffer. To prevent co-purification of host proteins, 20 mM of β-mercaptoethanol was added to the lysis buffer (discussed further in Section 7.4.1). Also, the amount of the Ni-NTA resin added to the protein fractions was increased to 2 ml and the protein-Ni-NTA interaction time was increased from 1 h to 2 h (Section 7.2.3).

Protein fractions from the modified protocol detailed above using only elution buffer 1 were subjected to SDS-PAGE (Figure 7.2). The gel lanes containing eluted protein (lane 3 – 5, Figure 7.2) contained a prominent band at approximately 40 kDa and much lighter bands from proteins of other molecular weights indicated that greater purification of the Nit 1 protein had been achieved with comparatively less non-specific binding than observed previously (Figure 7.1). Furthermore, the intensity of the band at around 40 kDa in the wash fraction was considerably less (lanes 6 – 7, Figure 7.2) which suggested that the increased amounts of resin used for the binding of the Nit1-His protein to the nickel ions and the increased protein-Ni-

NTA interaction time had considerably improved the binding of the His-tagged proteins to the resin.

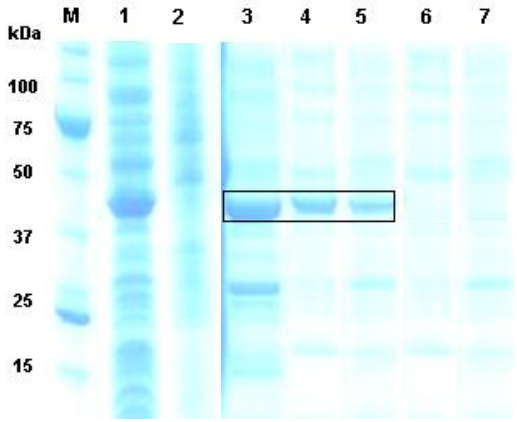


Figure 7.2 SDS-PAGE of protein fractions obtained during the modified procedure for purification of the Nit 1 protein from the total proteins of Origami™B(DE3) transformants. Lane M: molecular weight marker; Lane 1: total cellular protein prior to purification; Lane 2: flow through; Lanes 3 – 5: protein fractions eluted using elution buffer 1 (monomeric proteins); Lanes 6 – 7: wash fractions. Approximately 20 µl of the protein fractions were loaded per lane. The box indicates proteins of the expected molecular weight of Nit 1.

7.3.2 Identification of Nit 1 by mass spectrometry

The purified protein was identified to be the recombinant Nit 1 protein by mass spectrometric analyses performed as described in Section 7.2.6. The search performed by the Mascot software (Matrix Science, London) using the spectral data obtained from the nanoLC MS/MS analyses against the in-house created *T. reesei* database identified a total of 373 spectral matches from 21 peptides unique to the predicted *T. reesei* Nit 1 protein with an overall Mascot score of 5187 which indicated a highly significant match.

7.3.3 Activity of the purified recombinant Nit 1 protein

The purified recombinant Nit 1 eluted using elution buffer 1 was refolded (Section 7.2.4) and the concentration of the refolded protein fraction was determined by Bradford's assay (Bradford, 1976). The activity of the Nit 1 enzyme was determined from the purified protein fractions using the fluorometric nitrilase assay (Section 7.2.5). Since, considerable nitrilase activity was detected towards adiponitrile using the non-purified solubilized cell extracts

(Section 6.3.7.2), the initial assays using the purified recombinant Nit 1 were also carried out using adiponitrile as the substrate. The enzyme activities are represented as percentages relative to the nitrilase enzyme obtained from Codexis Pty Ltd (USA) which had a reported activity of 43 U/mg.

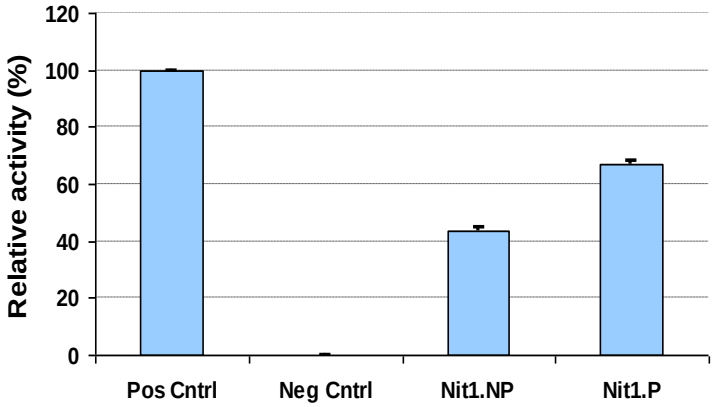


Figure 7.3 Nitrilase activity of the unpurified total protein (insoluble and soluble) of Origami™B(DE3) transformants expressing Nit 1 (Nit1.NP), and the purified Nit 1 enzyme (Nit1.P) relative to a purified commercial nitrilase (Pos Cntrl, Codexis, USA, activity 43 U/mg). The background fluorescence and absorbance of the negative control (Neg Cntrl; protein from cells transformed with vectors without *nit1* cDNA insert) was subtracted from all sample data. The fluorescence/ absorbance exhibited by the chemicals and buffers used in the assays was subtracted from the positive control data. Adiponitrile was used as substrate. Approximately 25 µg of total protein was used for the assays.

The nitrilase activity of the purified recombinant Nit 1 (~ 27 U/mg) was approximately 64% of the positive control (Figure 7.3), substantially higher than the nitrilase activity of the unpurified protein from the Origami transformants (approximately 42%, Section 6.3.7.2). This indicated that the eluted Nit 1 was in an active form and in a significantly more purified state, although total purity had not been achieved as some traces of other proteins in the purified fraction was evident in the SDS-PAGE gel (Figure 7.2).

7.3.4 Temperature and pH profiles of the purified recombinant Nit 1 nitrilase

The temperature and pH profiles of the purified Nit 1 enzyme were determined using fluorometric nitrilase activity assays with adiponitrile as substrate (Figure 7.4). The activity of

Nit 1 at pH 7 across the temperature range 20 °C – 60 °C is represented in Figure 7a as a percentage relative to the activity of the enzyme at 37 °C, pH 7 (the conditions used for all the nitrilase activity assays previously conducted in this work). Assays to determine the pH profile of Nit 1 were carried out at 40 °C, the temperature of optimal activity as determined by the temperature profile of the enzyme. Enzyme activities in Figure 7b are represented as a percentage relative to the activity at pH 7.5, the pH found to be optimal for Nit 1 in this work.

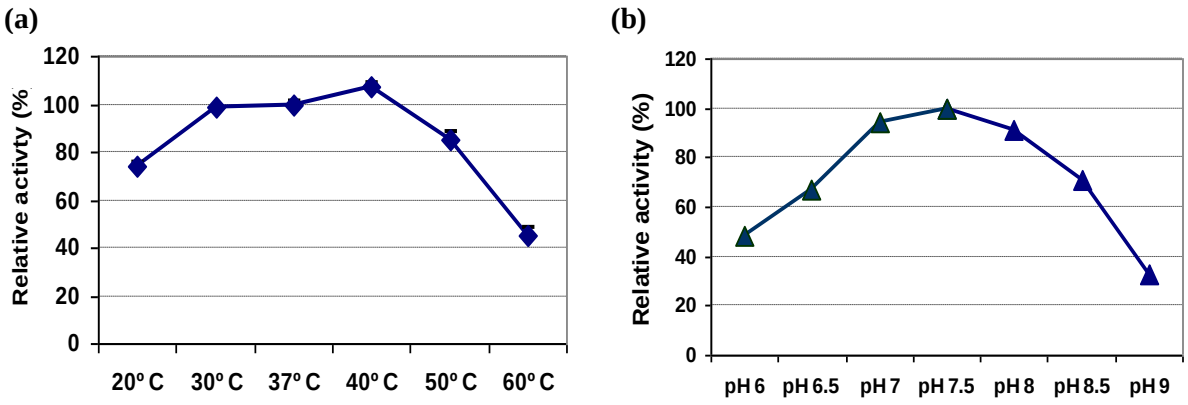


Figure 7.4 Enzyme activity profiles of the purified recombinant Nit 1 nitrilase determined using the fluorometric nitrilase activity assay (Section 7.2.5). (a) Temperature profile at pH 7. Activities are represented as percentages relative to the Nit 1 activity at 37 °C, pH 7. (b) pH profile at 40 °C. Activities are represented as percentages relative to the Nit 1 activity at 40 °C, pH 7.5. The assays were performed using 20 mM adiponitrile as substrate and 25 µg of protein.

The temperature at which the Nit 1 nitrilase exhibited highest activity in the assays was 40 °C. The Nit 1 activity decreased sharply at temperatures above 40 °C and reduced by 50% at 60 °C (Figure 7.4a). However, the activity only dropped to 75% at 20 °C, indicating reasonable activity at mild temperatures. A pH of 7.5 was found to be optimal for the Nit 1 activity (Figure 7.4b), although 90% activity was maintained at pH 7 and pH 8. The optimal temperature and pH range of the recombinant Nit1 enzyme as assessed in the assays was in line with most known fungal and bacterial nitrilases, as discussed further in section 7.4.

7.3.5 Substrate specificity of the purified recombinant Nit 1 nitrilase

The activity of the purified Nit 1 enzyme towards different nitrile substrates as determined by ammonia production in the fluorometric activity assay (Section 1.6.2.2; Section 7.2.5) is

shown in Table 7.1. The activity of the recombinant Nit 1 is represented as a percentage relative to the activity of Nit 1 towards adiponitrile, which based on earlier work (Chapter 4) is a preferred substrate for the enzyme.

Table 7.1 Substrate specificity of the purified recombinant *T. reesei* Nit 1 nitrilase. The activity of Nit 1 against the various nitrile substrates is expressed relative to the activity of Nit 1 towards its preferred substrate adiponitrile which had a specific activity of approximately 27 U/mg (section 7.3.3), determined according to ammonia production detected by the fluorometric assay (Section 7.2.5). Assays were carried out in triplicate and average relative activity is shown.

Substrate	Relative activity (%)
Adiponitrile	100
Benzonitrile	26
3-Cyanopyridine	17
Acetonitrile	18
Mandelonitrile	7

The purified Nit 1 enzyme clearly continued to display highest activity towards the aliphatic nitrile substrate adiponitrile. However, activity towards acetonitrile, the only other aliphatic nitrile included in the assays was only 18% of the activity towards adiponitrile. Activity towards the aromatic nitrile substrates benzonitrile, 3-cyanopyridine and mandelonitrile was only approximately a quarter or less of the activity of the enzyme towards adiponitrile. The activity of the recombinant Nit1 is yet to be tested on arylacetonitriles due to limitation in time and funds for this project.

7.4 Discussion

The expression of the *T. reesei nit1* gene in Origami^{TM}B (DE3) cells was evident in the work described in Chapter 6. Although the recombinant protein aggregated into inclusion bodies, a

functionally active nitrilase could be detected from the insoluble protein in the cells following protein solubilization and refolding (Section 6.3.7.2). In the work described in this chapter, purification of the Nit 1 protein via the attached His-tag enabled subsequent characterization of the recombinant enzyme. The purification technique and the attributes of the nitrilase are discussed below.

7.4.1 Purification of the recombinant Nit 1 enzyme

Polyhistidine tag (His-tag), one of the widely used affinity tags for recombinant protein purification, was used for the purification of the recombinant Nit 1 enzyme from *E. coli*. Although a much more extensive protein purification process involving Q-Sepharose and Sephacryl S-200 columns was used by Kaplan *et al.* (2011) to purify the recombinant *A. niger* nitrilase from *E. coli*, bacterial nitrilases have been successfully purified with a high degree of purity using His-tags (Heinemann *et al*., 2003; Kiziak *et al*., 2005). However, in the purification of Nit 1 described in this chapter, considerable contamination with non-target proteins was initially observed in the "purified" protein fractions eluted from the column using the His-tag purification system (Figure 7.1). This could have been due to non-specific binding of endogenous *E. coli* proteins containing metal binding sites or dispersed histidine residues to the Ni-NTA resin.

The issues of non-specific binding and co-purification of native proteins were addressed by the addition of 20 mM of imidazole and β-mercaptoethanol to the lysis buffer; imidazole was also added to the wash buffer (Section 7.3.1). Imidazole at low concentrations blocks the binding of dispersed histidine residues of native proteins by binding to the nickel ions immobilized in the Ni-NTA resin. Furthermore, co-purification of the endogenous proteins by formation of disulphide bonds with the target protein was prevented by the addition of β-mercaptoethanol (Thompson and O'Donnell, 1961). Even after the modifications to the purification protocol, a certain degree of non-specific binding was seen in the purified protein

fraction although the overall purity of the enzyme had considerably increased (Figure 7.2). Post purification, the protein was refolded to enable characterization of nitrilase activity via enzyme activity assays.

7.4.2 Characterization of the purified recombinant Nit 1 enzyme

The nitrilase activity assays carried out earlier in this work (Chapter 4, Chapter 6) were conducted at 37 °C, pH 7, conditions suitable for the activity of many known nitrilases (Table 1.2, Chapter 1), and also the conditions reportedly used by Codexis to assess the activity of their commercial purified nitrilase which served as a positive control in this work to which the activity of all samples had been compared. In the work described in this chapter, the temperature for optimal activity of the purified recombinant Nit 1 enzyme was established to be 40 °C, and the pH optimum was pH 7.5. Although the temperature and pH optima were slightly higher than those used for the previous assays (Chapter 4, Chapter 6) they were not markedly different, so the conditions used for previous assays should have been sufficient to provide a valid indication of the activity of the Nit 1 nitrilase in the samples at the time.

The temperature and pH optima of the Nit 1 enzyme were similar to that of many known nitrilases of fungi and bacteria (Table 1.2, Chapter 1), including the *A. niger* nitrilase (temperature optimum 45 °C, pH optimum 8) that was identified, purified and expressed recombinantly in *E. coli* by Kaplan *et al*. (2006a; 2011). Temperatures higher than 60 °C were not used for the investigation of the optimum temperature range of the Nit 1 enzyme as most of the known microbial nitrilases were reported to degrade at temperatures of 60 °C and above, which also appeared to be the case with the recombinant Nit 1 enzyme (Figure 7.4). Only a thermostable nitrilase from *P. abyssi* has been reported to have an optimum activity at a temperature of 100 °C (Mueller *et al*., 2006).

The investigation of the substrate specificities of the purified Nit 1 nitrilase (Section 7.3.5), confirmed that adiponitrile was the preferred substrate. Activity of the Nit 1 enzyme towards

any of the other substrates investigated in this work was distinctly lower (26% or less of the activity displayed towards adiponitrile, Table 7.1). This result was in line with the work performed earlier on the total cellular protein extracted from *T. reesei* grown in adiponitrile supplemented minimal medium (Chapter 4). Nitrilase activity of the mycelial protein was detected by HPLC only when using adiponitrile as substrate; no nitrilase activity was evident against benzonitrile, acetonitrile or 3-cyanopyridine (Section 4.3.3). Therefore, the Nit 1 enzyme appeared to have the same substrate specificity for adiponitrile when expressed by its natural host *T. reesei* or when expressed recombinantly in *E. coli.*

Adiponitrile is an aliphatic nitrile which is produced globally for the synthesis of adipic acid, an important aliphatic dicarboxylic acid, and hexamethylenediamine, both of which are used as precursors in the manufacture of nylon 66 and different resins (Musser, 2000). Since adipic acid occurs rarely in nature and there is an annual world-wide demand for approximately 2.5 billion kilograms of adipic acid annually, hydrolysis of adiponitrile to adipic acid is an industrially significant reaction. Besides being used in the production of nylon 66, adipic acid is also used in the manufacture of polymers such as polyurethane and hydrophilic drugs where it assists in the drug's pH controlled sustained release (Musser, 2000). Although, presently adipic acid is manufactured by a multistep reaction involving the nitric acid oxidation of cyclohexanol, cyclohexanone or a mixture of different ketone-alcohol oil, the ability of the Nit 1 enzyme to hydrolyze adiponitrile into adipic acid without the formation of the adipamide by-product (Section 4.3.3) could potentially provide a much greener alternative for adipic acid formation in the future.

The low activities of the Nit 1 nitrilase towards the other nitrile substrates tested in this work indicates that the enzyme does not have the broad substrate specificity that has been reported of other characterized fungal nitrilases (Martinkova *et al*., 2009; Winkler *et al*., 2009; Petrickova *et al*., 2012). Whereas activity towards adiponitrile suggests characterization of Nit 1 as an aliphatic nitrilase, activity against the aliphatic acetonitrile was substantially less

(18%, Table 7.1) and even lower than the activity towards the aromatic nitriles benzonitrile and 3-cyanopyridine. The low activity of the enzyme towards nitrile substrates other than adiponitrile may limit the applications for which the enzyme could be appropriate. Advanced protein engineering techniques and gene site saturation mutagenesis (GSSM) techniques have been used to improve both the substrate specificity and the enantio-selectivity of nitrilase enzymes (DeSantis *et al*., 2003; Xie *et al*., 2006; Kiziak *et al*., 2009) and will be discussed in Chapter 8.

7.5 Chapter summary

Purification of the His-tagged recombinant Nit 1 nitrilase from Origami^{TM}B (DE3) cells harboring the *T. reesi nit1* gene was achieved by Ni-NTA affinity chromatography. Initial purification attempts were moderately successful, although some non-specific binding to the Ni-NTA and co-elution of native *E. coli* proteins occurred. Addition of imidazole and β-mercaptoethanol to purification buffers reduced the incidence of the non-target proteins in the purified fractions and a more highly purified Nit 1 protein was attained. Nitrilase activity of the refolded purified Nit 1 was approximately 64% of that of a commercially purified nitrilase. The recombinant *T. reesei* nitrilase had a temperature optimum of 40 °C and pH optimum of 7.5, similar to the temperature and pH optima of many known nitrilases of fungal and bacterial origin. The activity of the recombinant nitrilase was highest against the aliphatic nitrile substrate adiponitrile, which could make it of interest for further characterization, improvement and possible application in industry *e.g.* for the production of nylon and resins. Suggestions for future research and development of the *T. reesei* nitrilase are discussed in Chapter 8.

Chapter 8: Thesis summary and future directions

Nitrilases are of great interest for the biocatalysis of nitriles in the manufacture of many fine chemicals and pharmaceuticals, and in biodegradation of toxic nitriles in the environment. Most of the nitrilases studied and utilized to date are of bacterial origin. However, a few fungal nitrilases have also been characterized and have exhibited a broad substrate range, high stability and chemo- and enantioselectivity (Section 1.4.3.1). With this as background, the research presented in the thesis explored *Trichoderma reesei*, an industrially important mesophilic filamentous fungus, as a source of nitrilases. The major outcomes of the work are summarized below with suggestions for future work.

8.1 Identification of putative nitrilase genes in the *T. reesei* genome

The genome sequence of the *T. reesei* wild-type strain QM6a provided the basis for *in silico* analysis to search for putative nitrilase genes. Predicted *T. reesei* proteins translated from the genome sequence were searched for amino acid sequence similarities to known enzymes of the nitrilase superfamily from other organisms (Chapter 3). A total of twelve predicted *T. reesei* proteins had sequence similarity to enzymes in the nitrilase superfamily (Table 3.1), and four had sequence similarities to enzymes of the nitrilase branch (nitrilases, cyanide hydratases and cyanide dihydratases). Further examination of the amino acid sequences of the predicted *T. reesei* enzymes was carried out by phylogenetic analysis, multiple sequence alignments and predicted structure alignments with known nitrilases, revealing that the putative *nit1* gene was the most promising candidate for nitrilase expression.

8.2 Endogenous expression of nitrilase(s) by *T. reesei*

To investigate expression of the putative nitrilase gene(s), the well characterized and high secreting mutant strain *T. reesei* Rut-C30 was cultivated in various media containing nitriles (adiponitrile, benzonitrile, acetonitrile and 3 cyanopyridine; Chapter 4). The predicted

nitrilase(s) were expected to be intracellular proteins based on *in silico* analysis, hence the mycelial protein was subjected to proteomic and enzyme analysis. Identification of the predicted Nit 1 protein was achieved by nanoLC ESI MS/MS analysis of spots from a 2D gel of mycelial protein of *T. reesei* Rut-C30 cultivated in minimal medium supplemented with adiponitrile; nitrilase activity of the mycelial protein against the substrate adiponitrile was verified by HPLC.

Since the broader aim of this project was the recombinant production of the *T. reesei* nitrilase in *E. coli* (Section 8.3), the next stage of the research was to amplify the cDNA of the nitrilase encoding gene by RT-PCR (Chapter 5). Northern blot analysis of RNA extracted from *T. reesei* Rut-C30 grown in the various nitrile-supplemented media revealed transcription of the *nit1* gene only when the fungus was grown in media supplemented with adiponitrile, which was in line with the proteomic and enzyme analysis carried out previously; however, *in vitro* reverse transcription of the RNA obtained from the same mycelia could not be achieved. Instead, the *nit1* cDNA was obtained by RT-PCR from the RNA of *T. reesei* Rut-C30 grown in nutrient rich "CLS" growth medium that did not contain nitriles. This suggested that the nitrilase was expressed constitutively, at least at a basal level, in *T. reesei*. Moreover, the healthier and quicker growth that *T. reesei* exhibited in the CLS medium may have allowed the subsequent isolation of higher quality RNA and concomitant success in obtaining cDNA via RT-PCR.

Sequencing of the *nit1* cDNA revealed an additional 10 nucleotides that were not expected from the predicted intron/ exon assignment of the *T. reesei nit1* gene in the JGI database. The 10 nucleotides originated from a small section of the first predicted intron in the gene as defined by JGI. The presence of these nucleotides in the cDNA obtained in this work revealed that they in fact represented an additional exon that had not previously been identified. Furthermore, the additional amino acid acids tryptophan (W), isoleucine (I), proline (P) and glycine (G) encoded by the newly discovered exon represent a highly conserved motif in

many known nitrilases, thus increasing the sequence similarity of the *T. reesei* predicted Nit 1 protein to known nitrilases above that which was predicted by the original JGI annotation of the gene sequence.

8.3 Recombinant expression of the *T. reesei nit 1* nitrilase gene in *E.coli*

To enable recombinant expression of the *T. reesei nit1* gene in *E. coli*, the amplified *nit1* cDNA was inserted into the pETDuet vector. Expression of the *nit1* gene was attempted in Origami™B(DE3) and Tuner™DE3 competent cells, which are known to boost the solubility of recombinant proteins. Cultivation of the transformants was carried out at both 37 and 22 °C in case the lower temperature could also enhance protein solubility. However, the recombinant Nit 1 protein could not be detected in the soluble protein of any of the *E. coli* transformants by proteomic analysis or activity based assays.

Northern blots and Dot blots of total RNA indicated transcription of *nit1* in the Origami™B(DE3) transformants grown at 22 °C, despite the inability to detect the Nit 1 protein in the soluble protein of these cells. Hence, nanoLC ESI MS/MS was used to analyze the insoluble protein of the Origami™B(DE3) transformant and the Nit 1 protein was identified. Furthermore, the effect of modifications in the IPTG concentrations on Nit 1 production was investigated wherein the production of the nitrilase enzyme was improved when the *E. coli* cells were induced with 0.4 mM IPTG. The solubilized protein fractions from the cells were refolded by dialysis and nitrilase activity was verified by fluorometric and colorimetric activity assays.

8.4 Purification of the recombinant Nit 1 nitrilase

The recombinant Nit 1 enzyme was purified from Origami™B (DE3) transformants harboring the pET-Nit1 plasmid using the His-tag protein purification system. Although, some non-specific binding and co-elution of native *E. coli* proteins was evident in the purified protein

fractions during initial purification attempts, a more highly purified Nit 1 was obtained by addition of imidazole and β-mercaptoethanol to the buffers. The activity of the purified recombinant Nit 1 protein was approximately 64% of the activity of a commercial purified nitirilase. The Nit 1 enzyme displayed optimal activity at 40 °C, pH 7.5, which was in line with many known microbial nitrilases. Determination of the substrate specificity of the recombinant Nit 1 enzyme revealed adiponitrile, an aliphatic nitrile, to be the most favoured substrate. The biocatalysis of adiponitrile is an industrially relevant mechanism in the manufacture of nylons and resins, making the *T. reesei* nitrilase of potential interest for future development and application in this area.

8.5 Future work

8.5.1 Reducing Nit 1 aggregation into inclusion bodies in *E. coli*

In the work performed by Kaplan *et al*. (2011) on recombinant expression of an *A. niger* nitrilase in *E. coli*, the recombinant nitrilase was reported to constitute almost half of the soluble proteins in the cell. However in this project, although steps were taken to enhance the solubility of the recombinant Nit 1 protein, the enzyme was still found to aggregate in inclusion bodies. Reducing the incidence of Nit 1 protein aggregation in *E. coli* would simplify the purification process and decrease losses of the recombinant protein during solubilization and refolding procedures.

Presently, there are several commercial expression systems available that employ fusion tags for enhanced production and solubility of recombinant proteins (*e.g.* pET NusA, pET Trx and pET Dsd; Novagen, USA). Furthermore, mutant *E. coli* strains such as C41(DE3) and C43(DE3), commercialized by Avidis (France), are known to assist in the expression of recombinant proteins in their soluble form. Also modifications to cultivation parameters such as growth pH, nutrients, dissolved oxygen and simultaneous over-expression of molecular chaperones are reported to prevent inclusion body formation (Sorensen and Mortensen, 2005).

8.5.2 Enhancing the Nit 1 nitrilase for future applications

Enzymes usually require improvement before their implementation on an industrial scale. The Nit 1 nitrilase exhibited optimal enzyme activity in relatively mild conditions (40 °C, pH 7.5). However, many industrial processes require enzymes to be active and stable under more unusual conditions *e.g.* high temperature and acidic/ alkaline. Some strategies to improve enzyme properties include rational protein design and directed evolution (Bornscheuer and Pohl, 2001; Hibbert and Dalby, 2005). Knowledge of protein structure and function can greatly enhance the capabilities of protein engineering. Since Nit 1 shares considerable sequence similarities with other known microbial nitrilases, a computational model of the Nit 1 protein structure could be created by protein threading which can further assist in the determination of the Nit 1 structure by X-ray crystallography. Elucidation of the Nit 1 structure could enable a better understanding of formation of oligomers and the positioning of the protein folds which might provide a possible explanation for the specificity of the Nit 1 enzyme towards adiponitrile. Furthermore, advanced protein engineering techniques and gene site saturation mutagenesis (GSSM) have been used previously to improve both the substrate specificity and the enantio-selectivity of the nitrilase enzymes (DeSantis *et al*., 2003; Kiziak and Stolz, 2009) and similarly could be applied to the Nit 1 protein.

8.5.3 Recombinant expression of Nit 1 in *T. reesei* Rut-C30

Expression of the *T. reesei* Nit 1 nitrilase in *E. coli* enabled characterization of the recombinant protein. However, large scale industrial production of recombinant proteins is typically carried out in yeast or filamentous fungi. Filamentous fungi are particularly attractive hosts due to their high secretion capacity; genes can be recombinantly expressed using an endogenous promoter and secretion signal, thereby enabling recovery of the target protein from the culture supernatant rather than requiring cell lysis. *Trichoderma reesei* Rut-C30 is a high secreting mutant strain that is widely used for the production of homologous

enzymes (chiefly cellulases) and various heterologous gene products. Therefore, the strain would form an excellent host for large scale recombinant expression of the homologous nitrilase in the future. To this end, vectors containing the *nit1* gene under the control of the strong inducible promoter and secretion signal of the main expressed cellobiohydrolase 1 (Harkki *et al*., 1991; de Faria *et al*., 2002) could be prepared for transformation into *T. reesei* Rut-C30 by protoplast (Penttilä *et al*., 1987) or biolistic bombardment (Te'o *et al*., 2002). Expression could also be attempted using constitutive promoters, such as the *pdc* (pyruvate decarboxylase) promoter that has recently been reported to produce increased amounts of recombinant proteins when compared to the *cbh1* promoter (Li *et al*., 2012).

8.6 Concluding remarks

In this work, the identification of a nitrilase gene in an industrially utilized filamentous fungus *T. reesei* Rut-C30 was achieved. The gene was found to encode a functionally active nitrilase with optimal activity at 40 °C, pH 7.5, and a preferred substrate of adiponitrile. The recombinant expression of the nitrilase gene in *E. coli* described in this work paves the way for future research in which the nitrilase could be further characterized, improved and developed for industrial applications. The project highlights the value of exploring well-known and currently utilized microoganisms as a source of novel enzymes.

References

Almatawah, Q. A., Cowan, D. A. 1999. Thermostable nitrilase catalysed production of nicotinic acid from 3-cyanopyridine. *Enzyme and Microbial Technology*, *25*, 718-724.

Almatawah, Q. A., Cramp, R., Cowan, D. A. 1999. Characterization of an inducible nitrilase from a thermophilic bacillus. *Extremophiles*, *3*, 283-291.

Andrade, J., Karmali, A., Carrondo, M. A., Frazão, C. 2007. Structure of amidase from *Pseudomonas aeruginosa* showing a trapped acyl transfer reaction intermediate state. *Journal of Biological Chemistry*, *282*, 19598-19605.

Arnold, D. 1995. Virus safety considerations for recombinant factor VIII (rFVIII, Kogenate). *Haemophilia*, *1*, 22-23.

Aro, N., Ilmén, M., Saloheimo, A., Penttilä, M. 2003. ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. *Applied and Environmental Microbiology*, *69*, 56-65.

Aro, N., Saloheimo, A., Ilmén, M., Penttilä, M. 2001. ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. *Journal of Biological Chemistry*, *276*, 24309-24314.

Arvas, M., Pakula, T., Lanthaler, K., Saloheimo, M., 2006. Common features and interesting differences in transcriptional responses to secretion stress in the fungi *Trichoderma reesei* and *Saccharomyces cerevisiae*. *BMC Genomics*, *7*, 32.

Arvas, M., Pakula, T., Smit, B., Rautio, J., Koivistoinen, H., Jouhten, P., Lindfors, E., Wiebe, M., Penttilä, M., Saloheimo, M. 2011. Correlation of gene expression and protein production rate - a system wide study. *BMC Genomics*, *12*, 616.

Asano, Y. 2002. Overview of screening for new microbial catalysts and their uses in organic synthesis, selection and optimization of biocatalysts. *Journal of Biotechnology*, *94*, 65-72.

Bailey, M. J., Nevalainen, K. M. H. 1981. Induction, isolation and testing of stable *Trichoderma reesei* mutants with improved production of solubilizing cellulase. *Enzyme and Microbial Technology*, *3*, 153-157.

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

Banerjee, A., Kaul, P., Banerjee, U. 2006. Enhancing the catalytic potential of nitrilase from *Pseudomonas putida* for stereoselective nitrile hydrolysis. *Applied Microbiology and Biotechnology*, *72*, 77-87.

Banerjee, A., Kaul, P., Sharma, R., Banerjee, U. C. 2003a. A high-throughput amenable colorimetric assay for enantioselective screening of nitrilase-producing microorganisms using pH sensitive indicators. *Journal of Biomolecular Screening*, *8*, 559-565.

Banerjee, A., Sharma, R., Banerjee, U. C. 2003b. A rapid and sensitive fluorometric assay method for the determination of nitrilase activity. *Biotechnology and Applied Biochemistry*,

37, 289-293.

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

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

Bartlett, J. S., Stirling, D., 2003. *PCR Protocols*, Humana Press, Totowa, NJ, pp. 3-6.

Bartling, D., Seedorf, M., Mithöfer, A., Weiler, E. W. 1992. Cloning and expression of an *Arabidopsis* nitrilase which can convert indole-3-acetonitrile to the plant hormone, indole-3-acetic acid. *European Journal of Biochemistry, 205*, 417-424.

Benz, P., Muntwyler, R., Wohlgemuth, R. 2007. Chemoenzymatic synthesis of chiral carboxylic acids via nitriles. *Journal of Chemical Technology and Biotechnology, 82*, 1087-1098.

Berbasov, D., Ellis, T.K., Soloshonok, V.A., in: Eusebio Juaristi, V. A. S. (Ed.), 2005. Enantioselective synthesis of beta- amino acids, Wiley- Interscience, New York, pp. 397-414.

Bhalla, T., Miura, A., Wakamoto, A., Ohba, Y., Furuhashi, K. 1992. Asymmetric hydrolysis of aminonitriles to optically active amino acids by a nitrilase of *Rhodococcus rhodochrous* PA-34. *Applied Microbiology and Biotechnology*, *37*, 184-190.

Bornscheuer, U.T., Pohl, M. 2001. Improved biocatalysts by directed evolution and rational protein design. *Current Opinion in Chemical Biology, 5*, 137-143.

Bork, P., Koonin, E. V. 1994. A new family of carbon-nitrogen hydrolases. *Protein Science, 3*, 1344-1346.

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

Braud, S., Moutiez, M., Belin, P., Abello, N., Drevet, P., Zinn-Justin., S., Courçon, M., Masson, C, Dassa., J, Charbonnier, J.B., Boulain, J.C., Ménez, A., Genet, R., Gondry, M. 2005. Dual expression system suitable for high-throughput fluorescence-based screening and production of soluble proteins. *Journal of Proteome Research, 4*, 2137-2147.

Brennan, B. A., Cummings, J. G., Chase, D. B., Turner, I. M., Nelson, M. J. 1996. Resonance Raman spectroscopy of nitrile hydratase, a novel iron-sulfur enzyme *Biochemistry, 35*, 10068-10077.

Breuer, M., Ditrich, K., Habicher, T., Hauer, B., Kesseler, M., Stürmer, R., Zelinski, T. 2004. Industrial methods for the production of optically active intermediates. *Angewandte Chemie International Edition, 43*, 788-824.

Burgess, C., O'Connell-Motherway, M., Sybesma, W., Hugenholtz, J., van Sinderen, D. 2004. Riboflavin production in *Lactococcus lactis*: Potential for in situ production of vitamin-enriched foods. *Applied and Environmental Microbiology, 70*, 5769-5777.

Cardamone, M., Puri, N. K., Brandon, M. R. 1995. Comparing the refolding and reoxidation of recombinant porcine growth hormone from a urea denatured state and from *Escherichia coli* Inclusion Bodies. *Biochemistry, 34,* 5773-5794.

Carrez, D., Janssens, W., Degrave, P., van den Hondel, C. A., Kinghorn, J.R, Fiers, W., Contreras, R. 1990. Heterologous gene expression by filamentous fungi: secretion of human interleukin-6 by *Aspergillus nidulans*. *Gene, 94,* 147-54.

Carter, G. L., Allison, D., Rey, M. W., Dunn-Coleman, N. S. 1992. Chromosomal and genetic analysis of the electrophoretic karyotype of *Trichoderma reesei*: mapping of the cellulase and xylanase genes. *Molecular Microbiology, 6,* 2167-2174.

Chae, H. J., Delisa, M. P., Cha, H. J., Weigand, W. A., Rao, G., Bentley, W.E. 2000. Framework for online optimization of recombinant protein expression in high-cell-density *Escherichia coli* cultures using GFP-fusion monitoring. *Biotechnology and Bioengineering, 69,* 275-285.

Chauhan, S., Wu, S., Blumerman, S., Fallon, R. D., Gavagan, J. E., DiCosimo, R., Payne, M.S. 2003. Purification, cloning, sequencing and over-expression in *Escherichia coli* of a regioselective aliphatic nitrilase from *Acidovorax facilis* 72W. *Applied Microbiology and Biotechnology, 61,* 118-122.

Chen, H., Xu, Z., Xu, N., Cen, P. 2005. Efficient production of a soluble fusion protein containing human beta-defensin-2 in *E. coli* cell-free system. *Journal of Biotechnology, 115,* 307-315.

Cherry, J. R., Fidantsef, A. L. 2003. Directed evolution of industrial enzymes: an update. *Current Opinion in Biotechnology, 14,* 438-443.

Chin, K.-H., Tsai, Y.-D., Chan, N.-L., Huang, K.-F., Wang, A. H., Chou, S. H. 2007. The crystal structure of XC1258 from *Xanthomonas campestris*: A putative procaryotic Nit protein with an arsenic adduct in the active site. *Proteins: Structure, Function, and Bioinformatics, 69,* 665-671.

Cho, C., Park, S., Nam, D. 2001. Production and purification of single chain human insulin precursors with various fusion peptides. *Biotechnology and Bioprocess Engineering, 6,* 144-149.

Clarke, H. T., Read, R. R. 1941. o-Tolunitrile and p-Tolunitrile. *Organic Syntheses, 1,* 514.

Collén, A., Saloheimo, M., Bailey, M., Penttilä, M., Pakula, T. M. 2005. Protein production and induction of the unfolded protein response in *Trichoderma reesei* strain *Rut-C30* and its transformant expressing endoglucanase I with a hydrophobic tag. *Biotechnology and Bioengineering, 89,* 335-344.

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

Dadd, M. R., Sharp, D. C., Pettman, A. J., Knowles, C. J. 2000. Real-time monitoring of nitrile biotransformations by mid-infrared spectroscopy. *Journal of Microbiology Methods, 41,* 69-75.

de Faria, F. P., Te'o, V. S. J., Bergquist, P. L., Azevedo, M. O., Nevalainen, K. M. H. 2002. Expression and processing of a major xylanase (XYN2) from the thermophilic fungus *Humicola grisea* var. *thermoidea* in *Trichoderma reesei*. *Letters in Applied Microbiology, 34*, 119-123.

de Marco, A. 2007. Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*. *Naure Protocols, 2*, 2632-2639.

Demain, A. L., 2007. The business of biotechnology. *Industrial Biotechnology*, GEN Publishing, Inc., New York, pp. 269-283.

Demolder, J. 1999. KEX2-like processing of glucoamylase-interleukin 6 and cellobiohydrolase-interleukin 6 fusion proteins by *Trichoderma reesei*. Abstract B38, 2nd European Conference on Fungal Genetics (ECFG2), Lunteren.

DeSantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z., Tomlinson, G., Huang, H., Tan, X., Bibbs, L., Chen, P., Kretz, K., Burk, M. J. 2003. Creation of a productive, highly enantioselective nitrilase through Gene Site Saturation Mutagenesis (GSSM). *Journal of the American Chemical Society, 125*, 11476-11477.

DeSantis, G., Zhu, Z., Greenberg, W. A., Wong, K., Chaplin, J., Hanson, S. R., Farwell, B., Nicholson, L. W., Rand, C. L., Weiner, D. P., Robertson, D. E., Burk, M. J. 2002. An enzyme library approach to biocatalysis: development of nitrilases for enantioselective production of carboxylic acid derivatives. *Journal of the American Chemical Society, 124*, 9024-9025.

DesRoches, C. L., 2012. Investigating the genes and pathways involved in the biosynthesis of indole-3-acetic acid in Fusarium graminearum. PhD thesis. *Department of Biological science*, University of Ottawa, Ottawa.

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

Dias, J. C. T., Rezende, R. P., Rosa, C. A., Lachance, M. A., Linardi, V. R. 2000.Enzymatic degradation of nitriles by a *Candida guilliermondii* UFMG-Y65. *Canadian Journal of Microbiology, 46*, 525-531.

Domsch, K. H., Gams, W., Anderson, T. H. 1980. *Compendium of soil fungi*, Vol 1, Academic Press, London.

Doyle, W. A., Smith, A., T 1996. Expression of lignin peroxidase H8 in *Escherichia coli*: folding and activation of the recombinant enzyme with Ca2+ and haem. *Biochemical Journal, 315*, 15-19.

Druzhinina, I. S., Kopchinskiy, A. G., Kubicek, C. P. 2006. The first 100 *Trichoderma* species characterized by molecular data. *Mycoscience, 47*, 55-64.

Druzhinina, I. S., Shelest, E., Kubicek, C. P. 2012. Novel traits of *Trichoderma* predicted through the analysis of its secretome. *FEMS Microbiology Letters, 27*,1574-6968.

Dubendorf, J. W., Studier, F. W. 1991. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of*

Molecular Biology, *219*, 45-59.

Duke, S. O. 2005. Taking stock of herbicide-resistant crops ten years after introduction. *Pest Management Science*, *61*, 211-218.

Durand, H., Clanet, M., Tiraby, G. r. 1988.Genetic improvement of *Trichoderma reesei* for large scale cellulase production. *Enzyme and Microbial Technology*, *10*, 341-346.

Dyer, W. E., Hess, F. D., Holt, J. S., Duke, S. O., 1993. *Horticultural Reviews*, John Wiley & Sons, Inc., New Jersey, pp. 367-408.

Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., McCafferty, J. 2004. Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *Biomedcentral Biotechnology*, *4*, 32.

Effenberger, F., Bohme, J. 1994. Enzyme-catalysed enantioselective hydrolysis of racemic naproxen nitrile. *Bioorganic and Medicinal Chemistry*, *2*, 715-721.

Elliott, S., Egrie, J., Browne, J., Lorenzini, T., Busse, L., Rogers, N., Ponting, I. 2004. Control of rHuEPO biological activity: The role of carbohydrate. *Experimental Hematology*, *32*, 1146-1155.

Fernandez-Castane, A., Caminal, G., Lopez-Santin, J. 2012. Direct measurements of IPTG enable analysis of the induction behavior of *E. coli* in high cell density cultures. *Microbial Cell Factories*, *11*, 58.

Ferrer-Miralles, N., Domingo-Espin, J., Corchero, J. L., Vazquez, E., Villaverde, A. 2009. Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, *8*, 17.

Fleming, F. 1999. Nitrile-containing natural products. *Natural Product Reports*, *16*, 597-606.

Forciniti, D. 1994. Protein refolding using aqueous two-phase systems. *Journal of Chromatography*, *A668*, 95-100.

Fowler, T., Berka, R. M., Ward, M. 1990. Regulation of the glaA gene of *Aspergillus niger*. *Current Genetics*, *18*, 537-545.

Freedonia, 2009. ReportLinker http://www.reportlinker.com/p0148002-summary/World-Enzymes-Market.html.

Freeman, W. M., Walker, S. J., Vrana, K. E. 1999. Quantitative RT-PCR: pitfalls and potential. *Biotechniques*, *26*, 112-122.

Freyssinet, G., Pelissier, B., Freyssinet, M., Delon, R. 1996. Crops resistant to oxynils: from the laboratory to the market. *Field Crops Research*, *45*, 125-133.

Gams, W., Bissett, J., in: Kubicek, C. P., Harman, G. E. (Eds.) 1998. *Trichoderma and Gliocladium*, Taylor & Francis Ltd., London, pp. 3-34.

García-Estrada, C., Fierro, F., Martín, J. F., in: Mendez-Vilas, A. (Ed.) 2010. Current Research, Technology and Education. *Topics in Applied Microbiology and Microbial Biotechnology*, Formatex, Badajoz, pp. 577-588.

Gavagan, J. E., DiCosimo, R., Eisenberg, A., Fager, S. K., Folsom, P.W., Hann, E.C., Schneider, R., Fallon, D. 1999. A Gram-negative bacterium producing a heat-stable nitrilase highly active on aliphatic dinitriles. *Applied Microbiology and Biotechnology*, *52*, 654-659.

Ghosh, A., Al-Rabiai, S., Ghosh, B. K.1982. Increased endoplasmic reticulum content of a mutant of *Trichoderma reesei* (RUT-C30) in relation to cellulase synthesis. *Enzyme and Microbial Technology*, *4*, 110-113.

Ghosh, A., Ghosh, B. K., Trimino-Vazquez, H., Eveleigh, D. E., Montenecourt, B. S. 1984. Cellulase secretion from a hyper-cellulolytic mutant of *Trichoderma reesei Rut-C30*. *Archives of Microbiology*, *140*, 126-133.

Glenn, M., Ghosh, A., Ghosh, B. K. 1985. Subcellular fractionation of a hypercellulolytic mutant, *Trichoderma reesei Rut-C30*: localization of endoglucanase in microsomal fraction. *Applied and Environmental Microbiology*, *50*, 1137-1143.

Godtfredsen, S. E., Ingvorsen, K., Yde, B., Anderson, O. 1985. *Biocatalysis in Organic Synthesis*, Elsevier, Amsterdam.

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

Gong, J. S., Lu, Z. M., Li, H., Shi, J. S., Zhou, Z. M., Xu, Z. H. 2012. Nitrilases in nitrile biocatalysis: recent progress and forthcoming research. *Microbial Cell Factories*, *11*, 142.

Gouka, R. J., Punt, P. J., van den Hondel, C. A. 1997. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Applied Microbiol Biotechnology*, *47*, 1-11.

Grigoriev, N. H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., Kuo, A., Minovitsky, S., Nikitin, R., Ohm, R. A., Otillar, R., Poliakov, A., Ratnere, I., Riley, R., Smirnova, T., Rokhsar, D., Dubchak, I. 2012. The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Research*, *40*, 26-32.

Guise, A., West, S., Chaudhuri, J. 1996. Protein folding in vivo and renaturation of recombinant proteins from inclusion bodies. *Molecular Biotechnology*, *6*, 53-64.

Gupta, V., Gaind, S., Verma, P. K., Sood, N., Srivastava, A. K. 2010. Purification and characterization of intracellular nitrilases from *Rhodococcus* sp. - potential role of periplasmic nitrilase. *African Journal of Microbiology Research*, *4*, 1148-1153.

Harkki, A., Mäntylä, A., Penttilä, M., Muttilainen, S., Bühler, R., Suominen, P., Knowles, J., Nevalainen, H. 1991. Genetic engineering of *Trichoderma* to produce strains with novel cellulase profiles. *Enzyme Microbial Technology*, *3*, 227-33.

Harkki, A., Uusitalo, J., Bailey, M., Penttilä, M., Knowles, J. K. C. 1989. A novel fungal expression system: secretion of active calf chymosin from the filamentous fungus *Trichoderma reesei*. *Nature Biotechnology*, *7*, 596-603.

Harper, D. B. 1976. Purification and properties of an unusual nitrilase from *Nocardia* N.C.I.B. 11216. *Biochemical Society Transactions*, *4*, 502-504.

Harper, D. B. 1977a. Fungal degradation of aromatic nitriles. Enzymology of C-N cleavage by *Fusarium solani*. *Biochemical Journal*, *167*, 685-692.

Harper, D. B. 1977b. Microbial metabolism of aromatic nitriles. Enzymology of C-N cleavage by *Nocardia* sp. (*Rhodochrous* group) N.C.I.B. 11216. *Biochemical Journal*, *165*, 309-319.

Harper, D. B. 1985. Characterization of a nitrilase from *Nocardia* sp. (*Rhodochrous* group) N.C.I.B. 11215, using p-hydroxybenzonitrile as sole carbon source. *International Journal of Biochemistry*, *17*, 677-683.

He, Y.-C., Xu, J.-H., Su, J.-H., Zhou, L. 2010. Bioproduction of glycolic acid from glycolonitrile with a new bacterial isolate of *Alcaligenes* sp. ECU0401. *Applied Biochemistry and Biotechnology*, *160*, 1428-1440.

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

Hewitt, L., McDonnell, J. M., 2004. Screening and optimizing protein production in *E. coli*, *Methods in Molecular Biology*, *278*, 1-16.

Hibbert E. G., Dalby P. A. 2005. Directed evolution strategies for improved enzymatic performance. *Microbial Cell Factories*, *4*, 29.

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

Hoyle, A. J., Bunch, A. W., Knowles, C. J. 1998.The nitrilases of *Rhodococcus rhodochrous* NCIMB 11216. *Enzyme and Microbial Technology*, *23*, 475-482.

Hung, C. L., Liu, J. H., Chiu, W. C., Huang, S. W., Hwang, J. K, Wang, W. C. 2007. Crystal structure of *Helicobacter pylori* formamidase AmiF reveals a cysteine-glutamate-lysine catalytic triad. *Journal of Biological chemistry*, *282*, 12220-12229.

Ilmen, M., Saloheimo, A., Onnela, M. L., Penttilä, M. E. 1997. Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. *Applied and Environmental Microbiology*, *63*, 1298-1306.

Ilmén, M., Onnela, M. L., Klemsdal, S., Keränen, S., Penttilä, M. 1996.Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus *Trichoderma reesei*. *Molecular genetics and genomics*, *253*, 303-314.

Ilmén, M., Thrane, C., Penttilä, M. 1996. The glucose repressor gene cre1 of *Trichoderma*: isolation and expression of a full-length and a truncated mutant form. *Molecular Genetics and Genomics*, *251*, 451-60.

Inoue, H., Nojima, H., Okayama, H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, *96*, 23-28.

Jain, D., Meena, V. S., Kaushik, S., Kamble, A., Chisti, Y., Banerjee, U.C. 2012. Production of nitrilase by a recombinant *Escherichia coli* in a laboratory scale bioreactor. *Fermentation*

Technology, *1*, 1-4.

Jallageas, J. C., Arnaud, A., Galzy, P. 1979. Nitrilases and amidases: Determination of activity by proton magnetic resonance spectrometry. *Analytical Biochemistry*, *95*, 436-443.

Jensen, S. E., Campbell, J. N. 1976. Amidase activity involved in peptidoglycan biosynthesis in membranes of *Micrococcus luteus* (*sodonensis*). *Journal of Bacteriology*, *127*, 319-326.

Johnson, D. V., Zabelinskaja-Mackova, A. A., Griengl, H. 2000. Oxynitrilases for asymmetric C-C bond formation. *Current Opinion in Chemical Biology*, *4*, 103-109.

Jones, D. T., Taylor, W. R., Thornton, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. *Computer applications in the biosciences : CABIOS*, *8*, 275-282.

Jouanneau, J. P., Lapous, D., Guern, J. 1991. In plant protoplasts, the spontaneous expression of defense reactions and the responsiveness to exogenous elicitors are under auxin control. *Plant Physiology*, *96*, 459-466.

Kabaivanova, L., Dobreva, E., Dimitrov, P., Emanuilova, E. 2005. Immobilization of cells with nitrilase activity from a thermophilic bacterial strain. *Journal of Industrial Microbiology and Biotechnology*, *32*, 7-11.

Kang, S., Kang, K., Lee, K., Back, K. 2007.Characterization of rice tryptophan decarboxylases and their direct involvement in serotonin biosynthesis in transgenic rice. *Planta*, *227*, 263-272.

Kaplan, O., Bezouska, K., Malandra, A., Vesela, A. B., Petříčková, A., Felsberg, J., Rinágelová, A., Křen, V., Martínková, L. 2011. Genome mining for the discovery of new nitrilases in filamentous fungi. *Biotechnology Letters*, *33*, 309-312.

Kaplan, O., Bezouska, K., Plihal, O., Ettrich, R., Kulik, N., Vaněk, O., Kavan, D., Benada, O., Malandra, A., Sveda, O., Veselá, A.B., Rinágelová, A., Slámová, K., Cantarella, M.,Felsberg, J., Dušková, J., Dohnálek, J., Kotik, M., Křen, V., Martínková, L. 2011. Heterologous expression, purification and characterization of nitrilase from *Aspergillus niger* K10. *Biomedcentral Biotechnology*, *6*, 11.

Kaplan, O., Vejvoda, V., Plihal, O., Pompach, P., Kavan, D., Bojarová, P., Bezouska, K., Macková, M., Cantarella, M., Jirků, V., Kren, V., Martínková, L. 2006a. Purification and characterization of a nitrilase from *Aspergillus niger* K10. *Applied Microbiology and Biotechnology*, *73*, 567-575.

Kaplan, O., Nikolaou, K., Charvátová-Pisvejcová, A., Martínková, L. 2006b. Hydrolysis of nitriles and amides by filamentous fungi. *Enzyme and Microbial Technology*, *38*, 260-264.

Kaplan, O., Vejvoda, V., Charvátová-Pisvejcová, A., Martínková, L. 2006c. Hyperinduction of nitrilases in filamentous fungi. *Journal of Industrial Microbiology and Biotechnology*, *33*, 891-896.

Kato, Y., Nakamura, K., Sakiyama, H., Mayhew, S. G., Asano, Y. 2000. Novel heme-containing lyase, phenylacetaldoxime dehydratase from *Bacillus* sp. Strain OxB-1: purification, characterization, and molecular cloning of the gene. *Biochemistry*, *39*, 800-809.

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

Kautto, L., Grinyer, J., Paulsen, I., Tetu, S., Pillai, A., Pardiwalla, S., Sezerman, U., Akcapinar, G. B., Bergquist, P., Te'o, J., Nevalainen, H. 2012. Stress effects caused by the expression of a mutant cellobiohydrolase I and proteasome inhibition in *Trichoderma reesei* Rut-C30. *New Biotechnology*, *30*, 183-191.

Kigawa, T., Yabuki, T., Matsuda, N., Matsuda, T., Nakajima, R., Tanaka, A., Yokoyama, S. 2004. Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *Journal of Structural and Functional Genomics*, *1*, 63-68.

Kim, J. S., Tiwari, M. K., Moon, H. J., Jeya, M., Ramu, T., Oh, D. K., Kim, I. W., Lee, J. K. 2009. Identification and characterization of a novel nitrilase from *Pseudomonas fluorescens* Pf-5. *Applied Microbiology and Biotechnology*, *83*, 273-283.

Kimani, S. W., Agarkar, V. B., Cowan, D. A., Sayed, M. F., Sewell, B. T. 2007. Structure of an aliphatic amidase from *Geobacillus pallidus* RAPc8. *Acta Crystallographica Section D Biological Crystallography*, *63*,1048-1058.

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

Kiziak, C., Klein, J., Stolz, A. 2007. Influence of different carboxy-terminal mutations on the substrate-, reaction- and enantiospecificity of the arylacetonitrilase from *Pseudomonas fluorescens* EBC191. *Protein Engineering Design and Selection*, *20*, 385-396.

Kiziak, C., Stolz, A. 2009. Identification of amino acid residues responsible for the enantioselectivity and amide formation capacity of the arylacetonitrilase from *Pseudomonas fluorescens* EBC191. *Applied and Environmental Microbiology*, *75*, 5592-5599.

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

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

Komeda, H., Hori, Y., Kobayashi, M., Shimizu, S. 1996. Transcriptional regulation of the *Rhodococcus rhodochrous* J1 nitA gene encoding a nitrilase. *Proceedings of the National Academy of Sciences*, *93*, 10572-10577.

Komeda, T., Sakai, Y., Kato, N., Kondo, K. 2002. Construction of protease-deficient Candida boidinii strains useful for recombinant protein production: cloning and disruption of proteinase A gene (PEP4) and proteinase B gene (PRBI). *Bioscience Biotechnology and Biochemistry*, *66*, 628-631.

Kopetzki, E., Schumacher, G., Buckel, P. 1989. Control of formation of active soluble or inactive insoluble baker's yeast a-glucosidase PI in *Escherichia coli* by induction and growth

conditions. *Molecular Genomics and Genetics, 216,* 149-155.

Kou, G., Shi, S., Wang, H., Tan, M., Xue, J., Zhang, D., Hou, S., Qian, W., Wang, S., Dai, J., Li, B., Guo, Y. 2007. Preparation and characterization of recombinant protein ScFv(CD11c)-TRP2 for tumor therapy from inclusion bodies in *Escherichia coli*. *Protein Expression and Purification, 52,* 131-138.

Kriechbaumer, V., Park, W. J., Piotrowski, M., Meeley, R. B., Gierl, A., Glawischnig, E. 2007. Maize nitrilases have a dual role in auxin homeostasis and beta-cyanoalanine hydrolysis. *Journal of Experimental Botany, 58,* 4225-4233.

Kubicek, C., Mikus, M., Schuster, A., Schmoll, M., Seiboth, B. 2009. Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina*. *Biotechnology for Biofuels, 2,* 19.

Kuhls, K., Lieckfeldt, E., Samuels, G. J., Kovacs, W., Meyer, W., Petrini, O., Gams, W., Börner, T., Kubicek, C.P. 1996. Molecular evidence that the asexual industrial fungus *Trichoderma reesei* is a clonal derivative of the ascomycete *Hypocrea jecorina*. *Proceedings of National Academy of Science U S A, 93,* 7755-7760.

Kurzatkowski, W., Torronen, A., Filipek, J., Mach, R. L., Herzog, P., Sowka, S., Kubicek, C.P. 1996. Glucose-induced secretion of *Trichoderma reesei* xylanases. *Applied Environmental Microbiology, 62,* 2859-65.

Layh, N., Parratt, J., Willetts, A. 1998. Characterization and partial purification of an enantioselective arylacetonitrilase from *Pseudomonas fluorescens* DSM 7155. *Journal of Molecular Catalysis B: Enzymatic, 5,* 467-474.

Laymon, R., Adney, W., Mohagheghi, A., Himmel, M., Thomas, S. 1996. Cloning and expression of full-length *Trichoderma reesei* cellobiohydrolase I cDNAs in *Escherichia coli*. *Applied Biochemistry and Biotechnology, 57-58,* 389-397.

Levy-Schil, S., Soubrier, F., Coq, A.-M. C.-L., Faucher, D., Crouzet, J., Pétré, D. 1995. Aliphatic nitrilase from a soil-isolated *Comamonas testosteroni* sp.: gene cloning and overexpression, purification and primary structure. *Gene, 161,* 15-20.

Li, J., Wang, J., Wang, S., Xing, M., Yu, S., Liu, G. 2012. Achieving efficient protein expression in *Trichoderma reesei* by using strong constitutive promoters. *Microbial Cell Factories, 11,* 84.

Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J., Tan, S. 2005. Comparison of affinity tags for protein purification. *Protein Expression and Purification, 41,* 98-105.

Liljeblad, A., Kanerva, L. T. 2006. Biocatalysis as a profound tool in the preparation of highly enantiopure Î²-amino acids. *Tetrahedron, 62,* 5831-5854.

Linardi, V. R., Dias, J. C. T., Rosa, C. A. 1996. Utilization of acetonitrile and other aliphatic nitriles by a *Candida famata* strain. *FEMS Microbiology Letters., 144,* 67-71.

Liu, T., Wang, T., Li, X., Liu, X. 2008. Improved heterologous gene expression in *Trichoderma reesei* by cellobiohydrolase I gene (*cbh1*) promoter optimization. *Acta Biochim*

Biophys Sin (Shanghai), *40*, 158-165.

Lundberg, K. S., Shoemaker, D. D., Adams, M. W. W., Short, J. M., Sorge, J. A., Mathur, E.J. 1991. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene*, *108*, 1-6.

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

Luo, H., Fan, L., Chang, Y., Ma, J., Yu, H., Shen, Z. 2010. Gene Cloning, Overexpression, and Characterization of the Nitrilase from *Rhodococcus rhodochrous* tg1-A6 in *E. coli*. *Applied Biochemistry and Biotechnology*, *160*, 393-400.

Luria, S. E., Burrous, J. W. 1957. Hybridization between *Eschericia coli* and *Shigella*. *Journal of Bacteriology*, *74*, 461-476.

Ma, B., Mayfield, M. B., Gold, M. H. 2001. The green fluorescent protein gene functions as a reporter of gene expression in *Phanerochaete chrysosporium*. *Applied and Environmental Microbiology*, *67*, 948-955.

Ma, L., Zhang, J., Zou, G., Wang, C., Zhou, Z. 2011. Improvement of cellulase activity in *Trichoderma reesei* by heterologous expression of a beta-glucosidase gene from *Penicillium decumbens*. *Enzyme and Microbial Technology*, *49*, 366-371.

Macauley-Patrick, S., Fazenda, M. L., McNeil, B., Harvey, L. M. 2005. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, *22*, 249-270.

Mach, R. L., Zeilinger, S. 2003. Regulation of gene expression in industrial fungi: *Trichoderma* . *Applied Microbiology and Biotechnology*, *60*, 515-522.

Mahadevan, R., Doyle, F. J. I. 2003. On-line optimization of recombinant product in a fed-batch bioreactor. *Biotechnology Progress*, *19*, 639-646.

Mandels, M., Parrish, F. W., Reese, E. T. 1962. Sophorose as an inducer of cellulose in *Trichoderma viride*. *Journal of Bacteriology*, *83*, 400-408.

Manulis, S., Haviv-Chesner, A., Brandl, M. T., Lindow, S. E., Barash, I. 1998. Differential Involvement of Indole-3-Acetic Acid Biosynthetic Pathways in Pathogenicity and Epiphytic Fitness of *Erwinia herbicola* pv. *gypsophilae*. *Molecular Plant-Microbe Interactions*, *11*, 634-642.

Maor, R., Haskin, S., Levi-Kedmi, H., Sharon, A. 2004. In Planta Production of Indole-3-Acetic Acid by *Colletotrichum gloeosporioides f. sp. aeschynomene*. *Applied and Environmental Microbiology*, *70*, 1852-1854.

Margolles-clark, E., Ilmén, M., Penttilä, M. 1997. Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. *Journal of Biotechnology*, *57*, 167-179.

Marron, A. O., Akam, M., Walker, G. 2012. Nitrile Hydratase Genes Are Present in Multiple Eukaryotic Supergroups. *PLoS ONE*, *7*, e32867.

Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., Chapman, J., Chertkov, O., Coutinho, P.M., Cullen, D., Danchin, E.G., Grigoriev, I.V., Harris, P., Jackson, M., Kubicek, C. P., Han, C. S., Ho, I., Larrondo, L. F., de Leon, A. L., Magnuson, J. K., Merino, S., Misra, M., Nelson, B., Putnam, N., Robbertse, B., Salamov, A.A., Schmoll, M.,Terry, A., Thayer, N., Westerholm-Parvinen, A., Schoch, C. L., Yao, J., Barabote, R., Nelson, M. A., Detter, C., Bruce, D., Kuske, C.R., Xie, G., Richardson, P., Rokhsar, D.S., Lucas, S.M., Rubin, E.M., Dunn-Coleman, N., Ward, M., Brettin, T.S. 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nature Biotechnology, 26*, 553-560.

Martinkova, L., Kaplan, V. 2010. Biotransformations with nitrilases. *Current Opinion in Chemical Biology*, *14*, 130-137.

Martinkova, L., Vejvoda, V., Kaplan, O., Kubac, D., Malandra, A., Cantarella, M., Bezouska, K., Kren, V. 2009. Fungal nitrilases as biocatalysts: Recent developments. *Biotechnology Advances*, *27*, 661-670.

Martinkova, L., Vejvoda, V., Kren, V. 2008. Selection and screening for enzymes of nitrile metabolism. *Journal of Biotechnology*, *133*, 318-236.

Mateo, C., Chmura, A., Rustler, S., van Rantwijk, F., Stolz, A., Sheldon, R.A. 2006. Synthesis of enantiomerically pure (S)-mandelic acid using an oxynitrilase-nitrilase bienzymatic cascade: a nitrilase surprisingly shows nitrile hydratase activity. *Tetrahedron: Asymmetry*, *17*, 320-323.

Mathew, C. D., Nagasawa, T., Kobayashi, M., Yamada, H. 1988. Nitrilase-catalyzed production of nicotinic acid from 3-cyanopyridine in *Rhodococcus rhodochrous* J1. *Applied Environmental Microbiology, 54*, 1030-1032.

Mathur, E. J., Toledo, G., Green, B. D., Podar, M., Richardson, T.H., Kulwiec, M., Chang, H.W. 2005. A biodiversity-based approach to development of performance enzymes. *Industrial Biotechnology*, GEN Publishing Inc, New York, pp. 283-287.

Maurer, A. J., in: Nadeau, J. (Ed.) 2012. *Introduction to experimental biophysics: Biological methods for physical scientists*, CRC Press, Boca Raton, FL, pp. 137-146.

Maurizi, M. R. 1992. Proteases and protein degradation in *Escherichia coli*. *Experientia*, *48*, 178-201.

Merten, O. W., Mattanovich, D., Cole, J., Lang, C., C., Larsson, G., Neubauer, P., Porro, D., Postma, P., Teixeira de Mattos, J., in: Mattanovich, D., Lang, C., Larsson, G. (Eds.), 2001. *Production of recombinant proteins with prokaryotic and eukaryotic cells*, Kluwer Academic Publishers, New York, pp. 339-346.

Mertens, N., Remaut, E., Fiers, W. 1995. Tight transcriptional control mechanism ensures stable high-level expression from T7 promoter-based expression plasmids. *Biotechnology (N Y)*, *13*, 175-179.

Middelberg, A. P. J. 1996. Large-scale recovery of recombinant protein inclusion bodies expressed in *Escherichia coli Journal of Microbiology*, *6*, 225-231.

Miettinen-Oinonen, A., Suominen, P. 2002. Enhanced production of *Trichoderma reesei*

endoglucanases and use of the new cellulase preparations in producing the stonewashed effect on denim fabric. *Applied and Environmental Microbiology*, *68*, 3956-3964.

Minshull, J., Ness, J. E., Gustafsson, C., Govindarajan, S. 2005. Predicting enzyme function from protein sequence. *Current Opinion in Chemical Biology*. *9*, 202-209.

Mitraki, A., King, J. 1989. Protein folding intermediates and inclusion body formation. *Nature Biotechnology*, *7*, 690-697.

Molins-Legua, C., Meseguer-Lloret, S., Moliner-Martinez, Y., Campins-Falcŏ, P. 2006. A guide for selecting the most appropriate method for ammonium determination in water analysis. *TrAC Trends in Analytical Chemistry*, *25*, 282-290.

Montenecourt, B. S., Eveleigh, D. E. in: Brown, R., Jurasek, L. (Eds.) 1979. Production and characterization of high yielding cellulase mutants of *Trichoderma reesei*. *Advanced Chemistry Society*, ACSSymposium, pp. 289-301.

Montenecourt, B. S., Nhlapo, S. D., Trimino-Vazquez, H., Cuskey, S., Schamhart, D. H., Eveleigh, D. E. 1981. Regulatory controls in relation to over-production of fungal cellulases. *Basic Life Sciences*, *18*, 33-53.

Mueller, P., Egorova, K., Vorgias, C. E., Boutou, E., Trauthwein, H., Verseck, S., Antranikian, G. 2006. Cloning, overexpression, and characterization of a thermoactive nitrilase from the hyperthermophilic archaeon *Pyrococcus abyssi. Protein Expression and Purification*, *47*, 672-681.

Murakami, T., Nojiri, M., Nakayama, H., Odaka, M., Yohda, M., Dohmae, N., Takio, K., Nagamune, T., Endo, I. 2000. Post-translational modification is essential for catalytic activity of nitrile hydratase. *Protein Science*, *9*, 1024-1030.

Murphy, R. A., Horgan, K. A. 2005. *Fungi*. John Wiley & Sons, Ltd, New York, pp. 113-143.

Musser, M. T. 2000. *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH Verlag GmbH and Co., KGaA, Weinheim.

Mylerova, V., Martinkova, L. 2003. Synthetic Applications of Nitrile-Converting Enzymes. *Current Organic Chemistry*, *7*, 1279-1295.

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

Nagasawa, T., Yamada, H. 1990. Large-scale bioconversion of nitriles into useful amides and acids in: Abramowicz, D. A. (Ed.), *Biocatalysis*, Van Nostrand Reinhold, New York, pp. 277-318.

Nagasawa, T., Yamada, H. 1995. Microbial production of commodity chemicals. *Pure and Applied Chemistry*, *67*, 1241-1256.

Nakai, T., Hasegawa, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Ueki, T., Nanba, H., Ikenaka, Y., Takahashi, S., Sato, M., Tsukihara, T. 2000. Crystal structure of N-carbamyl-D-

amino acid amidohydrolase with a novel catalytic framework common to amidohydrolases. *Structure*, *8*, 729-737.

Nakari-Setälä, T., Paloheimo, M., Kallio, J., Saloheimo, M. 2004. Genetic modification of carbon catabolite repression in the filamentous fungus *Trichoderma reesei* for improved protein production. *7th European Conference on Fungal Genetics*, Copenhagen, Denmark, 17 - 20 April. 133.

Nageshwar, Y. V. D., Sheelu, G., Shambhu, R. R., Muluka, H., Mehdi, N., Malik, M. S., Kamal, A. 2011. Optimization of nitrilase production from *Alcaligenes faecalis* MTCC 10757 (IICT-A3): effect of inducers on substrate specificity. *Bioprocess and Biosystems Engineering*, *34*, 515-523.

Nevalainen, K. M. H., Palva, E. T., Bailey, M. J. 1980. A high cellulase-producing mutant strain of *Trichoderma reesei*. *Enzyme and Microbial Technology*, *2*, 59-60.

Nevalainen, H. 1985 *Genetic Improvement of Enzyme Production in Industrially Important Fungal Strains*, Valtion Teknillinen Tutkimuskeskus, Technical Research Centre of Finland, Espoo Finland.

Nevalainen, H., Suominen, P., Taimisto, K. 1994. On the safety of *Trichoderma reesei*. *Journal of Biotechnology*, *37*, 193-200.

Nevalainen, K. M. H., Te'o, V. S. J. 2003. Enzyme production in industrial fungi- molecular genetic strategies for integrated strain improvement, in Dilip, K. A., George, G. K. (Ed.) *Applied Mycology and Biotechnology*, Elsevier, New York, pp. 241-259.

Nevalainen, K. M. H., Te'o, V. S. J., Bergquist, P. L. 2005. Heterologous protein expression in filamentous fungi. *Trends in Biotechnology*, *23*, 468-474.

Nishise, H., Kurihara, M., Tani, Y. 1987.Microbial synthesis of a tranexamic acid intermediate from dinitrile. *Agricultural and Biological Chemistry*, *51*, 2613-2616.

Novy, R., Morris, B., 2001. Use of glucose to control basal expression in the pET system *Innovations*, *13*, 8-10.

Nunberg, J. H., Meade, J. H., Cole, G., Lawyer, F. C., McCabe, P., Schweickart, V., Tal, R., Wittman, V.P., Flatgaard, J.E., Innis, M.A. 1984. Molecular cloning and characterization of the glucoamylase gene of *Aspergillus awamori*. *Molecular and Cellular Biology*, *4*, 2306-2315.

Nykänen, M. J., 2002. Protein secretion in *Trichoderma reesei*. Expression, secretion and maturation of cellobiohydrolase I, barley cysteine proteinase and calf thymosin in Rut-C30. PhD Thesis. University of Jyväskylä, Finland. pp. 107.

Nyyssonen, E., Penttila, M., Harkki, A., Saloheimo, A., Knowles, J. K., Keränen, S. 1993. Efficient production of antibody fragments by the filamentous fungus *Trichoderma reesei*. *Biotechnology (N Y)*, *11*, 591-595.

Olempska-Beer, Z. S., Merker, R. I., Ditto, M. D., DiNovi, M. J. 2006. Food-processing enzymes from recombinant microorganisms- a review. *Regulatory Toxicology and Pharmacology*, *45*, 144-158.

O'Reilly, C., Turner, P. D. 2003. The nitrilase family of CN hydrolysing enzymes - a comparative study. *Journal of Applied Microbiology*, *95*, 1161-1174.

Orgogozo, V., Rockman, M. V., Nelson, M., Fitch, D. A., 2011. Overlap Extension PCR: An Efficient Method for Transgene Construction, *Molecular Methods for Evolutionary Genetics*, Humana Press, New York, pp. 459-470.

Pace, H. C., Brenner, C. 2001. The nitrilase superfamily: classification, structure and function. *Genome Biology*, *2*, 1.

Paloheimo, M., Mäntylä, A., Kallio, J., Suominen, P. 2003. High-yield production of a bacterial xylanase in the filamentous fungus *Trichoderma reesei* requires a carrier polypeptide with an intact domain structure. *Applied Environmental Microbiology*, *69*, 7073-7082.

Paloheimo, M., Miettinen-Oinonen, A., Torkkeli, T., Nevalainen, H., Suominen, P., in: Suominen, P., T Reinikainen (Eds.), 1993. Enzyme production by *Trichoderma* using the cbh1 promoter. *Tricel 93 Symposium Workshop*, Finland, pp. 229-238.

Patel, R., Hanson, R., Goswami, A., Nanduri, V., Banerjee, A., Donovan, M. J., Goldberg, S., Johnston, R., Brzozowski, D., Tully, T., Howell, J., Cazzulino, D., Ko, R. 2003. Enzymatic synthesis of chiral intermediates for pharmaceuticals. *Journal of Industrial Microbiology & Biotechnology*, *30*, 252-259.

Patricelli, M. P., Cravatt, B. F. 2000. Clarifying the catalytic roles of conserved residues in the amidase signature family. *Journal of Biology Chemistry*, *275*, 19177-19184.

Patten, C. L., Glick, B. R. 1996. Bacterial biosynthesis of indole-3-acetic acid. *Canadian Journal of Microbiology*, *42*, 207-220.

Penttilä, M., Nevalainen, H., Rättö, M., Salminen, E., Knowles, J. 1987. A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene*, *61*, 155-164.

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, *20*, 3551-3567.

Peterson, R., Nevalainen, H. 2012. *Trichoderma reesei* RUT-C30- thirty years of strain improvement. *Microbiology*, *158*, 58-68.

Petersson, L., Carrio, M. M., Vera, A., Villaverde, A. 2004. The impact of dnaKJ overexpression on recombinant protein solubility results from antagonistic effects on the control of protein quality. *Biotechnology Letters*, *26*, 595-601.

Petrickova, A., Vesela, A. B., Kaplan, O., Kubac, D., Uhnáková, B., Malandra, A., Felsberg, J., Rinágelová, A., Weyrauch, P., Křen, V., Bezouška, K., Martínková, L. 2012. Purification and characterization of heterologously expressed nitrilases from filamentous fungi. *Applied Microbiology and Biotechnology*, *93*, 1553-1561.

Piotrowski, M. 2008. Primary or secondary? Versatile nitrilases in plant metabolism. *Phytochemistry*, *69*, 2655-2667.

Piotrowski, M., Schönfelder, S., Weiler, E.W. 2001. The *Arabidopsis thaliana* isogene NIT4 and its orthologs in tobacco encode Î²-Cyano-l-alanine Hydratase/Nitrilase. *Journal of Biological Chemistry, 276,* 2616-2621.

Podar, M., Eads, J. R., Richardson, T. H. 2005. Evolution of a microbial nitrilase gene family: a comparative and environmental genomics study. *BMC Evolution Biology, 5,* 42.

Polizeli, M. L., Rizzatti, A. C., Monti, R., Terenzi, H. F., Jorge, J. A., Amorim, D. S. 2005. Xylanases from fungi: properties and industrial applications. *Applied Microbiology and Biotechnology, 67,* 577-591.

Pollak, P., Romeder, G., Hagedorn, F., Gelbke, H.-P. 2000. Nitriles. *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim.

Prasad, S., Misra, A., Jangir, V., Awasthi, A., Raj, J., Bhalla, T. 2007. A propionitrile-induced nitrilase of *Rhodococcus* sp. NDB 1165 and its application in nicotinic acid synthesis. *World Journal of Microbiology and Biotechnology, 23,* 345-353.

Preiml, M., Hillmayer, K., Klempier, N. 2003. A new approach to β-amino acids: biotransformation of N-protected β-amino nitriles. *Tetrahedron Letters, 44,* 5057-5059.

Prinz, W. A., Åslund, F., Holmgren, A., Beckwith, J. 1997. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *Journal of Biological Chemistry, 272,* 15661-15667.

Punt, P. J., Dingemanse, M. A., Kuyvenhoven, A., Soede, R. D. M., Pouwels, P. H., van den Hondel, C. A. M. J. J. 1990. Functional elements in the promoter region of the *Aspergillus nidulans* gpdA gene encoding glyceraldehyde-3-phosphate dehydrogenase. *Gene, 93,* 101-109.

Punt, P. J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J., van den Hondel, C. 2002. Filamentous fungi as cell factories for heterologous protein production. *Trends in Biotechnology, 20,* 200-206.

Punt, P. J., Veldhuisen, G., van den Hondel, C. A. 1994. Protein targeting and secretion in filamentous fungi. A progress report. *Antonie Van Leeuwenhoek, 65,* 211-216.

Raczynska, J. E., Vorgias, C. E., Antranikian, G., Rypniewski, W. 2011. Crystallographic analysis of a thermoactive nitrilase. *Journal of Structural Biology, 173,* 294-302.

Rajan, S. S., Lackland, H., Stein, S., Denhardt, D. T. 1998. Presence of an N-terminal polyhistidine tag facilitates stable expression of an otherwise unstable N-terminal domain of mouse tissue inhibitor of metalloproteinase-1 in *Escherichia coli*. *Protein Expression and Purification, 13,* 67-72.

Rausch, T., Hilgenberg, W. 1980. Partial purification of nitrilase from Chinese cabbage. *Phytochemistry, 19,* 747-750.

Rawlings, N. D., Barrett, A. J. 2000. MEROPS: the peptidase database. *Nucleic Acids Research, 28,* 323-325.

Riley, G. L., Tucker, K. G., Paul, G. C., Thomas, C. R. 2000. Effect of biomass concentration

and mycelial morphology on fermentation broth rheology. *Biotechnology and Bioengineering*, *68*, 160-172.

Robertson, D. E., Chaplin, J. A., DeSantis, G., Podar, M., Madden, M., Chi, E., Richardson, T., Milan, A., Miller, M., Weiner, D. P., Wong, K., McQuaid, J., Farwell, B., Preston, L. A., Tan, X., Snead, M. A., Keller, M., Mathur, E., Kretz, P. L., Burk, M. J., Short, J. M. 2004. Exploring Nitrilase Sequence Space for Enantioselective Catalysis. *Applied and Environmental Microbiology*, *70*, 2429-2436.

Robinette, D., Matthysse, A. G. 1990. Inhibition by *Agrobacterium tumefaciens* and *Pseudomonas savastanoi* of development of the hypersensitive response elicited by *Pseudomonas syringae pv. phaseolicola*. *Journal of Bacteriology*, *172*, 5742-5749.

Rustler, S., Müller, A., Windeisen, V., Chmura, A., Fernandes, B. C. M., Kiziak, C., Stolz, A. 2007. Conversion of mandelonitrile and phenylglycinenitrile by recombinant *E. coli* cells synthesizing a nitrilase from *Pseudomonas fluorescens* EBC191. *Enzyme and Microbial Technology*, *40*, 598-606.

Rustler, S., Motejadded, H., Altenbuchner, J., Stolz, A. 2008. Simultaneous expression of an arylacetonitrilase from *Pseudomonas fluorescens* and a (S)-oxynitrilase from *Manihot esculenta* in *Pichia pastoris* for the synthesis of (S)-mandelic acid. *Applied Microbiology and Biotechnology*, *80*, 87-97.

Rustler, S., Stolz, A. 2007. Isolation and characterization of a nitrile hydrolysing acidotolerant black yeast-*Exophiala oligosperma* R1. *Applied Microbiology and Biotechnology*, *75*, 899-908.

Saaranen, M. J., Ruddock, L. W. 2012. Disulfide Bond Formation in the Cytoplasm. *Antioxididants and Redox Signaling*, *19*, 36-43

Sambrook, J., Russell, D. W., 2001. Molecular Cloning- A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.

Schmidt, R. C., Muller, A., Hain, R., Bartling, D., Weiler, E. W. 1996. Transgenic tobacco plants expressing the *Arabidopsis thaliana* nitrilase II enzyme. *The Plant Journal*, *9*, 683-691.

Schuster, A., Schmoll, M. 2010. Biology and biotechnology of *Trichoderma*. *Applied Microbiology and Biotechnology* , *87*, 787-799.

Seffernick, J. L., Samanta, S. K., Louie, T. M., Wackett, L. P., Subramanian, M. 2009. Investigative mining of sequence data for novel enzymes: a case study with nitrilases. *Journal of Biotechnology*, *143*, 17-26.

Seidl, V., Gamauf, C., Druzhinina, I., Seiboth, B., Hartl, L., Kubicek, C. 2008. The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85 kb (29 gene-encoding) region of the wild-type genome. *BMC Genomics*, *9*, 327.

Sewell, B. T., Thuku, R. N., Zhang, X., Benedik, M. J. 2005. Oligomeric Structure of Nitrilases: Effect of Mutating Interfacial Residues on Activity. *Annals of the New York Academy of Sciences*, *1056*, 153-159.

Sharma, N. N., Sharma, M., Bhalla, T. C. 2012. *Nocardia globerula* NHB-2 nitrilase

catalysed biotransformation of 4-cyanopyridine to isonicotinic acid. *AMB Express*, *2*, 25.

Shoemaker, S., Schweickart, V., Ladner, M., Gelfand, D., Kwok, S., Myambo, K., Innis, M. 1983. Molecular Cloning of Exo-Cellobiohydrolase I Derived from *Trichoderma reesei* Strain L27. *Nature Biotechnology*, *1*, 691-696.

Simossis, V. A., Heringa, J. 2003. The PRALINE online server: optimising progressive multiple alignment on the web. *Computational Biology and Chemistry*, *27*, 511-519.

Singh, R., Sharma, R., Tewari, N., Rawat, D. S. 2006. Nitrilase and its application as a 'green' catalyst. *Chemistry and Biodiversity*, *12*, 1279-87.

Snajdrova, R., Kristovä-Mylerovä, V., Crestia, D., Nikolaou, K., Kuzma, M., Lemaire, M., Galliene, E., Bolte, J., Bezouška, K., Křen, V., Martínková, L. 2004. Nitrile biotransformation by *Aspergillus niger*. *Journal of Molecular Catalysis B: Enzymatic*, *29*, 227-232.

Sorensen, H. P., Mortensen, K. K. 2005. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories*, *4*, 1.

Stalker, D. M., Malyj, L. D., McBride, K. E. 1988a. Purification and properties of a nitrilase specific for the herbicide bromoxynil and corresponding nucleotide sequence analysis of the bxn gene. *Journal of Biological Chemistry*, *263*, 6310-6314.

Stalker, D. M., McBride, K. E., Malyj, L. D. 1988b. Herbicide resistance in transgenic plants expressing a bacterial detoxification gene. *Science*, *242*, 419-423.

Steinborn, G., Boer, E., Scholz, A., Tag, K., Kunze, G., Gellissen, G. 2006. Application of a wide-range yeast vector (CoMedTM) system to recombinant protein production in dimorphic *Arxula adeninivorans*, methylotrophic *Hansenula polymorpha* and other yeasts. *Microbial Cell Factories*, *5*, 33.

Stevenson, D. E., Feng, R., Dumas, F., Groleau, D., Mihoc, A., Storer, A. C. 1992. Mechanistic and structural studies on *Rhodococcus* ATCC 39484 nitrilase. *Biotechnology and Applied Biochemistry*, *15*, 283-302.

Stevenson, D. E., Feng, R., Storer, A. C. 1990. Detection of covalent enzyme-substrate complexes of nitrilase by ion-spray mass spectroscopy. *FEBS Letters, 277*,112-114.

Stewart, E. J., Aslund, F., Beckwith, J. 1998. Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO Journal*, *17*, 5543-5550.

Studier, F., Rosenberg, A., Dunn, J., Dubendorff, J. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymology*, *189*, 60-89.

Studier, F. W., Moffatt, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, *189*, 113-130.

Swofford, D. L., 1991. PAUP: Phylogenetic Analysis Using Parsimony. Computer program and manual distributed by the Center for Biodiversity, Illinois Natural History Survey, Champaign, IL.

Takashima, S., Iikura, H., Nakamura, A., Masaki, H., Uozumi, T. 1996. Analysis of Cre1

binding sites in the *Trichoderma reesei* cbh1 upstream region. *FEMS Microbiology Letters*, *145*, 361-366.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, *28*, 2731-2739.

Tang, W., Sun, Z.-Y., Pannell, R., Gurewich, V., Liu, J.-N. 1997. An efficient system for production of recombinant urokinase-type plasminogen activator. *Protein Expression and Purification*, *11*, 279-283.

Teeri, T., Salovuori, I., Knowles, J. 1983. The molecular cloning of the major cellulase gene from *Trichoderma reesei*. *Nature Biotechnology*, *1*, 696-699.

Temin, H. M., Mizutani, S. 1970. RNA-dependent DNA polymerase in virions of Rous *sarcoma* virus. *Nature 226*, 1211-1213.

Te'o, V. J. S., Nevalainen, H. 2009. Multiple promoter platform for protein production. Patent number H WO/2009/076709.

Te'o, V. S. J., Bergquist, P. L., Nevalainen, K. M. H. 2002. Biolistic transformation of *Trichoderma reesei* using the Bio-Rad seven barrels Hepta Adaptor system. *Journal of Microbiological Methods*, *51*, 393-399.

Teter, S. A., Cherry, J. R.,2005. *American Institute of Chemical Engineers (AIChE) Annual Meeting Conference Proceedings*, Cincinnati, USA, pp. 12027-12033.

Thanka Christlet, T. H., Veluraja, K. 2001. Database analysis of O-glycosylation sites in proteins. *Biophysics Journal*, *80*, 952-960.

Thimann, K. V., Mahadevan, S. 1964. Nitrilase: I. Occurrence, preparation, and general properties of the enzyme. *Archives of Biochemistry and Biophysics*, *105*, 133-141.

Thompson, E. O., O'Donnell, I. J. 1961. Quantitative reduction of disulphide bonds in proteins using high concentrations of mercaptoethanol. *Biochimica et Biophysica Acta*, *53*, 447-449.

Thompson, J. D., Higgins, D. G., Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*, 4673-4680.

Thuku, R. N., Brady, D., Benedik, M. J., Sewell, B. T. 2009. Microbial nitrilases: versatile, spiral forming, industrial enzymes. *Journal of Applied Microbiology*, *106*, 703-727.

Uusitalo, J. M., Helena Nevalainen, K. M., Harkki, A. M., Knowles, J. K. C., Penttilä, M. E. 1991. Enzyme production by recombinant *Trichoderma reesei* strains. *Journal of Biotechnology*, *17*, 35-49.

Vejvoda, V., Kaplan, O., Klozova, J., Masak, J., Cejková, A., Jirků, V., Stloukal, R., Martínková, L. 2006. Mild hydrolysis of nitriles by *Fusarium solani* strain O1. *Folia Microbiol (Praha)*, *51*, 251-256.

Vejvoda, V., Kaplan, O., Bezouška, K., Martínková, L. 2006. Mild hydrolysis of nitriles by the immobilized nitrilase from *Aspergillus niger* K10. *Journal of Molecular Catalysis B: Enzymatic*, *39*, 55-58.

Vejvoda, V., Kaplan, O., Bezouška, K., Pompach, P., Šulc, M., Cantarella, M., Benada, O., Uhnáková, B., Rinágelová, A., Lutz-Wahl, S., Fischer, L., Křen, V., Martínková, L. 2008. Purification and characterization of a nitrilase from *Fusarium solani* O1. *Journal of Molecular Catalysis B: Enzymatic*, *50*, 99-106.

Vejvoda, V., Kubac, D., Davidova, A., Kaplan, O., Sulc, M., Sveda, O., Chaloupkova, R., Martinkova, L. 2010. Purification and characterization of nitrilase from *Fusarium solani* IMI196840. *Process Biochemistry*, *45*, 1115-1120.

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

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

Walsh, G., 2002. Proteins, Biochemistry and Biotechnology, Wiley, Chichester.

Wang, L., Ridgway, D., Gu, T., Moo-Young, M. 2005. Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. *Biotechnology Advances*, *23*, 115-129.

Wang, P., Sandrock, R. W., VanEtten, H. D. 1999. Disruption of the cyanide hydratase gene in *Gloeocercospora sorghi* increases its sensitivity to the phytoanticipin cyanide but does not affect its pathogenicity on the cyanogenic plant *Sorghum*. *Fungal Genetics and Biology*, *28*, 126-134.

Ward, M., Lin, C., Victoria, D. C., Fox, B. P., Fox, J. A., Wong, D. L., Meerman, H. J., Pucci, J. P., Fong, R. B., Heng, M. H., Tsurushita, N., Gieswein, C., Park, M., Wang, H. 2004. Characterization of Humanized Antibodies Secreted by *Aspergillus niger*. *Applied and Environmental Microbiology*, *70*, 2567-2576.

Ward, M., Wilson, L. J., Kodama, K. H., Rey, M. W., Berka, R. M. 1990. Improved production of chymosin in *Aspergillus* by expression as a glucoamylase-chymosin fusion. *Nature Biotechnology*, *8*, 435-440.

Waugh, D. S. 2011. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expression and Purification*, *80*, 283-293.

Webster, J., Weber, R., 2007. Introduction to Fungi, Cambridge University Press, London.

Wickens, M., Anderson, P., Jackson, R. J. 1997. Life and death in the cytoplasm: messages from the 3'end. *Current Opinion in Genetics and Development*, *7*, 220-232.

Williamson, D. S., Dent, K. C., Weber, B. W., Varsani, A., Frederick, J., Thuku, R. N., Cameron, R. A., van Heerden, J. H., Cowan, D. A., Sewell, B.T. 2010. Structural and biochemical characterization of a nitrilase from the thermophilic bacterium, *Geobacillus pallidus* RAPc8. *Applied Microbiology and Biotechnology*, *88*, 143-153.

Winkler, M., Kaplan, O., Vejvoda, V., Klempier, N., Martínková, L. 2009. Biocatalytic application of nitrilases from *Fusarium solani* O1 and *Aspergillus niger* K10. *Journal of Molecular Catalysis B: Enzymatic*, *59*, 243-247.

Wohlers, I., Stachelscheid, H., Borstlap, J., Zeilinger, K., Gerlach, J. R. C. 2009. The Characterization Tool: A knowledge-based stem cell, differentiated cell, and tissue database with a web-based analysis front-end. *Stem Cell Research*, *3*, 88-95.

Wyatt, J. M., Knowles, C. J. 1995. Microbial degradation of acrylonitrile waste effluents: the degradation of effluents and condensates from the manufacture of acrylonitrile. *International Biodeterioration and Biodegradation*, *35*, 227-248.

Xie, Z., Feng, J., Garcia, E., Bernett, M., Yazbeck, D., Tao, J. 2006. Cloning and optimization of a nitrilase for the synthesis of (3S)-3-cyano-5-methyl hexanoic acid. *Journal of Molecular Catalysis B: Enzymatic*, *41*, 75-80.

Xu, J., Nogawa, M., Okada, H., Morikawa, Y. 2000. Regulation of xyn3 gene expression in *Trichoderma reesei* PC-3–7. *Applied Microbiology and Biotechnology*, *54*, 370-375.

Yamada, H., Asano, Y., Hino, T., Tani, Y. 1979. Microbial utilisation of acrylonitrile. *Journal of Fermentation Technology*, *57*, 8-14.

Yamamoto, K., Oishi, K., Fujimatsu, I., Komatsu, K. 1991. Production of R-(-)-mandelic acid from mandelonitrile by *Alcaligenes faecalis* ATCC 8750. *Applied and Environmental Microbiology*, *57*, 3028-3032.

Yazbeck, D. R., Durao, P. J., Xie, Z., Tao, J. 2006. A metal ion-based method for the screening of nitrilases. *Journal of Molecular Catalysis B: Enzymatic*, *39*, 156-159.

Yin, J., Li, G., Ren, X., Herrler, G. 2007. Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *Journal of Biotechnology*, *127*, 335-347.

Yoon, S., Ahn, Y.-H., Han, K. 2001. Enhancement of recombinant erythropoietin production in CHO cells in an incubator without CO_2 addition. *Cytotechnology*, *37*, 119-132.

Zalkin, H., Smith, J. L. 1998. Enzymes utilizing glutamine as an amide donor. *Advances in Enzymology and Related Areas of Molecular Biology*, *72*, 87-144.

Zheng, Y.-G., Chen, J., Liu, Z.-Q., Wu, M.-H., Xing, L.-Y., Shen, Y.-C. 2008. Isolation, identification and characterization of *Bacillus subtilis* ZJB-063, a versatile nitrile-converting bacterium. *Applied Microbiology and Biotechnology*, *77*, 985-993.

Zhu, D., Mukherjee, C., Yang, Y., Rios, B. E., Gallagher, D. T., Smith, N. N., Biehl, E. R., Hua, L. 2008. A new nitrilase from *Bradyrhizobium japonicum* USDA 110: Gene cloning, biochemical characterization and substrate specificity. *Journal of Biotechnology*, *133*, 327-333.

Zou, G., Shi, S., Jiang, Y., van den Brink, J., de Vries, R., Chen, L., Zhang, J., Ma, L., Wang, C., Zhou, Z. 2012. Construction of a cellulase hyper-expression system in *Trichoderma reesei* by promoter and enzyme engineering. *Microbial Cell Factories 11*, 21.

Appendix 1:

Table S1: Accession numbers of the indentified *T. reesei* proteins and their closest identified sequence homolog.

***T. reesei* Protein**	***T. reesei* protein accession number (JGI database)**	**Closest sequence homolog (species)**	**NCBI protein accession number**
Nit 1	65190	*A. niger*	ABX75546
Nit 2	64996	*A. fumigatus*	XP747028
Nit 3	65106	*R. rhodochrous*	BAA02127
Nit 4	82197	*A. thaliana*	AEE77890
Nit 5	66766	*F. graminearum*	XP382167
Nit 6	54071	*B. bassiana*	EJP67905
Nit 7	5091	*C. elegans*	1EMS_B
Nit 8	4843	*F. graminearum*	XP388013
Nit 9	104803	*C. elegans*	1EMS_B
Nit 10	103960	*C. elegans*	1EMS_B
Nit 11	74366	*C. higginsianum*	CCF44222
Nit 12	121284	*V. dahliae*	EGY18146