Manuscript for review

Structural and biochemical characterization of a nitrilase from the thermophilic bacterium, Geobacillus pallidus RAPc8

Journal:	*Applied Microbiology and Biotechnology*
Manuscript ID:	Draft
Manuscript Category:	Original Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Williamson, Dael; University of the Western Cape, Department of Biotechnology Dent, Kyle; University of Cape Town, Molecular and Cell Biology Weber, Brandon; University of Cape Town, Electron Microscope Unit Varsani, Arvind; University of Canterbury, Biological Sciences Frederick, Joni; CSIR Biosciences, Enzyme Technologies Cameron, Rory; University of the Western Cape, Biotechnology van Heerden, Johan; University of Cape Town, Molecular and Cell Biology Cowan, Donald; University of the Western Cape, Institute for Microbial Biotechnology and Metagenomics Sewell, B. Trevor; University of Cape Town, Electron Microscopy Unit
Keyword:	nitrilase, Geobacillus pallidus

Structural and biochemical characterization of a nitrilase from the thermophilic bacterium, *Geobacillus pallidus* RAPc8

Dael S. Williamson[1,4], Kyle C. Dent[1,2], Brandon W. Weber[1], Arvind Varsani[1,5], Joni Frederick[2,3], Rory A. Cameron[4], Johan H. van Heerden[2], Donald A. Cowan[4], B. Trevor. Sewell[1*]

[1] Electron Microscope Unit, University of Cape Town, Rondebosch, Cape Town, 7701, South Africa

[2] Department of Molecular and Cell Biology, University of Cape Town, Rondebosch, Cape Town, 7701, South Africa

[3] Enzyme Technologies, CSIR Biosciences, Ardeer Road, Modderfontein, 1645, South Africa

[4] Institute for Microbial Biotechnology and Metagenomics, Department of Biotechnology, University of the Western Cape, Bellville, South Africa

[5] School of Biological Sciences, University of Canterbury, Ilam, Christchurch, 8140, New Zealand

* Corresponding Author
B.T. Sewell
Electron Microscope Unit
University of Cape Town
Rondebosch
Cape Town7701
South Africa
Tel: +27 21 650 2817
Fax: +27 21 689 1528
e-mail: trevor.sewell@uct.ac.za

Abstract

Geobacillus pallidus RAPc8 (NRRL: B-59396) is a moderately thermophilic gram-positive bacterium, originally isolated from Australian lake sediment. The *G. pallidus* RAPc8 gene expressing an inducible nitrilase was located and cloned using degenerate primers coding for well conserved nitrilase sequences, coupled with inverse PCR. The translated sequence showed higher percentage similarity to plant nitrilases than to many bacterial nitrilases. The nitrilase open reading frame was cloned into an expression plasmid and the expressed recombinant enzyme purified and characterized. The protein had a monomer molecular weight of 35,790 Da and the purified functional enzyme had an apparent molecular weight of ~600 kDa by size exclusion chromatography. In common with the plant nitrilases, the recombinant *G. pallidus* RAPc8 enzyme produced both acid and amide products from nitrile substrates. Electron microscopy and image classification showed complexes having crescent-like, “c-shaped”, circular and “figure-eight” shapes. Protein models suggested that the various complexes were composed of 6, 8, 10 and 20 subunits, respectively.

Introduction

Nitrilases (EC 3.5.5.1) convert nitriles to the corresponding carboxylic acids and ammonia. However several plant (Piotrowski *et al.*, 2001) and some bacterial nitrilases (O'Reilly and Turner, 2003; Brady *et al.*, 2006) have been identified which convert nitriles to both acid and amide products (Fernandes *et al.*, 2006). Nitrilases typically occur as homo-oligomers with a monomer size of around 40 kDa. Activity is usually dependant on subunit assembly, a process affected by temperature, pH, enzyme concentration, and in some instances the presence of a substrate (O'Reilly & Turner, 2003).

Members of the nitrilase superfamily occur in both eukaryotic and prokaryotic species (Pace and Brenner, 2001). Nitrilase-related enzymes are characterized by monomers having a conserved αββα-fold which associate in a consistent fashion to form dimers. In different members of the superfamily these dimers associate in different ways to form oligomeric complexes. In the case of nitrilases these oligomeric complexes are often spirals or helices. A further feature characterizing the nitrilases is the conserved Cys, Glu, Lys catalytic triad which is implicated in covalent catalysis in which the substrate binds to the cysteine (Pace and Brenner, 2001; Brenner, 2002). It has recently been suggested that this triad includes an additional, structurally conserved Glu residue which is not immediately apparent from sequence conservation, thereby forming a catalytic quartet (Kimani *et al.*, 2007; Thuku *et al.*, 2009). The nitrilase reaction is thought to be catalyzed via covalent thioimidate and thioester intermediates and the release of an amide is not generally observed. However, it has been noted that the difference between ammonia release and amide release from the tetrahedral intermediate formed following the hydrolysis of the thioimidate involves the breakage of

either the N-C bond or the S-C bond and that this may be dependent on rather small differences in the local electronic environment (Jandhyala *et al.*, 2005).

Members of the 'true' nitrilase branch of the superfamily exhibit amino acid sequence identity in the region of 30% and display considerable variation in substrate specificity (O'Reilly & Turner, 2003). In consequence, members of this branch have been subcategorized according to differences in catalytic properties. Cyanide degrading nitrilases (or cyanide dihydratases: CynD) function identically to nitrilase enzymes but efficiently hydrolyze only inorganic cyanide; as opposed to aliphatic and aromatic nitrilases which have a broad organic cyanide substrate range (O'Reilly & Turner, 2003). Cyanide hydratases (CHT) are a subgroup of the nitrilase branch and convert HCN to formamide only demonstrating the principle that these enzymes can function as pure nitrile hydratases.

Nitrilases are potentially useful industrial catalysts, especially in production of low volume, high cost enantiomeric fine chemicals (Brady *et al.*, 2004; Brady *et al.*, 2006). The enantio- and regio-selective characteristics of these biocatalysts can potentially be exploited to increase their applications in chemical synthesis (Brady *et al.*, 2004).

Previously a nitrilase from a moderately thermophilic gram-positive bacterium *Bacillus* sp. DAC521 had been purified and shown to have a relative molecular mass of 600 kDa (Almatawah *et al.*, 1999). SDS-PAGE showed the monomer size to be approximately 41 kDa, which suggested that the enzyme formed 15-mer functional complexes. These estimates were, however, potentially obscured by the co-purification of a putative GroEL homologue which interfered with relative mass estimates (Almatawah *et al.*, 1999).

We have identified and sequenced a putative nitrilase-encoding open reading frame (ORF) in a closely related organism; *G. pallidus* RAPc8 (Pereira, *et al.* 1998). Analysis of the 972 bp sequence confirmed the presence of functional motifs, including a putative Glu, Lys, Cys catalytic triad, that are characteristic of nitrilases. Here we report the cloning, expression and detailed structural characterisation of the purified recombinant *G. pallidus* RAPc8 nitrilase.

Material and Methods

Cloning of the *G. pallidus* RAPc8 nitrilase gene

A pair of degenerate primers was designed against nitrilase consensus regions identified in a multiple alignment of 10 bacterial, archaeal and eukaryotic nitrilases. Primer NIT1 (5'- TTY CCN GAR RYN TTY ATM CCN GGN TAY CC - 3') was designed against the conserved region FPEXFIPGYP. Primer NIT3 (5' - GAN CCR TCN CCN TCN CCC CA - 3') was designed against the conserved region WG(D/E)G(D/N)GS. Using *G. pallidus* RAPc8 genomic DNA as a template, these primers were used to PCR amplify an internal portion of the nitrilase gene. PCR reaction conditions were: 94°C, 4 minutes; 25 cycles of 94°C, 30 seconds; 60°C, 30 seconds; 72°C, 30 seconds, and 72°C, 5 minutes. PCR products were separated by gel electrophoresis and bands in the expected size range (~300 bp) were excised, purified and cloned into pGEM® -T Easy Vector Systems Kit (Promega, USA) to produce plasmid pNIT2.

The pNIT2 insert was excised and labelled using the random prime DIG DNA labelling and detection kit (Roche, Germany). The labelled probe was used in a Southern blot against *G. pallidus* RAPc8 genomic DNA digested with several common restriction enzymes. The

Southern blot showed hybridising bands at ~4 kb with both *Eco*RI and *Hind*III digested DNA.

The region containing the putative nitrilase gene was amplified using inverse PCR with two primers, NITin1 (5' – CCC AAT CGT TGT ACC GAA AG - 3') and NITin2 (5' – GAC GAA ACG ACG GAA CAA CT - 3'), corresponding to the sequence of the pNIT2 insert. 5 µg of *G. pallidus* RAPc8 genomic DNA was digested by either *Eco*RI or *Hind*III in a 30 µl reaction. The digested DNA was size fractionated by electrophoresis and fragments in the size range 3-5 kb were isolated, purified and self-ligated in a 20 µl overnight reaction containing approximately 50 ng DNA.

Five µl of the ligation reaction was used in a 50 µl PCR reaction using a proof reading *Pfu* DNA polymerase (Promega, USA). The reaction conditions were: 94°C, 4 minutes, 5 cycles of 94°C, 30 seconds, 50°C, 30 seconds and 72°C, 4 minutes, 5 cycles of 94°C, 30 seconds, 58°C, 30 seconds and 72°C, 4 minutes, and 72°C, 10 minutes. The reaction was finally cooled to 4°C for 10 mins. Following amplification, the reaction was incubated in the presence of *Taq* DNA polymerase in order to A tail the amplification product for downstream TA cloning. The ~4 kb products were electrophoresed, gel purified and cloned into pGEM® -T Easy Vector Systems Kit (Promega, USA).

Random sequencing of library clones identified two plasmids, pINE2 (from the *Eco*RI restriction digest reaction) and pINH2 (from the *Hind*III restriction digest reaction), each containing overlapping portions of the nitrilase gene. Sequencing of the entire insert from each plasmid showed that the full nitrilase gene sequence and flanking regions were contained within the two plasmids.

Flanking sequence primers were designed to amplify and clone the ORF into pET29a vector in-frame. Pal-F (5' – CAT ATG GAG GGG AAG AAT ATG TC – 3') and Pal-R (5' – GGA TCC TTA ATT TTT CCA CTC AAT ATG TGT – 3') were used with the proof reading Accuzyme DNA polymerase (Bioline) to amplify the nitrilase ORF from the wild type organism by a colony PCR (GeneBank Accession # of ORF DQ826045). The product was A-tailed and cloned into pGEM® -T Easy Vector Systems Kit (Promega, USA). The ORF was directionally cloned into pET29a into the NdeI site. The recombinant pET29a plasmid was electroporated into electrocompetent *E. coli* BL21(DE3)pLysS cells for expression.

Phylogenetic analysis of the gene sequence

Sixty nine nitrilase amino acid sequences were downloaded from Genbank (accession numbers shown on phylogenetic tree next to sequence ID: Figure 2) and aligned with the *G. pallidus* sequence using Clustal X (Thompson *et al.*, 1997). A neighbour joining tree of aligned sequences (1000 iterations) was drawn using MEGA ver4 (Tumura et al., 2007).

Expression and purification

The recombinant nitrilase gene (cloned in pET29a) was expressed in *E. coli* BL21(DE3)pLys S with 1 mM IPTG (Roche, Germany) induction. Cells were sedimented by centrifugation at 4000 x *g* for 20 min at 4°C and resuspended in 50 mM Tris-HCl, pH 8.0 containing a protease inhibitor cocktail (Roche, Germany). Cells were disrupted by sonication (Misonix 3000, US) and centrifuged at 20,000 x *g* for 30 min at 4°C. The supernatant was filtered through a 0.22 µm filter and subjected to anion exchange chromatography using a Hiprep 16/10 Q XL Column (GE-Healthcare) equilibrated with 50 mM Tris-HCl, 100 mM NaCl pH 8.0 and eluted with 50 mM Tris-HCl, 1 M NaCl pH 8.0 at 5 $ml.min^{-1}$.

Fractions containing the putative nitrilase, based on band patterning on SDS PAGE gels, were pooled and concentrated by ultrafiltation using a 50 ml Amicon stirred cell with a molecular weight cut-off of 10 kDa (Millipore, USA). The concentrated protein fraction was separated by gel filtration on a Sephacryl S-300 HR gel filtration column (GE-Healthcare) equilibrated with 50mM Tris-HCl, 200 mM NaCl pH 8.0.

Nitrilase assays

Ammonia Assay: The standard ammonia assay, as described in Almatawah *et al*. (1999), was used to measure activity during nitrilase purification. Substrates used (all at 5 mM) were benzonitrile, 4-cyanopyridine and 3-phenylproprionitrile. A unit of activity was defined as 1 μmole ammonia released per minute under standard assay conditions (at 37 ^{o}C in 50 mM Tris buffer, pH 8). All assays were performed in triplicate.

HPLC Assays: Reaction mixtures consisted of 750 μl of washed and induced whole cells added to 675 μl of a 50 mM Tris buffer, pH 8, and 75 μl of nitrile or amide substrate prepared in 100 % methanol (substrate final concentration 5 mM). The final reaction volume was 1.5 ml. Reactions were incubated at 60 °C for 10, 30 and 120 minutes with shaking (200 rev./min), after which they were stopped by addition of 0.2 ml acidified acetonitrile: 0.1% (v/v) trifluoroacetic acid (TFA). Stopped reactions were centrifuged at 13 000 x *g* for 10 min, filtered through a Cameo 25SS PES 0.22 μm filter (Osmonics) and 100 μl aliquots transferred into 2 ml HPLC vials containing 900 μl 1:1 acetonitrile (ACN) and aqueous 0.1% (v/v) TFA. Control reactions in which enzyme was omitted from the reaction, were included.

Substrates (from Sigma-Aldrich, USA) were made up as 100 mM stock solutions in methanol. Benzonitrile, benzamide, mandelonitrile, 4-cyanopyrodine, 4-chlorobenzonitrile, hydrocinnamonitrile, and the equivalent amides were analysed by reversed-phase HPLC (RP-HPLC) using a Waters 2690 Separations Module equipped with a 996 Photodiode Array Detector. A 100 x 4.6 mm Chromolith Performance SpeedROD RP–18e column (Merck, Germany) was eluted isocratically at a flow rate of 0.5 ml/min by 6% $TFA_{(aq)}$: 40% ACN at 28 °C, with sample injection volumes of 20 µl. The elution of substrates and nitrilase reaction products was monitored spectrophotometrically, and peak data extracted at 252 nm Integration and data analysis was done using Millennium Version 3.05.01 software (© 1998 Waters Corporation).

LC-Mass Spectral Analysis. Coupled HPLC and mass spectral analysis was performed using the Waters Integrity™ System with photodiode array detector and Thermabeam™ mass detector. Mass Spectral results of nitrile and amide standards were compared to internal library spectra from the Millennium Version 3.05.01 software (© 1998 Waters Corporation). Analyses were performed with an initial 15 min elution with 0.1% (v/v) formic acid at 0.2 $ml.min^{-1}$, followed by 50% methanol : 50% acetonitrile for 2 min at 0.3 $ml.min^{-1}$, 0.1% (v/v) formic acid for 5 min at 0.25 $ml.min^{-1}$ and 0.1% (v/v) formic acid for 8 min at 0.2 $ml.min^{-1}$. The column temperature was set at 40 °C.

Negative stain electron microscopy

Purified samples of *G. pallidus* RAPc8 nitrilase enzyme with a concentration of 0.5 $mg.ml^{-1}$ were pipetted onto glow discharged, carbon-coated copper grids and incubated at room temperature for 10 seconds. Grids was then blotted and washed twice prior to staining with 2% uranyl acetate and air-drying. Micrographs for image processing were recorded under

low-dose conditions on a LEO912 transmission electron microscope operating at 120 kV. Images were captured at 50000x magnification on a 2k x 2k CCD camera with 14µm pixels (Proscan, Germany). Captured images were examined using Boxer (Ludtke *et al.*, 1999). Ring-like particles were selected in 128 x 128 pixel boxes at a sampling rate of 2.27 Å/pixel. A total of 4098 images were captured. A total of 481 "figure 8" particles were selected in 256 x 256 pixel boxes. All selected particles were filtered and normalised to the same mean and standard deviation.

2D Averaging

Image processing was performed using SPIDER (Frank *et al*. 1996). The ring structures were classified into 30 averages using rotationally invariant k-means (Penczek *et al*. 1996) and the classes were refined using multi-reference alignments (Joyeux and Penczek, 2002) to exclude images that did not correlate. The "figure 8" images were averaged by reference-free alignment (Penczek *et al*. 1992) and this template was used to generate 5 class averages by rotationally invariant k-means and further refined by multi-reference alignment as described above.

Results

The isolated DNA sequence (Genbank ID: DQ826045) of the *G. pallidus* RAPc8 nitrilase was analysed and found to have a GC-content of 44% over the 972 bp ORF. The DNA melting temperature was calculated to be 82°C (Kibbe, 2007). Such (or even higher) values are typical of the thermophilic Bacillaceae (Takami et al., 2004). The amino acid sequence of the *G. pallidus* RAPc8 nitrilase had the highest protein sequence similarity of ~79 – 78 % to a nitrilases from *Geobacillus* sp. Y4.1MC1 and *Brevibacillus brevis* NBRC 100599

respectively. The *G. pallidus* RAPc8 nitrilase had ~63% amino acid similarity to nitrilases from *Natranaerobius thermophilus*, *Clostridium kluyveri* and *Clostridium difficile*. The plant nitrilases share between 46% and 53% amino acid identity to the *G. pallidus* RAPc8 nitrilase. Supplementary figure 1 shows a two dimensional graphical representation of pairwise amino acid sequence identities of various nitrilases which further elaborates this point..

Phylogenetic analysis of the nitrilases indicates that nitrilases from *Geobacillus* sp., *Clostridium* sp., *Fusobacterium* sp, *Dethiosulfovibrio peptidovorans*, *Selenomonas noxia*, *Natranaerobius thermophilus*, *Brevibacillus brevis*, *Microscilla marina*, *Xantobacter autotropicus*, *Synthtrophobacter fumaroxidans*, *Citreicella* sp., *Achromobacter piechaudaii*, *Janthinobacterium* sp., *Aspergillus fumigatus* and *Symomonas mobilis* group with nitrilases from plants (Figure 2).

Analysis of the translated ORF sequence using Prosite (http://www.expasy.org) identified key motifs that are generally associated with nitrilases. These included the characteristic Glu, Lys, Cys catalytic triad, found at positions 53, 144 and 178, respectively. The presence of these key motifs was consistent with the high levels of sequence homology (between 15 and 25%) evident in alignments of the ORF against other nitrilase sequences (Figure 1). The *G. pallidus* RAPc8 nitrilase monomer comprised 323 amino acids, with a calculated molecular weight of 35,790 Da and a theoretical pI of 6.16.

The enzyme was purified to homogeneity, as demonstrated by SDS-PAGE (Figure 3), by anion exchange chromatography followed by gel filtration chromatography. The nitrilase eluted from the calibrated gel filtration column with at an estimated native molecular weight of 600 kDa (data not shown), suggesting the functional multimeric form to be a 16-mer.

The hydrolysis of a range of nitriles was tested using the recombinant *G. pallidus* nitrilase (Table 1). Under standard reaction conditions, the enzyme was most active against 3-phenylpropionitrile. The relative rates of conversion appeared to be influenced more by the presence of electronegative and activating moieties than by steric effects, although 3-phenylpropionitrile is probably the least sterically hindered of all the substrates tested. No evidence of amide hydrolysis was detected.

Mass spectral analysis was used to confirm the nature of the nitrilase reaction products. With the exception of the product from 4-cyanopyridine, both the cognate amide and acid products were generated in nitrilase reactions. At all time intervals tested, 4-chlorobenzonitrile and mandelonitrile were converted to approximately equal amounts of amide and acid products. The relatively high yields of phenylpropionamide (87%), suggest that for this substrate attack by the second water molecule is relatively slow. While the assay values obtained using benzonitrile and benzamide as substrates remain apparently anomalous, this result suggests that the two water additions are not 'in series' but follow parallel pathways.

Negative stain EM showed heterogeneous isoforms for the *G. pallidus* RAPc8 nitrilase. The micrographs (Figure 4) showed that both crescent shaped, c-shaped and closed ring-like structures were present. Larger "figure-8" particles (shown by white arrows in Figure 4A) are also evident. It appears from the class averages shown in Figure 5A that there are 3 distinct forms of open and closed ring-like structures. These class averages suggested progressive elongation of the oligomers, possibly representing stages in the assembly of the *G. pallidus* RAPc8 nitrilase, by the addition of dimers along the helical ramp until they terminate in the apparently closed, circular structure forming a "lock-washer". The "figure-8"

structures appear to contain two of the "lock-washer" structures associated through the staggered ends.

The recent crystal structure of β-alanine synthase from *Drosophila melanogaster* (Lundgren *et al.*, 2008) shows, at atomic resolution, how a nitrilase homologue can form an eight-monomer helical ramp. On the basis of a comparison with this structure (Figure 5A: I, II and III) we hypothesise that the *G. pallidus* RAPc8 nitrilase structures represent a hexamer, an octamer and a decamer respectively. We were also able to identify classes which probably correspond to the structures I, II and III viewed from the side (Figure 5A, IV – VI).

A further precedent for the existence of c-shaped helical ramps in microbial nitrilases occurs in *R. rhodochrous* J1 (Thuku *et al.*, 2007). In this case the helical symmetry parameters of the C-terminal truncated nitrilase from *R. rhodochrous* J1 have an axial rise per dimer (Δz) of 15.8 Å and a rotation per dimeric subunit (Δφ) of -73.5° (Thuku *et al.*, 2007). If our conclusions concerning the number of subunits in each of the classes is correct, then there is sufficient information in the depicted views to estimate the axial rise per dimer (Δz) as well as the rotation per subunit (Δφ). The axial rise was calculated to be 15±1 Å by measuring the stagger in the side-views (Figure 5: IV and V). The rotation per subunit was calculated by measuring the outside angle of the top-views divided by the predicted number of dimers in the isoform (Figure 5: I and II). The two angles measured produced an average rotation angle of -75° ± 1°. The outer radius of the *G. pallidus* RAPc8 spiral was also measured from the top-views to be 47 ± 1 Å.

Discussion

The gene encoding the *G. pallidus* RAPc8 nitrilase was cloned and sequenced. Interestingly, this bacterial nitrilase showed greater sequence similarity to plant nitrilases than to many other bacterial nitrilases. The result supports the proposal by Pace and Brenner (2001) and Podar *et al* (2005) that horizontal gene transfer events between the eukaryotic and prokaryotic kingdoms as a consequence of ecological association account for the phylogenetic positions of nitrilase genes (Figure 2). Not only does prokaryotic nitrilase phylogeny not reflect the taxonomy of the host organism, but nitrilases are found amongst clusters of other genes which are separated by few or no intergenic nucleotides (Podar *et al.*, 2005). Such genomic structures are consistent with large transposed genetic elements.

The purified enzyme had an apparent mean molecular weight of 600 kDa as determined by size exclusion chromatography. With a molecular weight of 35,790 Da, an average composition of 16.8 (17) monomers per multimer is suggested. However, electron microscopy and image classification showed a range of structural forms in purified enzyme preparations, including crescent-like, “c-shaped”, circular and “figure-8” shapes. Structural models (Figure 6) suggested that these forms contained 6, 8, 10 and 20 subunits, respectively, equivalent to native molecular weights of 215 kDa, 286 kDa, 358 kDa and 716 kDa. Longer helices were not seen in the micrographs. The absence of such longer forms suggests that an interaction which occurs at the ten monomer, “lock-washer” stage causes termination of the helical ramp. This situation is analogous to the 14 monomer, self-terminating spirals seen in the cyanide dihydratase from *Pseudomonas stutzeri* AK61 (Sewell et al., 2003). The existence of the “figure 8” form is interesting. These structures can be formed using the same

inter-molecular interactions as occur in the spiral ramps by mating two "lock-washers" and thus form a completely closed structure.

In common with the plant nitrilases, the *G. pallidus* RAPc8 nitrilase produced both acid and amide products from aromatic nitrile conversions (Table 1). The relative rates of these conversions suggest that a combination of electron activating effects (enhanced electropositivity of the nitrile carbon by addition of electron-withdrawing groups) and steric hinderance (i.e., distance of the nitrile carbon from a bulky group) dictate the relative rates of reaction.

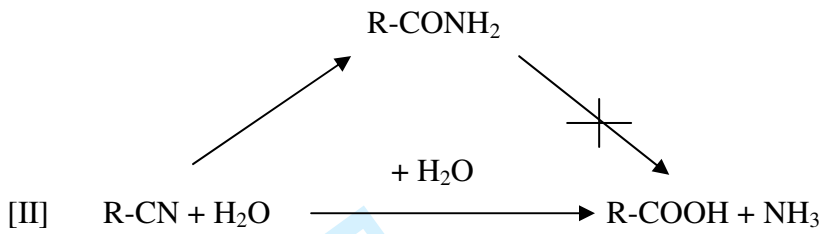
Product analysis data showed large variations in the ratios of acid to amide products. In most instances, these values could be dictated by the relative kinetics of the two water additions (i.e., the ratio k_1:k_2) in a sequential ('in series') reaction pathway [I]:

$$[I] \qquad R\text{-}CN + H_2O \xrightarrow{k_1} R\text{-}CONH_2 + H_2O \xrightarrow{k_2} R\text{-}COOH + NH_3$$

In such a case, the amide:acid ratio would be expected to decrease with reaction time, particularly with $k_1 > k_2$. We noted, however, that the amide:acid ratios were generally constant throughout the reaction period, an observation which is not consistent with an 'in series' reaction profile.

The apparently anomalous results obtained from use of the benzonitrile/benzamide substrate pair also suggests that the reaction pathway may not be 'in series'. The absence of any detectable hydrolysis of benzamide suggests that parallel reaction pathways exist [II], and that addition of the second water molecule is dependent on electronic events which are a

function of the first water addition. The difference between the two reaction routes is also thought to depend on which bond in the tetrahedral intermediate (following the postulated thioimidate complex) breaks (Jandhyala *et al.*, 2005).



A further interesting question is whether any insight can be gained as to why certain nitrilases demonstrate a tendency to form amide as the dominant product. The closest homologue to the tetrahedral intermediate to have been visualized is the cacodylate adduct of XC1258 from *Xanthomonas campestris*. (Chin et al 2007) (pdb:2e11). Simple homology modelling suggests the residues adjacent to the tetrahedral adduct in the case of *G. pallidus* RAPc8 nitrilase are Y59, which is widely conserved, and W179, which occurs immediately following the active site cysteine and is characteristic of the nitrilases. While the location of these residues provide no obvious solution to a mechanism for 'product specificity', we speculate that either the hydrophobicity of the aromatic residue side chains or the electronegativity of the tyrosine hydroxyl could potentially influence the immediate electronic environment of the tetrahedral intermediate and dictate the relative rates of N-C bond or S-C bond cleavage.

Acknowledgements

DW was supported by a NRF and DAAD bursary. KD was supported by a Harry Crossley fellowship. Much of the work was funded by the Carnegie Corporation of New York through

the Joint UCT/UWC Masters programme in structural biology. DAC also thanks the South African National Research Foundation (Institutional Research Development Program) for financial support. AV was supported by the Carnegie Corporation of New York.

References

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

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

Brady D, Dube N, Pettersen R (2006) Green Chemistry: Highly selective biocatalytic hydrolysis of nitrile compounds. *S Afr J Sci* **102**, 339-344.

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

Chin KH, Tsai YD, Chan NL, Huang KF, Wang AH, Chou SH (2007) The crystal structure of XC1258 from *Xanthomonas campestris*: a putative prokaryotic Nit protein with an arsenic adduct in the active site. *Proteins* **69**, 665-671.

Effenberger F, Oβwald S (2001) Enantioselective hydrolysis of (RS)-2-fluoroarylacetonitriles using nitrilase from *Arabidopsis thaliana*. *Tetrahedron: Asymmetry* **12**, 279-285.

Fernandes BCM, Mateo C, Kiziak C, Chmura A, Wacker J, van Rantwijk F, Stolz A, Sheldon RA (2006) Nitrile hydratase activity of a recombinant nitrilase. *Adv Synth Cataly* **348**, 2597-2603.

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

Jandhyala D, Willson RC, Sewell BT, Benedik MJ (2005) Comparison of cyanide-degrading nitrilases. *Appl Microbiol Biotechnol* **68**, 327–335.

Joyeux L, Penczek PA (2002) Efficiency of 2D alignment methods. *Ultramicroscopy*, **92**, 33-46.

Kibbe WA (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucl Acids Res* **35**(Web Server issue): W43-6.

Kimani SW, Agarkar VB, Cowan DA, Sayed MF, Sewell BT (2007) Structure of an aliphatic amidase from *Geobacillus pallidus* RAPc8. *Acta Crystallogr D Biol Crystallogr* **63**, 1048-1058.

Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct Biol* **128**, 82–97.

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

O'Reilly C, Turner PD (2003) The nitrilase family of CN hydrolysing enzymes - a comparative study. *J Appl Biol* **95**, 1161-1174.

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

Pereira RA, Graham D, Rainey FA, Cowan DA (1998) A novel thermostable nitrile hydratase. *Extremophiles* **2**, 347-357.

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

Penczek PA, Zhu J, Frank J (1996) A common-lines based method for determining orientations for N>3 particle projections simultaneously. *Ultramicroscopy* **63**, 205–218.

Piotrowski M, Schönfelder S, Weiler EW (2001) The *Arabidopsis thaliana* isogene NIT4 and its orthologs in tobacco encode β-cyano-L-alanine hydratase/nitrilase. . *Biol Chem* **276**, 2616-2621.

Podar M, Eads JR, Richardson TH (2005) Evolution of a microbial nitrilase gene family: a comparative and environmental genomics study. *BMC Evol Biol* 5,doi:10.1186/1471-2148-5-42.

Sambrook J, Fritsch E, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. 2nd Edn. Cold Spring Harbor Laboratory Press, NY

Sewell BT, Berman MN, Meyers PR, Jandhyala D, Benedik MJ (2003) The cyanide degrading nitrilase from *Pseudomonas stutzeri* AK61 is a two-fold symmetric, 14-subunit spiral. *Structure* **11**, 1413-1422.

Sewell BT, Thuku RN, Zhang X, Benedik MJ (2005) The oligomeric structure of nitrilases: the effect of mutating interfacial residues on activity. *Ann NY Acad Sci* **1056**, 153-159.

Takami H, Takaki Y, Chee GJ, Nishi S, Shimamura S, Suzuki H, Matsui S, Uchiyama, I (2004) Thermoadaptation trait revealed by the genome sequence of thermophilic *Geobacillus kaustophilus*. *Nucl Acids Res* **32**, 6292-6303.

Thompson JD, Gibson TJ, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl Acids Res* **25**, 4876-4882.

Thuku RN, Weber BW, Varsani A, Sewell BT (2007) Post-translational cleavage of recombinantly expressed nitrilase from *Rhodococcus rhodochrous* J1 yields a stable, active helical form. *FEBS J* **274**, 2099–2108.

Thuku RN, Brady D, Benedik MJ, Sewell BT (2009) Microbial nitrilases: versatile, spiral forming, industrial enzymes. *Appl Microbiol* **106**, 703-727.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol **24**, 1596–1599

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

Table and figure legends

Table 1
Substrate specificity of *G. pallidus* RAPc8 nitrilase. Values are normalised to benzonitrile. Mass spectral analysis was used to confirm the identities of the amide and acid products.

Figure 1
Alignment of the *G. pallidus* RAPc8 nitrilase amino acid sequence to closest relatives. The closely related sequences are *Zea mays* (ZMAY), *Pseudomonas fluorescens* Pf5 (PFLU), *Photobacterium profundum* 3TCK (PPRO) and *Arabidopsis thaliana* (NIT4). Conserved residues are denoted by '*' and putative active site residues of the characteristic Glu, Lys, Glu, Cys catalytic quartet are indicated in grey. Residues found within 11 Å of the active Cys are denoted by '#'. Alignment was performed using GenTHREADER (Jones, 1999).

Figure 2
A neighbour joining tree of 69 known nitrilase protein sequences aligned with the *G. pallidus* RAPc8 nitrilase sequence using Clustal X (Thompson *et al.*, 1997) and drawn using MEGA ver 4 (Tumara et al 2007).

Figure 3
SDS-PAGE of *G. pallidus* RAPc8 nitrilase during purification. Lane M, marker; lane 1, uninduced *E. coli* BL21; lane 2, IPTG induced cells; lane 3, cell free supernatant; lane 4, anion exchange chromatography fraction; lane 5, gel filtration fraction.

Figure 4
Transmission electron micrographs of purified *G. pallidus* RAPc8 nitrilase. **A**; enzyme preparation "figure-8" structures (white arrows); **B** preparation in which closed rings are the major form.

Figure 5
Class averages representing common particle views produced by iterative classification and alignment of isolated particle images. **A** depicts top (I-III) and corresponding side views (IV-VI) of three distinct isoforms showing the open and closed ring-like structure. These possibly represent a hexamer, octamer and decamer (or above) respectively. **B** shows the projections determined from the deposited co-ordinates of □-alanine synthase from *D. melanogaster* (Lundgren *et al.*, 2008) generated using SPIDER (Frank *et al.*, 1996). **C** shows the "figure 8" class average which is likely to represent two decamers which have associated in opposite orientations.

Figure 6
Oligomers of *G. pallidus* RAPc8 nitrilase interpreted from 2D averaging and constructed by applying the axial shift and helical twist determined from the images, to a dimeric homology model of the nitrilase. **A** Top (I-III) and corresponding side views (IV-VI) of three distinct isoforms are represented by a hexamer, octamer and decamer. **B** Model of the "figure-8" form. Image visualised using UCSF Chimera (Pettersen *et al.*, 2004; Goddard *et al.*, 2005).

Supplementary Figure 1
A two-dimensional graphic representation of pairwise amino acid aequence identities of various nitrilases. Each block is coloured representing the percentage sequence identity on a scale which ranges from dark blue (0%) to brown (100%).

Table 1

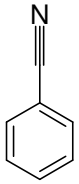
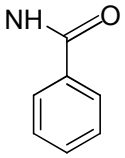
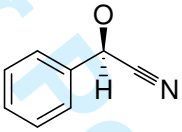
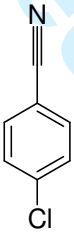
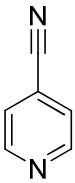
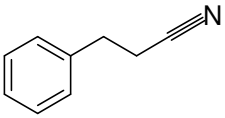
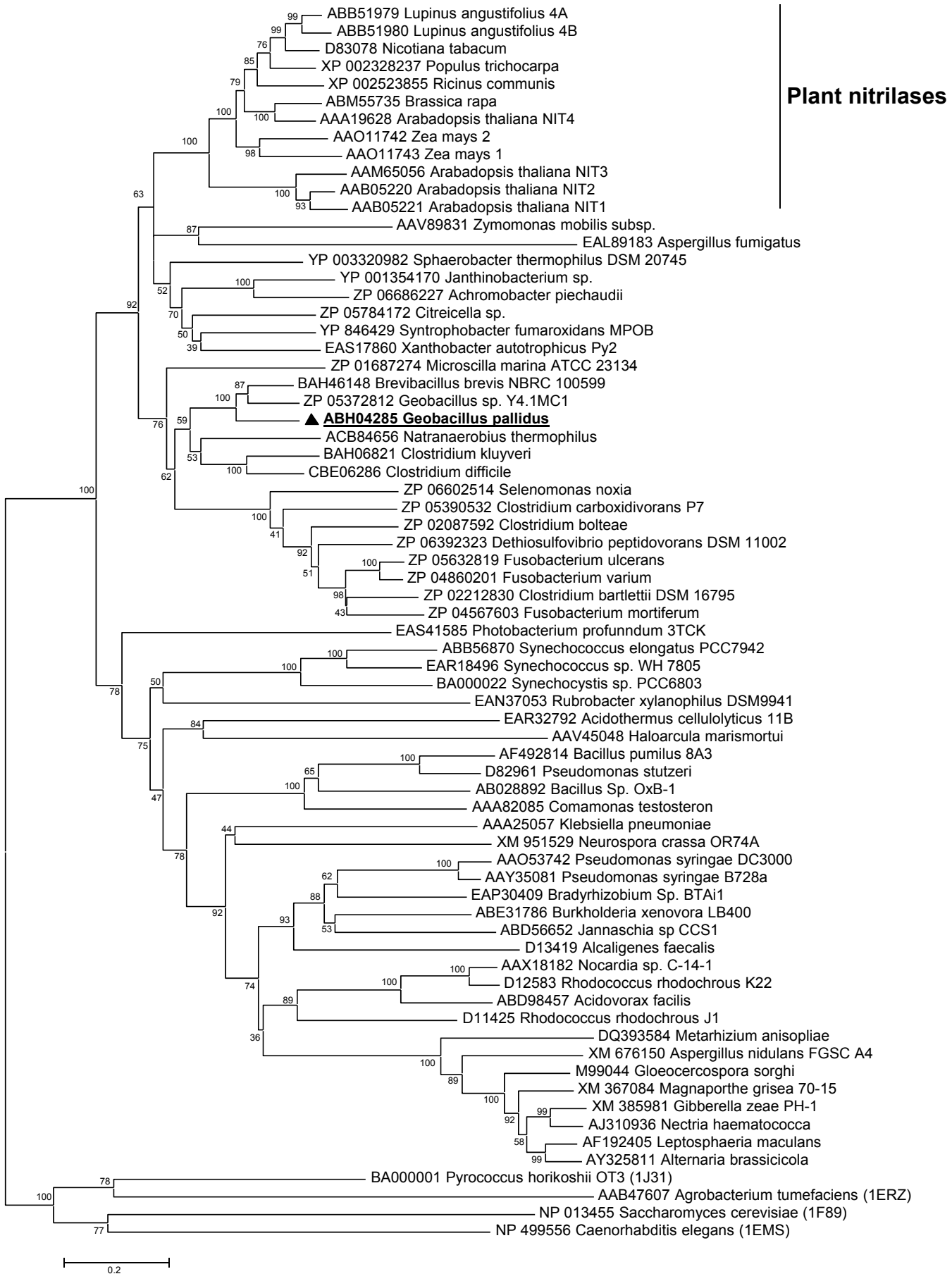
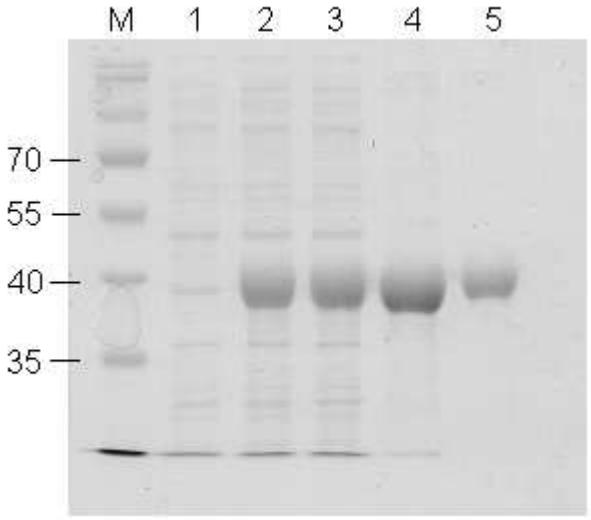
Compound	Structure	Relative rate of substrate conversion	Acid : Amide	
Benzonitrile		100	63	37
Benzamide		0	N/A	
Mandelonitrile		316	55	45
4-Cholorobenzonitrile		329	48	52
4-Cyanopyridine		1138	100	0
3-Phenylpropionitrile		1650	13	87

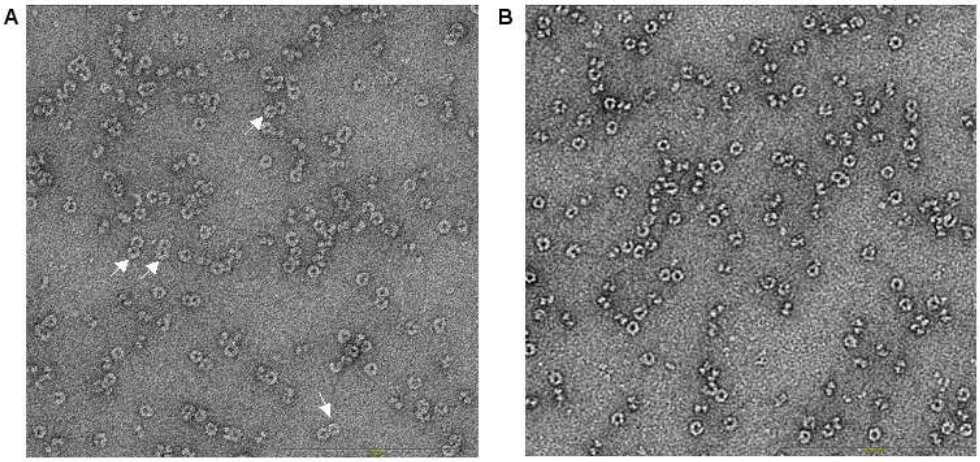
Figure 1

```
                                                        <D>                  *   *   **
                                                                                   ###
GPAL   1:----------------MEGKNMSNRAQK-VKVAVIQASSVIM---------DRDATTKKAVSLIHQAAEKGAKIVVFPEA
ZMAY   1:MALVTSGSGADQVIAEVAMNGGADPSATXVRATVVQASTIFH---------DTPATLDKAERLIAEAAGYGSQLVVFPEA
PFLU   1:---------------------------MPKSVVAALQIGALPEG---------KAATLEQILSYEAAIIEAGAQLVVMPEA
PPRO   1:------------------------------MIVKVAITQK-------PPVLLDLKSSLNKAVEIMNEVSDMGAQLVVFPEA
NIT4   7:TSHMTAAPQTNGHQIF--PEIDMSAGDSSSIVRATVVQASTVFY-------DTPATLDKAERLLSEAAENGSQLVVFPEA
2VHH  42:TSAKDIAEQNGFDIKGYRFTAREEQTRKRRIVRVGAIQNSIVIPTTAPIEKQREAIWNKVKTMIKAAAEAGCNIVCTQEA
                                     <--b1-->               <--------a1------->  <b2>

                 <--C->                                <D>   <F>                   *
          #  ####                                          ###                    #
GPAL  55:FIPAYPRGLSFGTTIGSRSAEGRKDWYRYWSN----SVAVPDETTEQLGEAARKAGVYLVIGVTERDNEFSGGTLYCSVL
ZMAY  72:FIGGYPRGSTFGFGISISIINPKDKGKEAFRRYHASAIDVPGPEVTRLAAMAAKYKVFLVMGVIE----REGYTLYCSVL
PFLU  46:LLGGYPKGEGFGTQLGYRLPEGREAFARYFAN----AIEVPGVETDALAALSARTGANLVLGVIE----RSGSTLYCTAL
PPRO  45:FLPGYPSWIWRLRPGGDMALGNKIHTKLRNN-----AVDIASGGLDSICEAAAKLNLVVVIGMNEIDSEFSGSTLFNTVV
NIT4  78:FIGGYPRGSTFELAIGSRTAKG-RDDFRKYHASAIDVPG---PEVERLALMAKKYKVYLVMGVIE----REGYTLYCTVL
2VHH 122:WTMPFAFCTREKF---------PWCEFAEEA-------ENGPTTKMLAELAKAYNMVIIHSILERDME-HGETIWNTAV
                           <-----a2---->          <----a3---->   <-b3->          <---b4-

                    ***     <*C>    *     **   <F>      *     <-----A---->
                     ##########  #                  #######               #############
GPAL 131:FFDSDGQLLGKHRKLKPTAAERIVWGEGDGSTLPVFDTPYGRIGALICWENYMPLARAAMYAQGIQIYIAPTA----DAR
ZMAY 148:FFDPLGRYLGKHRKLMPTALERIIWGFGDGSTIPVYDTPLGKIGALICWENKMPLLRTALYGKGIEIYCAPTA----DSR
PFLU 118:YFDPQQGLSGKHRKLMPTGTERLIWGKGDGSTLPVLDTQVGRVGAVICWENMMPLLRTAMYAQGIEVWCA----PTVDER
PPRO 120:VIDANGKIVNRHRKIMPTNPERMVWGFGDGSGLRVVDTSVGRIGCLICWENYMPLARYSLFTQDIDIYIAPTW----DSG
NIT4 150:FFDSQGLFLGKHRKLMPTALERCIWGFGDGSTIPVFDTPIGKIGAAICWENRMPSLRTAMYAKGIE-----IYCAPTADS
2VHH 184:VISNSGRYLGKHRKNHIPRMESTYYMEG-NTGHPVFETEFGKLAVNICYGRHHPQNWMMFGLNGAEIVFNPSATIGRLSE
         ->     <-b5->          <b6>        <b7>          <-----a5----->    <b8>         <

         <-----A----->               <-------C------>
          ##                 ###                  ####
GPAL 207:ETWQSTIRHIALEGRCFVLSANQYVTKDMYPKDLACYDELASSPEIMS---RGGSAIVGPLGEYVAEPVFGKEDIIIAEL
ZMAY 224:PVWQASMTHIALEGGCFVLSANQFCRRKDYPPPPEYEFAGLGEEPSADTVVCPGGSVIISPSGEVLAGPNYEGEALITAD
PFLU 194:EMWQVSMRHIAHEGRCFVVSACQVQASPEELGLEIANWPAQRPL-------IAGGSVIVGPMGDVLAGPLVGRAGLISAQ
PPRO 196:DSWIASMNHIAREGGCWVLSTATALQGEDIPESFPERDNLFPAEEWI----NPGDAVVIKPFGGIIAGPLHREKGILYSD
NIT4 225:RETWLASMTHIALEGGCFVLSANQFCRRKDYPSPPEYMFSGSEESLTPDSVVCAGGSSIISPLGIVLAGPNYRGEALITA
2VHH 264:PLWSIEARNAAIANSYFTVPINRVGTEQFPNEYTSGDGNKAHKEFGP----FYGSSYVAAPDGSRTPSLSRDKDGLLVVE
         -----a6----->   <b9>  <b10>                  <b11>   <b12>    <b13>    <--b14->

                              <-------C------><----------A---------->
GPAL 284:DMKQIAYSQFDFDPVGHYARPDVFKLLVNKEKKTTIEWKN--------------------
ZMAY 304:LDLGEIVRAKFDFDVVGHYSRPEVLRLVVNDQPQLPV---SFTSAAERTPAAKSDIDTKSY
PFLU 267:IDTADLVRARYDYDVVGHYARPDVFELTVDQRPRPGV---RFT------------------
PPRO 272:IDLGAARDSRKALDVAGHYNRPDIFHFEVDRRTQPPI---KFIDDSNGSD-----------
NIT4 305:DLDLGDIARAKFDFDVVGHYSRPEVFSLNIREHPR----KAVSFKTSKVMEDESV------
2VHH 340:LDLNLCRQVKDFWG-FRMTQRVPLYAESFKKASEHGFKPQIIKET-----------------
          <---a7-->          <--a8->             <b12>
```

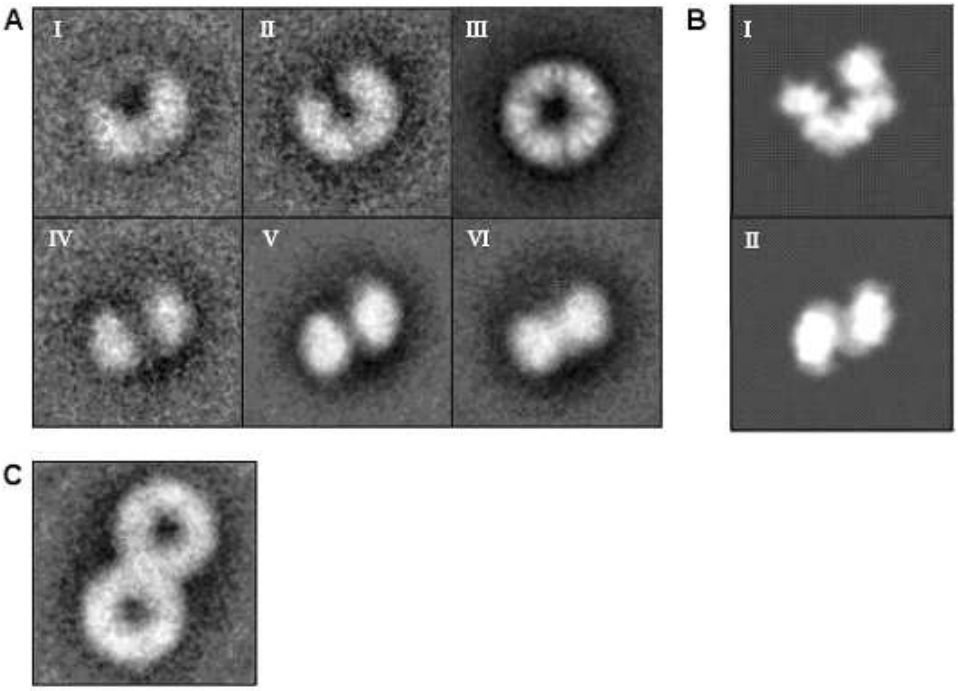

Plant nitrilases
ABB51979 Lupinus angustifolius 4A
ABB51980 Lupinus angustifolius 4B
D83078 Nicotiana tabacum
XP 002328237 Populus trichocarpa
XP 002523855 Ricinus communis
ABM55735 Brassica rapa
AAA19628 Arabadopsis thaliana NIT4
AAO11742 Zea mays 2
AAO11743 Zea mays 1
AAM65056 Arabadopsis thaliana NIT3
AAB05220 Arabadopsis thaliana NIT2
AAB05221 Arabadopsis thaliana NIT1
AAV89831 Zymomonas mobilis subsp.
EAL89183 Aspergillus fumigatus
YP 003320982 Sphaerobacter thermophilus DSM 20745
YP 001354170 Janthinobacterium sp.
ZP 06686227 Achromobacter piechaudii
ZP 05784172 Citreicella sp.
YP 846429 Syntrophobacter fumaroxidans MPOB
EAS17860 Xanthobacter autotrophicus Py2
ZP 01687274 Microscilla marina ATCC 23134
BAH46148 Brevibacillus brevis NBRC 100599
ZP 05372812 Geobacillus sp. Y4.1MC1
▲ ABH04285 Geobacillus pallidus
ACB84656 Natranaerobius thermophilus
BAH06821 Clostridium kluyveri
CBE06286 Clostridium difficile
ZP 06602514 Selenomonas noxia
ZP 05390532 Clostridium carboxidivorans P7
ZP 02087592 Clostridium bolteae
ZP 06392323 Dethiosulfovibrio peptidovorans DSM 11002
ZP 05632819 Fusobacterium ulcerans
ZP 04860201 Fusobacterium varium
ZP 02212830 Clostridium bartlettii DSM 16795
ZP 04567603 Fusobacterium mortiferum
EAS41585 Photobacterium profunndum 3TCK
ABB56870 Synechococcus elongatus PCC7942
EAR18496 Synechococcus sp. WH 7805
BA000022 Synechocystis sp. PCC6803
EAN37053 Rubrobacter xylanophilus DSM9941
EAR32792 Acidothermus cellulolyticus 11B
AAV45048 Haloarcula marismortui
AF492814 Bacillus pumilus 8A3
D82961 Pseudomonas stutzeri
AB028892 Bacillus Sp. OxB-1
AAA82085 Comamonas testosteron
AAA25057 Klebsiella pneumoniae
XM 951529 Neurospora crassa OR74A
AAO53742 Pseudomonas syringae DC3000
AAY35081 Pseudomonas syringae B728a
EAP30409 Bradyrhizobium Sp. BTAi1
ABE31786 Burkholderia xenovora LB400
ABD56652 Jannaschia sp CCS1
D13419 Alcaligenes faecalis
AAX18182 Nocardia sp. C-14-1
D12583 Rhodococcus rhodochrous K22
ABD98457 Acidovorax facilis
D11425 Rhodococcus rhodochrous J1
DQ393584 Metarhizium anisopliae
XM 676150 Aspergillus nidulans FGSC A4
M99044 Gloeocercospora sorghi
XM 367084 Magnaporthe grisea 70-15
XM 385981 Gibberella zeae PH-1
AJ310936 Nectria haematococca
AF192405 Leptosphaeria maculans
AY325811 Alternaria brassicicola
BA000001 Pyrococcus horikoshii OT3 (1J31)
AAB47607 Agrobacterium tumefaciens (1ERZ)
NP 013455 Saccharomyces cerevisiae (1F89)
NP 499556 Caenorhabditis elegans (1EMS)
0.2



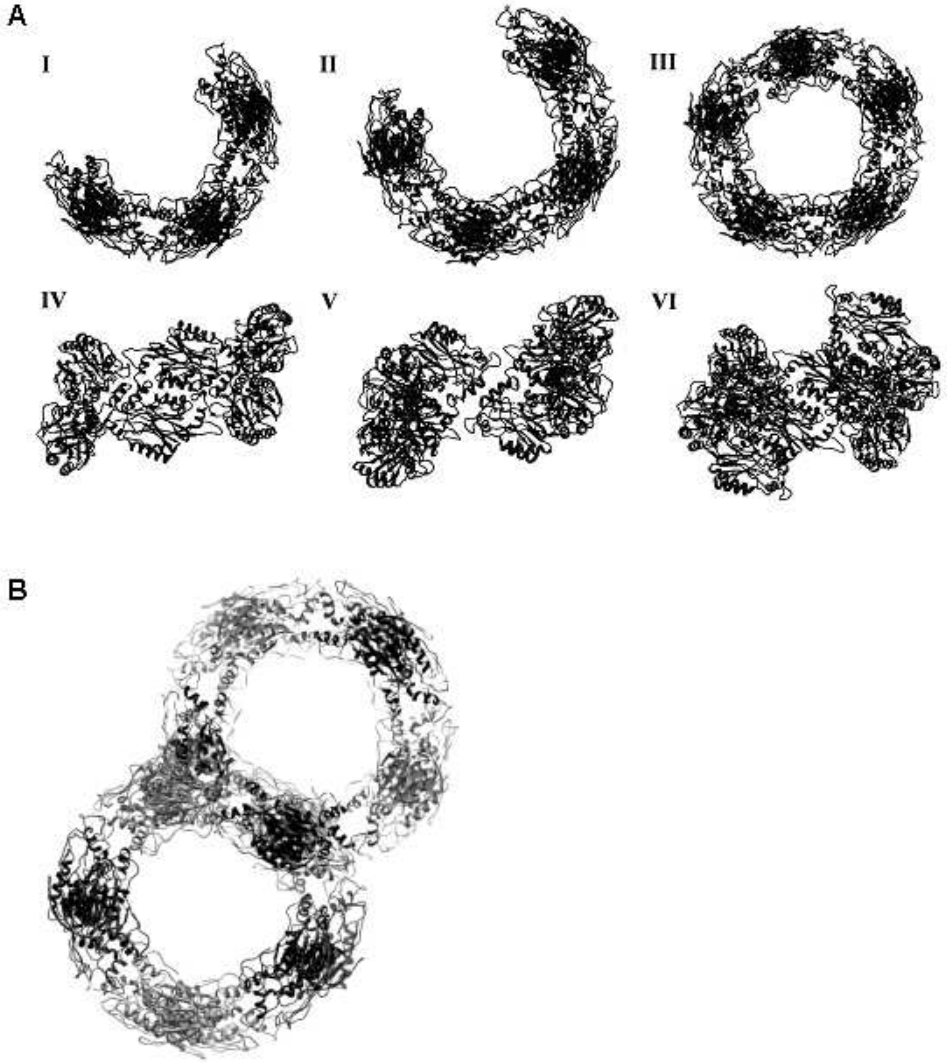
13x12mm (600 x 600 DPI)



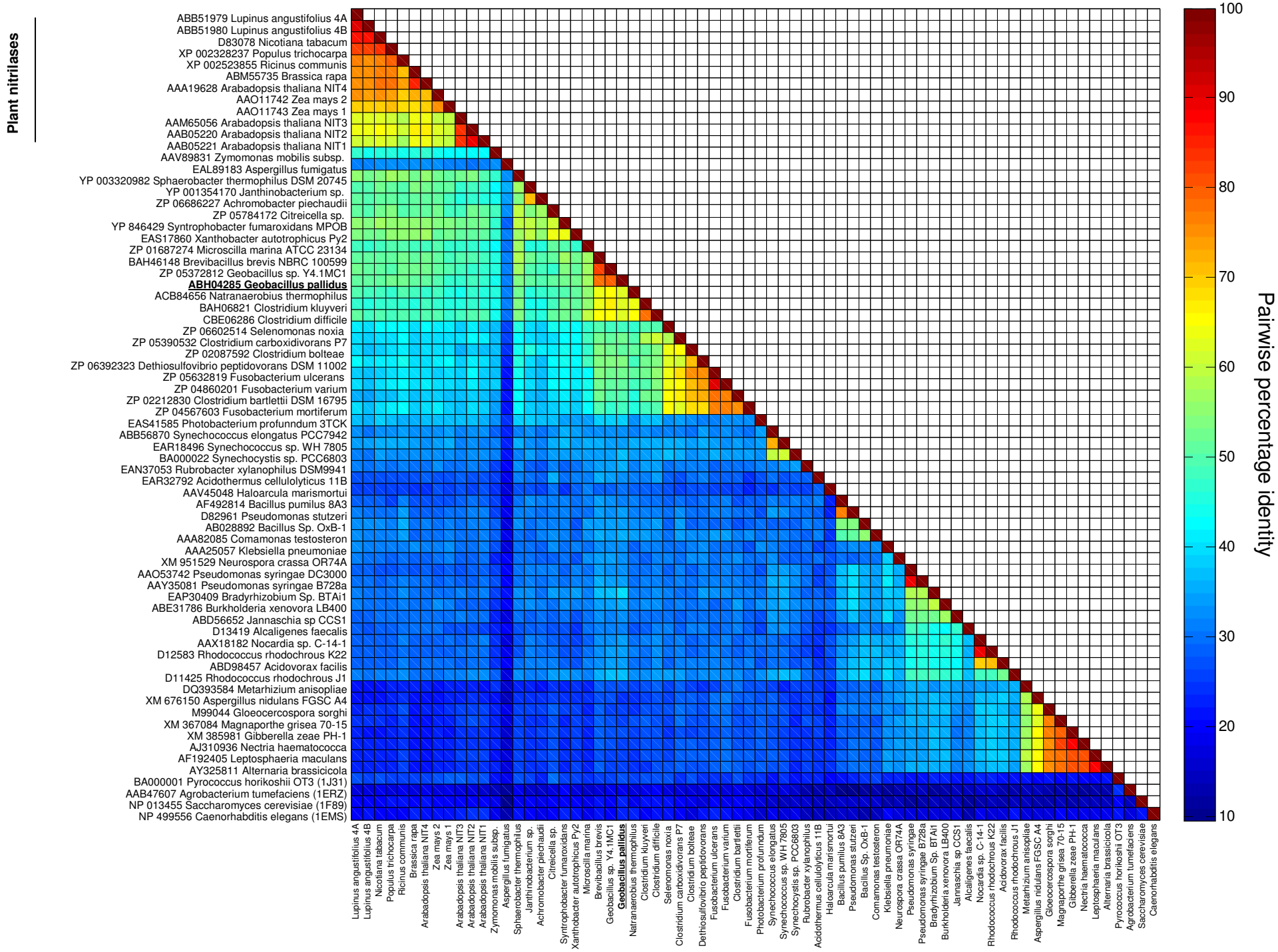
32x15mm (600 x 600 DPI)



23x17mm (600 x 600 DPI)



24x27mm (600 x 600 DPI)


Plant nitrilases
ABB51979 Lupinus angustifolius 4A
ABB51980 Lupinus angustifolius 4B
D83078 Nicotiana tabacum
XP 002328237 Populus trichocarpa
XP 002523855 Ricinus communis
ABM55735 Brassica rapa
AAA19628 Arabadopsis thaliana NIT4
AAO11742 Zea mays 2
AAO11743 Zea mays 1
AAM65056 Arabadopsis thaliana NIT3
AAB05220 Arabadopsis thaliana NIT2
AAB05221 Arabadopsis thaliana NIT1
AAV89831 Zymomonas mobilis subsp.
EAL89183 Aspergillus fumigatus
YP 003320982 Sphaerobacter thermophilus DSM 20745
YP 001354170 Janthinobacterium sp.
ZP 06686227 Achromobacter piechaudii
ZP 05784172 Citreicella sp.
YP 846429 Syntrophobacter fumaroxidans MPOB
EAS17860 Xanthobacter autotrophicus Py2
ZP 01687274 Microscilla marina ATCC 23134
BAH46148 Brevibacillus brevis NBRC 100599
ZP 05372812 Geobacillus sp. Y4.1MC1
ABH04285 Geobacillus pallidus
ACB84656 Natranaerobius thermophilus
BAH06821 Clostridium kluyveri
CBE06286 Clostridium difficile
ZP 06602514 Selenomonas noxia
ZP 05390532 Clostridium carboxidivorans P7
ZP 02087592 Clostridium bolteae
ZP 06392323 Dethiosulfovibrio peptidovorans DSM 11002
ZP 05632819 Fusobacterium ulcerans
ZP 04860201 Fusobacterium varium
ZP 02212830 Clostridium bartlettii DSM 16795
ZP 04567603 Fusobacterium mortiferum
EAS41585 Photobacterium profunndum 3TCK
ABB56870 Synechococcus elongatus PCC7942
EAR18496 Synechococcus sp. WH 7805
BA000022 Synechocystis sp. PCC6803
EAN37053 Rubrobacter xylanophilus DSM9941
EAR32792 Acidothermus cellulolyticus 11B
AAV45048 Haloarcula marismortui
AF492814 Bacillus pumilus 8A3
D82961 Pseudomonas stutzeri
AB028892 Bacillus Sp. OxB-1
AAA82085 Comamonas testosteron
AAA25057 Klebsiella pneumoniae
XM 951529 Neurospora crassa OR74A
AAO53742 Pseudomonas syringae DC3000
AAY35081 Pseudomonas syringae B728a
EAP30409 Bradyrhizobium Sp. BTAi1
ABE31786 Burkholderia xenovora LB400
ABD56652 Jannaschia sp CCS1
D13419 Alcaligenes faecalis
AAX18182 Nocardia sp. C-14-1
D12583 Rhodococcus rhodochrous K22
ABD98457 Acidovorax facilis
D11425 Rhodococcus rhodochrous J1
DQ393584 Metarhizium anisopliae
XM 676150 Aspergillus nidulans FGSC A4
M99044 Gloeocercospora sorghi
XM 367084 Magnaporthe grisea 70-15
XM 385981 Gibberella zeae PH-1
AJ310936 Nectria haematococca
AF192405 Leptosphaeria maculans
AY325811 Alternaria brassicicola
BA000001 Pyrococcus horikoshii OT3 (1J31)
AAB47607 Agrobacterium tumefaciens (1ERZ)
NP 013455 Saccharomyces cerevisiae (1F89)
NP 499556 Caenorhabditis elegans (1EMS)
Lupinus angustifolius 4A
Lupinus angustifolius 4B
Nicotiana tabacum
Populus trichocarpa
Ricinus communis
Brassica rapa
Arabadopsis thaliana NIT4
Zea mays 2
Zea mays 1
Arabadopsis thaliana NIT3
Arabadopsis thaliana NIT2
Arabadopsis thaliana NIT1
Zymomonas mobilis subsp.
Aspergillus fumigatus
Sphaerobacter thermophilus
Janthinobacterium sp.
Achromobacter piechaudii
Citreicella sp.
Syntrophobacter fumaroxidans
Xanthobacter autotrophicus Py2
Microscilla marina
Brevibacillus brevis
Geobacillus sp. Y4.1MC1
Geobacillus pallidus
Natranaerobius thermophilus
Clostridium kluyveri
Clostridium difficile
Selenomonas noxia
Clostridium carboxidivorans P7
Clostridium bolteae
Dethiosulfovibrio peptidovorans
Fusobacterium ulcerans
Fusobacterium varium
Clostridium bartlettii
Fusobacterium mortiferum
Photobacterium profunndum
Synechococcus elongatus
Synechococcus sp. WH 7805
Synechocystis sp. PCC6803
Rubrobacter xylanophilus
Acidothermus cellulolyticus 11B
Haloarcula marismortui
Bacillus pumilus 8A3
Pseudomonas stutzeri
Bacillus Sp. OxB-1
Comamonas testosteron
Klebsiella pneumoniae
Neurospora crassa OR74A
Pseudomonas syringae
Pseudomonas syringae B728a
Bradyrhizobium Sp. BTAi1
Burkholderia xenovora LB400
Jannaschia sp CCS1
Alcaligenes faecalis
Nocardia sp. C-14-1
Rhodococcus rhodochrous K22
Acidovorax facilis
Rhodococcus rhodochrous J1
Metarhizium anisopliae
Aspergillus nidulans FGSC A4
Gloeocercospora sorghi
Magnaporthe grisea 70-15
Gibberella zeae PH-1
Nectria haematococca
Leptosphaeria maculans
Alternaria brassicicola
Pyrococcus horikoshii OT3
Agrobacterium tumefaciens
Saccharomyces cerevisiae
Caenorhabditis elegans
100
90
80
70
60
50
40
30
20
10
Pairwise percentage identity