# Genome Mining, Phylogenetic and Structural Analysis of Bacterial Nitrilases for the Biodegradation of Nitrile Compounds

**Richa Salwan** ( richaihbt332@gmail.com )
College of Horticulture and Forestry https://orcid.org/0000-0003-2597-5552
**Vivek Sharma**
Chandigarh University
**Surajit Das**
National Institute of Technology Rourkela

**Short Report**

**Keywords:** Nitrile compounds, Biodegradation, Docking, Substrate specificity, Catalytic, superfamily

**Posted Date:** June 1st, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-561386/v1

# Abstract

Microbial nitrilases play vital role in biodegradation of nitrile-containing contaminants in pollutant and effluents treatments in chemical and textile industries as well as the biosynthesis of IAA from tryptophan in plants. However, the lack of structural information hinders the correlation of its activity and substrate specificity. Here, we have identified bacterial genomes for nitrilases bearing unassigned functions including hypothetical, uncharacterized, or putative role. The genomic annotations revealed four predicted nitrilases encoding genes as uncharacterized subgroup of the nitrilase superfamily. Further, the annotation of these nitrilases revealed relatedness with nitrilase hydratases and cyanoalanine hydratases. The characterization of motif analysis of these protein sequences, predicted a single motif of 20-28 aa, and glutamate (E), lysine (K) and cysteine (C) residues as a part of catalytic triad along with several active site residues. The structural analysis of the modeled nitrilases revealed geometrical and close conformation of α-helices and β-sheets arranged in a sandwich structure. The catalytic residues constituted the substrate binding pocket and exhibited the wide nitrile substrate spectra for both aromatic and aliphatic nitriles containing compounds. The aromatic amino acid residues Y159 in active site were predicted to show importance for substrate specificity. The substitution of non-aromatic alanine residue in place of Y159 completely disrupted the catalytic activity for indole-3-acetonitrile. The present study reports several uncharacterized nitrilases which have not been reported so far for their role in the biodegradation of pollutants, xenobiotics which could find applications in industries.


# Introduction

Nitrile (− C ≡ N) containing organic compounds such as cyanoglycosides, cyanolipids, ricinine and phenylacetonitrile alongwith other synthetic derivatives are abundant in the environment (Gong et al., 2012). Most of these compounds are toxic, mutagenic and carcinogenic due to the presence of cyano group in their structure. The enzyme catalyzing transformation of these nitrile organic (− C ≡ N) compounds into their corresponding amides and carboxylic acids such as nicotinic acid, glycolic acid and other is used for industrial production (Shaw et al., 2003; Mylerova and Martinkova, 2003; Banerjee et al., 2004; Panova et al., 2007). The biodegradation of toxic nitrile compounds from environmental wastes and contaminants using nitrile-converting enzymes is one of the efficient methods to convert them into their corresponding acid and ammonia (Gong et al., 2012). For the transformation of nitrile compounds, two types of pathways for the enzymatic break down of nitriles have been identified. The first pathway involves nitrilases (E.C. 3.5.5.1), catalyse the transformation of nitriles compounds to their corresponding acid. The second pathway involves nitrile hydratases, converts nitriles into amides and subsequently, amidases (E.C.3.5.1.4) assisted hydrolysis of amides into ammonia and carboxylic acids. The respective enzymes for both pathways have been reported form bacteria, fungi and plants (Egelkamp et al., 2017; Martínková et al., 2009; Rustler and Stolz, 2007; Piotrowski 2008).

For the transformation of toxic nitrile compounds from environmental wastes, nitrilases are preferred over chemical methods, due to their ecofriendly, efficient and cost-effective nature (Kumar et al., 2011). In additions, the nitrilases are also of particular interest due to their potential use in food additives (Tang et

al., 2019), surface modification of polyacrylonitrile (PAN) fibers (Zhua et al., 2013), pharmaceutical and agrochemicals sectors (Zhang et al., 2019), degradation of bromoxynil, ioxynil, acrylonitrile (Shen et al., 2020) contaminated waste bodies such as near to gold mining, electroplating industries and production of indole acetic acid (IAA) from tryptophan (Salwan et al., 2020). Since, microbial nitrilases can transform different nitriles into their carboxylic acids or amides therefore, have immense biotechnological applications in the detoxification of nitriles, cyanide and for the biosynthesis of plant growth regulators (Rucka et al., 2019; Howden et al., 2009; Kiziak et al., 2005; O'Reilly and Turner, 2003). The nitrilase family was initially proposed in 2001 by Pace and Brenner. At present, nitrilases have been categorized into 13 branches with different activities and substrate specificities but only first branch is known to catalyze nitrile hydrolysis, whereas 2–9 branches hydrolyse amides, carbamates or perform reverse amidase reaction and branch 10 reported for tumor-suppressor activity in humans and mice (Jones et al., 2021; Pace and Brenner, 2001; Brenner, 2002; O'Reilly and Turner, 2003). At structural level, the nitrilases constitute ααββ sandwich with a twisted helical structure and conserved catalytic tetrad of Glu-Lys-Glu-Cys (Brenner, 2002; Nagasawa et al., 2000; Thuku et al., 2007; Jones et al., 2021).

With the advent of genome and metagenome sequencing, presently over 5000 nitrilase encoding genes have been deposited in the GenBank database (Shen et al., 2020). Despite the initial reports on nitrilases from plants, several nitrilases have been characterized from bacteria and fungi (Thimann and Mahadevan, 1964; Hook and Robinson, 1964; Gong et al., 2012). Most of the bacterial nitrilases have been explored for their role in chemical synthesis or degradation but the biological role of nitrilases is largely unknown (DeSantis et al., 2002). Moreover, the lack of structural data limits the correlation of amino acid sequence, specificity and activity of nitrilases towards substrates (Podar et al., 2005). Therefore, considering the importance of nitrilases, the present study involves the distribution and diversity of bacterial nitrilases at genomic level, their phylogenetic and structural characterization for motifs/ domains, and interaction with substrates. Here, we have described *in silico* annotation of nitrilases followed by the structure prediction of promising biocatalysts to propose their role in degradation of aliphatic and aromatic nitrile containing compounds.

# Material And Methods/computational Methods

### Retrieval of microbial genomes

The complete bacterial genomes of *Acidovorax* sp. MR-S7 BANP01, *Acidovorax oryzae* RS-1 (AFPT01), *Acinetobacter* sp. ATCC27244 (ABYN01), *Bacillus cereus* W (ABCZ02), *B. cereus* G9241 (AAEK01), *Bradyrhizobium diazoefficiens* SEMIA 5080 (ADOUO2), *Cupriavidus necator* B9 (FMSH01), *C. necator* 5 (CAIGK01), *Pyrococcus horikoshii* UBA8834 (DUJN01), *Rhodococcus rhodochrous* BKS6-46 (AGVW02), *R. opacus* PD630 (AGVD01), *R. rhodochrous* TRN7 (FBUK01), *Sphingomonas* sp. KC8 (AFMP01), *Sphingomonas* sp. S17 (AFGG01) and *Sphingomonas* sp. LH128 (ALVC01) were obtained from National Center for Biotechnology Information (NCBI) and annotated using RAST server. Besides this, previously sequenced genomes of *Streptomyces* sp. AC30, *Streptomyces* sp. AC40 and *Bacillus* sp. SBA 12 were also used for nitrilase prediction (Salwan et al., 2020a and b). All the genomes annotated for the presence of

nitrilases and genes encoding nitrilases were deduced for corresponding amino acid sequence using of ExPASy translation tool (https://web.expasy.org/translate/).

### Analysis of amino acid sequence

The amino acid sequences were subjected to pBLAST analysis. The sequences with unassigned functions including uncharacterized, putative or hypothetical were selected and compared in order to identify new nitrilases. The amino acid composition analysis of these nitrilases was done to determine molecular weight, pI value, acidic, basic, polar, and non-polar amino acids as well as GRAVY and Aliphatic index using Expasy ProtParam tool.

### Phylogenetic analysis

The neighbor-joining phylogenetic tree was made by using Mega 7 (Kumar et al., 2016). The multiple alignments of sequences retrieved from NCBI database were performed using CLUSTALW program (http://www.genome.jp/tools/clustalw/). The sequences were submitted to InterPro Scan for determining family and domain. Phylogenetic tree was constructed by retrieving amino acid sequences from NCBI BLAST tool.

### Domain analysis

The different domains were predicted by submitting amino acid sequence of nitrilases in Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de/) (Letunic et al., 2012) as well as NCBI-Conserved Domain Database (NCBI-CDD) (Marchler-Bauer et al., 2015). The N-terminal signal peptide was predicted using SignalP Server to determine extracellular or intracellular nature of the protein (Petersen et al., 2011) (www.cbs.dtu.dk/services/SignalP/). The catalytic triad, active site residues, conserved domains and sites were predicted using NCBI-CDD. Multiple Expectation maximizations for Motif Elicitation (MEME) was used to determine conserved motifs in the sequence (Bailey et al., 2009).

### Structural features

The catalytic domain of nitrilase was modeled using user-template approach based the nitrilase from *Mus musculus* nitrilase-2 (PDB code, 2W1V_A) in SWISS MODEL (http://swissmodel.expasy.org/) and 3D structures of the predicted models were prepared in PYMOL (Schwede et al., 2003). The quality of final models was assessed with QMEANclust (Benkert et al., 2009). PROCHECK analysis of the 3D models was performed to check the stereo-chemical qualities (Laskowski et al., 1993). The secondary structures were predicted using POLYVIEW Server (http://polyview.cchmc.org/).

### Docking analysis

The molecular docking analysis of nitrilase from *Acidovorax* sp. MS-R7 was also done to predict stable ligand-protein interaction (https://www.dockingserver.com). Different substrates including idole-3-acetonitrile, α-picolinic acid amide, phenylacetonitrile, 2-cyanopyridine, crotonolide, butenedinitrile,

acrylonitrile, malononitrile hexonamide and benzamide were docked with the PDB structure to determine interactions of modeled protein and respective ligands.

# Results And Discussion

### Annotation of genomes for nitrilases

In this study, a total of 16 genomes were annotated using RAST server for nitrilase encoding genes. The genomes size ranged from 1.7 to 9.15 Mbp size with GC content 35.2 to 70%. The presence of total predicted genes in the genome varies from 1900 in *Pyrococcus horikoshii* UBA8834 to 9063 in *Bradyrhizobium diazoefficiens* SEMIA 5080 (Table 1). The genome mining revealed highest number of nitrilase encoding genes in *Bradyrhizobium diazoefficiens* SEMIA 5080 followed by *Acidovorax* sp. MR-S7, *A. oryzae* RS-1, *Acinetobacter* sp. ATCC27244, *Cupriavidus necator* B9, *C. necator* 5, *Rhodococcus opacus* PD630, *Sphingomonas* sp. S17, *Sphingomonas* sp. LH128 and *Streptomyces* AC40, whereas the genome of *Bacillus cereus* W, *B. cereus* G9241, *Bacillus* sp. SBA12, *Pyrococcus horikos*hii UBA8834, *Rhodococcus rhodochrous* BKS6-46, *R. rhodochrous* TRN7, *Sphingomonas* sp. KC8 and *Streptomyces* sp. AC30 lacks nitrilase encoding genes (Table 1). Nitrilase predicted in *Streptomyces* AC40 was only 119 aa long possibly because of the draft genome sequence (Salwan et al., 2020). Further, the identification of amino acid sequences revealed grouping of nitrilases to C-N hydrolase, amidase, glutamine amidotransferase, nitrile hydratase and periplasmic nitrile proteins (Table S1). Among all, four nitrilases showing identity towards uncharacterized subgroup of the nitrilase superfamily were characterized to identify newer sources of nitrilases. The size of nitrilase varies from 208–345 amino acids and shared 24–37% sequence identity to various nitrilases (Table 2). Further, the evolutionary and phylogenetic history of nitrilase revealed relatedness of uncharacterized nitrilases AcNit, As7nit, CnB9 and Cn5 with nitrilases belonging to C_N hydrolase domain and grouped as separate cluster (Fig. 1).

Table 1
Characteristics of genomes and annotation based on RAST server

| Organism | Genome Id | Size (bp) | GC Content | Number of Contigs | Number of Coding Sequences | Number of RNAs | Nitrilases |
|---|---|---|---|---|---|---|---|
| *Acidovorax* sp. MR-S7 | BANPO1 | 5,007,754 | 68.3 | 130 | 4918 | 55 | 2 |
| *Acidovorax oryzae* RS-1 | AFPT | 5,522,282 | 68.7 | 156 | 5323 | 44 | 1 |
| *Acinetobacter* sp. ATCC27244 | ABYNO | 3,330,366 | 39.4 | 255 | 3333 | 64 | 2 |
| *Bacillus cereus* W | ABCZO2 | 5,240,573 | 35.5 | 102 | 5482 | 136 | 0 |
| *Bacillus cereus* G9241 | AAEKO1 | 5,934,942 | 35.2 | 207 | 6316 | 155 | 0 |
| *Bradyrhizobium diazoefficiens* SEMIA 5080 | ADOUO2 | 9,085,533 | 64 | 13 | 9063 | 53 | 6 |
| *Cupriavidus necator* B9 | FMSH | 7,303,706 | 66.3 | 550 | 7389 | 57 | 2 |
| *Cupriavidus necator* ISOLATE 5 | CAIGK | 7,191,616 | 64.5 | 59 | 6865 | 56 | 1 |
| *Pyrococcus horikoshii* UBA8834 | DUJNO1 | 1,724,643 | 41.8 | 8 | 1900 | 47 | 0 |
| *Rhodococcus rhodochrous* BKS6-46 | AGVW | 6,213,641 | 67.4 | 609 | 6195 | 53 | 0 |
| *Rhodococcus opacus* PD630 | AGVD | 9,149,864 | 67 | 491 | 9091 | 55 | 3 |
| *Rhodococcus rhodochrous* TRN7 | FBUKO1 | 4,871,006 | 70.2 | 173 | 5330 | 59 | 0 |
| *Sphingomonas* sp. KC8 | AFMP | 4,074,265 | 63.7 | 70 | 4036 | 49 | 0 |
| *Sphingomonas* sp. S17 | AFGGO1 | 4,268,406 | 65.7 | 62 | 4035 | 52 | 1 |
| *Sphingomonas* sp. LH128 | ALVCO1 | 6,457,516 | 64.9 | 658 | 6793 | 56 | 2 |

Table 2
Identification of nitrilases retrieved from the annotated genomes using pBLAST tool in NCBI

| Sequence ID | Amino acids | % Identity | Nearest neighbour | Identification |
|---|---|---|---|---|
| CAIGKG010000010 *Cupriavidus necator* 5 (Cn5Nit) | 273 | 34 | Chain A, Nitrilase Pyrococcus abyssi GE5, to and | Figure 003879: Uncharacterized subgroup of the nitrilase superfamily |
| | | 32 | Chain A, Hydrolase *Xanthomonas campestris* pv. *campestris* | |
| | | 32 | Chain A, Amidase, *Nesterenkonia* sp. 10004 | |
| | | 27 | to Chain A, Hydrolase, Carbon-nitrogen Family *Staphylococcus aureus* subsp. *aureus* COL | |
| BANP01 *Acidovorax* sp. MR-S7 (AcNit) | 271 | 37 | Chain A, Hydrolase *Xanthomonas campestris* pv. *campestris* | Figure 003879: Uncharacterized subgroup of the nitrilase superfamily |
| | | 31 | Chain A, Amidase *Nesterenkonia* sp. 10004 | |
| | | 26 | Chain A, Hydrolase, Carbon-Nitrogen Family *Staphylococcus aureus* subsp. *aureus* COL] | |
| | | 26 | Chain A, N-carbamyl-d-amino Acid Amidohydrolase *Agrobacterium* sp. KNK712. | |
| ABYN01000224 *Acinetobacter* sp. ATCC 27244 (As7Nit) | 274 | 26 | putative carbon-nitrogen family hydrolase from *Staphylococcus aureus* | Figure 003879: Uncharacterized subgroup of the nitrilase superfamily |
| | | 25 | Amidase from *Nesterenkonia* sp. 10004 | |
| | | 24 | N-carbamoyl-D-amino acid amidohydrolase *Agrobacterium tumefaciens* | |
| FMSH01000197 *Cupriavidus necator* B9 (Cn9Nit) | 276 | 29 | putative carbon-nitrogen family hydrolase from *Staphylococcus aureus* | Figure 003879: Uncharacterized subgroup of the nitrilase superfamily |
| | | 32 | CN-hydrolase superfamily protein *Xanthomonas campestris* pv. *campestris* | |

The comparative amino acid composition revealed molecular weight and pI in the range ~ 24–54 kDa and 4.76–7.81, respectively which is closer to the previously characterized aliphatic or aromatic nitrilases. The lower pI value is probably due to the presence of higher contents of acidic amino acids (30%) and lower contents of basic amino acids (25%), and − 0.032 GRAVY and 93 aliphatic indexes. Besides this, more number of negatively charged amino acids over the surface, higher content of non-polar (55%) and polar amino acids (22%) and less Pro (4%) and more Gly (9%) residues in nitrilases may provide flexibility to the protein structures.

## Domain and motif analysis

The domain analysis of nitrilases AcNit, As7Nit, Cn5Nit and Cn9Nit revealed a conserved catalytic domain C-N Hydrolase of ~ 252–297 aa belonging to nitrilase superfamily. The signal peptide cleavage site was not predicted as per SignalP server and Interpro Scan (http://www.cbs.dtu.dk/services/SignalP-4.1/; https://www.ebi.ac.uk/interpro/search/sequence/) thus, revealing its intracellular location. Three residues located at Glu, Lys and Cys (EKC) formed a catalytic triad centre, although the location of amino acid is variable. The triad is conserved where C residue plays important role in maintaining its activity and K and E residues mediate acid-base catalysis for nitrilase (Raczynska et al., 2011). The crystal structures of various nitrilases had been resolved in eukaryotes including animals, plants and yeasts, as well as prokaryotic nitrilases (6MG6 and 3WUY) (Shen et al., 2020). It has been reported that the putative substrate binding pocket formed of hydrophobic and hydrophilic residues lies near the surface in nitrilase of *Synechocystis* sp. strain PCC6803 (Shen et al., 2020). Further characterization of these four nitrilases revealed presence of single motif consisted of 28 aa although the position is variable among all (Table 3; Fig. 2b). Catalytic residues were found conserved in all the predicted proteins (Table 3).

Table 3
Annotation of genes encoding nitrilases and prediction of motifs, domains, active sites, secondary structures and other conserved sites based on InterProScan and NCBI-CDD

| Sequence ID | Motifs | Domain | Catalytic triad | Active sites | SS | Genes for xenobiotic degradation |
|---|---|---|---|---|---|---|
| ABYN01000224<br><br>*Acinetobacter* sp. ATCC 27244<br><br>(AcNit) | 21–49 | CN-Hydrolase 2-254 | E41, K115, C157 | E41, T99, K115, F119, E132, C157, Y158, L160, R161, A182 | 9H, 17S | K00362, K00627<br><br>K00622, K00791<br><br>K00930, K00626<br><br>K00982 |
| BANP01<br><br>*Acidovorax* sp. MR-S7<br><br>(As7Nit) | 20–48 | CN-Hydrolase<br><br>1-298 | E40, K113, C158 | E40, N96, K113, F117, E128, C158, Y159, L161, R162, A183 | 9H, 18S | K00362, K00627<br><br>K00364, K00625<br><br>K00791, K00930<br><br>K00363, K00626 K00984, K00982 |
| CAIGKG010000010<br><br>*Cupriavidus necator* (Cn5Nit) | 26–54 | CN-Hydrolase<br><br>8-263 | E46, K119, C159 | E46, N102, K119, F123, E134, C159, Y160, L162, R163, A185 | 8H, 15S | K00362, K00625<br><br>K00791, K00626 |
| FMSH01000197<br><br>*Cupriavidus necator* B9<br><br>(Cn9Nit) | 29–57 | CN-Hydrolase 10–265 | E49, K122, C162 | E49, N105, K122, F126, E137, C162, Y163, L165, R166, A188 | 9H, 17S | K00625<br><br>K00791<br><br>K00626 |

## Structural characterization of identified nitrilases

The secondary structure analysis revealed presence of 30% helix, 31% sheets and 2% loops or intrinsically disordered regions, forming a 4-layer α-β-β-α sandwich structure as per SCOP classification (Table 2; Fig. 2a). Ramachandran plot depicted 88–92% amino acid residues in most favored regions, 7.5–10.4% amino acid residues in additional allowed regions, and 0.4–1.3% amino acid residues in the disallowed conformations (Fig. 3). The 3D structure of 4 nitrilases prepared by taking best matches which involves 32–38% identity with nitrilase of *Mus musculus* nitrilase-2 (PDB: 2W1V_A), and 26% identity with nitrilase

of *Staphylococcus aureus* subsp. *aureus* (PDB: 3P8K) and 31% identity with nitrilase of *Nesterenkonia* sp. 10004 (3HKX_A). The modeling and superimposition of modeled structures AcNit, As7Nit, Cn5Nit and Cn9Nit with the template 2W1V_A revealed root mean square deviation (RMSD) 1.7Å, 0.093, 0.137 and 0.156, respectively. The modeled 3D structures of nitrilases gave QMean score 0.65, 0.67, 0.63 and 0.64, indicating good quality for the model to represent the enzymatic protein as the values fall within the prescribed limits. The geometry and topology of the predicted model was close to the 2W1V_A template. The crystal structures of nitrilase from *Arabidopsis thaliana* (PDB: 6I00), *Caenorhabditis elegans* (PDB:1EMS), *Helicobacter pylori* (PDB: 6MG6), *Mus musculus* (PDB: 2W1V), *Pyrococcus abyssi* (PDB:3KLC), *Saccharomyces cerevisiae* (PDB: 4H5U and 1F89), and *Synechocystis* sp. (PDB: 3WUY) has already been reported. However, with most of the nitrilases, the crystal structure had been resolved, came from eukaryotes such as animals, plants and yeasts, and only the nitrilases (6MG6 and 3WUY) came from prokaryotes. The catalytic triad EKC appeared distantly in primary structure and forms substrate binding pocket along with other residues, a characteristic of the nitrilase superfamily (Fig. 4). The structural superimposition of modeled nitrilases showed almost identical structural coordination of the catalytic triads with the selected template 2W1V_A. AcNit displays a relatively big substrate-binding pocket formed of aromatic and hydrophobic residues E40, N96, K113, F117, E128, C158, Y159, L161, R162, and A183 (Fig. 4), to accommodate both aliphatic and aromatic nitrile substrates and displays the broadest nitrile substrate spectrum. Aromatic residues in the substrate-binding pocket of AcNit are possibly involved in stabilizing aromatic nitrile substrates. Similarly, crystallographic studies have reported substrate binding pocket composed of hydrophobic (EKHEC) and hydrophilic residues (YFNYWWPMVF) in *Synechocystis* sp. strain PCC6803 (Zhang et al., 2014). Further, the involvement of amino acid N118 and E142 has been reported in stabilizing the active site E53 and K135, respectively. Similarly, aromatic nitrile substrates are stabilized by involving W170 and F202 and other residues PGSMVGQIF are responsible for making substrate-binding loop (Zhang et al., 2014; Raczynska et al., 2011). Previous studies have also revealed formation of oxyanion hole by involving phenylalanine which play major role in recognizing substrates (Shen et al., 2020). Therefore, the selected nitrilase AcNit contain catalytic C158 which acts as a nucleophile, E40 activates the sulfhydryl group of C158 by acting as a base, and K113 stabilizes intermediates formed in the reaction. The conserved residues N96 and E128 also help in providing stability and activating catalytic triad by making hydrogen bonds with E40 and K113, respectively.

## Protein ligand interaction based on docking

Nitrilases are known for hydrolytic activity against aromatic, aliphatic and arylacetonitrile substrates. The predicted protein ligand interaction of AcNit displays higher binding affinities − 5.79 kcal/mol for indole-3-acetonitrile, -4.56 kcal/mol for α-picolinic acid amide, 4.29 kcal/mol for phenylacetonitrile, -4.15 kcal/mol for 2-cyanopyridine, -3.32 kcal/mol for crotonolide, -3.24 kcal/mol for butenedinitrile, -2.73 kcal/mol for acrylonitrile, and − 2.68 kcal/mol for malononitrile compared to + 196.77 Kcal/mol for hexonamide and + 107 kcal/mol for benzamide. The ineteraction between proteins and ligands takes place only when $\Delta G$ or Gibbs free energy change is negative. The negative $\Delta G$ determines the stable interaction of proteins with the ligands. Similarly, the nitrilase of *L. aggregata* DSM 13394 is reported do not act on aromatic and heterocyclic nitriles but showed high preference towards arylacetonitriles hydrolysis like

iminiodiacetonitrile (Zhang et al., 2012). Other nitrilase *Comamonas testosterone*, *Geobacillus pallidus* rapc8, *Nocardia globerula* NHB-2, *Pseudomonas aeruginosa* 10145, *Rhodococcus rhodochrous* J1, *Streptomyces* sp. MTCC 7546 have been reported for aliphatic, aromatic and arylaectonitrile substrates and reported suitable for various industrial applications (Nigam et al., 2009; Levy-Schil et al., 1995; Williamson et al., 2010; Harper et al., 1985; Alonso et al., 2008; Kobayashi et al., 1989).

To investigate whether aromatic amino acid Y159 in nitrilase plays important role for determining substrate specificity, saturation mutagenesis was performed. The introduction of non-aromatic alanine residue in place of Y159 completely disrupted the catalytic activity for indole-3-acetonitrile which is indicated by the occurrence of + 5.69 kcal/mol binding energy for IAN (Fig. 4d). It is known that strong binding of protein and the ligand depends upon the accuracy of binding energy. Lower is the binding energy; stronger is the affinity for binding substrates. Therefore, the presence of hydrophilic residues in the substrate binding pocket justifies the activity of modeled nitrilase AcNit towards aromatic substrates also.

### Gene prediction for xenobiotic degradation and secondary metabolites

Various genes responsible for degradation of xenobiotics including benzoate, bisphenol, fluorobenzoate, furfural, naphthalene, aminobenzoate, styrene, atrazine, xylene, caprolactam, chloroalkane and chloroalkene, cytochrome P450 and steroids were predicted. Similarly, other gene operons involved in the biosynthesis of phenylpropanoid, stilbenoid, diarylheptanoid and gingerol, flavanoid, flavones, isoflavanoid, indole alkaloid, isoquinoline alkaloid, acridone, penicillin and cephalosporin, monobactam and clavulanic acid were predicted. The presence of these genes suggests the role of nitrilases in biodegradation of pollutants and xenobiotic compounds. It could also prove useful for production.

## Conclusions

Nitrilases due to their broad substrate specificities have been explored for their role in the transformation of nitrile containing toxic, mutagenic and carcinogenic organic compounds into their corresponding amides and carboxylic acids derivatives in industrial production. Nitrilases also play vital role in IAA production from tryptophan dependent pathway. Microbes are promising source of nitrilases and have several industrial, agricultural and environmental applications. Among all nitrilases, AcNit could also serve as a strong template to carry protein engineering for enhanced production and expression for suitable industrial applications. Further, characterization and detailed crystallographic studies could also demonstrate their biological role by *in vitro* analysis and scope for novel enzymes.

## Declarations

### Acknowledgement

RS is thankful to Science Academies for providing Summer Research Fellowship and acknowledges College of Horticulture and Forestry, Neri for providing necessary infrastructure and National Institute of

Technology, Rourkela for kind assistance during the completion of this work. VS is thankful to Chandigarh University for providing infrastructural facilities.

### Funding

The authors are thankful to SEED Division, Department of Science and Technology, GOI for providing financial benefits (SP/YO/125/2017) and (SEED-TIASN-023-2018) during the completion of this work.

### Conflicts of interest/Competing interests

The authors declared no conflict of interests.

### Ethics approval

Not applicable

### Consent to participate

NA

### Consent for publication

Not applicable

### Availability of data and material

Not applicable

### Code availability

Not applicable

### Authors' contributions

RS and VS were involved in the execution of computational work, MS writing and editing. SD assisted in finalizing the work.

## References

1. Alonso FOM, Oestreicher EG, Antunes OAC. (2008). Production of enantiomerically pure D-phenylglycine using Pseudomonas aeruginosa 10145 as biocatalyst. Braz J Chem Eng.25(1):1–8.
2. Benkert P, Künzli M, Schwede T. (2009). QMEAN server for protein model quality estimation. Nucleic Acids Res. 37(Web Server issue):W510-4. doi: 10.1093/nar/gkp322.
3. O'Reilly C., Turner P.D. (2003). The nitrilase family of CN hydrolysing enzymes – a comparative study, J. Appl. Microbiol. 95:1161–1174.

4. DeSantis G, Zhu Z, Greenberg WA, Wong K, Chaplin J, Hanson SR, Farwell B, Nicholson LW, Rand CL, Weiner DP, Robertson DE, Burk MJ. (2002). An enzyme library approach to biocatalysis: development of nitrilases for enantioselective production of carboxylic acid derivatives. *J Am Chem Soc.***124**:9024–9025. doi: 10.1021/ja0259842
5. Egelkamp, R., Friedrich, I., Hertel, R. et al. (2020). From sequence to function: a new workflow for nitrilase identification. Appl Microbiol Biotechnol 104:4957–4970. https://doi.org/10.1007/s00253-020-10544-9
6. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 31(13):3784-3788. doi:10.1093/nar/gkg563
7. Gong JS, Lu ZM, Li H, Shi JS, Zhou ZM, Xu ZH. (2012). Nitrilases in nitrile biocatalysis: recent progress and forthcoming research. Microb Cell Fact 11:142. doi:10.1186/1475-2859-11-142
8. Harper DB. (1985). Characterization of a nitrilase from Nocardia sp. (Rhodochrous group) N.C.I.B. 11215, Using p-hydroxybenzonitrile as sole carbon source. Int J Biochem. 17(6):677–683.
9. Hook RH, Robinson WG. (1964). Ricinine nitrilase: II. Purification and properties. *J Biol Chem.* **239**:4263–4267.
10. Howden A.J., Harrison C.J., Preston G.M. A conserved mechanism for nitrile metabolism in bacteria and plants. Plant J. 2009;57:243–253
11. Letunic I, Doerks T, Bork P. (2012). SMART 7: recent updates to the protein domain annotation resource, Nucleic Acids Research, 40(D1):D302–D305 https://doi.org/10.1093/nar/gkr931
12. Kiziak C., Conradt D., Stolz A., Mattes R., Klein J. Nitrilase from Pseudomonas fluorescens EBC191: cloning and heterologous expression of the gene and biochemical characterization of the recombinant enzyme. Microbiology. 2005;151:3639–3648
13. Kobayashi M, Nagasawa T, Yamada H. (1989). Nitrilase of *Rhodococcus rhodochrous* J1. Febs J.182: 349–356.
14. Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. Journal of applied crystallography 26(2):283-291
15. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993). PROCHECK- a program to check the stereochemical quality of protein structures. J App Cryst 26:283-291
16. Levy-Schil S, Soubrier F, Crutz-Le Coq AM, et al. (1995). Aliphatic nitrilase from a soil-isolated Comamonas testosteroni sp.: gene cloning and overexpression, purification and primary structure. Gene 161(1):15–20.
17. Marchler-Bauer A et al. (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res.43(D)222-6.
18. Martínková L, Uhnáková B, Pátek M, Nešvera J, Křen V. (2009). Biodegradation potential of the genus *Rhodococcus*. Environ Int. **35**:162–177. doi: 10.1016/j.envint.2008.07.018
19. Mylerova V, Martinkova L. (2003). Synthetic applications of nitrile-converting enzymes. *Curr Org Chem* **7**:1279–1295. doi: 10.2174/1385272033486486.

20. Nigam VK, Khandelwal AK, Gothwal RK, et al. (2009). Nitrilase-catalysed conversion of acrylonitrile by free and immobilized cells of *Streptomyces* sp. J Biosci 34(1):21–26.
21. Petersen TN, Brunak S, von Heijne G, Nielsen H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 29;8(10):785-6. doi: 10.1038/nmeth.1701. PMID: 21959131.
22. Piotrowski M. (2008). Primary or secondary? Versatile nitrilases in plant metabolism. Phytochemistry. **69**:2655–2667.doi: 10.1016/j.phytochem.2008.08.020.
23. Podar M, Eads JR, Richardson TH (2005) Evolution of a microbial nitrilase gene family: a comparative and environmental genomics study. BMC Evol Biol 5: 42
24. Raczynska JE, Vorgias CE, Antranikian G, et al. (2011). Crystallographic analysis of a thermoactive nitrilase. J Struct Biol. 173(2):294–302.
25. Rucká L, Chmátal M, Kulik N, Petrásková L, Pelantová H, Novotný P, Pˇríhodová R, Pátek M, and Martínková L (2019). Genetic and Functional Diversity of Nitrilases in Agaricomycotina. Int. J. Mol. Sci. 2019, 20, 5990; doi:10.3390/ijms20235990
26. Rustler S, Stolz A. (2007). Isolation and characterization of a nitrile hydrolysing acidotolerant black yeast- *Exophiala oligosperma* R1. Appl Microbiol Biotechnol. **75**:899–908. doi: 10.1007/s00253-007-0890-3.
27. Salwan R, Sharma V, Sharma A, Singh A. (2020a). Molecular imprints of plant beneficial Streptomyces sp. AC30 and AC40 reveal differential capabilities and strategies to counter environmental stresses. Microbiol Res.235:126449. doi: 10.1016/j.micres.2020.126449.
28. Salwan R, Sharma V. (2020b). Genome wide underpinning of antagonistic and plant beneficial attributes of *Bacillus* sp. SBA12. Genomics 112(4): 2894-2902
29. Schwede T, Kopp J, Guex N, Peitsch MC. (2003). SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res. 1;31(13):3381-5. doi: 10.1093/nar/gkg520.
30. Shaw NM, Robins KT, Kiener A. (2003). Lonza: 20 years of biotransformations. Adv Synth Catal. 345:425–435. doi: 10.1002/adsc.200390049.
31. Shen JD, Cai X, Liu ZQ, Zheng YG (2020): Nitrilase: a promising biocatalyst in industrial applications for green chemistry, Critical Reviews in Biotechnology, DOI: 10.1080/07388551.2020.1827367
32. Shen JD, Cai X, Liu ZQ, Zheng YG. (2021). Nitrilase: a promising biocatalyst in industrial applications for green chemistry. Crit Rev Biotechnol. 41(1):72-93. doi: 10.1080/07388551.2020.1827367.
33. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016 Jul;33(7):1870-4. doi: 10.1093/molbev/msw054. Epub 2016 Mar 22. PMID: 27004904.
34. Tang CD, Ding PJ, Shi HL, et al. (2019). One-pot synthesis of phenylglyoxylic acid from racemic mandelic acids via cascade biocatalysis. J Agric Food Chem. 67(10): 2946–2953.
35. Thimann KV, Mahadevan S. (1964). Nitrilase. I. Occurrence, preparation, and general properties of the enzyme. Arch Biochem Biophys. 105:133–141. doi: 10.1016/0003-9861(64)90244-9.

36. Williamson DS, Dent K, Weber B, et al. (2010). Structural and biochemical characterization of a nitrilase from the thermophilic bacterium, *Geobacillus pallidus* RAPc8. Appl Microbiol Biotechnol. 88(1):143–153.
37. Zhang L, Yin B, Wang C, et al. (2014). Structural insights into enzymatic activity and substrate specificity determination by a single amino acid in nitrilase from *Syechocystis* sp. PCC6803. J Struct Biol. 188(2): 93–101
38. Zhang Q, Wu ZM, Hao CL, et al. (2019). Highly regio-and enantioselective synthesis of chiral intermediate for pregabalin using one-pot bienzymatic cascade of nitrilase and amidase. Appl. Microbiol. Biot. 103:5617–5626.
39. Zhang, C.S., Zhang, Z.J., Li, C.X., Yu, H.L., Zheng, G.W., & Xu, J.H. (2012). Efficient production of (R)-o-chloromandelic acid by deracemization of o-chloromandelonitrile with a new nitrilase mined from *Labrenzia aggregata*. Applied Microbiology and Biotechnology, 95(1), 91–99. doi:10.1007/s00253-012-3993-4

# Figures

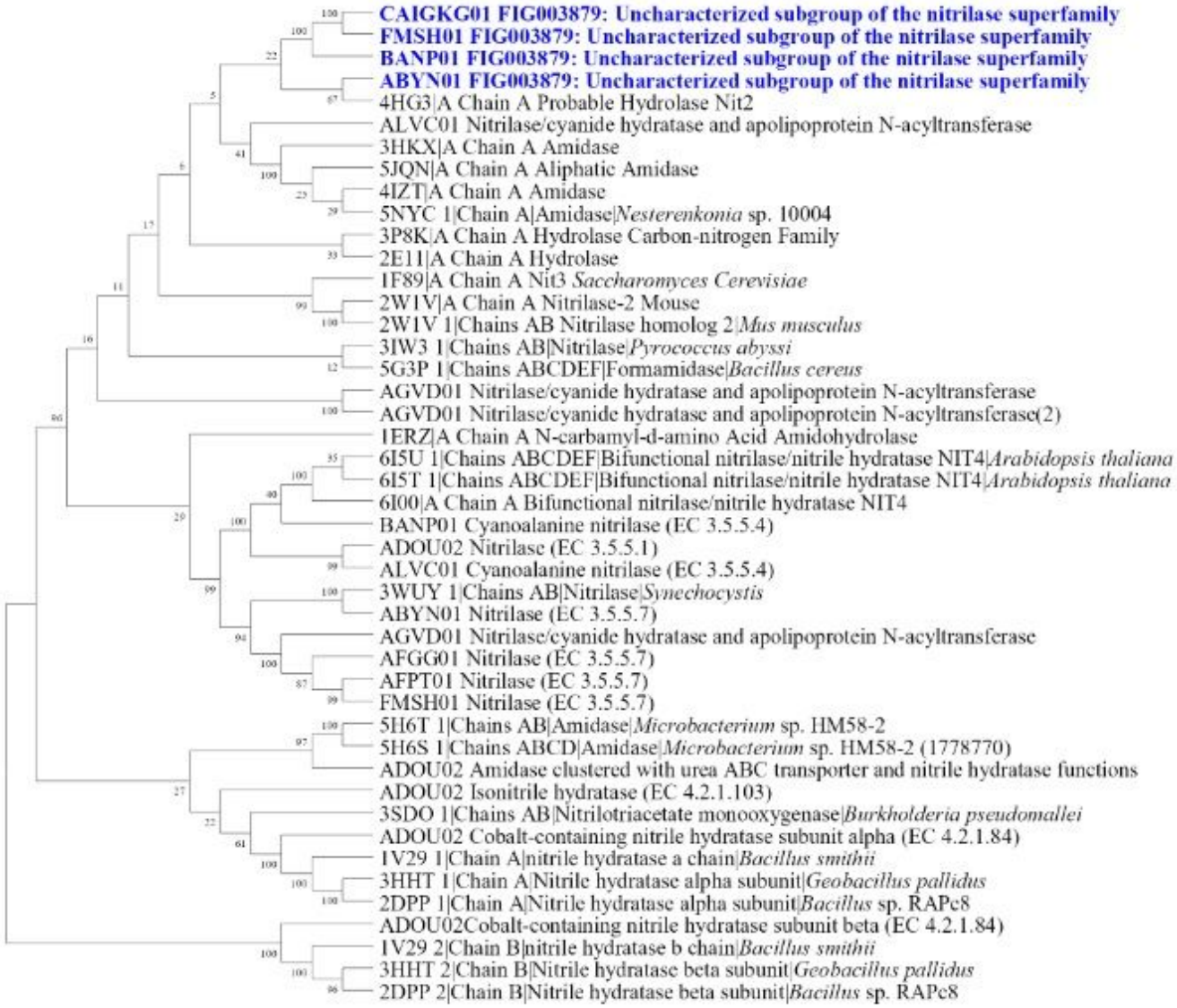


**Fig. 1**

## Figure 1

Phylogenetic tree depicting evolutionary study of nitrilase retrieved from NCBI and PDB. The tree was prepared in MEGA version 7.0. Boot strap values are shown next to branches.

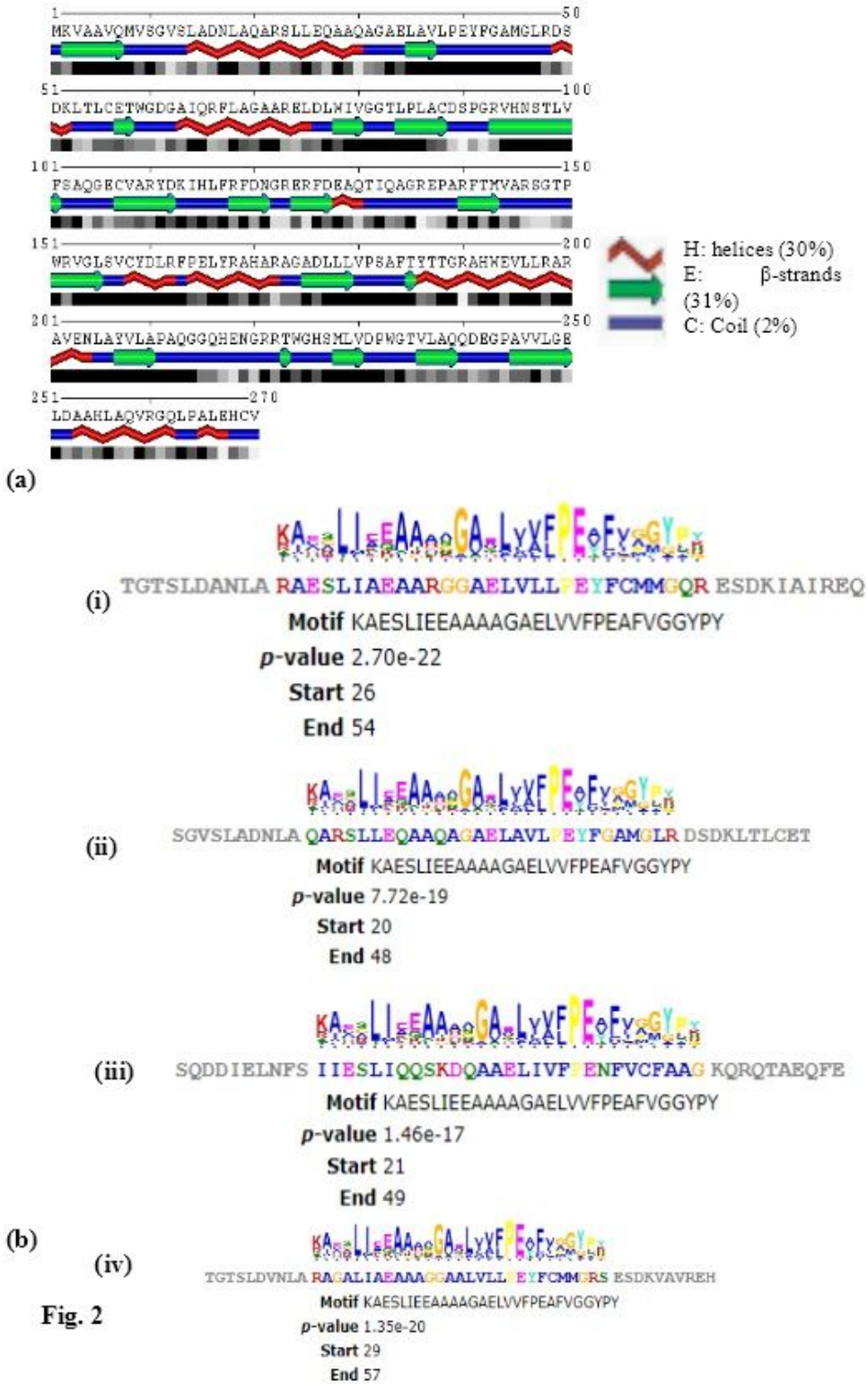


Fig. 2

## Figure 2

(a) Prediction of secondary structures of nitrilase: α-helices and β-strands expressed in helices and arrows, respectively as depicted in POLYVIEW; (b) A comparative location of Motif's for nitrilase (i) AcNit, (ii) As7Nit, (iii) Cn5Nit, and Cn9Nit obtained from MEME.

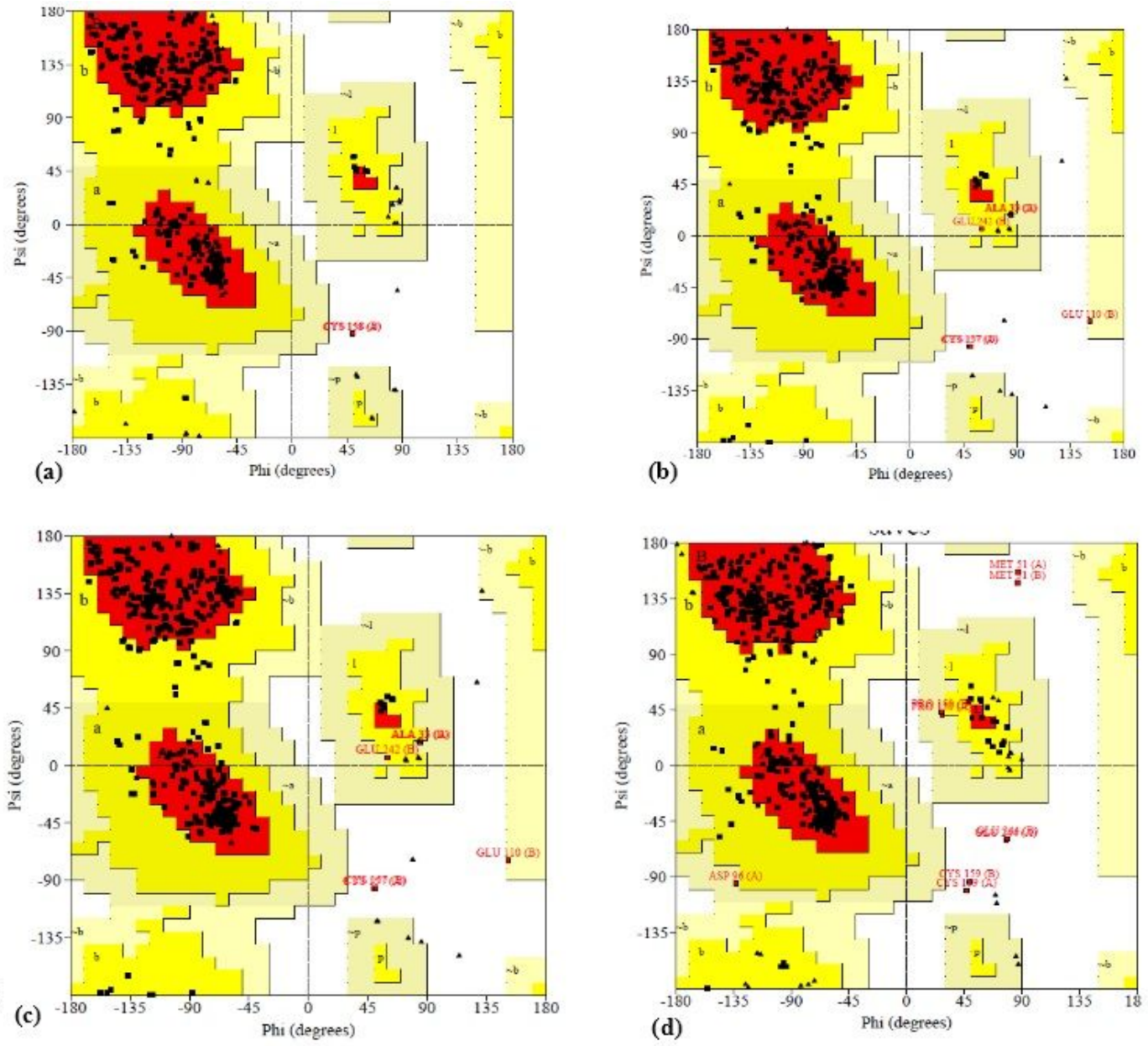


Fig. 3

## Figure 3

Ramachandran plot for the predicted 3D model of nitrilase showing distribution of amino acid residues in the most favored regions and additional allowed regions as black dots, and in the generously allowed and disallowed conformations in red fonts.

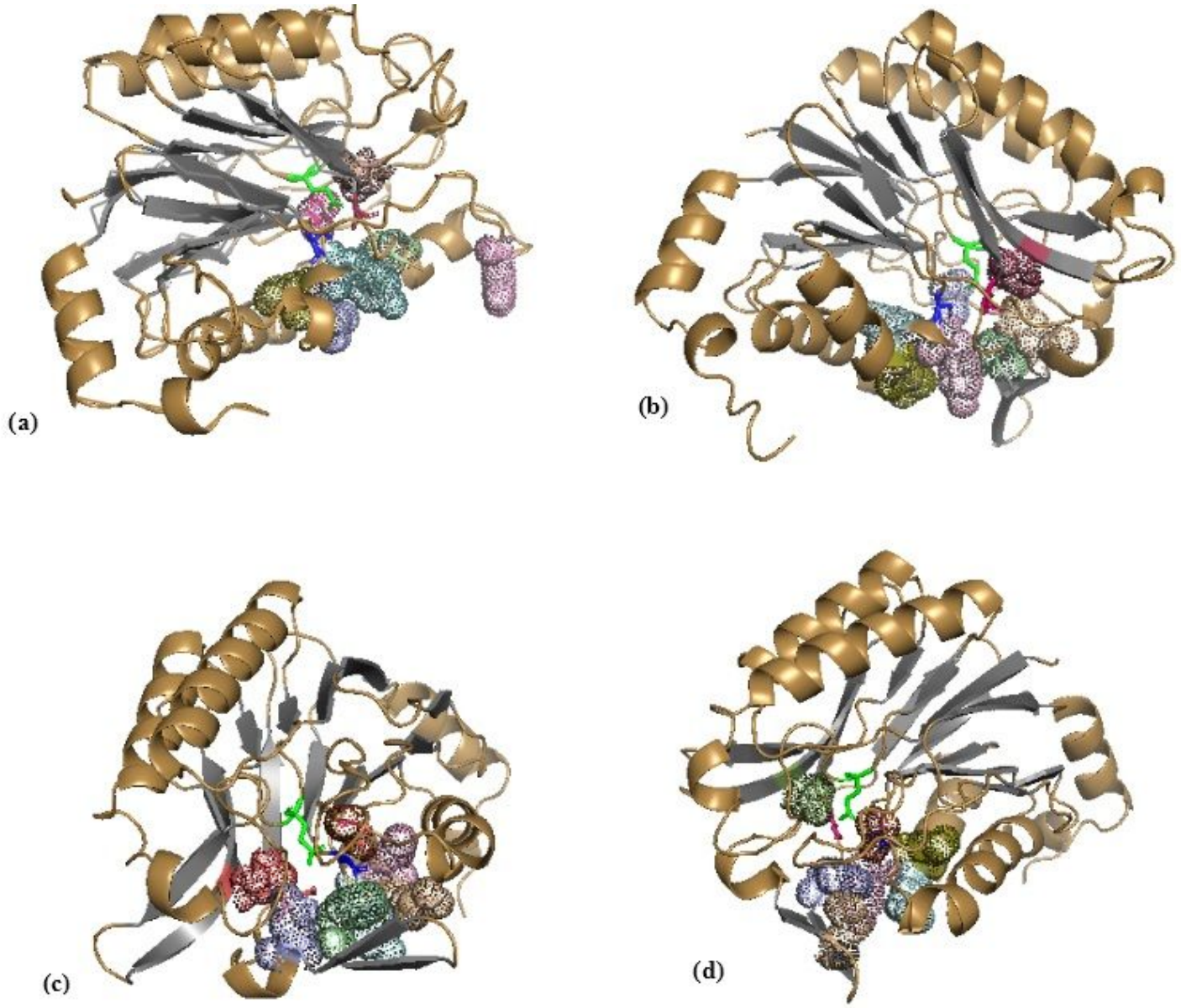


Fig. 4

## Figure 4

Model structure of nitrilases (a) AcNit, (b) As7Nit, (c) Cn5Nit, and (d) Cn9Nit generated using the SWISS-MODEL with 2W1V_A as template. The molecule shows catalytic triad E, K and C in sticks and other active site residues N96, F117, E128, Y159, L161, R162 and A183 as dotted spheres. The graphics were prepared in PYMOL.

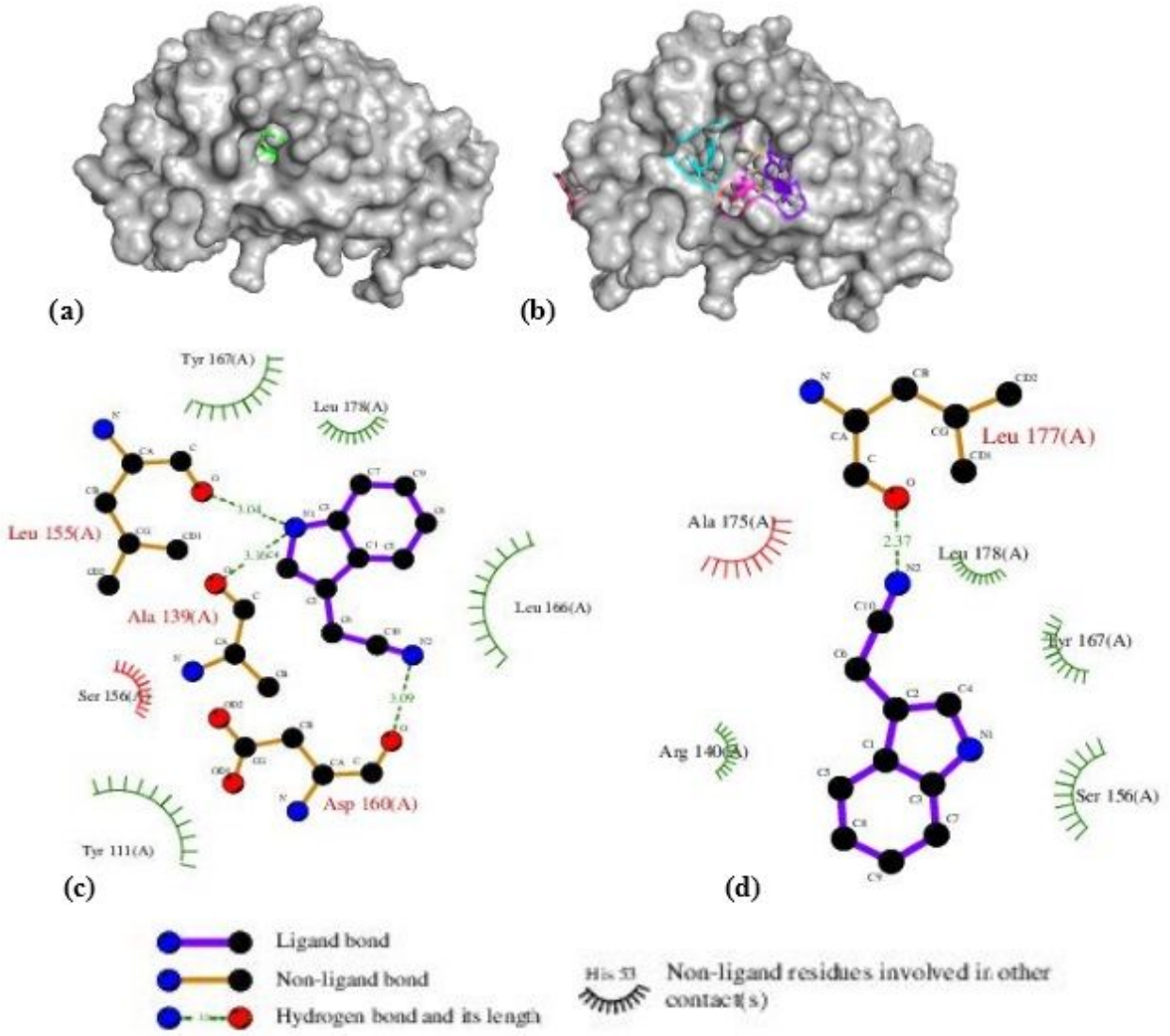


**Fig. 5**

## Figure 5

(a) Surface view of catalytic residues; (b) active site residues forming substrate binding pocket; (c) interaction of AcNit with substrate IAN; and (d) interaction of mutated AcNit at position Y159 to A159 with IAN showing no hydrogen bonding.