Journal of Biological Engineering Research and Review, 2016; 3(2): 08-18
ISSN: 2349-3232 Online
DOI: 10.5281/zenodo.15337568

Research Article

In Silico Identification and Comparative Analyses of Active Sites of Copper Containing Nitrite Reductase (CuNiR) in Fungal and Bacterial Spp.

Utpal Kumar Adhikari, M. Mizanur Rahman *

Department of Biotechnology and Genetic Engineering, Faculty of Applied Science and Technology,

Islamic University, Kushtia-7003, Bangladesh

*Email: mmrahmanbtg79@hotmail.com

ARTICLE INFO:

Article History:
Received: 18/07/2016
Revised: 24/08/2016
Accepted: 25/08/2016
Available Online: 12/09/2016

Keywords:
*Nir*K gene,
CuNiR,
Denitrification,
phylogenetic tree,
molecular model,
active site

Abstract: The *Nir*K gene encoding the copper-containing nitrite reductase (CuNiR) has a key regulatory role in the global denitrification process. In this study, a comparative analysis between fungal and bacterial species have been done to understand the comparison and identification of common amino acid residues of the active sites which may be involved in the denitrification mechanism in both fungal and bacterial species. The sequences were retrieved from NCBI database and further analyzed using different Bioinformatics tools and servers. The amino acid composition, physico-chemical properties and subcellular localization showed a distinctive result for fungal and bacterial species. There was no significant sequence similarity found among these species, but the phylogenetic tree showed that they have come from a common ancestor gene. The active site analyses showed that the amino acid residues namely Valine, Leucine, Threonine, Arginine, Glycine, Tyrosine, Serine, Proline, Asparagine and Glutamine were conserved in both the fungal and bacterial species, but the dissimilarity of amino acid composition (%) exists among these most conserved amino acid residues. It is expected that this information will be helpful for further research on the molecular mechanism of the denitrification process in the wet laboratory.

INTRODUCTION

Denitrification process plays an important role in the global nitrogen cycle. It involves several reaction cycles. It produces some gaseous components such as nitrous oxide (N_2O) which is hazardous for the environment as well as for the human and animal beings [1]. The pathway mainly completes the following reaction $NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$, each step is catalyzed by the specific reductase enzyme namely nitrate (NO_3^-) reductase, nitrite (NO_2^-) reductase, nitric oxide (NO) reductase and nitrous oxide (N_2O) reductase, respectively [2, 3]. These enzymes encompass the assimilatory and dissimilatory denitrification pathway and contribute to the environment with the help of various denitrifying bacteria and their respective genes. A key enzyme of the denitrification process is copper-containing nitrite reductase (CuNiR) encoded by the *Nir*k gene and responsible for the conversion of nitrite (NO_2) to nitric oxide (NO) which is the most critical step in the denitrification process [4]. The enzymatic or molecular mechanism of heme-containing nitrite reductase is well-known [5] but the exact molecular mechanism of CuNiR enzyme is still unknown to the researchers [4, 6]. The denitrification usually occurs in the denitrifying prokaryotic bacterial species [2], but it is also found in eukaryotic fungal [7] and archaeal species [8]. The *Nir*K encoding CuNiR has been found in several bacterial species, including *Phaeobacter inhibens*, *Burkholderia oklahomensis*, *Moraxella catarrhalis* [4, 9, 10] and fungal species, including *Aspergillus oryzae* [7], *Fusarium oxysporum* [11], and *Fusarium lichenicola*. There are blue and green colored species are available among the CuNiR containing bacterial species [6, 12]. It is supposed that the color of the protein is regulated by the Met144, Tyr171 and Thr171 residues [13].

The microorganisms containing the CuNiR are naturally found in low oxygen containing environments [14]. The subcellular localization of the bacterial CuNiR is mainly periplasm [2]. The fungal denitrification process occurs in the mitochondria and at the same time it combines the electron transport chain (ETC) of mitochondria to produce ATP for the function [15, 16]. The identification of subcellular localization in a species is problematic and time consuming process, but various Bioinformatics tools has made it easy for the researcher [17, 18]. The diversity of *Nir*K encoding CuNiR enzyme in bacterial species has seen in agricultural soils as well as in other ecosystems [19, 20]. The multiple sequence alignment and diverse phylogenetic analysis, among fungal and bacterial species show the significant sequence identity and evolutionary relationship, but do not ascertain the actual

relationship due to the lack of information about fungal genetic component [16].

The CuNiR enzyme is comprised of three indistinguishable monomers that closely interconnected together [21]. Every CuNiR monomer consists of one type I and one type II copper site which are distinguished through the relevant spectroscopic signatures [22, 23]. These two important sites are attached for fast electron transmission using a Cysteine-Histidine bridge [24]. The type I copper atoms is buried within the protein's N-terminal domain and each one is synchronized by His95 and His145, Cys136, and Met150 residues [24, 25]. The type II copper is extremely related to nitrite reduction and organized by a water molecule and three Histidine ligands (His100, His135 and His306) whereas two Histidine residues, His100 and His135 are provided by N-terminal and His306 is provided by C-terminal domain [4, 22]. The distance between the type I and type II copper site is approximately 12.5 Å and are connected through Cys136 and His135 ligand [21, 22].

The experiment in the native structure of *Alcaligenes faecalis* CuNiR revealed that nitrite binds through two active oxygen molecules and forms hydrogen bond interaction with the side chain of Asp98 that associates with His255 via solvent- bridges and these interactions are required for the catalysis [4, 24, 25]. The role of Asp92 and His254 residues have been investigated in the *Alcaligenes xylosoxidans* CuNiR using point mutation, structure determination and enzymatic activity measurement. In this species Asp92 creates hydrogen bonds to the water molecules at the nitrite binding sites which are also mediated by His254 [13]. The metalloenzymes especially CuNiR enzyme completes complex catalytic mechanism for the denitrification process and the mechanism has been inspected in several species using different techniques [26]. But the homology modeling and the analyses of ligand binding sites have only been done for *Pseudomonas chlororaphis* subsp. *Aureofaciens* [27].

Therefore, the active site information on CuNiR enzyme is very essential for understanding the catalytic or molecular mechanism of denitrification process. Various wet lab based experiment has been conducted for a few species to understand the catalytic or molecular mechanism but not for other CuNiR containing species. In that case, we have intended to conduct a comparative research between fungal and bacterial species for understanding the evolutionary relationships, various physico-chemical properties, structure and amino acid composition in the active sites of CuNiR enzyme using various Bioinformatics tools and servers.

MATERIALS AND METHODS

Retrieval of target sequences

In the present study, six Fasta formatted protein sequences of fungi and bacteria encoding copper-containing nitrite reductase (*Nir*k gene) were retrieved from NCBI gene bank database for *in silico* analysis. The selected fungi sequences are from *Aspergillus oryzae* (ACZ26344.1), *Fusarium oxysporum* (ABU88100.1), *Fusarium lichenicola* (ADB82651.1) and the bacterial species are *Phaeobacter inhibens* DSM 17395 (AFO93407.1), *Burkholderia oklahomensis* (AIO69883.1) and *Moraxella catarrhalis* (AIK01138.1).

Target sequence analyses of fungal and bacterial species

The target sequences of fungal and bacterial species were subjected to ExPASy ProtParam server [28] for physico-chemical characterization and determination of amino acid composition in the protein sequences. The ExPASy ProtParam is an organized server for various physico-chemical properties of protein such as molecular weight, theoretical isoelectric point (pI value), instability index, aliphatic index, grand average of hydropathy (GRAVY) value, etc. The sequences were also subjected to CELLO v2.5 [29], WOLF PSORT [18] and PSORTb v3.0 server [17] for the prediction of subcellular location and SBASE server for domain prediction [30]. The important five motifs were anticipated among the selected fungal and bacterial species using the MEME (Multiple Em for Motif Elicitation) server version 4.10.1 [31] and MAST server (Motif Alignment and Search Tool) version 4.10.1 [32]. These motifs were further used for the identification of protein family using the Pfam 27.0 server at the Sanger Institute [33].

Secondary structural feature prediction

The secondary structural features of CuNiR encoding *Nir*k gene of fungi and bacteria were predicted using the two important and reliable servers PSIPRED [34] and SOPMA [35].

Multiple sequence alignment and phylogenetic tree construction

Multiple sequence alignment was done by exposing the target sequences of fungi and bacteria to the ClustalW server [36] and the alignment file was reconstructed using the BoxShade server. The homologous sequences were searched using the NCBI Blastp 2.2.32+ to know the evolutionary relationships among bacterial and fungal species. The 49 homologous sequences with 65% to 97% sequence identity encoding CuNiR (*Nir*k gene) were searched using the 6 selected fungal and bacterial species. The designated sequences were aligned in ClustalW server and phylogenetic tree were reconstructed using the MEGA 6.0 software [37] based on the Neighbor-Joining statistical method [38] and Poison model including bootstrap analysis through 1000 replications.

Protein model building, refinement, validation and deposition to PMDB database

The three dimensional protein model of *Nir*k gene encoding CuNiR of fungal and bacterial species was predicted by using the I-TASSER server [39], an iterative threading assembly refinement algorithm for protein structure prediction. This server employs the PPA (Profile-Profile threading Alignment) developed by iterative implementation of the threading assembly refinement program and secondary structure [40]. This server predicts about five models for each query sequence and the best protein model can be identified based on the C-score, TM-score, number of decoys and RMSD value. The selected best models were then subjected to the Swiss-PdbViewer software version 4.1.0 [41] for energy minimization to eliminate bad contacts among the atoms using the steepest descent and conjugate gradient technique.

The minimization process was performed in Vacuo through the GROMOS96 43B1 parameters set.

The energy minimized models were further evaluated by the PDBsum server [42] using PROCHECK [43]. This server uses the Phi/ Psi Ramachandran plot for backbone conformation of the protein model. The protein models were also subjected to the QMEAN server [44], SAVES server (http://nihserver.mbi.ucla.edu/SAVES) for Verify 3D [45] and ERRAT [46] result and ProSA server [47] for calculating the Z-score. The energy minimized and validated protein models were finally deposited to the PMDB (Protein Model Data Base) (http://mi.caspur.it/PMDB) [48] database with the unique PMDB identifier number.

Active sites and function annotation of predicted protein models

The active sites of the modeled protein were identified using the Castp (Computed Atlas of Surface Topography of Protein) servers [49]. Castp predicts active sites as the surface pocket and cavities that contain binding site residues. The protein contains different pockets and cavities with different areas and volumes. The Castp calculate the active pockets or binding sites using the pocket algorithm of the alpha shape theory. The functions were predicted by the I-TASSER server. This server predicts the function based on the local and global relationship to the template proteins in Protein Database (PDB) with well-known structure and functions [50].

RESULTS AND DISCUSSION

Analyses of primary structural sequences

Six revised full length sequences from fungal and bacterial species encoding CuNiR (*Nir*K gene) were retrieved from the NCBI database for *in silico* comparative analyses. The essential and non-essential amino acid compositions were predicted using the ExPASy ProtParam server.

Compositions of essential and non-essential amino acids were calculated as shown in the Table-S1. The amino acids V (Valine) and L (Leucine) were found high percentages among of all essential amino acid compositions in the both bacterial and fungal species. On the other hands, M (Methionine) and W (Tryptophan) were found as low percentages among of all essential amino acid compositions in the both bacterial and fungal species. Similar percentages of essential amino acid composition were found in the both bacterial and fungal species, whereas, percentages of composition of non-essential amino acids were found higher in the bacterial species as compared to fungal species, especially G (Glycine) and A (Alanine). At the same time, it has also been found a lower difference between the amino acids P (Proline) and C (Cysteine) in both the bacterial and fungal species. So, the results revealed that there is significant difference exists between the *Nir*k gene encoding bacterial and fungal species.

The physico-chemical properties of the selected fungal and bacterial species were also calculated using the ExPASy ProtParam server. The predicted results are shown in Table-1. The molecular weight of the selected fungal species was ranged from 49814.5 Da to 51695.6 Da whereas the value ranged from 43644.9 to 54241.2 for the bacterial species. The isoelectric point or theoretical pI value ranged from 5.02 to 6.52 among the selected fungal and bacterial species. Theoretical pI of a protein typically depends on amino acid sequences. This property is a pH indicator of the protein and it occurs when the protein contains zero net charge. The bacterial species *P. inhibens* and *B. oklahomensis* contains the lowest (5.02) and highest (6.52) value, respectively. The results clearly exhibit the weakly acidic nature of the protein for both the selected fungal and bacterial species.

Instability index is the organized form of dipeptides which specifically depends upon the length of the protein or enzymes. The selected fungal species contain an instability index value from 39.70 to 45.98 while the bacterial species contain 31.34 (*M. catarrhalis*) to 36.94 (*P. inhibens*). The fungal species *F. lichenicola* encompasses highest (45.98) instability index value. The Instability index value above 40 confirms the instability nature but the value less than 40 reveals the stability nature of the overall 3D structure of the protein [28]. So, the results revealed that the selected fungal species are instable (except *A. oryzae*) and the bacterial species are stable in nature.

Table 1: Physico-chemical properties of CuNiR encoding *Nir*k gene of fungal and bacterial species

Species name	No. of Amino acid	Molecular weight	Theoretical pI	Instability index	Aliphatic index	GRAVY
A. oryzae	452	49814.5	6.36	39.70	78.27	-0.316
F. oxysporum	496	53704.9	6.35	41.41	78.99	-0.263
F. lichenicola	483	51695.6	6.14	45.98	77.52	-0.220
P. inhibens	407	43654.9	5.02	36.94	78.67	-0.188
B. oklahomensis	518	54241.2	6.52	32.27	78.38	-0.103
M. catarrhalis	408	43644.9	5.52	31.34	73.92	-0.318

The aliphatic index value expresses the relative value of the protein, which is occupied by the aliphatic side chains. The aliphatic index value controls the thermo-stability and structural stability of the enzyme. In this *in silico* sequence analysis the aliphatic index value ranged from 73.92 (*M. catarrhalis*) to 78.99 (*F. oxysporum*) for the fungal and bacterial species. The thermo-stability increases with the increasing of the aliphatic index value. Therefore, high aliphatic index values deduce that CuNiR proteins of these fungal and bacterial species will be capable to exhibit their effectiveness in extensive temperature due to their stable properties.

The grand average hydropathy (GRAVY) value for a protein designate the hydrophobicity or hydrophilicity nature

on the basis of positive and negative value, respectively [29]. The negative GRAVY values ranged from -0.103 (*B. oklahomensis*) to -0.318 (*M. catarrhalis*) and revealed that all of the selected fungal and bacterial CuNiR proteins analyzed in this study are hydrophilic in nature.

The subcellular location within fungi or bacteria provides essential information about the cellular function and possible interaction partners of the protein [30]. The subcellular locations of the selected fungal species (*A. oryzae, F. oxysporum, F. lichenicola*) were extracellular and mitochondrial as predicted by CELLO v2.5 and other subcellular localization prediction server WOLF PSORT predicted the location in mitochondria and extracellular position with a significant confidence score. The subcellular locations for the bacterial species (*P. inhibens*, *B. oklahomensis*, and *M. catarrhalis*) were periplasmic as predicted by PSORTb v3.0 which is also confirmed by the CELLO v2.5 server. So, the prediction result of different servers ensures the several locations for fungal and bacterial species.

In domain analysis, the fungal species *A. oryzae*, *F. oxysporum*, and *F. lichenicola* exhibits multi-copper oxidase type-1 and multi-copper oxidase type-2 like domains with a 99% confidence score by the SBASE sever whereas the bacterial species *P. inhibens* shows 100% confidence score for these two domains. The other two bacterial species *B. oklahomensis* and *M. catarrhalis* reveals three domains namely, multi-copper oxidase type-1, multi-copper oxidase type-2 and cytochrome C class-1 like domain with a 99% confidence score. The domain sites were not in the similar positions in the sequences for all of these selected bacterial and fungal species.

Five conserved motifs from the selected fungal and bacterial sequences were predicted using the MEME program is shown in Table-2. The Pfam analysis showed motif 1 and motif 2 were found to be interrelated with Cu-oxidase_3 and Cu-oxidase_ 2 domain structure, respectively. But the motif 3, motif 4 and motif 5 did not exhibit any correlation with any recognized domain structure. The motif 1 and motif 4 are present in six selected fungal and bacterial species while the motif 2, motif 3 and motif 5 are present in three fungal species (*A. oryzae*, *F. oxysporum*, and *F. lichenicola*) and two bacterial species (*B. oklahomensis* and *M. catarrhalis*) (Figure-S1). So, the motif 2, 3, and 5 are not present in *P. inhibens* but the MAST result showed that there is repeated motif (motif 1) present in *P. inhibens*. The motif is important for the protein and the relationship or variation among the motifs can play an important role for the protein metabolism, protein function or in protein structure [31, 32]

Table 2: List of predicted motif sequences among CuNiR sequences of fungal and bacterial species and the respective families

Motif	Best Possible Match (motifs)	Width	Sites	Family
1.	TFNGSVPGPFIRVREGDIVELSLTNKDDSKMPHNIDCHAFTGPGGG	46	6	Cu-oxidase_3
2.	ETKTFRFKLLYPGLYVYHCAVAPVGMHIANGMYGLMYVQPE	41	5	Cu-oxidase_ 2
3.	DKPLKAKVGDTVRIFFGNGGPNLTSSFHIIGTHFDKVYRDG	41	5	Not found
4.	GIQTWSVPCGGSTIVDFKMEVPGTYTLVDHAIFR	34	6	Not found
5.	LPPVDKEYYVMQSEFYHEPPEVDDDGQRS	29	5	Not found

Secondary structural features analysis

The secondary structural features are important to estimate about the protein three dimensional structures. In this *in silico* study, the secondary structural features for the fungal and bacterial species were predicted by the PSIPRED and SOPMA server are shown in Table-3. The predicted results showed that, the random coils are dominated to over alpha helix and extended strands in the selected fungal and bacterial species. In PSIPRED, the random coils were ranged from 60.85% (*A. oryzae*) to 68.94% (*F. lichenicola*) for fungal species whereas the value was 59.95% (*P. inhibens*) to 64.29% (*B. oklahomensis*) in bacterial species but in SOPMA, the random coils were ranged from 49.34% (*A. oryzae*) to 55.07% (*F. lichenicola*) for fungal species and 48.84% (*B. oklahomensis*) to 50.74% (*M. catarrhalis*) for bacterial species. The PSIPRED showed a very less amount of alpha helix while the SOPMA showed significant amount of alpha helix for both fungal and bacterial species. So, the result revealed that there are significant difference exists in the PSIPRED and SOPMA predicted results. But when the three dimensional structures were predicted, the structures showed the alpha helix, extended strands and random coils according to the PSIPRED prediction results.

Multiple sequence alignment and phylogenetic tree analysis

Multiple sequence alignment among the selected fungal and bacterial species of CuNiR encoding *Nir*k gene were performed using ClustalW server and reconstructed by the BoxShade server. The reconstructed multiple sequence alignment file is shown in Figure-S2. The total alignment score was 13916 for these sequences. Multiple sequence alignment exhibits that fungal species *A. oryzae* with *F. oxysporum*, *F. lichenicola*, *P. inhibens*, *B. oklahomensis*, and *M. catarrhalis* were similar 53.76%, 53.54%, 20.39%, 29.65% and 30.39%, respectively. In similarly, *F. oxysporum* with *F. lichenicola*, *P. inhibens*, *B. oklahomensis*, *M. catarrhalis* were similar 77.02%, 22.11%, 27.42% and 30.14%, respectively while *F. lichenicola* with *P. inhibens*, *B. oklahomensis*, *M. catarrhalis* were similar 22.36%, 28.78% and 30.39%, respectively. The bacterial species *P. inhibens* with *B. oklahomensis*, *M. catarrhalis* were similar 23.59% and 19.66%, respectively and finally *B. oklahomensis* show 61.76% similarity with *M. catarrhalis*. A significant similarity is found between the fungal species *F. oxysporum* and *F. lichenicola* but there is no significant similarity exists among the bacterial species. So, the results revealed that very lowest similarity exists between the fungal and bacterial species.

Table 3: The secondary structural features of *Nir*K gene encoding CuNiR of fungal and bacterial species

Species name	PSIPRED			SOPMA		
	Alpha-helix (%)	Extended strand (%)	Random coil (%)	Alpha-helix (%)	Extended strand (%)	Random coil (%)
A. oryzae	7.07	32.08	60.85	21.90	28.76	49.34
F. oxysporum	5.85	30.64	63.51	14.92	30.24	54.84
F. lichenicola	2.90	28.16	68.94	16.15	28.78	55.07
P. inhibens	15.23	24.82	59.95	21.87	28.01	50.12
B. oklahomensis	15.83	19.88	64.29	25.10	26.06	48.84
M. catarrhalis	15.45	22.79	61.76	21.32	27.94	50.74

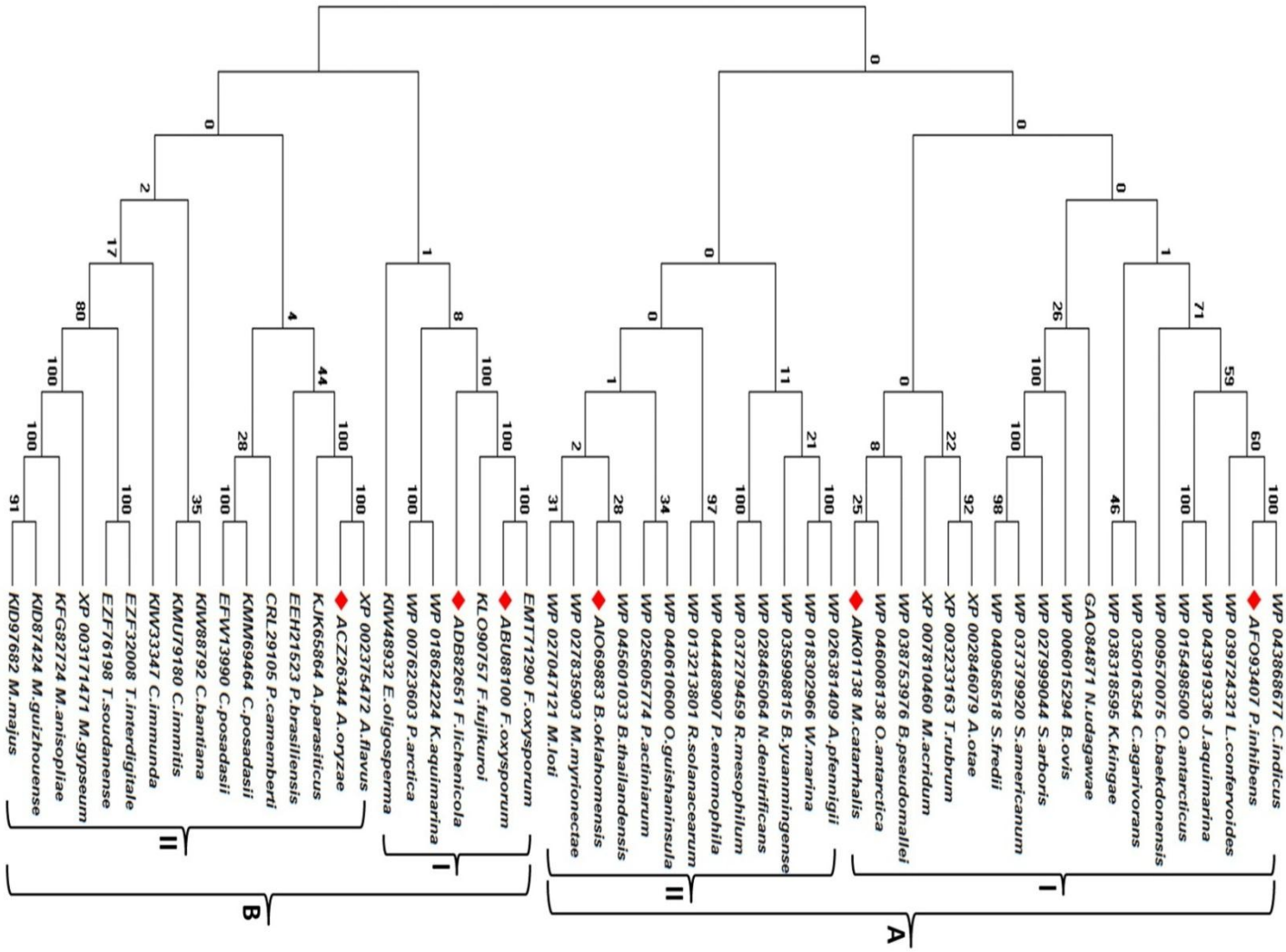

Figure 1: The phylogenetic tree among the fungal and bacterial *Nir*k gene encoding CuNiR enzyme sequences. The red colored bullets indicate the selected sequences of our study. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = 48.70947419 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap tests (1000 replicates) are shown next to the branches. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved 55 amino acid sequences. Evolutionary analyses were conducted in MEGA6.

The phylogenetic tree was constructed among the 6 reviewed sequences and 49 homologous sequences (65% to 97% sequence identity) of fungal and bacterial species of *Nir*k encoding CuNiR enzyme. The retrieved Fasta formatted sequences were subjected to multiple sequence alignment by ClustalW server and the evolutionary relationships were analyzed by Neighbour-Joining method and bootstrap value through the MEGA 6.0 software. The phylogenetic tree clearly

showed that the selected species are mainly divided into two major Clusters namely A and B. Cluster A is further divided into two other sub-clusters namely I and II and encloses mainly selected bacterial species namely, *P. inhibens*, *B. oklahomensis*, and *M. catarrhalis* but the Cluster A also containing four fungal species namely *Arthroderma otae, Metarhizium acridum, Neosartorya udagawae* and *Trichophyton rubrum*. Cluster B is also distributed into two sub-clusters namely I and II and consists of primarily selected fungal species *A. oryzae, F. oxysporum* and *F. lichenicola* and at the same time containing two bacterial species namely *kangiella aquimarina* and *Paraglaciecola arctica*. The bacterial species *P. inhibens* exhibited high similarity with *Celeribacter indicus,* whereas *M. catarrhalis* exhibited relationship with *Oleispira antarctica* within sub-cluster I of major Cluster A. The other bacterial species *B. oklahomensis* showed a relationship with *Burkholderia thailandensis*, *Maritalea myrionectae* and *Mesorhizobium loti* within sub-cluster II of major Cluster A (Figure-1). The fungal species *F. oxysporum* and *F. lichenicola* showed similarity with *Fusarium oxysporum* f. sp. *cubense race* 4, *Fusarium fujikuroi*, and bacterial species, namely *Paraglaciecola arctica* and *Kangiella aquimarina* with significant bootstrap value within sub-cluster I of major Cluster B. These two bacterial species *P. arctica* and *K. aquimarina* have come from two different orders, namely Alteromonadales and Oceanospirillales, respectively, whereas the fungal species *F. oxysporum* and *F. fujikuroi* are from Hypocreales order, but positioned in the same cluster. The last fungal species *A. oryzae* exposed high similarity with *Aspergillus flavus*, *Aspergillus parasiticus*, *Penicillium camemberti* that have come from the same Eurotiales order and *Coccidioides posadasii* from Onygenales order within sub-cluster II of major Cluster B (Figure-1). So, the phylogenetic tree analysis results showed significant evolutionary relationships among the selected fungal and bacterial species with the predicted CuNiR homologous sequences and ensured that they have come from a common ancestor gene.

Three-dimensional structure prediction of CuNiR enzymes

There is no three-dimensional structures available in the Protein Data Bank (PDB) for our selected bacterial and fungal species and in that case we intended to build the three dimensional model of CuNiR enzymes. The protein three-dimensional models provide necessary information about the active sites and molecular and biological function of the enzymes.

The three-dimensional structures were predicted using the I-TASSER server, which construct the model using multiple threading methods. This server predicts several templates for protein modeling using multiple threading programs, namely, MUSTER, FFAS-3D, SPARKS-X, HHSEARCH2, Neff-PPAS, HHSEARCH I, wdPPAS, HHSEARCH, cdPPAS, and pGenTHREADER. I-TASSER server uses the templates, which have significant Z-score. The templates used in this protein model prediction are shown in Table-4.

Table 4: List of top ten templates used by I-TASSER for 3D structure predictions of CuNiR enzyme

Species name	Templates
A. oryzae	2dv6A, 1n70A, 2dv6A, 1kbvA, 2dv6A, 1kbvA, 2dv6A, 1kbvA, 2dv6A, 1kbvA
F. oxysporum	2dv6A, 1n70A, 2dv6A, 1kbvA, 2dv6A, 1kbvA, 2dv6A, 1kbvA, 2dv6A, 2bw4A
F. lichenicola	2dv6A, 1mzzA, 2dv6A, 3ziyA, 2dv6A, 1kbvA, 2dv6A, 1kbvA, 2dv6A,, 2bw4A
P. inhibens	2dv6A, 1kbvA, 2dv6A, 1mzzA, 2dv6A, 2xwzA, 2dv6A, 2bw4A, 2bw4A, 1kbvA
B. oklahomensis	3ziyA, 3ziyA, 2zooA, 3ziyA, 3ziyA, 3ziyA, 3ziyA, 2zooA, 3ziyA, 1j9qA
M. catarrhalis	3ziyA, 3ziyA,2zooA, 3ziyA, 3ziyA, 3ziyA, 3ziyA, 3ziyA, 3ziyA, 4knsA

Table 5: Model evaluation data for the predicted structures of the six CuNiR enzymes of fungal and bacterial species.

Species Name	C-Score	Exp. TM-Score	Exp. RMSD (Å)	No. of Decoys	Cluster Density
A. oryzae	-0.89	0.60±0.14	9.1±4.6	1414	0.1059
F. oxysporum	-1.55	0.52±0.15	10.9±4.6	765	0.0545
F. lichenicola	-0.73	0.62±0.14	8.9±4.6	1776	0.1231
P. inhibens	-1.05	0.58±0.14	9.2±4.6	1364	0.0941
B. oklahomensis	-0.75	0.62±0.14	9.1±4.6	1306	0.0980
M. catarrhalis	1.51	0.92±0.06	3.7±2.6	2400	1.0000

This server predicts maximum five models for each protein sequence. The model quality is estimated quantitatively by C-score (Confidence score) based on the significance of threading template alignments and convergence parameters of the structure assembly simulations. The structural similarity between template and predicted model are measured by the TM-score (Template Modeling Score) and RMSD when the native structure is known [33]. In this *in silico* analysis, I-TASSER server predicts five models for each *Nir*K encoding CuNiR enzyme sequence. The best protein models were selected and predicted based on the maximum C-score, TM-score (TM-score> 0.5) and number of decoys which were further used for evaluation and validation purposes. The C-score of the predicted model were ranged from -0.73 (*F. lichenicola*) to 1.51 (*M. catarrhalis*) whereas the TM-score were ranged from 0.52±0.15 (*F. oxysporum*) to 0.92±0.06 (*M. catarrhalis*). The selected fungal and bacterial species contain the RMSD (Root Mean Square Deviation) score from 3.7±2.6

(*M. catarrhalis*) to 10.9±4.6 (*F. oxysporum*) are shown in Table-5.

The predicted models were initiated in correct topology based on the C-score, TM-score, RMSD and number of decoys. The C-score ranges from -5 to 2 is deliberated as a significant score, but the highest score implies a model with a greater confidence. Generally, a TM-score >0.5 indicates a model of correct topology and a TM-score < 0.17 means a random similarity and these cutoff does not depends on the protein length [34]. According to the server fixed criteria the C-score for the protein model of *A. oryzae*, *F. lichenicola*, *P. inhibens*, and *B. oklahomensis* contains low score, but the TM-score, RMSD and number of decoys showed the significance score for all the selected protein models. So, by analyzing all the results it can be said that the predicted models can be used for further purposes.

The selected I-TASSER predicted protein models were then subjected to Swiss-PdbViewer software for energy minimization purposes and to remove the bad contacts among the atoms. The energy (Kj/Mole) value before and after energy minimization for each *Nir*k gene encoding CuNiR enzymes are shown in Table-6.

Table 6: Model evaluation scores and energy of the model of before and after energy minimization

Species name	ProSA (Z-Score)	ERRAT	Verify3D (%)	Energy Before Minimization (Kj/Mol)	Energy After Minimization (Kj/Mol)
A. oryzae	-4.22	65.000	69.03	-471.984	-14346.392
F. oxysporum	-3.81	74.021	64.52	696.105	-17069.479
F. lichenicola	-4.85	64.119	74.74	-854.455	-17411.957
P. inhibens	-5.55	81.818	75.68	-9585.484	-16127.387
B. oklahomensis	-6.99	80.315	92.66	-10603.396	-18351.020
M. catarrhalis	-8.27	85.930	94.12	-11244.015	-16325.650

Validation of the predicted three-dimensional structures and deposition to PMDB database

The validation of these six fungal and bacterial protein models was carried out by performing PROCHECK for complete symmetrical study along with the stereochemical exactness of the modeled protein structure through evaluating residue by residue geometry and comprehensive structural geometry. Ramachandran plots using PROCHECK showed the variability in most favored regions and also in additionally allowed, generously allowed and disallowed region which was determined for each refined 3D model. The general maximum likelihood of finding residues of protein (more than 90%) in the core region as revealed by Ramachandran plot suggests better stereochemical quality [35]. The Ramachandran plot statistics for the selected protein model are shown in Table-7 and the Figure-S3.

Table 7: Ramachandran plot statistics of six predicted protein model of fungal and bacterial species

Species name	Ramachandran plot statistics			
	Most Favored Region	Additional Allowed Region	Generously Allowed Region	Disallowed Region
A. oryzae	54.8	33.1	6.3	5.8
F. oxysporum	61.0	29.5	5.4	4.1
F. lichenicola	57.0	30.0	6.9	6.1
P. inhibens	68.4	25.7	3.3	2.7
B. oklahomensis	74.3	20.5	3.3	1.9
M. catarrhalis	75.9	20.8	1.5	1.8

The PROCHECK analysis results designate that a moderately low percentage of amino acid residues has phi/psi angles in the disallowed regions of bacterial species namely *P. inhibens*, *B. oklahomensis*, and *M. catarrhalis* but the fungal species contain high disallowed regions. The percentage of amino acid residues in the core or allowed region were found to be 94.2%, 95.9%, 93.9%, 97.4%, 98.1% and 98.2% for *A. oryzae, F. oxysporum, F. lichenicola, P. inhibens, B. oklahomensis,* and *M. catarrhalis*, respectively, whereas the amino acid residues in disallowed regions were found to be 5.8% (22), 4.1% (17), 6.1% (24) for fungal species *A. oryzae, F. oxysporum, F. lichenicola*, respectively, and 2.7% (9), 1.9% (8), 1.8% (6) for bacterial species *P. inhibens, B. oklahomensis*, and *M. catarrhalis*, respectively, as shown in Table-7. So, the predicted protein model showed quite significant stereochemical quality.

The predicted protein models were further evaluated by the QMEAN server. The QMEAN Z-score were predicted as 0.387(-4.53), 0.396 (-4.34), 0.529 (-2.78), 0.456 (-3.76), 0.469 (-3.48) and 0.562 (-2.48) for *A. oryzae, F. oxysporum, F. lichenicola, P. inhibens, B. oklahomensis*, and *M. catarrhali*, respectively. The predicted significant QMEAN score ensures the quality of *Nir*k gene encoding CuNiR enzymes.

ProSA was used to identify potential errors in the 3D model of a protein. The ProSA server displayed two important features of the structure, one is z-score and the other is residues energy plot. The negative value of the z-score indicates the accuracy of the predicted protein model. The Z-score of the predicted model lies between -3.81(*F. oxysporum*) to -8.27 (*M. catarrhalis*) are shown in Table-7 which indicated the high accuracy of the protein model. The Verify_3D is used to estimate the accuracy of a predicted model which can be

assessed by comparing between 3D profile and its own amino acid sequence (1D). It also compares the predicted structure with a good protein structure [36]. The results revealed that the protein model of *B. oklahomensis* and *M. catarrhalis* showed above 90% score, whereas the other fungal and bacterial species showed acceptable score for the protein model (Table-6). The result indicates the protein model accuracy. ERRAT evaluate the predicted protein model by analyzing the information of non-bonded interactions between miscellaneous atom types and finally determine the results by comparing with highly refined protein structures [37]. The results showed that the value lies between 64.119 (*F. lichenicola*) to 85.930 (*M. catarrhalis*) which are quite significant score for the protein model quality (Table-6).

After completing the evaluation processes the protein models were deposited in the PMDB database and assigned to a unique identifier number for *A. oryzae* (PM0080330), *F. oxysporum* (PM0080331), *F. lichenicola* (PM0080332), *P. inhibens* (PM0080333), *B. oklahomensis* (PM0080334), and *M. catarrhalis* (PM0080335). The predicted final protein models of *Nir*K encoding CuNiR enzyme of fungal and bacterial species are shown in Figure-2.

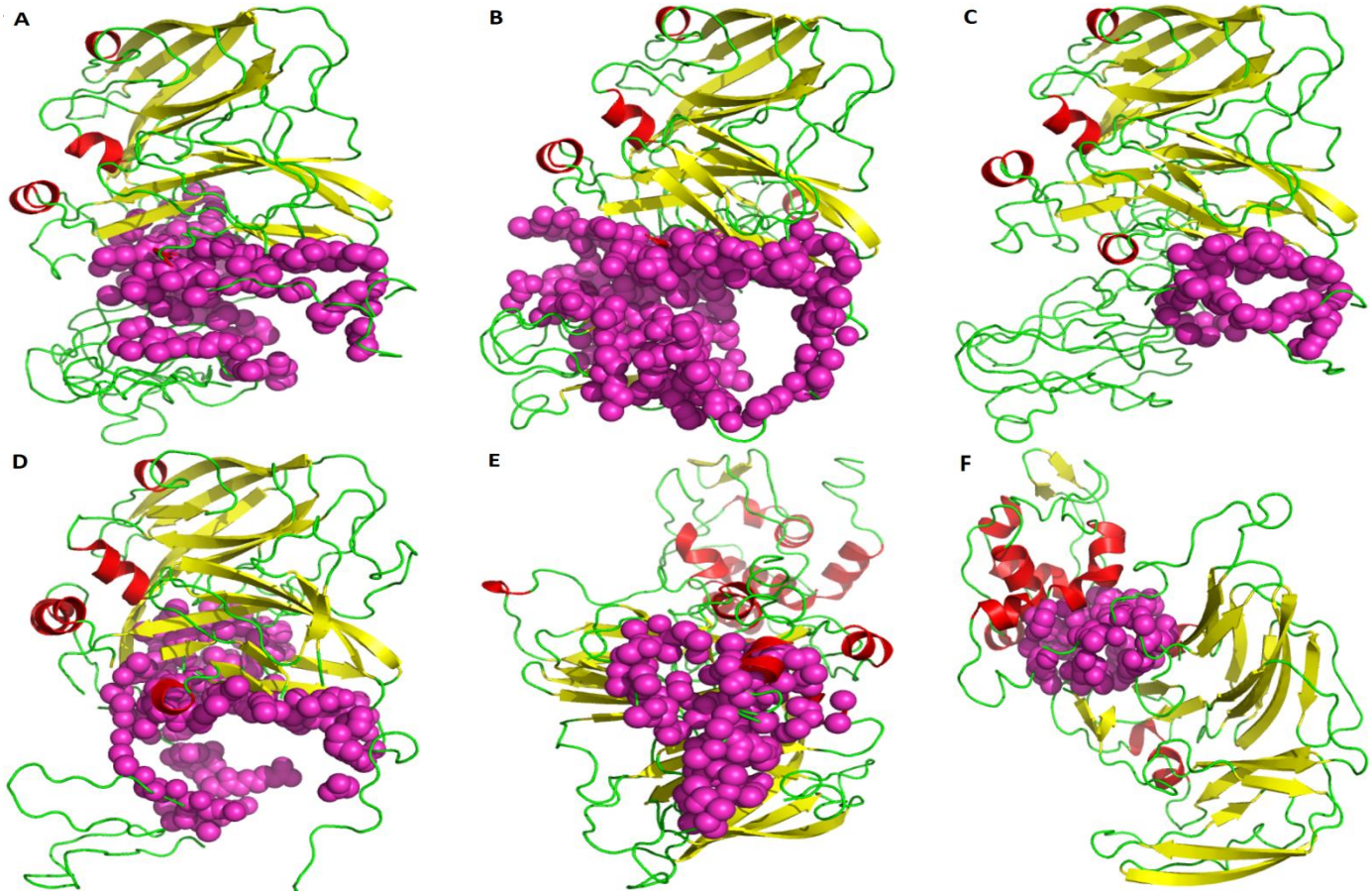

Figure 2: Three-dimensional structures and active sites of the CuNiR protein visualized by PyMOL tools. (A) *A. oryzae* (B) *F. oxysporum* (C) *F. lichenicola* (D) *P. inhibens* (E) *B. oklahomensis* (F) *M. catarrhalis*. Helices, sheets and coils are colored in red, yellow and green, respectively and the active sites are shown in magenta color.

Active sites and function analyses of *Nir*K encoding CuNiR enzymes

In this *in silico* analysis the predicted models were medium resolution models and resolutions were ranged from 2.6 to 4.6 Å RMSD. These medium resolution models can be used for the identification of the central locations of functionally essential residues including binding or active sites of the protein [38]. The active sites of *Nir*K encoding CuNiR enzymes in fungal and bacterial species were predicted using Castp (Computed Atlas of Surface Topography of Proteins) server. This online server predicts surface topography and measures the surface area and volume of all the expected pockets and cavities on the three dimensional structure of protein.

The Castp server predicted 60, 71, 69, 66, 79 and 80 pockets and cavities for the CuNiR enzymes of *A. oryzae, F. oxysporum, F. lichenicola, P. inhibens, B. oklahomensis,* and *M. catarrhalis*, respectively. But the protein active sites which have the largest surface area and volume were taken under consideration for comparative analysis because the large pockets with high volume most probably contain the enzyme binding site [39]. The volumes were 4577.7, 8484, 1727.2, 5620.2, 1955.4, 1303.9 and the surface area 1989.9, 3151, 671.4, 1875.7, 987.7, and 924.8 which were respectively denoted to *A. oryzae, F. oxysporum, F. lichenicola, P. inhibens, B. oklahomensis*, and *M. catarrhalis*. The active sites of the protein are made up of amino acid residues. So, the amino acid residues were taken under considerations which were

present in all the selected fungal and bacterial CuNiR enzymes. The active site information was visualized and computed by using the PyMOL Castp plugin tool. The most conserved amino acid residues of the predicted protein models were assembled to make a graph for comparing among the active site residues (Figure-3). In this study, the active sites contain most conserved essential amino acid residues Val (V), Leu (L), Thr (T), and Arg (R) and non-essential amino acid residues Gly (G), Tyr (Y), Ser (S), Pro (P), Asn (N) and Gln (Q) (Table-S2). The active sites also contain high quantity of Ala (A) residues for all bacterial species, but not present in fungal species *F. lichenicola*. The selected fungal species contain high quantity of Lys (K) but is absent in bacterial species *M. catarrhalis*.

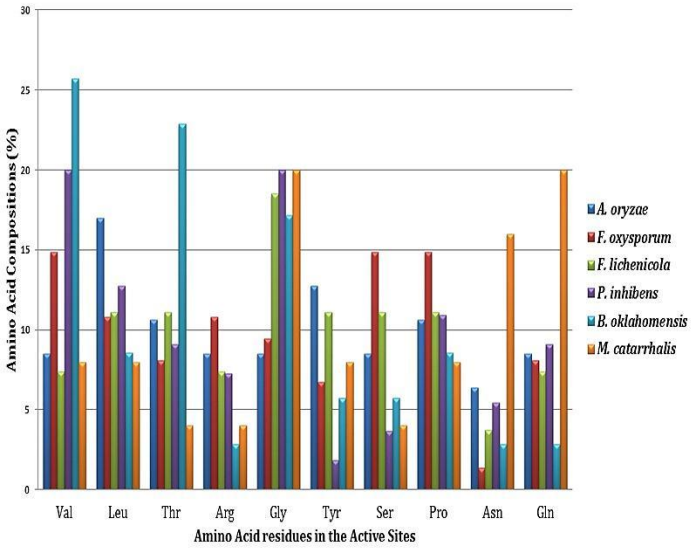

Figure 3: Most conserved amino acid composition in the active sites of *Nir*k encoding CuNiR enzyme. The different colored bars represent the six selected CuNiR sequences. The X-axis represents the amino acid residues while the Y-axis represents the percentage of amino acid composition in the active sites of each of the CuNiR sequences.

The given chart (Figure-3) compares the amino acid composition and revealed that there are distinctive amount of amino acid exists in the active sites of predicted fungal and bacterial *Nir*K encoding protein three dimensional structures. The further analysis of the active sites of fungal species showed 230 amino acid residues while the bacterial species include 178 amino acid residues in the active sites. The most conserved amino acid residues in the fungal species have 64.36% effect while the bacterial species have 65.18% effect to the other amino acid residues in the active sites. So, the effect of other non-conserved amino acid of the fungal and bacterial species is 35.64% and 34.82%, respectively in the formation of active sites which are nearly similar value for fungal and bacterial species. The results accomplish that, a significant variance exists between the fungal and bacterial CuNiR enzyme in the amino acid composition of active sites. At the same time it can be assumed that the predicted most conserved amino acid residues can contribute to the formation of active sites. These amino acid residues can also contribute to the molecular mechanism of action of *Nir*K encoding CuNiR enzyme in both fungal and bacterial species.

I-TASSER server predicted the gene ontology (GO) terms for the query protein based on the functional homology score. Each CuNiR protein model is associated with multiple GO terms, but the GO terms were taken under considerations which have highest functional homology scores. Each GO term denotes a particular function. The CuNiR proteins of *A. oryzae, F. oxysporum, F. lichenicola, P. inhibens, B. oklahomensis*, and *M. catarrhalis* were associated with GO terms GO: 0005507, G0: 0050421, GO: 0055114 and GO: 0042128 known to be copper ion binding, nitrite reductase (NO-forming) activity, oxidation-reduction process, and nitrate assimilation, respectively with significant score (Table-S3). These predicted functions are highly associated and significant with the CuNiR enzymes.

CONCLUSION

In this study, comparative analyses have been carried out to identify the relationship between CuNiR containing fungal and bacterial species. The comparative analyses revealed the distinctive result in amino acid composition and instability index, but they are acidic, stable at high temperature and have good interaction with water molecules. All of these species contain two common domains and motifs as their distinctive characteristics. Protein models contain high quantity of random coils as predicted by the secondary structural features. Phylogenetic tree or evolutionary relationship analysis showed a good relationship among selected species and other predicted species from NCBI databases. The results also exhibited 10 most conserved amino acid residues in the active sites of CuNiR enzymes for both fungal and bacterial species. So, the results accomplish that an extensive wet lab based research on this most conserved amino acid residues is needed to gain a better understanding of the molecular mechanism of denitrification.

SUPPLEMENTARY INFORMATION: Table S1-3 and Figure S1-3 are provided as supplementary file. This is available online at http://biologicalengineering.in/July-Dec-2016-Vol-3-Issue-2/.

COMPETING INTEREST

The authors have declared that no competing interest exists.

ACKNOWLEDGEMENT

The authors reply to thanks to the Professor A.K.M Akhtarul Islam, Department of English, Islamic University, Kushtia for editing and checking the spelling and grammatical error of this manuscript.

REFERENCES

1. Maeda K, Spor A, Edel-Hermann V, et al. N_2O production, a widespread trait in fungi. Sci Rep 2015;5:9697.
2. Zumft WG Cell biology and molecular basis of denitrification. Microbiol Mol Biol Rev 1997;61:533–616.
3. Francis C a, Beman JM, Kuypers MMM New processes and players in the nitrogen cycle: the microbial ecology of

anaerobic and archaeal ammonia oxidation. ISME J 2007;1:19–27.
4. Li Y, Hodak M, Bernholc J Enzymatic Mechanism of Copper-Containing Nitrite Reductase. Biochemistry 2015;54:1233–1242.
5. Fülöp V, Moir JW, Ferguson SJ, Hajdu J The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd1. Cell 1995; 81:369–377.
6. Abraham ZH, Lowe DJ, Smith BE Purification and characterization of the dissimilatory nitrite reductase from *Alcaligenes xylosoxidans* subsp. *xylosoxidans* (N.C.I.M.B. 11015): evidence for the presence of both type 1 and type 2 copper centres. Biochem J 1993; 2:587–593.
7. Nakanishi Y, Zhou S, Kim S-W, et al. A Eukaryotic Copper-Containing Nitrite Reductase Derived from a *Nir*K Homolog Gene of *Aspergillus oryzae*. Biosci Biotechnol Biochem 2010;74:984–991.
8. Cabello P Nitrate reduction and the nitrogen cycle in archaea. Microbiology 2004;150:3527–3546.
9. Adman ET, Godden JW, Turley S The structure of copper-nitrite reductase from Achromobacter cycloclastes at five pH values, with NO_2^- bound and with type II copper depleted. J Biol Chem 1995; 270:27458–27474.
10. Murphy ME, Turley S, Adman ET Structure of nitrite bound to copper-containing nitrite reductase from *Alcaligenes faecalis*. Mechanistic implications. J Biol Chem 1997;272:28455–28460.
11. Kobayashi M, Shoun H The copper-containing dissimilatory nitrite reductase involved in the denitrifying system of the fungus *Fusarium oxysporum*. J Biol Chem 1995; 270:4146–4151.
12. Yamamoto K, Uozumi T, Beppu T The blue copper protein gene of *Alcaligenes faecalis* S-6 directs secretion of blue copper protein from *Escherichia coli* cells. J Bacteriol 1987; 169:5648–5652.
13. Ellis MJ, Prudêncio M, Dodd FE, et al. Biochemical and crystallographic studies of the Met144Ala, Asp92Asn and His254Phe mutants of the nitrite reductase from *Alcaligenes xylosoxidans* provide insight into the enzyme mechanism. J Mol Biol 2002; 316:51–64.
14. Seib KL, Wu H-J, Kidd SP, et al. Defenses against Oxidative Stress in *Neisseria gonorrhoeae*: a System Tailored for a Challenging Environment. Microbiol Mol Biol Rev 2006; 70:344–361.
15. Takaya N, Kuwazaki S, Adachi Y, et al. Hybrid respiration in the denitrifying mitochondria of *Fusarium oxysporum*. J Biochem 2003;133:461–5.
16. Kim S-W, Fushinobu S, Zhou S, et al. Eukaryotic nirK Genes Encoding Copper-Containing Nitrite Reductase: Originating from the Protomitochondrion? Appl Environ Microbiol 2009;75:2652–2658.
17. Yu NY, Wagner JR, Laird MR, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 2010;26:1608–1615.
18. Horton P, Park K-J, Obayashi T, et al. WoLF PSORT: protein localization predictor. Nucleic Acids Res 2007;35:W585–W587.
19. Sharma S, Aneja MK, Mayer J, et al. Diversity of Transcripts of Nitrite Reductase Genes (*nir*K and *nir*S) in Rhizospheres of Grain Legumes. Appl Env Microbiol 2007;71:2001–2007.
20. Lund MB, Smith JM, Francis CA Diversity, abundance and expression of nitrite reductase (*nir*K)-like genes in marine thaumarchaea. ISME J 2012;6:1966–1977.
21. Nojiri M, Xie Y, Inoue T, et al. Structure and function of a hexameric copper-containing nitrite reductase. Proc Natl Acad Sci 2007;104:4315–4320.
22. Boulanger MJ, Kukimoto M, Nishiyama M, et al. Catalytic Roles for Two Water Bridged Residues (Asp-98 and His-255) in the Active Site of Copper-containing Nitrite Reductase. J Biol Chem 2000;275:23957–23964.
23. Lawton TJ, Sayavedra-Soto LA., Arp DJ, Rosenzweig AC Crystal Structure of a Two-domain Multicopper Oxidase: implications for the evolution of multicopper blue proteins. J Biol Chem 2009;284:10174–10180.
24. Leferink NGH, Han C, Antonyuk S V, et al. Proton-coupled electron transfer in the catalytic cycle of *Alcaligenes xylosoxidans* copper-dependent nitrite reductase. Biochemistry 2011;50:4121–4131.
25. Boulanger MJ, Murphy MEP Crystal structure of the soluble domain of the major anaerobically induced outer membrane protein (AniA) from pathogenic *Neisseria*: a new class of copper-containing nitrite reductases. J Mol Biol 2002;315:1111–1127.
26. Wijma HJ, Jeuken LJC, Verbeet MP, et al. A Random-sequential Mechanism for Nitrite Binding and Active Site Reduction in Copper-containing Nitrite Reductase. J Biol Chem 2006; 281:16340–16346.
27. Adhikari UK, Rahman F, Ahmmed M, et al. Homology Modeling and Structural Assessment of CuNiR of *Pseudomonas Chlororaphis* Subsp . *Aureofaciens Nir*k. J Biol Eng Res Rev 2016;3:6–15.
28. Gasteiger E, Hoogland C, Gattiker A, et al. Protein Identification and Analysis Tools on the ExPASy Server. Proteomics Protoc Handb 2005.
29. Yu C, Lin C, Hwang J Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci 2004;1402–1406.
30. Pongor S, Skerl V, Cserzo M, et al. The SBASE protein domain library, release 2.0: a collection of annotated protein sequence segments. Nucleic Acids Res 1993;21:3111–3115.
31. Bailey TL, Johnson J, Grant CE, Noble WS The MEME Suite. Nucleic Acids Res 2015;43:W39–W49.
32. Bailey TL, Gribskov M Combining evidence using p-values: application to sequence homology searches. Bioinformatics 1998; 14:48–54.
33. Finn RD, Bateman a., Clements J, et al. Pfam: the protein families database. Nucleic Acids Res 2014;42:D222–D230.
34. McGuffin LJ, Bryson K, Jones DT The PSIPRED protein structure prediction server. Bioinformatics 2000;16:404–405.
35. Geourjon C, Deléage G SOPMA: significant improvements in protein secondary structure prediction by consensus

prediction from multiple alignments. Comput Appl Biosci 1995;11:681–684.
36. Thompson JD, Higgins DG, Gibson TJ CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.
37. Tamura K, Stecher G, Peterson D, et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 2013;30:2725–2729.
38. Saitou N, Nei M The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evo 1987;4:406–425.
39. Yang J, Yan R, Roy a, et al. The I-TASSER Suite: Protein structure and function prediction. Nat Methods 2015;12:7–8.
40. Yang J, Zhang Y I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res 2015;43:W174–W181.
41. Guex N, Peitsch MC SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis 1997;18:2714–2723.
42. Laskowski RA PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. Nucleic Acids Res 2004;33:D266–D268.
43. Laskowski RA, Rullmannn JA, MacArthur MW, et al. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996 ;8:477–486.
44. Benkert P, Kunzli M, Schwede T QMEAN serverforprotein model quality estimation. Nucleic Acids Res 2009;37:W510–W514.
45. Lüthy R, Bowie JU, Eisenberg D Assessment of protein models with three-dimensional profiles. Nature 1992;356:83–85.
46. Colovos C, Yeates TO Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 1993; 2:1511–1519.
47. Wiederstein M, Sippl MJ ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35:W407–W410.
48. Castrignano T The PMDB Protein Model Database. Nucleic Acids Res 2006;34:D306–D309.
49. Dundas J, Ouyang Z, Tseng J, et al. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res 2006;34:W116–W118.
50. Roy A, Kucukural A, Zhang Y I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 2010;5:725–738.
51. Guruprasad K, Reddy B V, Pandit MW Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng 1990;4:155–161.
52. Kyte J, Doolittle RF A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–132.
53. Yu C-S, Cheng C-W, Su W-C, et al. CELLO2GO: A Web Server for Protein Subcellular Localization Prediction with Functional Gene Ontology Annotation. PLoS One 2014;9:e99368.
54. Shiba K Natural and artificial peptide motifs: their origins and the application of motif-programming. Chem Soc Rev 2010; 39:117–126.
55. Saito H, Kashida S, Inoue T, Shiba K The role of peptide motifs in the evolution of a protein network. Nucleic Acids Res 2007;35:6357–6366.
56. Zhang Y, Skolnick J Scoring function for automated assessment of protein structure template quality. Proteins Struct Funct Bioinforma 2004;57:702–710.
57. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM Stereochemical quality of protein structure coordinates. Proteins Struct Funct Genet 1992;12:345–364.
58. Arakaki AK, Zhang Y, Skolnick J Large-scale assessment of the utility of low-resolution protein structures for biochemical function assignment. Bioinformatics 2004;20:1087–1096.
59. Liang J, Edelsbrunner H, Woodward C Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein Sci 1998;7:1884–1897

About Corresponding Author

Dr. M. Mizanur Rahman is Professor, Chairman and Senior Researcher of the Dept. of Biotechnology and Genetic Engineering, Islamic University, Kushtia-7003, Bangladesh. He obtained his PhD degree from School of Crop Science, Padova University (Present Department name DAFNAE - Department of Agronomy Food Natural Resources Animals and Environment). He is the associate editor of the Journal of Applied Science and Technology, Islamic University and reviewer of the 11 international journals. His research interests are bioinformatics with microbiology such as analysis biodiversity of dynamic microbial population, denitrification and plant growth promoting microorganism, phytoremediation, bioremediation, and antibacterial or antifungal activity of natural plant products. He has been published more than 50 papers in the national and international journals.