MOLECULAR MODELING AND IN SILICO CHARACTERIZATION OF NITROGEN METABOLITE REPRESSOR NmrA OF OPPORTUNISTIC HUMAN PATHOGEN Aspergillus fumigatus

Muhammad Tausif Chaudhry*, Raheela Chaudhry

Address(es):
CMD, National Physical and Standards Laboratory, PCSIR, Islamabad, Pakistan.

*Corresponding author: tausif_chaudhry@yahoo.com

ARTICLE INFO
Received 15. 11. 2018
Revised 15. 12. 2018
Accepted 29. 12. 2018
Published 1. 4. 2019

Regular article

ABSTRACT

The most prevalent airborne human pathogenic fungus Aspergillus fumigatus is ubiquitous in soil and like other ascomycotous fungi it can utilize a variety of nitrogen sources. Nitrogen metabolite repression (NMR) may induce virulence in A. fumigatus as it is pathogenic fungi but the mechanism of regulation and its key components are not identified. The present study focuses on the molecular modeling and in silico characterization of the main player of this regulation, the NmrA of A. fumigatus. Phytochemical and structural characterization using various sequence and structure based predictors and quality assessment of the proposed two- and three-dimensional models were carried out. The characteristic motifs such as glycine-rich NAD(P)-binding motif (GxxGxGxG) and altered active site motif (HxxxK) were located in NmrA along with DNA-binding residues (T11, R39, D40 and A45). The results obtained using bioinformatics tools indicated that the protein was hydrophilic in nature, stable in vitro and had very low disorder probability. Based on the quality score, the proposed secondary and tertiary structures were correct and extremely good to represent NmrA of A. fumigatus. Phylogenetic analysis signifies its close relation with NMR regulatory protein of opportunistic human pathogens A. lentulus and A. novofumigatus.

Keywords: Aspergillus fumigatus; conserved motif; in silico; nitrogen metabolite repression; NmrA; phylogenetic analysis; tertiary structure

INTRODUCTION

The opportunistic human pathogen, Aspergillus fumigatus is a ubiquitous saprophytic fungus, which is associated with various diseases of the respiratory tract. Its airborne spores (conidia) are highly dispersible and significantly more hydrophobic than those of other Aspergillus species (Kwon-Chung and Sugui, 2013). Invasive aspergillosis caused by colonization of these conidia in patients with immune defects causes more than 0.4 million deaths each year (Bongomin et al., 2017). Likewise, its colonization in respiratory airways develops allergic bronchopulmonary aspergillosis (ABPA) with serious consequences in about 5 million asthmatic people worldwide (Denning et al., 2014).

Like other ascomycotous fungi, A. fumigatus grows in soil recycling nitrogen and carbon by degrading organic debris rich in poly saccharides and proteinaceous substrates (Latge 1999; Abad et al., 2010). Nitrogen limitation is reported to induce the expression of virulence genes in various pathogenic fungi (Lopez-Berges, 2010). In filamentous fungi, nitrogen metabolite repression (NMR) occurs when less-preferred nitrogen sources (such as nitrate) are simultaneously available in the medium with energetically favored compounds (ammonium or glutamine) but the pathway for non-preferred nitrogen sources remains inactive (Tudzynski, 2014). NMR is a global regulatory switch which controls activation of enzyme and repressing coding genes for degradation of non-preferred nitrogen sources (Magasanik and Kaiser, 2002). As reported, the major GATA transcription factor AreA and its co-repressor NmrA play a vital role in nitrogen regulatory network (Tudzynski, 2014). Inactivation of these regulatory proteins results in inability of fungal growth on nitrogen sources except for ammonium and glutamine (Wilson and Arst, 1998).

NmrA is among the highly conserved regulatory proteins among Aspergillus species (Lamb et al., 2003). The regulator gene nmrA has been characterized in two Aspergillus species; A. nidulans (Andrianopoulos et al., 1998) and A. flavus (Han et al., 2016). NmrA of A. nidulans was monomeric in solution when over-expressed in E. coli and the protein was crystallized into trigonal, monoclinic and hexagonal forms (Nichols et al., 2001). Interestingly, instead of carbon catabolite repressor CreA, NmrA controls the ability of both GATA transcription factors AreA and AreB to sense the status of carbon metabolism (Macios et al., 2012). NmrA is a member of the short-chain dehydrogenase reductase (SDR) superfamily. Like other SDR family proteins, it binds to dinucleotides NAD+ and NADP+. However, it has a decreased affinity for the reduced forms NADH and NADPH that indicates its possible role in redox signaling, and no affinity for nitrogen metabolite signal molecules ammonium and glutamine (Lamb et al., 2003). In addition to preventing activation of AreA, nmrA in A. flavus appears to be involved in aflatoxin biosynthesis, conidiation, sclerotia formation, invasive virulence and susceptibility to rapamycin and methanesulfonate.

As evident from the comparative genomic analysis of sequenced genera, A. fumigatus genes are different than those commonly found in other pathogenic fungi i.e. Aspergillus, Candida and Cryptococcus (Nierman et al., 2005). Nitrogen limitation may induce expression of virulence genes in A. fumigatus like other pathogenic fungi. However, nitrogen metabolite repression in A. fumigatus has not yet characterized. Moreover, the components of this regulation and their structures are not identified. The present study is designed to propose the two- and three-dimensional structures of NmrA of A. fumigatus using homology modeling and its in silico characterization using various bioinformatics tools. The structural validation and quality of the proposed model was assessed by Ramachandran plot and some neural network based predictors. Conservations and characteristic motifs were also located in NmrA. Important physicochemical parameters were computed using protein sequence and structural features like disorder, foldability and bonds of the proposed structures were predicted. In addition, the evolutionary distances were computed and phylogenetic analysis was conducted for nineteen related Aspergillus species.

MATERIAL AND METHODS

Sequence analysis and alignments

Protein sequence of NmrA (Accession No. KEG 83585) was retrieved from complete genome sequence of A. fumigatus var. RP-2014 available at GenBank (http://www.ncbi.nlm.nih.gov/genbank) with accession No. GCA 000731615. This sequence was compared with a library of database of sequences using the most widely used bioinformatics program for protein sequence searches, basic local alignment tool for protein, BLASTP (Altschul et al., 1990), at NCBI server to get most similar protein sequences. Similarly, protein sequences of NmrA...
from eighteen other Aspergillus species were retrieved. NCBI’s conserved domain database (CDD; Marchler-Bauer et al., 2017) was used to annotate functional units and conservations in NmrA. Pair-wise and multiple sequence alignments were made with CLUSTAL X 2.1 (Larkin et al., 2007). Sequence identity was determined using BioEdit 7.2.5 (Hall, 1999).

Secondary and tertiary structure modeling
A specific distribution of α-helices, β-sheets and 3_10 helices in the secondary structure of NmrA was predicted by SSPro8 at SCRATCH protein predictor server (Cheng et al., 2005). Similarly, tertiary structure was predicted by 3DPro at SCRATCH server. The final three-dimensional structures of NmrA were produced using the program VMD 1.9.3 (Humphrey et al., 1996).

Physicochemical characterization and structural features
Physicochemical parameters like molecular weight, isoelectric point (pl), molar attenuation coefficient at 280 nm, instability index, aliphatic index and grand average of hydropathicity were computed by ProtParam (Gasteiger et al., 2003). Structure features like disordered residues, disorder probability, number of domains and presence of disulfide bonds were determined by SCRATCH (Cheng et al., 2005). Unfoldability was predicted by FoldIndex (Prilusky et al., 2005). Mean hydrogen bond energy and percentage of packing defects were determined by VADAR 1.8 (Willard et al., 2003). A structural based neural network predictor DISPLAR (Tjong and Zhou, 2007) was used to predict NmrA residues that contact DNA.

Identification of NmrA accessory proteins
NmrA protein sequences from nineteen Aspergillus species were retrieved from NCBI protein database (www.ncbi.nlm.nih.gov/protein). To predict protein-protein functional interactions and to calculate confidence score of all the interactions, STRING 10.5 (Szklarczyk et al., 2017) was used.

Phylogenetic analysis
The evolutionary history was inferred using the neighbor-joining method of Saito and Nei (1987) based on 1000 replications. The evolutionary distances among NmrA from nineteen Aspergillus species including A. fumigatus were computed using the Poisson-correction (PC) method (Zuckerkandl and Pauling, 1965) with the units of the number of amino acid substitutions per site using MEGA7 (Kumar et al., 2016).

RESULTS AND DISCUSSION
In the most prevalent airborne pathogenic fungus A. fumigatus, NmrA is repressor of the GATA transcriptional factor AreA, which regulates several genes required for utilization of less-preferred nitrogen sources such as nitrate (Tudzynski, 2014). Conversely, the E3IP protein MeA is an activator of NmrA (Wong et al., 2007). These three regulatory proteins NmrA, AreA and MeA are highly conserved in filamentous fungi (Wagner et al., 2010). NmrA of A. fumigatus has not been characterized on molecular bases and the regulatory components are not identified. The present study elucidates the structural modeling and in silico characterization using bioinformatics tools.

Secondary structure of NmrA
As predicted by SSPro8 (Cheng et al., 2005), NmrA of A. fumigatus consisted of 11 α-helices, 11 β-sheets, and two 3_10 helices (Figure 1). Conserved amino acid residues in the selected nineteen Aspergillus species are highlighted in black. Moreover, dimlucoamide NADP⁺ binding site, active site and regulatory sites are marked in the protein sequence. As reported, the characteristic nucleotide-binding motif in β-domain SRDs is GXXGXXG (Murichak et al., 1999). This glycin-rich motif was located in NmrA (amino acid residues 9–15). Similarly, the SDR characteristic motif HxxxK (residues 133–137) was present instead of FxxK, which is associated with enzyme active site (Stammers et al., 2001). NmrA binding to the regulator AreA blocks it from binding the DNA.

Tertiary structure of NmrA
A tertiary structure of NmrA of A. fumigatus is shown in Figure 2. The overall structure of NmrA consisted of two domains. The N-terminal domain contains six parallel β-sheets (strands β1 to β-6) interconnected by α-helices (α-1 to α-5, residues 1–156) and a 3_10 helix. The C-terminal domain contains five β-sheets (strands β7 to β-11), six α-helices (α-6 to α-11) and two 3_10 helices. A structurally diverse loop region of 32 amino acid residues between α-9 and α-10 (residues 235–285) was also located. This region was found to be disordered to varying degrees in NmrA of A. nidulans and C-terminal domain was required for full biological activity (Stammers et al., 2001).

Physicochemical characterization
The physicochemical parameters of NmrA of A. fumigatus as determined by ProtParam are shown in Table 1. At pH 5.92 i.e. pl of the protein, net charge on NmrA is zero. Instability index is used to determine whether the proteins will be stable in test tube. Values <40 indicated that protein was fairly stable (Guruprasad et al., 1990). The instability index 37.28 indicated that mra was stable in vitro. As defined, aliphatic index (AI) of a protein is the relative volume occupied by aliphatic side chains of non polar and hydrophobic alanine, isoleucine, leucine proline and valine (Iki, 1980). 106 of total 311 amino acid residues (~34 %) were aliphatic in NmrA and an index of 37.28 was fairly expected for this protein of mesophilic origin. Grand average of hydropathy (GRAVY) values are calculated by dividing sum of hydropathy values of all the amino acid residues by total number of residues (Gasteiger et al., 2003). As poitens with GRAVY values ≤0 are hydrophilic, a value of -0.38 indicated the hydrophilic nature of NmrA.
Table 1 Global sequence features of NmrA of A. fumigatus computed by ProtParam.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Calculated Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid residues</td>
<td>311</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>34210.578 kDa</td>
</tr>
<tr>
<td>Isoelectric point (pI)</td>
<td>5.92</td>
</tr>
<tr>
<td>Molar attenuation coefficient (ε) at 280 nm</td>
<td>31400 M⁻¹ cm⁻¹</td>
</tr>
<tr>
<td>Instability index</td>
<td>37.28</td>
</tr>
<tr>
<td>Aliphatic index (AI)</td>
<td>78.75</td>
</tr>
<tr>
<td>Grand average of hydropathy index (GRAVY)</td>
<td>-0.338</td>
</tr>
</tbody>
</table>

Structural features

Structural features of NmrA are shown in Table 2. A wide range of biological processes such as gene regulation, DNA replication and repair are controlled by protein-DNA interactions. Four amino acid residues in NmrA sequence were predicted by DISPLAR (Tjong and Zhou, 2007) to bind DNA. These were T11, R39, D40 and A45. These residues are also highlighted in Figure 1. Other 196 residues did not participate in DNA binding and 111 residues were buried inside protein molecule. As NmrA of A. fumigatus does not contain any cysteine residues, there were no disulfide bonds. Due to the presence of only one disordered residue, the disorder probability was very low (0.32 %). As determined by VADAR (Willard et al., 2003), z score values for mean hydrogen bond energy, percentage of packing defects and percentage of bad bond length were 0.98, 1.93 and 0.49, respectively.

Table 2 Structural features of NmrA of A. fumigatus determined by SCRATCH, FoldIndex and VADAR.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Calculated Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of domains</td>
<td>1</td>
</tr>
<tr>
<td>Disorder probability</td>
<td>0.32 %</td>
</tr>
<tr>
<td>Disulfide bonds</td>
<td>0</td>
</tr>
<tr>
<td>Packing defects</td>
<td>88</td>
</tr>
<tr>
<td>Unfolding ability</td>
<td>0.126</td>
</tr>
<tr>
<td>Free energy of folding</td>
<td>-288.51 kJ mol⁻¹</td>
</tr>
</tbody>
</table>

Structure validation

The overall quality of NmrA structure was assessed by the distribution of two torsion angles φ and ψ, usually called the Ramachandran plot (Hollingsworth and Karplus, 2010) and the plot calculations were computed by PROCHECK. As shown in Figure 3A, about 83 % of amino acid residues (221 of 311) lie in the most favored or core regions (shown in red) that confirmed good quality of the proposed structure. Also, 12 % and 4.5 % residues were in additional allowed regions (shown in yellow) and generously allowed regions (shown in light yellow), respectively. Only two residues, L27 and A308 lie in disallowed region as highlighted in plot. As predicted by the neural network based predictor ProQ, the Levitt-Gerstein (LG) score (Levitt and Gerstein, 1998) and MaxSub score (Siew et al., 2000) were 4.053 and 0.352, respectively. LG score >4 indicated extremely good model and MaxSub score >0.1 shows correct model (Cristobal et al., 2001). Thus, the proposed structure was correct and extremely good to represent NmrA of A. fumigatus. Similarly, score computed by QMEAN was 0.703 with z score value -1.08 as shown in comparison with non-redundent set of protein data bank (PDB) structures (Figure 3B), which indicates good quality of the protein structure.

Identification of NmrA accessory proteins

The accessory interacting proteins accompany and assist the primary protein for its function or help it to fold, stabilize and locate in the right place. Eleven accessory proteins of NmrA (KEY83585) of A. fumigatus were identified by STRING database. The confidence interval map of these proteins is shown in Figure 4. As indicted by Table 3 these proteins were involved in biosynthesis and metabolism of amino acids and secondary metabolites in A. fumigatus. Confidence score of the accessory protein glutamate synthase Glt1 (KEY80835) as identified was 0.424.

Table 3 KEGG pathways of the accessory proteins of NmrA of A. fumigatus identified by STRING. The colors correspond to Figure 4.

<table>
<thead>
<tr>
<th>Pathway ID</th>
<th>Description</th>
<th>Count of protein set</th>
</tr>
</thead>
<tbody>
<tr>
<td>01100</td>
<td>Metabolic pathways</td>
<td>10</td>
</tr>
<tr>
<td>01110</td>
<td>Biosynthesis of secondary metabolites</td>
<td>9</td>
</tr>
<tr>
<td>01230</td>
<td>Biosynthesis of amino acids</td>
<td>8</td>
</tr>
<tr>
<td>00400</td>
<td>Phenylalanine, Tyrosine and Tryptophan biosynthesis</td>
<td>4</td>
</tr>
<tr>
<td>00250</td>
<td>Alanine, Aspartate and Glutamine metabolism</td>
<td>2</td>
</tr>
<tr>
<td>00260</td>
<td>Histidine metabolism</td>
<td>2</td>
</tr>
<tr>
<td>00540</td>
<td>Glycine, Serine and Threonine metabolism</td>
<td>2</td>
</tr>
</tbody>
</table>
Phylogenetic analysis

Figure 5 represents the phylogenetic relationship of NmrA of A. fumigatus and other nitrogen metabolite regulatory proteins from eighteen Aspergillus species. Phylogenetically close species were also opportunistic human pathogens that cause invasive aspergillosis with high mortality rates i.e. A. lentulus and A. novofumigatus (Balajee et al., 2005; Peláez et al., 2013) with 91.3 and 89 % sequence identity, respectively. NmrA of Histoplasma capsulatum, a dimorphic fungus and causative agent of pulmonary histoplasmosis had only 45.5 % sequence identity and was taken as out group. Three other clusters were evident from the figure. The most important fungi of these clusters were A. nigre, A. nidulans and A. wentii with 58.4–75.5 % sequence identity with A. fumigatus NmrA.

Figure 5 Consensus phylogenetic tree of NmA of selected Aspergillus species. Accession numbers are indicated for each species. NmA of H. capsulatum (EGC45405) serves as outgroup. The tree is drawn to scale so that the scale bar indicates percent divergence (distance) in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA7.

CONCLUSION

Based on the quality score, the proposed structure was correct and extremely good to represent NmA of A. fumigatus. Structural and physicochemical characterization indicated that the protein was hydrophilic in nature, stable in vitro and had very low disorder probability. Characteristic dinucleotide binding motif, SDR motif, DNA binding residues and regulatory site were also located. The protein had high sequence identity with phylogenetically close opportunistic human pathogens A. lentulus and A. novofumigatus.

REFERENCES


http://dx.doi.org/10.1099/mmm.0.044867-0


