Malik et al., J Proteomics Bioinform 2017, 10:8
DOI: 10.4172/jpb.1000441

Journal of
Proteomics & Bioinformatics

Research Article Open Access

Structural Model of TaNAM-B1 Transcription Factor from Wheat (*Triticum aestivum*) Insight into the Nutritional Grain Quality

Murtaza Malik, Saurabh Pandey, Kaushlendra Tripathi, Rashmi Jain and Tanushri Kaul*

Nutritional Crop Improvement Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

Abstract

Wheat is a propitious crop which renders several health benefits, but the modern wheat lacks essential micronutrients like zinc and iron. In this paper, various bioinformatics tools and molecular modeling approaches have been embodied for the development and exploration of structure and various properties, respectively. The *NAM-B1* gene is a NAC transcription factor which plays an essential role in the translation machinery. It is also responsible for maintaining the nutritional seed grain quality in wheat. *NAM-B1* gene is accountable for the Fe and Zn concentration in wheat and is also held responsible for delayed senescence in the crop by three weeks. In this study, we have developed the structural model of the *NAM-B1* protein using the structure of 3ULX as template by Modeller 9.12. The resultant model was further processed and refined using various tools such as PROCHECK, ProSA, Verify3D and RMSD. As a result, the model developed was found to be conceivable with 62% a.a sequence identity with the template. Investigation revealed that conserved region found was accountable of the response generated during the improvement of the nutritional quality of grain. TaNAM-B1 plays an ingenious role in leaf senescence and abiotic stress tolerance in plants.

Keywords: *NAM-B*; NAC transcription factor; Wheat; Structural modeling

Introduction

Wheat, a member of the *Poaceae* family is a cardinal food crop. Wheat conjointly with other staple cereals contributes to high amount of calories and nutrients. Hexaploid wheat (*Triticum aestivum L.* AABBDD) is the essential food crop for more than 1/3rd of the global population. It is an important source of major micronutrients. The present day wheat consisting of around 42 chromosomes lacks several important micronutrients like zinc and iron [1].

The *NAM-B1* gene, a NAC transcription factor was the first gene to be held accountable for the variation in grain protein content. This was identified using a map-based positional cloning approach in wheat. The allele for high grain protein content was identified in a wild emmer accession [2]. Wheat plays an essential role in the translation machinery. It is also responsible for maintaining the nutritional seed grain quality in wheat [2]. This gene plays an efficient role in various processes such as development, DNA Binding, leaf senescence and protection against stresses [3].

NAM-B1 delayed senescence by three weeks affected the Fe and Zn concentrations in grain [4]. The Sequence comparison between hexaploid and tetra ploid wheat varieties, reflected that most modern wheat varieties have non-functional protein either due to deletion, frame shift mutation within the first intron or a thymine residue insertion at position [5]. Tetraploid wheat consists of *NAM-B1* functional orthologous copy on chromosome 6A characterizes as Gpc-A1. One paralogous copy is present on chromosome 2B, represented as Gpc-B2. Hexaploid wheat Gpc-1 has two functional orthologous copies on chromosomes 6A and 6D (Gpc-D1) and two paralogous Gpc-2 copies that are present on chromosomes 2B (Gpc-B2) and 2D (Gpc-D2). The admittance of the wild emmer *NAM-B1* gene into modern wheat cultivars result in near-isogenic lines with up to 20% higher nutrition value especially Zn and Fe.

TaNAM-B1 in bread wheat is accountable for their role in seed and grain developmental processes, etc. but little is known about their structure. Development of 3D model of the same would help us in better understanding and enhancing the grain and nutritional quality of crop. In the present investigation, in silico analysis and homology modelling studies TaNAM-B1 from Bread wheat was reported as the three dimensional structures of this gene were not developed. Hence for the better understanding of the structural feature and molecular functions, the model structure for the respective gene was constructed. The model TaNAM-B1 was validated with PROCHECK, PROSAII, Verify3D, RMSD. Homology modeling of TaNAM-B1 could prove beneficial in functional characterization of plant response to improve the nutritional grain quality and leaf senescence.

Materials and Methods

Screening protein sequence

The sequence of the *Triticum aestivum* TaNAM-B1 was retrieved from NCBI database URL http://www.ncbi.nlm.nih.gov/ with accession number AIZ97665.1. Template selection was done using BLASTp for the query against PDB [6]. The target subsequently selected was the x ray crystal structure with PDB code 3ULX. Sequence alignments were completed employing ClustalW. ScanProsite was used to identify motif.

Alignment and phylogenetic analysis

All the sequences of APX were aligned using ClustalW http://www.ebi.ac.uk/Tools/msa/ method to find out the similarity present among the sequences of the same family. Phylogenetic analysis of

***Corresponding author:** Tanushri Kaul, Nutritional Crop Improvement Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi, 110067, India, Tel: +91-9999966205; E-mail: tanushri@icgeb.res.in

Received June 27, 2017; **Accepted** August 02, 2017; **Published** August 07, 2017

Citation: Malik M, Pandey S, Tripathi K, Jain R, Kaul T (2017) Structural Model of TaNAM-B1 Transcription Factor from Wheat (*Triticum aestivum*) Insight into the Nutritional Grain Quality. J Proteomics Bioinform 10: 198-201. doi: 10.4172/jpb.1000441

the sequences was done MEGA (version 4.0.02) [7], using UPGMA method. Each node was tested using the bootstrap approach by taking 5,000 replications. The number is indicated in percentages against each node. The branch lengths were drawn to scale indicated.

Motif analysis

Protein motifs analysis was done by online (Multiple Expectation Maximization for Motif Elicitation) tool version 3.5.7 method by using width of motif 25 and 50 residues respectively and number of 7 motifs, remaining rest of default parameters [8].

Secondary structure prediction

Analysis of secondary structure prediction of the TaNAM-B1 protein sequences we employed several programs based on different algorithms such as PSSFinder, prediction using Markov chains, PSIPRED and EMBL (http://www.ebi.ac.uk/services/structures) protein structure tool.

Structure modeling and validation

Structure was modeled by method of homology SWISS-MODEL and by 'threading' method (I-TASSER). Energy minimization was done employing the Swiss-PDB-Viewer with potential forces [9]. The 3D models were checked by Verify 3D. Steric and conformational were assessed employing Ramachandran plot server ProCheck and ProSA-Web. In RMSD, corresponds to distance between two objects [7]. This value showed the degree to which two 3D structures are similar. The RMSD value between the template 3ULX and our model structure was calculated employed MOE.

Results and Discussion

A total of 18 full-length amino acid sequences of NAM-B1 from different plants were considered for multiple sequence alignment. Multiple sequence alignment highlighted in blue colour represents the conserved sequence of amino acid residues among homologues of NAM-B1 families. This conservation is related with dissimilarities which are adequate for carrying out variations at the structural and functional levels (Figure 1). For phylogenetic analysis among *NAM-B1* from various plants, a rooted tree was generated using the respective sequences (Figure 2). The phylogenetic analysis of *NAM-B1* with different plants indicated two clusters: Cluster A, and cluster B. Cluster A showed two different sub-cluster (a) and (b). Its tree upon the expansion, give *NAM-B1* in *Aegilops speltoides*, *Aegilops longissima*, *Aegilops bicornis*, *Triticum aestivum*, *Aegilops tauschii*, *Triticum timopheevii*, *Triticum dicoccoides, Hordeum vulgare, Brachypodium distachyon, Oryza sativa, Setaria italic, Zea mays, Sorghum bicolor, Vigna gracilis, Manihot esculenta and Manihot nana Müll. Arg* which indicated orthologous and homologous linkage amongst themselves. This signified that *NAM-B1* protein family is strictly conserved from ancestral plants. A broad exploration of protein motifs and their function was done using MEME tools, which recognized various conserved motifs in the *NAM-B1* (Table 1 and Figure 3). Protein motifs studies also confirm similar essentials for improvement of this *NAM-B1* family. Motifs consisting of the signature are well conserved and having substitutions which do not affect their activity. A total of seven protein motifs: 1, 2, 3, 4, 5, 6 and 7 were clearly observed amongst 18 different *NAM-B1* sequences present widely amongst plants. In all plant, motif-1 was most commonly observed and was functionally related to Casein kinase II phosphorylation site, N-myristoylation, Protein kinase C, LDL-receptor class B and NAC domain properties. Motif-2 present, was functionally related to N-myristoylation site,

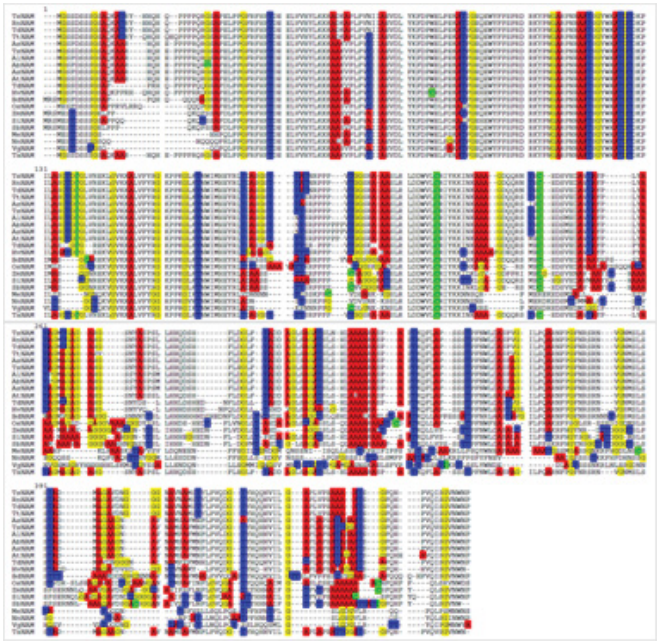

Figure 1: Multiple sequences alignment of deduced amino acid from TaNAMB-1 and other homologues protein sequences.

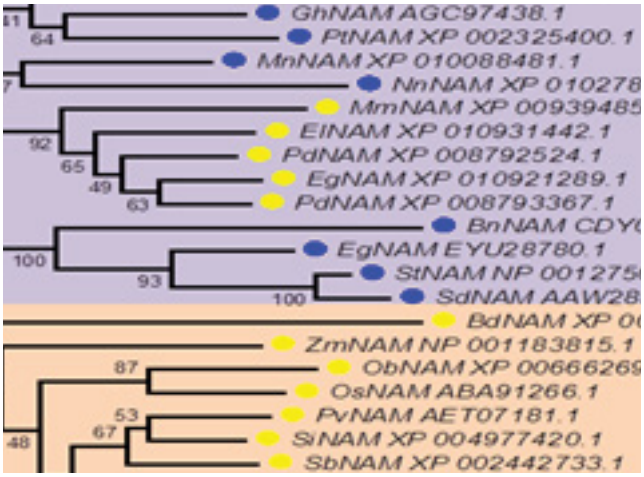

Figure 2: Phylogenetic tree analysis showing similarity of TaNAMB-1 with other plants using MEGA software version 4.0.

NAC and NAM profile. Motif-3 was related to CK2_PHOSPHO_SITE Casein kinase II phosphorylation site, NAC domain profile isoforms. Motif 4 represented Protein kinase C phosphorylation site, NAC. Beside this, Motif-5, 6 and 7 were most commonly related to Casein kinase II, N-glycosylation site, and Histidine-rich region with respect. With a high probability, 7% alpha helices and 15% beta sheets were found in the molecule. The existence of other secondary structure elements and their length could not be determined unambiguously (Figure 4). The length of TaNAM-B1 protein sequences was found to be 404aa residues by sequence search using BLAST identified with PDB ID: 3ULX. The *OsNAM-B1* (PDB ID: 3ULX) showed 59.2% sequence identity to the query sequence with an e-value of 4e-69. ScanProsite server predicted that the fragment ranging from 34- 205 residues was NAC domain, ID PS51005. The sequence alignment revealed that the *NAM-B1* Transcription DNA binding stress responsive domain was

Motif	Best Possible Motif	Width	Function
M1	IAEVDLYKFDPWELPEKATFGEQEW YFFSPRDRKYPNGARPNRAATSGYW	50	Casein kinase II phosphorylation site, N-myristoylation, Protein kinase C, LDL-receptor class B, NAC domain
M2	CGLVREKVGVKKALVFYRGKPPK GLKTNWIMHEYRLTDAS	40	N-myristoylation site, NAC and NAM profile
M3	PRQRGSAPELPPGFRFHPTDEELV VHYLKKKAAKVPLPVTI	41	CK2_PHOSPHO_SITE Casein kinase II phosphorylation site, NAC domain profile
M4	AAASLRLDDWVLCRIYKKINK	21	Protein kinase C phosphorylation site, NAC
M5	DQQRSTECEDSVEDAV	16	Casein kinase II
M6	GILPQARNFPGFNRSRNVGNM	21	N-glycosylation site
M7	LPVQDGTYHQHHVIL	15	Histidine-rich region

Table 1: Multilevel consensus and their functions for the MEME defined motifs.

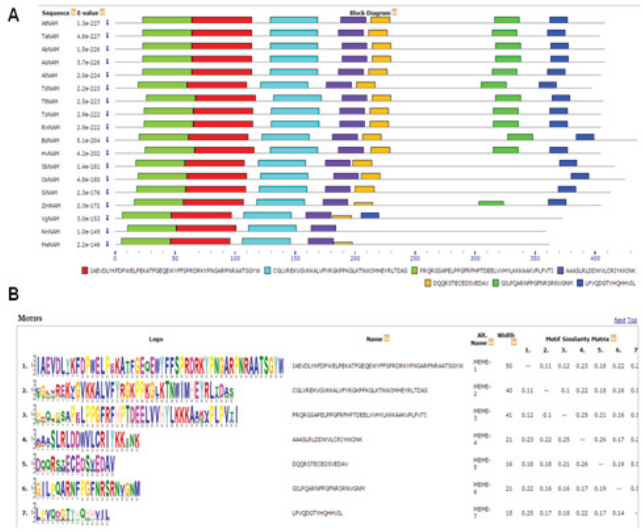

Figure 3: Conserved motif sequences in the deduced amino acid sequences of NAM-B1 from wheat and other homologues.

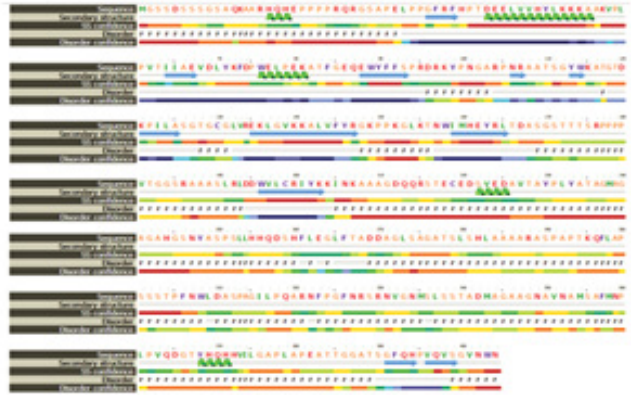

Figure 4: Secondary structure prediction of the beta strand and alpha helix topology of TaNAMB1 protein sequences.

conserved both in TaNAM-B1 as target and OsNAM-B1 as template (Figure 5). These highly conserved transcription factor performs dual function under biotic and abiotic stress responses generated during plant growth and development. The structure was then predicted and taken further for validation. The stereochemical quality and accuracy of the predicted TaNAM-B1 model was evaluated by procheck tool using Ramachandran plot. It revealed 93% residues were found in

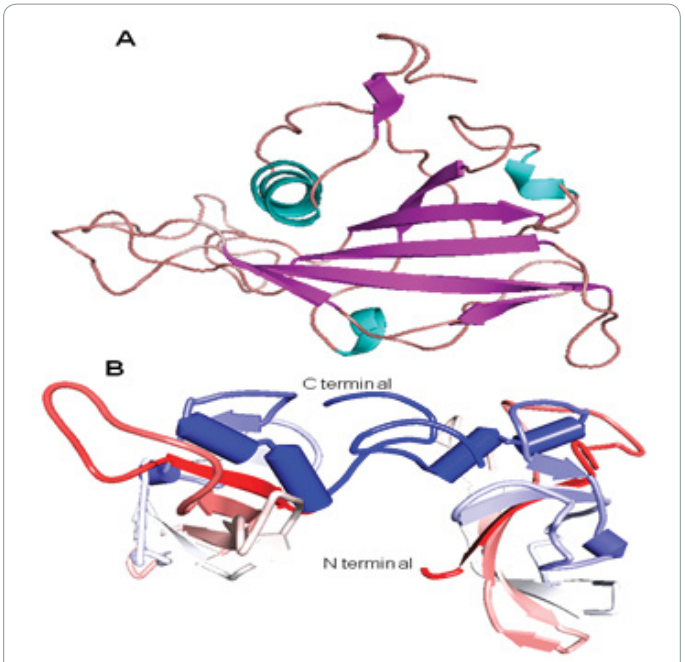

Figure 5: Schematic cartoon representation: (A) Structural Modeling of TaMAB1 visualized by PyMOL; and (B) Representation of N-C terminal using Discovery Studio V.4.1.

most favored region, 17% residues in additionally allowed region, 5% residues in generously allowed and 3% residues were found in the disallowed region following the Ramachandran plot. This model was compared with the Swiss Model and 95.0% residues were found in the most favored region, 3% residues in additionally allowed region and 2% residues in generously allowed region. Consequently the model structure was further produced employing Molecular Modeling and Simulations tools (MOE) which was used for the analysis. To further analyze the overall model quality of target (TaNAM-B1) and template (*OsNAM-B1*), we compared using ProSA web analysis server and this further revealed a Z score value measure of model quality and total energy of the both structures such as template and target as -4.74 of OsNAM-B1 and -4.95 of TaNAM-B, respectively. The Z-score analysis revealed a decent accuracy of our TaNAM-B1. Verify3D was used for determination of the compatibility of an atomic model (3D) with its own residues by assigning a structural class based on its alpha, beta, and loop, polar, nonpolar. Therefore, the comparison result revealed 79.89% of the residues had a score ≥0.2 for a best quality of our TaNAM-B1 model. The degree of structure similarity was measured using root-mean-square distance (RMSD) between corresponding pair of atoms as template (1.483 Å) and target (1.32 Å). The RMSD examination of

the wheat TaNAM-B1 model was measured using its template and employing web server tools like SuperPose.

Conclusion

NAC transcription factor *NAM-B1* of the NAC family is associated with the grain protein content. It is key factor involved in the regulation of gene expression and plays a significant role in the improvement of the nutritional quality of seed. This study focused on the development of the 3D structure of TaNAM-B1 which would facilitate further study in research. This model developed would act as a good source for functional analysis of crystal structure derived experimentally. The motifs analyzed during the study can be further altered using genome editing tools, such as CRISPR-cas9 technology. This would help us in various genomic advancements resulting in better molecular function in seed grain quality, nutritional value and abiotic stress tolerances to crops.

References

1. Borrill P, Connorton JM, Balk J, Miller AJ, Sanders D, et al. (2014) Biofortification of wheat grain with iron and zinc: integrating novel genomic resources and knowledge from model crops. Front Plant Sci 5: 53.
2. Distelfeld A, Avni R, Fischer AM (2014) Senescence, nutrient remobilization, and yield in wheat and barley. J Exp Bot 65: 3783-3798.
3. Chatzav M, Peleg Z, Ozturk L, Yazici A, Fahima T, et al. (2010) Genetic diversity for grain nutrients in wild emmer wheat: potential for wheat improvement. Ann Bot 105: 1211-1220.
4. Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science 314: 1298-1301.
5. Hagenblad J, Asplund L, Balfourier F, Ravel C, Leino MW (2012) Strong presence of the high grain protein content allele of NAM-B1 in Fennoscandian wheat. Theoretical Applied Genetics 125: 1677-1686.
6. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.
7. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biol Evol 24: 1596-1599.
8. Pandey S, Negi YK, Marla SS, Arora S (2011) Comparative insilico analysis of ascorbate peroxidase protein sequences from different plant species. J Bioengineer Biomedical Sci 1: 2.
9. Pandey S, Negi YK, Chinreddy S, Sathelly K, Arora S, et al. (2014) Modeling and phylogenetic analysis of cytosolic ascorbate peroxidase (OsAPX1) from rice reveal signature motifs that may play a role in stress tolerance. Bioinformation 10: 119.