Frontiers in Medical Science Research

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

In silico analysis on structure and DNA binding mode of AtNF -YB8 in Arabidopsis thaliana L

Jubaili Norah Ahmed H, Liang-jianLi, Rui Hu, Wan-Jun wang

School of Life Science and Engineering, Southwest Jiaotong University, Chengdu 610031, China
**Corresponding Author*

ABSTRACT. *Nuclear factor Y (NF-Y) is a heterotrimeric, DNA-binding transcription factor (TF) that is conserved in all eukaryotes, which consists of three subunits, NF-YA, NF-YB, and NF-YC, and binds to the CCAAT box in the promoter regions of its target genes to regulate their expression. AtNF-YB8, a transcription factor from Arabidopsis thaliana, plays an important role in many biological processes in plants. In order to understand the structure and DNA binding model of AtNF-YB8, the 3D structure model of AtNF-YB8 was constructed and docked with the eleven target DNA. Also, the decomposition of the binding energy on a per-residue basis in the AtNF-YB8/DNA07 complex by MD simulations showed that the residues Lys-34, Lys-96, Lys-101 and Lys-110 were probably for specific recognition of DNA.*

Keywords: *Arabidopsis thaliana, AtNF-YB8, DNA binding, Molecular modeling, Molecular dynamics, Transcription factor*

1. Introduction

Transcription factors (TFs) are known as proteins that recognize specific DNA sequences in the control regions of target genes and thus negatively or positively regulate their expression, and different TFs can form various groups to combinatorially and synergistically play context-dependent regulatory roles [1-3]. Nuclear factor Y (NF-Y), an evolutionarily conserved trimeric TF, is widespread in plants, animals, and other eukaryotes [4-8], and is also termed CCAAT Binding Factor (CBF) or Heme Activator Protein (HAP). The NF-Y complex consists of the subunits NF-YA (CBF-B/HAP2), NF-YB (CBF-A/HAP3), and NF-YC (CBF-C/HAP5), all of which are necessary for binding to the CCAAT box [9-10]. All subunits contain evolutionarily-conserved DNA binding and subunit interaction domains to form heterotrimeric complexes [11-14].

Over the past few years, NF-Y has attracted considerable interest, and multiple lines of evidence reveals that NF-Y plays important roles in many biological

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

processes both in plants and animals. Each subunit contains an evolutionarily conserved HAP domain that is critical for subunit interactions and DNA binding [11, 15-16]. NF-Y recognizes CCAAT sequence mostly through the conserved C-terminus of NF-YA although all three subunits are required for DNA binding activity [17]. NF-YA and NF-YC each possess large glutamine-rich activation domains outside of the HAP domain [18-19]. Located within the HAP domains of NF-YB and NF-YC is a histone-fold motif that is found in all core histones and that mediates dimerization and contributes to DNA binding [20-22]. NF-YB and NF-YC subunits dimerize to form a complex that associates with NF-YA [11]. Once the trimer forms, it can bind specifically to CCAAT boxes, which are cis-regulatory elements found in either orientation in 25–30% of eukaryotic promoters, and typically located between 60 and 100 bp upstream of the transcriptional start site [23-24].

AtNF-YB8, a member of the NF-YB family from Arabidopsis thaliana, plays an important role in plant development. Like other transcription factors, AtNF-YB8 is a transcription activator consisting of a conserved N-terminal DNA binding domain and a C-terminal activation domain [25]. To disclose the interaction between the AtNF-YB8 with DNA was very important to understanding the regulatory mechanism at the molecular level, however, the corresponding binding mode of AtNF-YB8 with DNA was not clear. To date, the crystal structure of AtNF-YB8 has not been resolved. In the study, the AtNF-YB8 protein sequence was subjected to molecular modeling, docking and dynamics simulations in order to analyze its structural features and interactions with DNA.

2. Materials and methods

2.1 Detect binding site of transcription factor (TFBS)

Essential genes are indispensable for maintaining the survival of all kinds of organisms, and they can be regarded as cornerstones of synthetic biology. Because of the important significance of essential genes, we performed a molecular simulation between transcription factor (here we used NF-Y) of an essential gene in Arabidopsis thaliana and its binding site (TFBS). A workflow about how we obtained our data is show in Figure 1. AT1G01040 (Accession id in NCBI) is an essential gene in A. thaliana, which was obtained from supplement file of reference [26]. The gene information of AT1G01040 is further retrieved from NCBI (National Center for Biotechnology Information). It is in chromosome NC_003070. The nucleotide sequence of gene AT1G01040 was extracted according to the gene position following obtained 2000 bases located in upstream of gene AT1G01040. JASPAR (http://jaspar.genereg.net/) is a database, which comprises a curated, non-redundant set of profiles for eukaryotes, and they can be used to define transcription factor binding sites for eukaryotes [27]. In this work, we used JASPAR to detect and define the binding sites of transcription factor for gene AT1G01040.

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

2.2 Modeling of AtNF-YB8

The NCBI protein database (http://www.ncbi.nlm.nih.gov/protein/) was used to search the amino acid sequence of the AtNF-YB8 protein. The BLAST server (http://blast.ncbi.nlm.nih.gov) was used to search the template for the protein. We applied XXXXX (PDB ID: XXXX) as the template for the AtNF-YB8 protein, and the homology of amino acid sequence was aligned. Homology modeling of mus musculus AtNF-YB8 protein was carried out using modeller9.20 package (https://salilab.org/modeller/).

2.3 Molecular dynamics simulations

The AmberTools 15 program was used for MD simulations of the AtNF-YB8 structure. The forcefield “leaprc.ff14SB” was used for the DNA. The system was placed in a rectangular box (with a 10.0 Å boundry) of TIP3P water using the “SolvateOct” command with the minimum distance between any solute atoms. Equilibration of the solvated complex was done by carrying out a short minimization (5000 steps of each steepest descent and conjugate gradient method), 500 ps of heating, and 500 ps of density equilibration with weak restraints using the sander module. At last, 20 ns of MD simulations were carried out.

2.4 Validation of model

The stereo chemical quality of AtNF-YB8 structure was checked by inspection of the Psi/Phi Ramachandran plot using PROCHECK [28].

2.5 Modeling of DNA

The 3D structures of the double helix DNA were built by SYBYL-X 2.0. The nucleic acid type was selected as DNA-duplex and the DNA canonical form was selected as B-form helix. The constructed DNA model was then subjected to MD simulation using the same method as that used for the initial AtNF-YB8 model.

2.6 Docking of AtNF-YB8 with DNA

Molecular docking studies were performed to investigate the binding mode between the AtNF-YB8 protein and the DNA using the ZDOCK server (http://zdock.umassmed.edu/). For docking, the default parameters were used as described in the ZDOCK server. The top ranked pose as judged by the docking score was subject to visually analyze using PyMoL 1.7.6 software (http://www.pymol.org/).

2.7 Binding free energy and energy decomposition per residue calculations

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

The binding free energies (ΔGbind in kcal/mol) were calculated using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method, implemented in AmberTools 15. Moreover, to identify the key protein residues responsible for the ligands binding process, the binding free energy was decomposed on a per-residue basis. For each complex, the binding free energy of MM/GBSA was estimated as follows:

ΔGbind= Gcomplex–Gprotein–Gligand

where ΔGbind is the binding free energy and Gcomplex, Gprotein and Gligand are the free energies of complex, protein, and ligand, respectively.

3. Results and discussion

3.1 Detect binding site of transcription factor

We finally obtained 11 TFBS sequences based on JASPAR database for essential gene AT1G01040. Their detailed information was listed in the following Table 1. There are 4 out of 11 TFBS locating on the positive strand of chromosome NC_003070, and the remaining ones are located on the complement strand of chromosome NC_003070. The length of TFBS is range from 12 to 15, and GC content range from 38.4% to 66.6% in our result.

Essential genes are a subset of all genes. They are necessary for maintaining the several for organisms, and those genes also provide fundamental functions [29]. In addition, essential gene can also be regarded as the most basic material for constructing minimal metabolism network [30]. Therefore, the molecular simulation between transcription factor and its binding site can help scientific immunity make an insight into the essential genes' function and transcription. Essential genes in whole-genome scale, obtained through wet-lab methods, are widely available for dozens of prokaryotes and several eukaryotic organisms. They can be accessed from open access database with any register such as DEG database [31] and OGEE database [32], which provide source of essential genes from different species and necessary information of essential genes. We can further perform more molecular simulations based on those data in the future. There is no doubt that some union binding model can be found via more and more molecular simulations between TF and TFBS of essential genes.

Table 1 Detailed information of the 11 TFBS

No.	***Matrix ID[a]***	***Sequence ID[b]***	***Start***	***End***	***Strand***	***Predicted sequence***
DNA 01	*MA1224.1*	*NC_003070.9:21121-23120*	*28*	*39*	-	*GTGACCGACTTA*
DNA 02	*MA1218.1*	*NC_003070.9:21121-23120*	*27*	*39*	+	*TTAAGTCGGTCAC*

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

DNA 03	*MA1217.1*	*NC_003070.9:21121-23120*	*27*	*41*	-	*CAGTGACCGACTTAA*
DNA 04	*MA1224.1*	*NC_003070.9:21121-23120*	*1538*	*1549*	-	*TGGACCGACTCG*
DNA 05	*MA1224.1*	*NC_003070.9:21121-23120*	*292*	*303*	-	*GACGCCTACAAA*
DNA 06	*MA1224.1*	*NC_003070.9:21121-23120*	*1057*	*1068*	-	*AACGCCGAAAAC*
DNA 07	*MA1218.1*	*NC_003070.9:21121-23120*	*1056*	*1068*	+	*TGTTTTCGGCGTT*
DNA 08	*MA1218.1*	*NC_003070.9:21121-23120*	*291*	*303*	+	*TTTTGTAGGCGTC*
DNA 09	*MA1218.1*	*NC_003070.9:21121-23120*	*637*	*649*	-	*TAAAGTCGATCGA*
DNA 10	*MA1218.1*	*NC_003070.9:21121-23120*	*1537*	*1549*	+	*ACGAGTCGGTCCA*
DNA 11	*MA1224.1*	*NC_003070.9:21121-23120*	*1074*	*1085*	-	*GCCATCGACTTT*

aMatrix IDs that matching TFBS in JASPAR database; bnumber 9 in NC_003070.9 denotes the assembled version is 9 of chromosome NC_003070.

3.2 Refinement of AtNF-YB8 model

In order to assess the stability of the model and find the energetically favorable structure for further docking study, the constructed model of AtNF-YB8 was subjected to MD simulation. Structural stability of the constructed model during the MD simulation was examined using RMSD. The RMSD plot for the protein backbone atoms with reference to the initial structure, as a function of time, was shown in Fig. 1. The first 10 ns were considered as an equilibration period followed by a plateau where the model became stable in the RMSD range 2□3 Å.

Distinguishing the flexible regions of the protein could help to understand the stability of protein. To examine the flexible regions of the model, the average root mean square fluctuation (RMSF) plot for the Cα-atoms with respect to the residues was generated. Those residues deviating more than 2.5 Å were considered as highly flexible elements of the protein [25]. As shown in Fig. 2, the high fluctuations mainly occurred in the loop regions (L1: residue 1-3, L2: residue 95-96, L3: residue 104-110). The low fluctuations mainly occurred in the α-helix and β-sheet regions. The regions with RMSF values lower than 2.0 Å clearly showed that the protein structural core was well-constructed and the model was stable in the course of MD simulation.

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

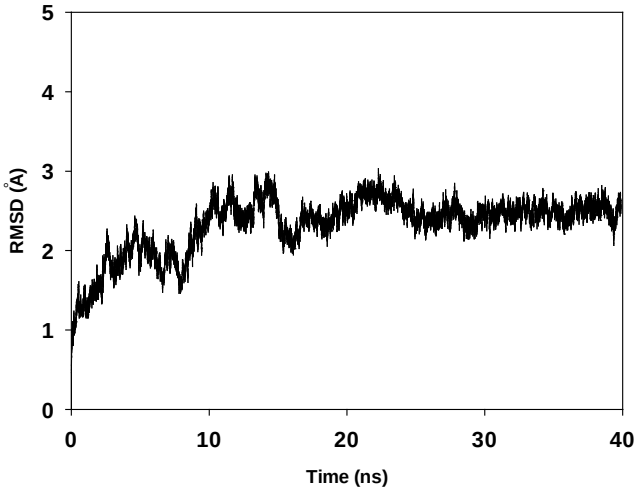


Fig. 1 The root-mean-square deviations (RMSDs) of all the atoms of AtNF-YB8 with respect to its initial structure as function of time.

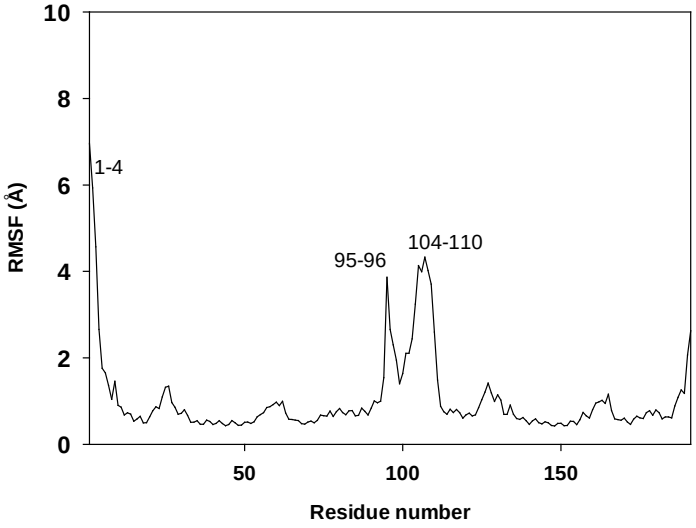


Fig. 2 RMSF of residues of the protein AtNF-YB8 during the 40 ns simulation.

3.3 Validation of AtNF-YB8 model

The stereo chemical quality of the AtNF-YB8 model was checked by PROCHECK based on the Ramachandran plot. The results summarized in Fig. S1

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

showed 96.6% of the residues in the most favored regions and only 3.4 % in additional allowed regions.

3.4 Modeling of DNA for AtNF-YB8 binding

The eleven DNA motif for AtNF-YB8 binding was subjected to 3D structure modeling. The B-form of the DNA conformation was selected to mimic the in vivo conformation. MDsimulation of the DNA models were carried out to gently relax the structures. The RMSD plots of DNA during MD simulation showed that the RMSD increased and hovered around 2 Å, indicating equilibration of the system (Fig. S2-S12). The equilibrated structures (Fig. S13-S23) from the simulation was selected for further docking studies.

3.5 Docking and MD simulation of AtNF-YB8 with DNA

The molecular docking studies were performed to investigate the binding mode between the AtNF-YB8 and the eleven DNA structures using the ZDOCK server, and the ZDOCK Scores were shown in Table 2. From the Table 2, we can see that the score between AtNF-YB8 and the DNA07 was the hignest. The interaction between the AtNF-YB8 (green) and the DNA07 (rose red) was shown in Fig. 3. One hydrophobic interaction was observed between the residues Leu-95, Phe-98, Trp-99, Met-103, Met-149 and Leu-157 of the AtNF-YB8 and the nucleotides DT-5, DT-6, DC-7, DG-8, DA-22, DA-23 and DA-24 of the DNA07, forming a strong hydrophobic binding (Fig. 4). Detailed analysis showed that the residue Trp-99 of the AtNF-YB8 formed CH-π interactions with the nucleotides DA-22 and DA-23 of the DNA07. Importantly, six hydrogen bond interactions were shown between the residue Arg-21 of the AtNF-YB8 and the DG-8 of the DNA07 (bond length: 2.6 and 3.2 Å), the Leu-95 of the AtNF-YB8 and the DT-5 of the DNA07 (bond length: 3.2 Å), the Trp-99 of the AtNF-YB8 and the DA-23 of the DNA07 (bond length: 2.7 Å), the Lys-101 of the AtNF-YB8 and the DA-22 of the DNA07 (bond length: 3.1 Å), the Asn-106 of the AtNF-YB8 and the DC-10 of the DNA07 (bond length: 2.9 Å), which were the main binding affinity between the AtNF-YB8 and the DNA07. All these interactions helped the DNA07 to anchor in the binding site of the AtNF-YB8.

In conclusion, the above molecular simulations give us rational explanation of interaction between the AtNF-YB8 and the DNA07, which provided valuable information for further study of binding sites between the AtNF-YB8 and the DNA07.

Table 2. ZDOCK Scores between the AtNF-YB8 protein and the DNA.

Complex	*ZDOCK Scores*
AtNF-YB8/DNA01	*1689.635*
AtNF-YB8/DNA02	*1424.969*
AtNF-YB8/DNA03	*1559.652*

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

AtNF-YB8/DNA04	*1631.995*
AtNF-YB8/DNA05	*1493.927*
AtNF-YB8/DNA06	*1630.386*
AtNF-YB8/DNA07	***1797.666***
AtNF-YB8/DNA08	*1579.085*
AtNF-YB8/DNA09	*1467.120*
AtNF-YB8/DNA10	*1601.353*
AtNF-YB8/DNA11	*1539.807*

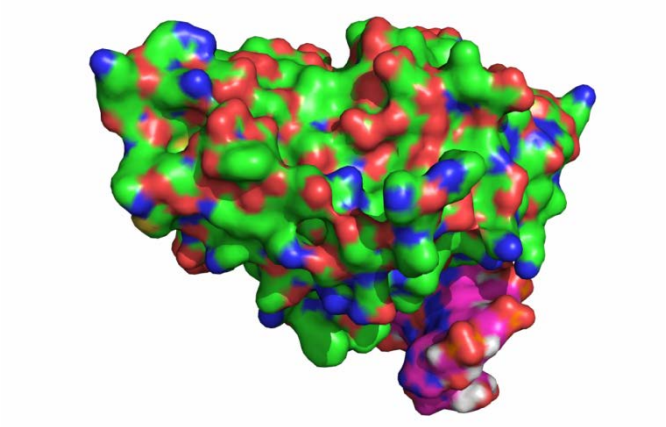


Fig. 3 Total view of the interactions between the AtNF-YB8 (green) and DNA07 (rose red).

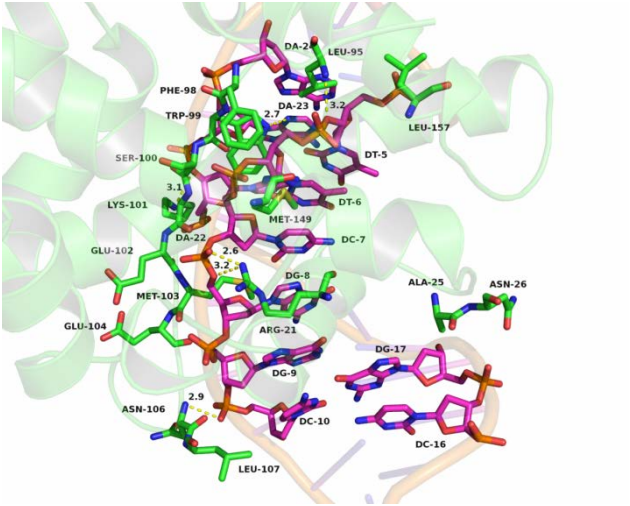


Fig. 4 Detailed view of the interactions between the AtNF-YB8 (green) and DNA07

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

(rose red).

3.6 Molecular dynamics results

To explore the potential binding mode between DNA07 and AtNF-YB8, molecular docking and molecular dynamics simulations were performed using the ZDOCK and AmberTools15 software package. The preferential binding mechanism of AtNF-YB8 with DNA07 was determined by 40-ns molecular dynamics simulations based on the docking results. To explore the dynamic stability of the models and to ensure the rationality of the sampling strategy, the root-mean-square deviation (RMSD) values of the protein backbone based on the starting structure along the simulation time were calculated and plotted in Fig. 5. As shown in Fig. 5, the protein structures of the two systems were stabilized during the simulation.

The root mean square fluctuations (RMSF) of the residues of the whole protein in the AtNF-YB8/DNA07 complex and in the free AtNF-YB8 were calculated to reveal the flexibility of these residues. The RMSF of these residues are shown in Fig. 6, clearly depicting different flexibilities in the binding site of AtNF-YB8 in the presence and absence of DNA07. Most of the residues in the AtNF-YB8 binding site that bind with DNA07 show a small degree of flexibility with a RMSF of less than 2 Å when compared with free AtNF-YB8, indicating that these residues seem to be more rigid as a result of binding to DNA07. However, the residues at the positions 95-110 in the DNA07 binding site that bind with AtNF-YB8 showed a big degree of flexibility with the RMSF values reached to 4 Å when compared with the free protein, indicating that these residues seem to be more flexible as a result of binding to DNA07.

To gain more information about the residues surrounding the binding site and their contribution to the whole system, the electrostatic, Van der Waals, solvation, and total contribution of the residues to the binding free energy were calculated with the MMGBSA method. The summations of the per residue interaction free energies were separated into Van der Waals (ΔEvdw), solvation (ΔEsol), electrostatic (ΔEele) and total contribution (ΔEtotal). In the AtNF-YB8/DNA07 complex, the residues Lys-34, Lys-96, Lys-101 and Lys-110 have strong electrostatic (ΔEele) contributions, with the ΔEele of < -10.0 kcal/mol (Fig. 7). Detailed analysis showed that the residues Lys-34, Lys-96, Lys-101 and Lys-110 were all close to the phosphate group of the DNA07, forming electrostatic interactions. Moreover, the residues Lys-34, Lys-96, Lys-101 and Lys-110 also formed four strong hydrogen bond interactions with the “O” atoms of the DNA07, with the bond lengths of 2.0 Å, 2.0 Å, 2.0 Å and 2.1 Å, respectively (Fig. 8). Besides, the residues Ser-16 and Ser-100 have moderate electrostatic (ΔEele) contribution, with the ΔEele of < -7.0 kcal/mol (Figure 3). In fact, the residues Ser-16 and Ser-100 oriented the phosphate group of the DNA07, forming two strong hydrogen bond interactions (bond length: 2.3 Å and 2.1 Å). In addition, residues Ala-13, Arg-17, Lys-96, Lys-101 and Met-103 with the ΔEvdw of < -2.5 kcal/mol, have strong Van der Waals interaction with the ligand because of the close proximity between the residues and the DNA07. Overall, the majority of the decomposed energy interaction originated from Van der

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

Waals interactions, mainly through hydrophobic interactions, such as Ile-12, Ala-13, Ile-15, Leu-95, Met-103 and Leu-107. In addition, the total binding free energy for the AtNF-YB8/DNA07 complex calculated according to the MMGBSA approach, and the an estimated ΔGbind of -37.8 kcal/mol was found for DNA07, suggesting that DNA07 can strongly bind to the binding site of the AtNF-YB8.

In summary, the above molecular dynamics simulations give us rational explanation of the interactions between DNA07 and AtNF-YB8, which provided valuable information for further study of binding mode between the AtNF-YB8 and the DNA07.

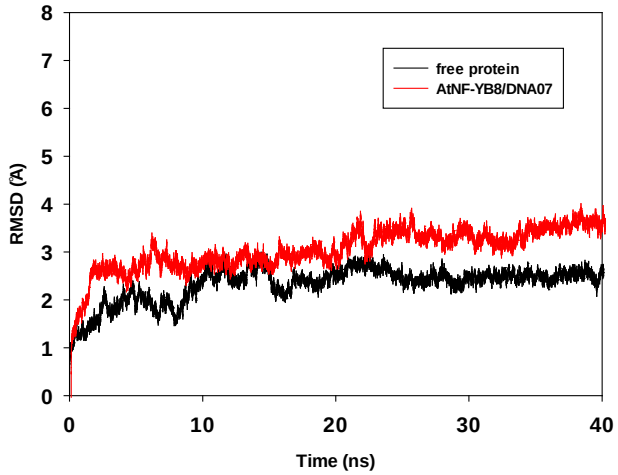
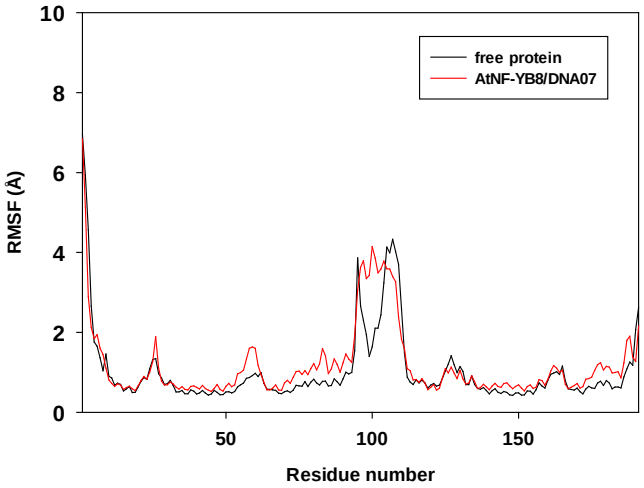


Fig. 5 The root-mean-square deviations (RMSDs) of all the atoms of AtNF-YB8/DNA07 complex with respect to its initial structure as function of time.



ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

Fig. 6 RMSF of residues of the whole protein in AtNF-YB8/DNA07 complex and free AtNF-YB8 during the 40 ns simulation.

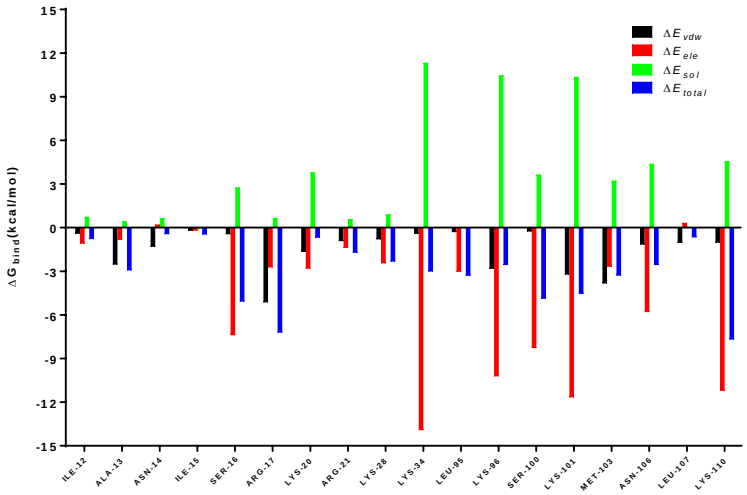


Fig. 7 Decomposition of the binding energy on a per-residue basis in the AtNF-YB8/DNA07 complex.

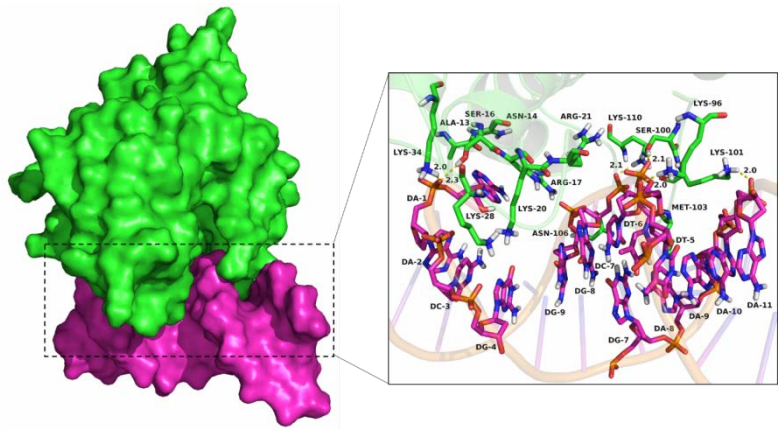


Fig. 8 The predicted binding mode of DNA07 in AtNF-YB8 binding pocket obtained from MD simulation.

Conclusions

In conclusion, despite AtNF-YB8 was shown to regulate gene expression

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

involved in plant development, the corresponding regulating mechanism has not been elucidated. In this study, the 3D structure model of AtNF-YB8 was constructed and docked with eleven target DNA motif using the ZDOCK server. The ZDOCK Scores between AtNF-YB8 and the DNA07 (5'-TGTTTTCGGCGTT-3') was the hignest. Decomposition of the binding energy on a per-residue basis in the AtNF-YB8/DNA07 complex showed that the residues Lys-34, Lys-96, Lys-101 and Lys-110 were probably for specific recognition of DNA. Meanwhile, the residues Ala-13, Arg-17, Lys-96, Lys-101 and Met-103 might be the key residues for affinity with DNA. The study provided the in silico framework to understand the interactions of AtNF-YB8 with DNA at the molecular level.

References

[1]A.J. Warren(2002) . Eukaryotic transcription factors, Current opinion in structural biology, no.12,pp.107-114.
[2]G. Stampfel, T. Kazmar, O. Frank, S. et al(2015) , Transcriptional regulators form diverse groups with context-dependent regulatory functions, Nature, vol, 528, no.7580, 147-151.
[3]G. Li, H. Zhao, L. Wang, Y. Wang, et al(2018), The animal nuclear factor Y: an enigmatic and important heterotrimeric transcription factor, American journal of cancer research, no. 8, pp.1106-1125.
[4]T. Laloum, S. De Mita, P. Gamas, et al(2013). CCAAT-box binding transcription factors in plants: Y so many?, Trends in plant science, no. 18, pp.157-166.
[5]R.D. Welch, C. Guo, M. Sengupta, et al(2017). Rev-Erb co-regulates muscle regeneration via tethered interaction with the NF-Y cistrome, Molecular metabolism, no.6, pp.703-714.
[6]G. Begum, M. Otsu, U. Ahmed, Z. et, al(2018). NF-Y-dependent regulation of glutamate receptor 4 expression and cell survival in cells of the oligodendrocyte lineage, Glia,no. 66, pp.1896-1914.
[7]A. Dorn, J. Bollekens, A. Staub, et al(1987), A multiplicity of CCAAT box-binding proteins, Cell, no. 50, pp. 863-872.
[8]M. Morey, S.K. Yee, T. Herman,et al(2008). Coordinate control of synaptic-layer specificity and rhodopsins in photoreceptor neurons, Nature, no. 456, pp.795-799.
[9]M. Nardini, N. Gnesutta, G. Donati, et al(2013). A. Fossati, C. Vonrhein, D. Moras, C. Romier, M. Bolognesi, R. Mantovani, Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination, Cell , no.152, pp. 132-143.
[10]H. Zhao, D. Wu, F. Kong, et al(2016). The Arabidopsis thaliana Nuclear Factor Y Transcription Factors, Frontiers in plant science , no.7, pp.2045.
[11]Sinha S , Maity S N , Lu J , et al(1995). Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3.Proceedings of the National Academy of Sciences, vol. 92, no.5, pp.1624-1628.
[12]S. Sinha, I.S. Kim, K.Y. Sohn,et al(1996). Three classes of mutations in the A subunit of the CCAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex, Molecular and cellular biology ,

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

no.16, pp.328-337.
[13]V. Calvenzani, B. Testoni, G. Gusmaroli, et al(2012). Interactions and CCAAT-binding of Arabidopsis thaliana NF-Y subunits, PloS one, no. 7, pp.42902.
[14]C. Ren, Z. Zhang, Y. Wang, et al(2016), Genome-wide identification and characterization of the NF-Y gene family in grape (vitis vinifera L.), BMC genomics , no.17, pp.605.
[15]R. Hooft van Huijsduijnen, X.Y. Li, D. Black, et al. Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y (CP-1/CBF) subunits, The EMBO journal, no. 9, pp.3119-3127.
[16]X.Y. Li, R. Mantovani, R. Hooft van Huijsduijnen,et al(1992), C. Benoist, D. Mathis, Evolutionary variation of the CCAAT-binding transcription factor NF-Y, Nucleic acids research, no. 20, pp.1087-1091.
[17]R. Mantovani(1999), The molecular biology of the CCAAT-binding factor NF-Y, Gene, no. 239, pp. 15-27.
[18]Coustry F , Maity S N , De Crombrugghe B(1995) . Studies on Transcription Activation by the Multimeric CCAAT-binding Factor CBF. Journal of Biological Chemistry, vol.270, no.1, pp.468-475.
[19]A. de Silvio, C. Imbriano, R. Mantovani(1999). Dissection of the NF-Y transcriptional activation potential, Nucleic acids research, vol.27, no.13, 2578-2584.
[20]G. Arents, E.N. Moudrianakis(1995), The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization, Proceedings of the National Academy of Sciences of the United States of America , vol.92, no.24, pp.11170-11174.
[21]A.D. Baxevanis, G. Arents, E.N. Moudrianakis, et al(1995), A variety of DNA-binding and multimeric proteins contain the histone fold motif, Nucleic acids research, vol. 23, no.14, pp.2685-2691.
[22]Luger K , Mader A W , Richmond R K , et al(1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature, vol.389, no.6648, pp.251-260.
[23]P. Bucher(1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences, Journal of molecular biology, vol.212, no.4, pp.563-578.
[24]R. Mantovani(1998). A survey of 178 NF-Y binding CCAAT boxes, Nucleic acids research vol.26, no.5, pp.1135-1143.
[25]Wei Q , Ma C , Xu Y , et al(2017). Control of chrysanthemum flowering through integration with an aging pathway. Nature Communications, vol.8, no.1, pp.829.
[26]Lloyd J P , Seddon A E , Moghe G D , et al(2015). Characteristics of Plant Essential Genes Allow for within- and between-Species Prediction of Lethal Mutant Phenotypes. The Plant Cell, vol.27, no.8, pp.2133-2147.
[27]Khan A , Oriol Fornés, Stigliani A , et al(2017). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Research, vol. 77, no.21, pp.43.
[28]R.A. Laskowski, M.W. MacArthur, D.S. Moss,et al(1993). PROCHECK: a program to check the stereochemical quality of protein structures., J Appl Crysth, vol,26 no.2, pp. 283-291.
[29]M.J. Lajoie, S. Kosuri, J.A. Mosberg, et al(2013), Probing the limits of genetic

ISSN 2618-1584 Vol. 1, Issue 2: 01-14, DOI: 10.25236/FMSR.20190201

recoding in essential genes, Science, vol.342, no.6156, pp. 361-363.
[30]Y.N. Ye, B.G. Ma, C. Dong, et al(2016), A novel proposal of a simplified bacterial gene set and the neo-construction of a general minimized metabolic network, Sci Rep, no.6, pp. 35082.
[31]H. Luo, Y. Lin, F. Gao,et al(2014). DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements, Nucleic Acids Res, no. 42, pp. D574-580.
[32]W.H. Chen, G. Lu, X. Chen, et al(2017), OGEE v2: an update of the online gene essentiality database with special focus on differentially essential genes in human cancer cell lines, Nucleic Acids Res, no. 45, pp.D940-D944.