In silico analysis of mutations in PITX3 gene

Abida Arshad
Department of Zoology,
PMAS Arid Agriculture University,
Rawalpindi, Pakistan.
abida.arshad6@gmail.com

Muhammad Arshad
Department of Bioinformatics & Biotechnology,
International Islamic University,
Islamabad, Pakistan.
m.arshad@iiu.edu.pk

Rashda Abbasi
Institute of Biomedical and Genetic Engineering,
Islamabad, Pakistan.
r.abbasi@daad-alumni.de

Nafees Ahmad
Institute of Biomedical and Genetic Engineering,
Islamabad, Pakistan
nafravian1@yahoo.com

Christian MK Sieber
Institute of Bioinformatics and Systems Biology,
Helmholtz Center Munich-German Research Center for
Environmental Health,
Neuherberg, Germany.
christian.sieber@helmholtz-muenchen.de

***Abstract*-PITX3 belongs to a class of heomeodomain transcription factors involved in the development of dopaminergic neurons and ocular lens. Despite a great degree of homology, the mutation in human and mouse *Pitx3* gene exhibit differences in the range and extent of phenotypic effects. The current study was designed to predict the effect of mutations in the mouse and human *PITX3* gene using *in silico* tools. We used publically available bioinformatics tools to identify the secondary structure, functional domains, three-dimensional structure and DNA binding residues. Analysis of functional domains in the PITX3 revealed a lack of OAR domain in the G219fs mutation and in the mouse eyeless mutation. There was no difference in the functional motifs of the S13N and K111E mutation compared to the wild-type PITX3. However, an additional helix-turn-helix (HTH) domain is predicted in K111E mutation. Comparison of three-dimensional structures of the wild-type and mutant proteins did not show significant differences except 220delG. The eyeless mouse mutant protein exhibited a very different structure compared to the wild-type mouse Pitx3. Our results indicate that three-dimensional structure of the protein is a good predictor of the in vitro and in vivo behavior of the PITX3 protein and provides guidelines for performing the functional assays of the mutant proteins.**

Keywords-Parkinson; mutation; domains.

I. INTRODUCTION

PITX3 is a homeodomain transcription factor belonging to a bicoid class of homeobox gene family. It is mapped to chromosome 10 and is a homolog of mouse Pitx3. PITX3 comprises of four exons and encodes a protein of 302aa, consisting of 60aa DNA binding homeodomain and 14aa OAR domain. The OAR domain acts as an intra-molecular switch for the activity of these transcription factors [1]. The homeodomain contains an important lysine residue which is important for the recognition of TAA(T/G)CC motif present in the promoter region of PITX3 target genes [2, 3].

PITX3 is very specifically expressed in the ocular lens and midbrain dopaminergic neurons; thus has a very important role in the development of lens and maintenance of dopaminergic neurons. Any ablation in the PITX3 results in the abnormal development or complete loss of ocular lens and loss of dopaminergic neurons-a hallmark of Parkinson's disease (PD). Various single nucleotide polymorphisms (SNPs) and mutations in *PITX3* gene have been associated with loss of dopaminergic neurons (PD) and varying degree of ocular defects (Microphthalmia or Anophthalmia) [4-6]. So far, three mutations in *PITX3* gene have been reported in various studies that affect either its N- or C-terminal region. The only N-terminal mutation (S13N) detected in a family of autosomal dominant cataract is a single nucleotide change [7]. The other two mutations that affect the C-terminal region and present in the OAR domain include 17bp duplication (657-673 dup17; G219fs) and a single nucleotide deletion (650delG) [7-9]. These mutations within and around the OAR domain possibly affect the interaction of PITX3 with other proteins and result in various defects including posterior polar cataract and anterior segment mesenchymal dysgenesis (ASMD; 8]. The most severe phenotype was observed in patients with 650delG, as this is the only mutation reported in homozygous condition [6]. These patients exhibit microphthalmia and neurological deficits like mental retardation, weak reflexes and increased muscle tone.

In addition to the human, mutations in the mouse *Pitx3* gene have also resulted in ocular phenotype without lens (aphakia and eyeless;10, 11]. However, they also show dopaminergic deficits. The reason for this discrepancy in the range of phenotypic effects of mutations is not known.

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

In the present study, we have analyzed the effect of known mutations in the human and mouse *Pitx3* gene on the structure of the protein and their possible effects on the DNA and protein interaction using bioinformatics tools.

II. MATERIAL AND METHODS

The amino acid sequence of wild-type PITX3 protein was retrieved from Ensembl database (www.ensembl.org). Various sequence analysis tools were used based on their free on web availability, user-friendly approach and their increasing use in bioinformatics. To identify the localization of functional domains in the derived sequence we used Interproscan (http://www.ebi.ac.uk/Tools/pfa/iprsca;) 12]. The prediction of intrinsically disordered regions and identification of week linkages was done by GlobPlot (http://globplot.embl.de; 13]. We used web-based tool; DiANNA for the calculation of long distance interactions by disulfide bonds that might contribute to structural stability of protein (http://clavius.bc.edu/~clotelab/DiANNA; 14]. The prediction of tertiary structure was performed by threading approach using I-Tasser (http://zhanglab.ccmb.med.umich.edu/I-TASSER; 15]. The 3-D structures were evaluated by RAMPAGE server (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php; 16]. BindN was used to predict the DNA binding residues (http://bioinformatics.ksu.edu/bindn/; 17]

III. RESULT

In this study, we used the wild-type PITX3 amino acid sequence to identify the functional domain of the protein. InterproScan identified a 60aa long homeodomain and an OAR domain consisting of 14aa (Figure 1). This wild-type sequence was compared with the mutant proteins sequences S13N, G219fs and 220delG (c. 650delG). Interproscan analysis showed lack of OAR domain in the G219fs. Further, we also included K111E mutant for the comparison; this mutant was cloned and used for *in vitro* studies by Sakazume *et al* [18]. Analysis of these mutant protein revealed that K111E mutant contains an additional homeodomain. This domain covers the region of the mutation and lies between residues 99-118 and 84-95 (Figure 1). We then analyzed this sequence for the interdomain segments using GlobPlot. It identifies the order/globularity and disorder within the protein sequence by Russel-Linding scale. The amino acid position identified to be resistant to mutation are 16-58, 134-163, 177-194, 203-231, 255-260 and 278-292.

Further, we identified the long distance interactions in the wild-type as well as three mutant proteins. DiANNA predicted two disulfide linkages in the wild-type sequence between CYS-33, CYS-129 and CYS-237, CYS-260. Pitx3-K111E mutant also showed the similar pattern. The 220delG mutant showed CYS-33, CYS-236 and CYS-129, CYS-293 linkages. However, S13N mutant sequence resulted in linkage between CYS-260, CYS-294 in addition to CYS-33, CYS-129. The mutant G219fs revealed quite different pattern in the linkages. This sequence has seven cysteine residues compared to wild-type and other mutant sequences with five cysteine residues. Three disulfide linkages detected in G219fs mutant are at position CYS-33, CYS-273; CYS-221, CYS-236 and CYS-241, CYS-288.

The three-dimension structures of the proteins were predicted by I-TASSER server. This program generated five models, out of which the one with a greater value of parameters (C-score, Exp.TM-Score, Exp. RMSD, number of decoys and Cluster density) is selected as the best (Figure 2). The more prominent difference in the structure was observed in 220delG. This structure showed aberrant folding of the protein that also affected the structure of homeodomain. In addition to the human PITX3 structures we also generated and compared wild-type mouse Pitx3 and a mutant mouse (eyeless) protein (Figure 3). The eyeless mutant has an insertion of G at position 416 of Pitx3 cDNA and result in G139fs, causing addition of novel 121 amino acids at C-terminal. The wild-type and mutant PITX3 structures were compared through evaluation results which showed differences in the number of amino acids in favored, allowed and outlier regions (Table 1).

BindN could not detect any difference in the binding sites of the wild-type and mutant PITX3 sequences except for the G219fs that showed 84 binding residues compared to the 41 in wild-type sequence; however, these additional binding residues lie in the C-terminal region of the protein (Figure 4).

IV. DISCUSSION

PITX3 is a transcription factor necessary for the development of dopaminergic neurons and ocular lens. Interestingly, no mutation in this gene has been reported so far explaining the loss of dopaminergic neurons in humans; although, certain nucleotide polymorphisms has been associated with the Parkinson's disease. However, all the known human *PITX3* mutations are reported in context of anterior ocular defects.

Contrary to such situation, the two natural mouse mutants (*aphakia* and *eyeless*) exhibit ocular as well as dopaminergic deficits and therefore, considered as model for PD [10]. Why no mutation in *PITX3* that result in dopaminergic deficit detected in human so far is not known.

In the present study, we tried to predict the structural and thus the functional impact of known mutation in the human and mouse *Pitx3*. Both organisms show 99% homology in the nucleotide as well amino acid sequence of this gene and have similar expression pattern as well [7]. Therefore, it is expected that this gene has similar functional role and impact in both humans and mouse. We analyzed N-terminal (S13N) and C-terminal (G219fs and 220delG) mutations in human PITX3. The 220delG mutant showed most significant difference in the

3-D structure of the protein affecting the homeodomain as well. These results are consistent with the phenotype observed in the 220delG human mutants [8]. The G219fs mutation resulted in an ablation of OAR domain. This domain acts as an intermolecular switch and helps in protein-protein interaction [1, 19]. The *in vitro* experiments done by Sakazume *et al* (2007) showed that the interaction between DNA and S13N through homeodomain is not different from that of the wild-type PITX3 [18]. However, they observed alteration in the binding of G219fs mutant protein with the DNA. These observations could be due to changes in the OAR domain. On the basis of our results and as reported by Sakazume *et al* (2007), it is evident that G219fs can interacts with DNA;

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

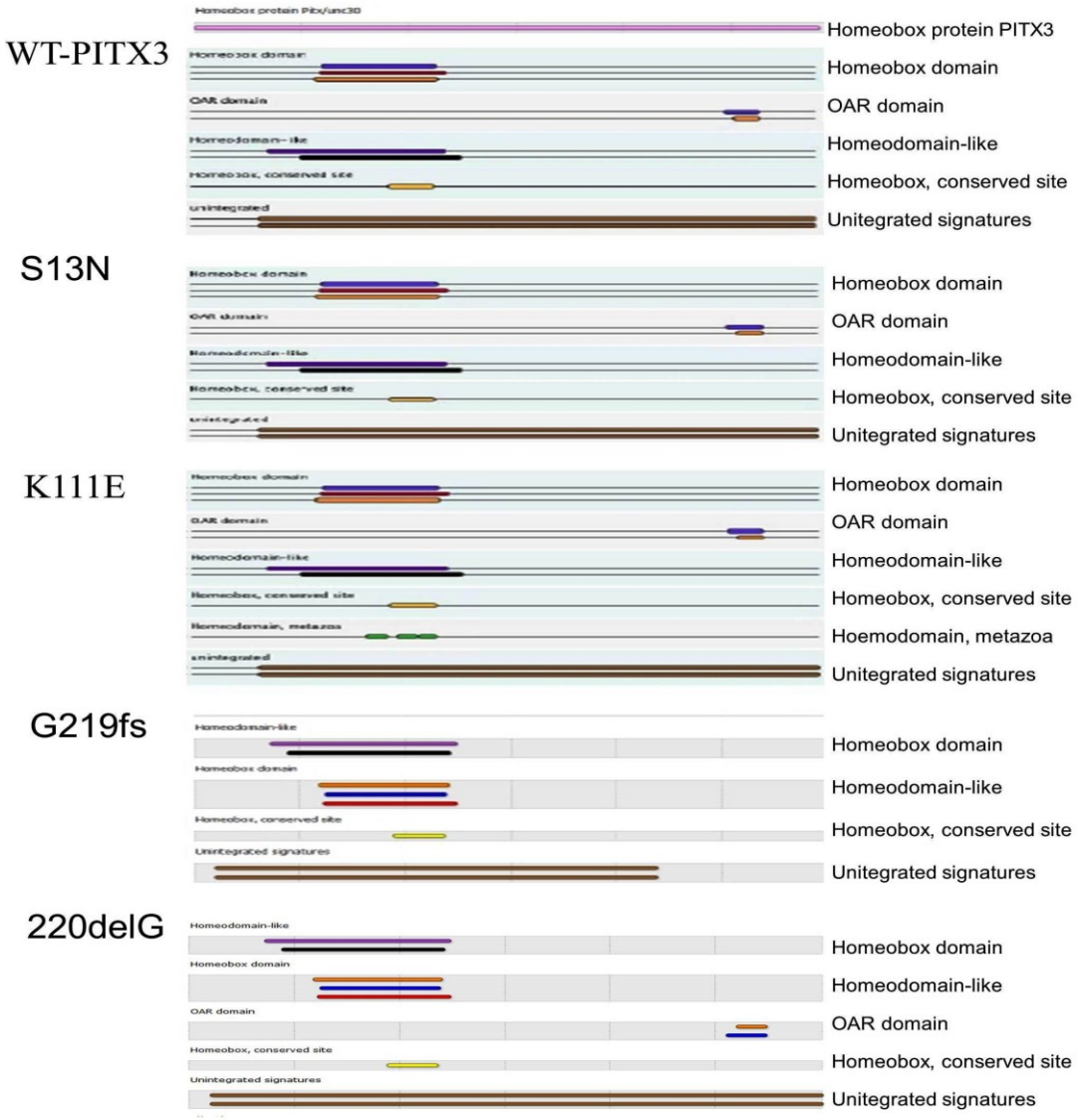

Figure 1: Comparison of predicted functional domains of wild-type and mutated PITX3 protein. The wild-type structure has a homeodomain box and OAR domain. Similar domains are present in the S13N and K111E and 220delG. However, G219fs mutant lacks OAR domain. The K111E has an additional helix turn helix (HTH) loop represented in green boxes. This HTH may have an extended DNA binding.

Table 1: The values of favoured, allowed and outlier region in the protein structures as predicted by RAMPAGE.

Evaluation of residue	human PITX3	S13N	K111E	G219fs	220delG	mouse Pitx3	eyeless (G139fs)
Favored	189 (63.0%)	231 (77%)	192 (64.0%)	217 (70.9%)	224 (73.4%)	235 (78.3%)	235 (91.1%)
Allowed	74 (24.7%)	51 (17%)	76 (25.3%)	49 (16.0%)	48 (15.7%)	43 (14.3)	12 (4.7 %)
Outliers	37 (12.3)	18 (6.0%)	32 (10.7%)	40 (13.1)	33 (10.8%)	22 (7.3)	11 (4.3%)

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

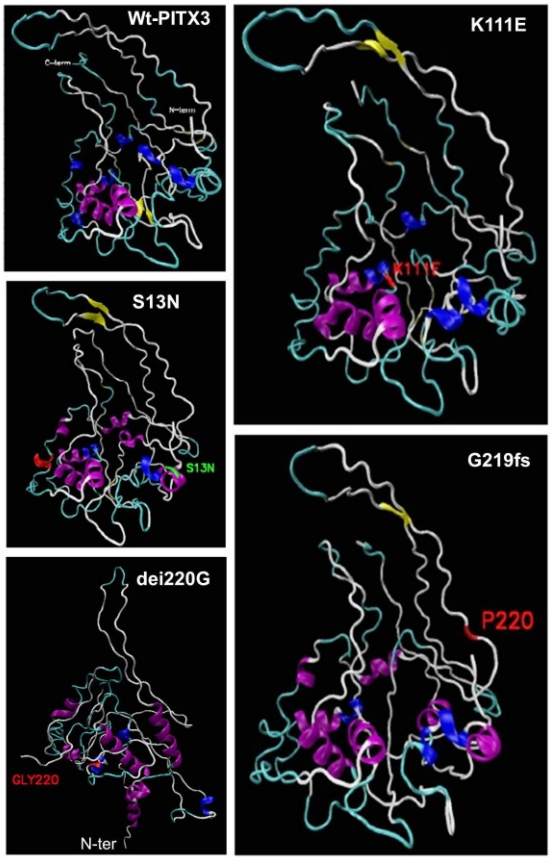

Figure 2: 3D-Structures of wild-type and mutant human PITX3 protein as predicted by I-Tasser. The structures are shown in new cartoon style via VMD software and the colors are in accordance with the secondary structure. The mutant residues are marked with different colors and labeled. No gross abnormality in the helices are observed except in the 220delG mutant, where the helices moved apart disturbing the OAR as well as homeodomain that possibly can influence the binding capabilities of this mutant with DNA and protein.

however, it fails to form the binding complex with other factors like wild-type protein and therefore, results in lack of gene regulatory activity [18]. This notion is further supported by the *in vivo* effect of this mutation, as families exhibiting this mutation show more severe phenotype compares to the S13N mutation [7].

Further, a mutation in the homeodomain box (K111E), that was generated experimentally by Sakazume *et al* (2007) was also analyzed in order to compare the *in silico* and *in vitro* data [18]. In this study, an additional homeodomain box is predicted. However, this helix turn helix (HTH) loop may have different stereo-chemical requirements for the binding, therefore, unable to display the typical binding pattern of the homeobox factors.

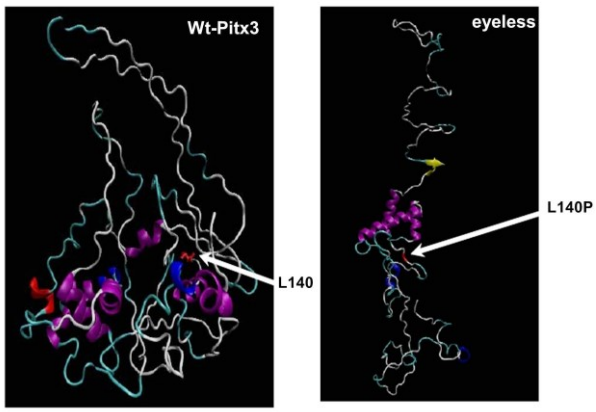

Figure 3: Predicted three-dimensional structures of wild-type and eyeless mouse Pitx3 protein. The eyeless mutant has an insertion of G at position 416 of cDNA. This mutant protein has quite different 3-D structure as the insertion changes the open reading frame and result in 121 new amino acids at C-terminal. The structures are shown in new cartoon style via VMD software and the colors are in accordance with secondary structure.

Figure 4: Prediction of wild-type and mutant PITX3 residues that can potentially bind with DNA. No significant difference in the binding residues was observed in the wild-type and mutant proteins in the homeodomain box. However, G219fs mutant protein show increased binding residues in the C-terminal region. Binding residues are labeled with '+' and in red color; non-binding residues are labeled with '-' and in green. Confidence level are also mentioned; 0 (lowest), 9 (highest).

978-1-4799-7294-4/14/$31.00 ©2014 IEEE

The most interesting results in our modeling data was that of the mouse mutant (eyeless). That protein showed a much distorted structure of the protein. Similar to human mutants, the mouse mutant protein also reserved the homeodomain; however, again the C-terminal part including the OAR domain was disturbed in this Pitx3 mutation. Although, the *in vitro* analysis for this mutation has not been done in any study but the *in vivo* data showed both ocular and dopaminergic deficits (Rosemann *et al.,* 2010) [10]. These results support the fact that the gene regulatory role of this transcription factor depends not only on homeodomain but also the OAR domain. However, one crucial question is that why human PITX3 mutants do not exhibit dopaminergic phenotype compared to the mouse. One explanation could be that the mutation in homeobox or with severe effect are deleterious in human and result in prenatal death. Moreover, this could be justified by not only with the position of mutation itself but with the extent of changes it render on 3-D structure of the protein; as 220delG (c.650delG) mutant showed more broader phenotype including microphthalmia and neurologic deficits [6]. These structural changes not only influence the interaction of the protein with the DNA but also with the co-factors. Thus, the folded structure of the protein determines the temporal range of functional and phenotypic effects of the mutation.

In conclusion, our *in silico* analysis on different mutations in *PITX3* gene is in concordance to the available *in vitro* and *in vivo* data, and provides a good guideline for the *in vitro* analysis of the mutations specifically in the PITX3 and in other proteins in general.

REFERENCES

[1] Brouwer A, ten Berge D, Wiegerinck R, Meijlink F: The OAR/aristaless domain of the homeodomain protein Cart1 has an attenuating role in vivo. Mech Dev 2003, 120:241-252.

[2] Driever W, Nusslein-Volhard C: The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. Nature 1989, 337:138-143.

[3] Ahmad N, Aslam M, Muenster D, Horsch M, Khan MA, Carlsson P, Beckers J, Graw J: Pitx3 directly regulates Foxe3 during early lens development. Int J Dev Biol, 57:741-751.

[4] Fuchs J, Mueller JC, Lichtner P, Schulte C, Munz M, Berg D, Wullner U, Illig T, Sharma M, Gasser T: The transcription factor PITX3 is associated with sporadic Parkinson's disease. Neurobiol Aging 2009, 30:731-738.

[5] Liu J, Sun QY, Tang BS, Hu L, Yu RH, Wang L, Shi CH, Yan XX, Pan Q, Xia K, Guo JF: PITX3 gene polymorphism is associated with Parkinson's disease in Chinese population. Brain Res 2011, 1392:116-120.

[6] Bidinost C, Matsumoto M, Chung D, Salem N, Zhang K, Stockton DW, Khoury A, Megarbane A, Bejjani BA, Traboulsi EI: Heterozygous and homozygous mutations in PITX3 in a large Lebanese family with posterior polar cataracts and neurodevelopmental abnormalities. Invest Ophthalmol Vis Sci 2006, 47:1274-1280.

[7] Semina EV, Ferrell RE, Mintz-Hittner HA, Bitoun P, Alward WL, Reiter RS, Funkhauser C, Daack-Hirsch S, Murray JC: A novel homeobox gene PITX3 is mutated in families with autosomal-dominant cataracts and ASMD. Nat Genet 1998, 19:167-170.

[8] Berry V, Yang Z, Addison PK, Francis PJ, Ionides A, Karan G, Jiang L, Lin W, Hu J, Yang R, et al: Recurrent 17 bp duplication in PITX3 is primarily associated with posterior polar cataract (CPP4). J Med Genet 2004, 41:e109.

[9] Burdon KP, McKay JD, Wirth MG, Russell-Eggit IM, Bhatti S, Ruddle JB, Dimasi D, Mackey DA, Craig JE: The PITX3 gene in posterior polar congenital cataract in Australia. Mol Vis 2006, 12:367-371.

[10] Rosemann M, Ivashkevich A, Favor J, Dalke C, Holter SM, Becker L, Racz I, Bolle I, Klempt M, Rathkolb B, et al: Microphthalmia, parkinsonism, and enhanced nociception in Pitx3 (416insG) mice. Mamm Genome 2010, 21:13-27.

[11] Varnum DS, Stevens LC: Aphakia, a new mutation in the mouse. J Hered 1968, 59:147-150.

[12] Zdobnov EM, Apweiler R: InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics 2001, 17:847-848.

[13] Linding R, Russell RB, Neduva V, Gibson TJ: GlobPlot: Exploring protein sequences for globularity and disorder. Nucleic Acids Res 2003, 31:3701-3708.

[14] Ferre F, Clote P: DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res 2005, 33:W230-232.

[15] Zhang Y: I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008, 9:40.

[16] Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC: Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 2003, 50:437-450.

[17] Wang L, Brown SJ: BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. Nucleic Acids Res 2006, 34:W243-248.

[18] Sakazume S, Sorokina E, Iwamoto Y, Semina EV: Functional analysis of human mutations in homeodomain transcription factor PITX3. BMC Mol Biol 2007, 8:84.

[19] Norris RA, Kern MJ: Identification of domains mediating transcription activation, repression, and inhibition in the paired-related homeobox protein, Prx2 (S8). DNA Cell Biol 2001, 20:89-99

978-1-4799-7294-4/14/$31.00 ©2014 IEEE