

Article Type: Research Article

**Received:** 24/12/2021 **Published:** 26/01/2022



**DOI:** 10.46718/JBGSR.2021.10.000250

# Wilms's Tumor Gene Mutations: Loss of Tumor Suppresser Function:

## **A Bioinformatics Study**

Rukhshan Khurshid¹, Uzma Jabbar²\*, Mazhar Mushtaq³, Asmat Saleem⁴ and Attiya Abbasi⁵

<sup>1</sup>Department of Biochemistry, Shalamar Medical and Dental College. Pakistan

<sup>2</sup>Department of Biochemistry, Fatima Memorial Medical and Dental College. Pakistan

<sup>3</sup>Department of Basic Medical Sciences, Sulaiman Al Rejhi University. Saudi Arabia

International Center for Chemical and Biological Sciences, University of Karachi. Pakistan

<sup>5</sup>Department of Chemistry, University of Karachi. Pakistan

\*Corresponding author: Uzma Jabbar, Department of Biochemistry, Fatima Memorial Medical and Dental College. Pakistan

#### **ABSTRACT**

**Introduction:** Mutation in the Wilms's Tumor (WT1) gene product has been detected in both sporadic and familial cases suggesting that alteration in WT1 may disrupt its normal function. The study aims to find the protagonist amino acid in WT1 proteins by mutating these residues with other amino acids.

Material and Methods: The 3D modeling approach by MODELLER 4 was utilized to build a homology of WT1 proteins. Quality of the WT1 model was verified by predicting 10 models of WT1 and hence selecting the best one. Stereochemistry of model was evaluated by PROCHECK. Mutational studies were done by WHAT IF. Five human WT1 mutations were modeled which were Lys371→Ala371, Ser415→Ala415, Cys416→Ala416, His434→Asp434 and His434→Arg434.

Result: Based on active side of WT1 protein and its role in DNA binding mutation. No significant change was observed when Lys371 was mutated to Ala371, Ser415 was mutated to Ala415. Significant change was observed in Cys416 mutated to Ala416. In mutant Ala416, loss of coordination with the metal ion Zn was also predicted. In case of Mutants His434→Asp434, there was a loss of coordination of metal ion (Zn203) with mutant Asp434. In case of mutant His434→Arg434, there was a loss of Zn203 coordination with Arg434. His434 does not interact directly with any DNA base, whereas mutated Arg434 is predicted to interact directly with DNA base.

**Conclusion:** It is concluded that mutation of amino acid residue Cys416→Ala416, His434→Asp434 and His434→Arg434 may lose the proto-oncogenic function of WT1.

**Keywords:** WT1 protein, MODELLER9.0, Mutation, Active side residues

#### Introduction

WT1, is a protein, which in humans is encoded by the WT1 gene on chromosome 11p13. The WT1 is responsible for the normal kidney development. Mutations in this gene are reported to develop tumors and developmental abnormalities in the genitourinary system. Conversion of proto-oncogenic function to oncogenic in WT1 has also been documented cause of various hematological malignancies.

Multifaceted protein of WT1 gene has transcriptional factor activity [1]. It regulates the expression of insulinlike growth factor and transforming growth factor system, implicated in breast tumorigenesis [2]. A main function of WT1 is to regulate transcription, which control the

expression of genes involved in the process of proliferation and differentiation [3]. In wide range of tumor, WT1 is shown to be predisposing factor for cancer, therefore it has become hot target in research to find out it's inhibitor which can be safely used as a treatment of cancer. It can induce apoptosis in embryonic cancer cell, presumably through the withdrawal of required growth factor survival signal [4]. WT1 is involved in the normal tissue homeostasis and as an oncogene in solid tumors, like breast cancer [5]. Increased expression of WT1 is related with poor prognosis in breast cancer [6]. A number of hypotheses are postulated for the relationship of WT1 with tumorigenesis. Acceding to one of the hypothesis, elevated levels of WT1 in tumors may be related with increased proliferation because normally WT1

have a role with apoptosis [7,8]. Another study proposed that WT1 can alter many genes of the the family of BCL2 [9,10] and also have a role to regulate with Fas-death signaling pathway [11]. Furthermore, it is suggested that WT1 can encourage cell proliferation by up-regulation of protein cyclin D1 [12].

A group of workers hypothesized that WT1 has been observed in the vasculature of some tumour types [13] and its expression may be related with angiogenesis especially in endometrial cancer [14]. Another hypothesis based on the fact that WT1 is a main regulator of the epithelial/ mesenchymal balance and may have a role in the epithelialto-mesenchymal transition of tumor cells [3]. Expression of WT1 is higher in Estrogen Receptor (ER) positive than in ER negative tumors. It is therefore possible that WT1 not only interact with ER alpha, but it may orchestrate its expression [15]. A study, on triple negative breast cancers [7] has shown that high WT1 levels associate with poor survival due to increased angiogenesis [16,17], altered proliferation/ apoptosis [10,11], and induction of cancer- epithelial-tomesenchymal transition [4]. In breast tumors, WT1 is mainly related with a mesenchymal phenotype and increased levels of CYP3A4 [18]. A mutation in the zinc finger region of WT1 protein has been identified in the patients that abolished its DNA binding activity [19]. A study also observed that the mutation in the WT1 gene product has been detected in both sporadic and familial cases suggesting that alteration in WT1 may disrupt its normal function [20]. Bioinformatics approaches are being utilized to resolve the biological problems. Efforts start with the prediction of 3D structures. To achieve the aim, study was designed to view 3D structure of WT1, tumor suppressor protein predicted by homology modeling and to study the role of crucial residues in WT1 proteins by mutating these residues with other amino acids.

#### **Material and Methods**

#### **Structure Prediction**

3D structure of WT1 was taken as target of human WT1. Figure 1 shows the normal interaction of WT1 with DNA strands based on the crystal structure of a zinc finger protein.

The binding of protein-DNA complex involves four zinc finger binding domain in the C-terminal/Zn finger region of WT1. These are;

(i) Cys325, Cys330 and His339, His343: figure 2; yellow highlighted.

(ii) Cys355, Cys360 and His373, His377: figure 3; yellow highlighted.

(iii)Cys385, Cys388 and His397, His401: figure 4; yellow highlighted.

(iv)Cys416, Cys421 and His434, His435: figure 5; yellow highlighted.

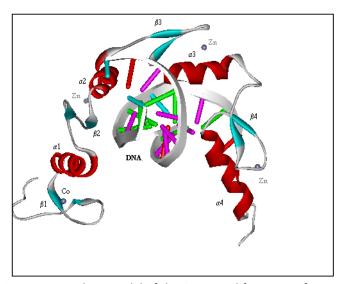


Figure 1: Homology model of the C-terminal fragment of Human Wilms Tumor protein with bound DNA strands based on the crystal structure of a zinc finger protein, znf268, (PDB; id:1aay) and a five-finger protein, GLI (PDB; id:2gli).

The 449 amino acid sequences of WT1 were used for homology modeling. Sequences of WT1 were retrieved from Swiss Prot Data Bank in FASTA format [21]. The best suitable templates were used for 3D-structure prediction. The retrieved amino acid sequences of WT1 were subjected to BLAST [22]. Templates were retrieved on the base of query coverage and identity. The 3D structures were predicted by MODELLER 9.0 [23] that is the requirement of 3D structure building of target protein. Tools including stereochemistry and Ramchandran plots were used for the structure evaluation [24]. Identification of Template was carried out, and Sequence Alignment was carried out by using FASTA, BLAST. Quality of the WT1 model was verified. Stereochemistry of model was evaluated by PROCHECK [25]. Mutational studies were done by WHAT IF [26]. Five human WT1 mutants are modeled. These were: Lys371→Ala371, Ser415 $\rightarrow$ Ala415, Cys416 $\rightarrow$ Ala416, His434 $\rightarrow$ Asp434 and His434→Arg434.

#### **Results and Discussion**

The study was largely based on active side of the WT1 and its role in DNA binding mutation. Zinc finger binding domain interact selectively and non-covalently. This zinc finger-binding domain is the classical zinc finger domain, in which two conserved cysteine and histidine co-ordinate a zinc ion at the active site.

#### Cys416→Ala416 MUTANT

Significant change was observed in Cys416 mutated to Ala416. In mutant Ala416, reduction in the Van der Waal's contact between the amino acids. Loss of coordination with the metal ion Zn was also predicted (Figure 6 A and B).

Figure 2.					Figure 3.				
10	20	30	40	50	10	20	30	40	50
MGSDVRDLNA	LLPAVPSLGG	GGGCALPVSG	AAQWAPVLDF	APPGASAYGS	MGSDVRDLNA	LLPAVPSLGG	GGGCALPVSG	AAQWAPVLDF	APPGASAYGS
60	70	80	90	100	60	70	80	90	100
LGGPAPPPAP	PPPPPPPHS	FIKQEPSWGG	AEPHEEQCLS	AFTVHFSGQF	LGGPAPPPAP	PPPPPPPHS	FIKQEPSWGG	AEPHEEQCLS	AFTVHFSGQF
110	120	130	140	150	110	120	130	140	150
TGTAGACRYG	PFGPPPPSQA	SSGQARMFPN	APYLPSCLES	QPAIRNQGYS	TGTAGACRYG	PFGPPPPSQA	SSGQARMFPN	APYLPSCLES	QPAIRNQGYS
160	170	180	190	200	160	170	180	190	200
TVTFDGTPSY	GHTPSHHAAQ	FPNHSFKHED	PMGQQGSLGE	QQYSVPPPVY	TVTFDGTPSY	GHTPSHHAAQ	FPNHSFKHED	PMGQQGSLGE	QQYSVPPPVY
210	220	230	240	250	210	220	230	240	250
GCHTPTDSCT	GSQALLLRTP	YSSDNLYQMT	SQLECMTWNQ	MNLGATLKGV	GCHTPTDSCT	GSQALLLRTP	YSSDNLYQMT	SQLECMTWNQ	MNLGATLKGV
260	270	280	290	300	260	270	280	290	300
AAGSSSSVKW	TEGQSNHSTG	YESDNHTTPI	LCGAQYRIHT	HGVFRGIQDV	AAGSSSSVKW	TEGQSNHSTG	YESDNHTTPI	LCGAQYRIHT	HGVFRGIQDV
310	320	330	340	350	310	320	330	340	350
RRVPGVAPTL	VRSASETSEK	RP <mark>FMCAYPGC</mark>	NKRYFKLSHL	<b>QMHSRKH</b> TGE	RRVPGVAPTL	VRSASETSEK	RPFMCAYPGC	NKRYFKLSHL	QMHSRKHTGE
360	370	380	390	400	360	370	380	390	400
KPYQCDFKDC	ERRFSRSDQL	KRHQRRHTGV	KPFQCKTCQR	KFSRSDHLKT	KPYQCDFKDC	ERRFSRSDQL	KRHQRRHTGV	KP <mark>FQCKTCQR</mark>	KFSRSDHLKT
410	420	430	440		410	420	430	440	
HTRTHTGKTS	EKPFSCRWPS	CQKKFARSDE	LVRHHNMHQR	NMTKLQLAL	HTRTHTGKTS	EKPFSCRWPS	CQKKFARSDE	LVRHHNMHQR	NMTKLQLAL
Figure 4					Figuro F				
Figure 4.					Figure 5.				
10	20	30	40	50	10	20	30	40	50
10 MGSDVRDLNA	LLPAVPSLGG	GGGCALPVSG	AAQWAPVLDF	APPGASAYGS	10 MGSDVRDLNA	LLPAVPSLGG	GGGCALPVSG	AAQWAPVLDF	APPGASAYGS
10 MGSDVRDLNA 60	LLPAVPSLGG 70	GGGCALPVSG 80	AAQWAPVLDF 90	APPGASAYGS 100	10 MGSDVRDLNA 60	LLPAVPSLGG 70	GGGCALPVSG 80	AAQWAPVLDF 90	APPGASAYGS 100
10 MGSDVRDLNA 60 LGGPAPPPAP	LLPAVPSLGG 70 PPPPPPPPHS	GGGCALPVSG 80 FIKQEPSWGG	AAQWAPVLDF 90 AEPHEEQCLS	APPGASAYGS 100 AFTVHFSGQF	10 MGSDVRDLNA 60 LGGPAPPPAP	LLPAVPSLGG 70 PPPPPPPPHS	GGGCALPVSG 80 FIKQEPSWGG	AAQWAPVLDF 90 AEPHEEQCLS	APPGASAYGS 100 AFTVHFSGQF
10 MGSDVRDLNA 60 LGGPAPPPAP 110	LLPAVPSLGG 70 PPPPPPPPHS 120	GGGCALPVSG 80 FIKQEPSWGG 130	AAQWAPVLDF 90 AEPHEEQCLS 140	APPGASAYGS 100 AFTVHFSGQF 150	10 MGSDVRDLNA 60 LGGPAPPPAP 110	LLPAVPSLGG 70 PPPPPPPPHS 120	GGGCALPVSG 80 FIKQEPSWGG 130	AAQWAPVLDF 90 AEPHEEQCLS 140	APPGASAYGS 100 AFTVHFSGQF 150
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210	LLPAVPSLGG 70 PPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT	LLPAVPSLGG 70 PPPPPPPPS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT	LLPAVPSLGG 70 PPPPPPPPS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260	PPPPPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW	PPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310	PPPPPPPPSQA 170 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310	LLPAVPSLGG 70 PPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL	PPPPPPPPSQA 170 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL	APPGASAYGS 100  AFTVHFSGQF 150  QPAIRNQGYS 200  QQYSVPPPVY 250  MNLGATLKGV 300  HGVFRGIQDV 350  QMHSRKHTGE
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL 360	PPPPPPPPSQA 170 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK 370	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC 380	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL 390	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE 400	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL 360	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK 370	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC 380	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL 390	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE 400
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL 360 KPYQCDFKDC	PPPPPPPPSQA 170 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK 370 ERRFSRSDQL	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC 380 KRHQRRHTGV	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL 390 KPFQCKTCQR	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE 400	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL 360 KPYQCDFKDC	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK 370 ERRFSRSDQL	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC 380 KRHQRRHTGV	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL 390 KPFQCKTCQR	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE 400
10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL 360 KPYQCDFKDC 410	PPPPPPPPSQA 170 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK 370 ERRFSRSDQL 420	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC 380	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL 390 KPFQCKTCQR 440	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE 400 KFSRSDHLKT	10 MGSDVRDLNA 60 LGGPAPPPAP 110 TGTAGACRYG 160 TVTFDGTPSY 210 GCHTPTDSCT 260 AAGSSSSVKW 310 RRVPGVAPTL 360 KPYQCDFKDC 410	LLPAVPSLGG 70 PPPPPPPPHS 120 PFGPPPPSQA 170 GHTPSHHAAQ 220 GSQALLLRTP 270 TEGQSNHSTG 320 VRSASETSEK 370 ERRFSRSDQL 420	GGGCALPVSG 80 FIKQEPSWGG 130 SSGQARMFPN 180 FPNHSFKHED 230 YSSDNLYQMT 280 YESDNHTTPI 330 RPFMCAYPGC 380 KRHQRRHTGV 430	AAQWAPVLDF 90 AEPHEEQCLS 140 APYLPSCLES 190 PMGQQGSLGE 240 SQLECMTWNQ 290 LCGAQYRIHT 340 NKRYFKLSHL 390	APPGASAYGS 100 AFTVHFSGQF 150 QPAIRNQGYS 200 QQYSVPPPVY 250 MNLGATLKGV 300 HGVFRGIQDV 350 QMHSRKHTGE 400 KFSRSDHLKT

Figure 2-5: Amino Acid sequence of human Wilms Tumor protein. Highlighted C-terminal fragment are the zinc binding domain of the protein.

Cys416 is located at the domain interface with its polar side chain completely buried (0.00 Å). Replacement of this amino acid may account for considerable changes in the interior of protein (Table 1). We have predicted the possible changes that arise due to the mutation of Cys to Ala by molecular modeling experiments. Amino acids, Pro419, Ser420, Cys421, His434 and some atoms of His438 (ND1, NE2, CD2 and CE1) are present near Cys416. Zinc (Zn203) is also present in the vicinity (1.82 Å) of Cys416 (Figure 6 A and B). The mutated residue, Ala is also predicted to remain buried (0.00 Å) in the interior of protein. Significant

change is observed however, in the surrounding area of the mutated Ala416. Only a few atoms of His434 (CD2 and NE2) and His438 (CE1) were seen in the surrounding. This may reduce the Van der Waal's contacts between the respective amino acids. The loss of coordination with the metal ion, zinc was also predicted as the distance is increased from 1.82 Å to3.12 Å. It is therefore predicted that Cys416 plays a vital role in the interaction with other amino acid residues as well as in the metal coordination. It is observed that there is a possibility of loss of these interactions in case of Cys416 replacement.

Figure 6A

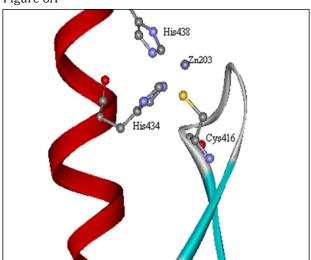
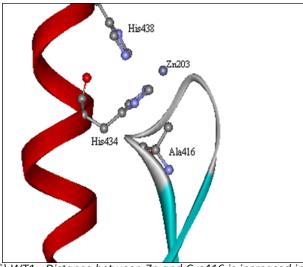


Figure 6B



**Figure 6 A and B**: Wild Type (Cys416) and mutated (Ala416) WT1. Distance between Zn and Cys416 is increased in mutated (Ala416) model. Cys416 is predicted to be found in the vicinity of His434 and His438 which are implicated in catalysis (6A) while Ala416 can only interact with His434 and not with His438 in the mutated model (6B).

#### His434→Arg434 MUTANTS

In case of mutant His434 $\rightarrow$ Arg434, there was a loss of zn203 coordination with Arg434. His434 does not interact directly with any DNA base, whereas mutated Arg434 is predicted to interact directly with DNA base, A1. This suggests that change might effect on the DNA binding

proved that hydrogen bonding network close to the zincbinding motif plays a significant role in stabilizing the coordination of the zinc metal ion to the protein [23]. The mutated amino acid, Arg434 also moved considerably form buried to relatively exposed environment (2.28 Å to 5.35 Å). Presence of positively charged Arg on the surface could

Figure 7A

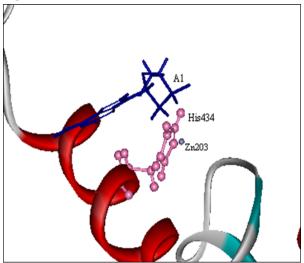
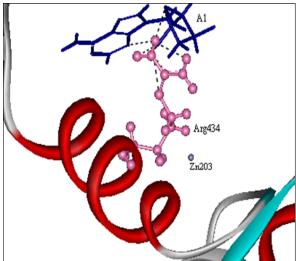


Figure 7B



**FIGURE 7 A and B:** Wild Type (His434) and mutated (Arg434) WT1. Distance between Zn and Arg434 is increased in mutated model. Arg434 is predicted to bind DNA base A1 (B) while His434 in the original model (A) show no bonding with DNA base.

pattern, Figure 7 A and B.

In case of mutation of His434→Arg434, the distance between the mutated Arg and zinc (Zn203) was increased from 2.28 Å to5.00 Å suggesting that there could be a loss of coordination with the metal ion. Mutational studies

account for additional interaction of the protein with other proteins or with the surrounding water molecules. His434 does not interact directly with any DNA base whereas mutated Arg434 is predicted to interact directly with DNA base Adenine, A1. (Figure 7 A and B). This suggests that the change might cause the DNA binding pattern.

**TABLE 1:** Comparison of surface accessibilities (Å) of the wild type and mutated residues and those in the vicinity of the mutated residues in the five WT1 mutants

Amino acid	Wild type	Mutant
	(Lys371)	(Ala371)
Leu370	12.93	13.28
Lys/Ala371	47.04	33.66
	(Ser415)	(Ala 415)
Cys416	0.00	0.00
Ser/Ala415	7.96	7.62
	(Cys416)	(Ala 416)
His434	8.96	0.53
Cys/Ala416	0.00	0.00
	(His434)	(Asp 434)
Glu430	14.83	0.00
His/Asp434	8.96	0.53
	(His434)	(Arg434)
Glu430	14.83	0.00
His/Arg434	8.96	4.43

Lys371→Ala371 and Ser415→Ala415 MUTANT

No significant change was observed when Lys371 was mutated to Ala371, and Ser415 was mutated to Ala415. It is observed in this mutation that the change that arise in the overall structure and surrounding amino acid residues (Table 1). Lys371 is present on the surface (accessibility = 47.04 Å) of the WT1 molecule. It was observed that the internal protein structure was not affected considerably, as Lys371 is present on the outer most surface of the protein. In the original model, Lys371 stacks against thymine. It also forms a water-mediated contact with side chain hydroxyl of Ser367. Although, Ala371 also stacks against the same DNA base but the distance is slightly altered. The hydrogen bond between Ala371 and Ser367 has not been predicted in the mutated model. It has been demonstrated that mutation within finger 2 and 4 abolished sequence specific binding of WT1 to DNA bases [19]. The mutation of the corresponding lysine in a peptide could reduce its affinity for DNA seven folds [27]. On the other hand, it is reported [28] that a surface mutation would not cause a significant change in the internal structure of protein. However, the replacement of a basic polar residue with a non-polar one could account for a reduction in polarity. The modeling studies of Lys to Ala mutation do not however support this finding and require further analysis.

Mutation of Ser415→Ala415 in the WT1 model (Table 1). Ser415 is located near the active center of WT1. It has been demonstrated that Ser415 makes a water-mediated

contact with phosphate of DNA base, guanine [20]. In our predicted model of WT1, Ser415 makes two water (numbers 516 and 568) mediated contacts. Mutation of this Ser with Ala resulted in the loss of one of these contacts leading to the loss of binding. The replacement of relatively polar residue, Ser to a non-polar one, Ala could account for this reduced interaction. This is also evident by a slight decrease in the accessibility of Ala (Ser415 = 7.96 Å; Ala415 = 7.61 Å).

### His434→Asp434 MUTANTS

In case of mutants His434→Asp434, there was a loss of coordination of metal ion (zn203) with mutant Asp434. Glu430 move from relatively exposed to completely buried environment. His434 is also present at the active center of WT1. We predicted two mutants; His434→Asp434 and His434→Arg434 mutants by molecular modeling (Table 1). In case of His434→Asp434 mutation, the water mediated contact is lost. The distance between mutated Asp and zinc (Zn203) was also increased from 2.28 Å to 3.57 Å suggesting that there could be a loss of coordination with the metal ion as well. The amino acid Glu340 that is present near His434 also moved considerably form relatively exposed to completely buried environment (14.83 Å to 00.0 Å).

#### **Conclusion**

It is concluded that mutation of amino acid residue Cys416→Ala416, His434→Asp434 and His434→Arg434 of WT1 may lose its function to regulate the function of genes by binding to specific parts of DNA. Besides the mutation of above-mentioned amino acid residue, the role of WT1 in cell growth, cell differentiation, apoptosis and tumor suppressor function is also lost.

#### **Conflict of Interest**

Authors declares no conflict of interest.

#### **References:**

- Hohenstein P, Hastie ND (2006) The many facets of the Wilms' tumour gene, WT1. Human molecular genetics. 15(2): 196–201.
- 2. Gupta S, Joshi K, Wig JD, Arora SK (2009) High frequency of loss of allelic integrity at Wilms' tumor suppressor gene-1 locus in advanced breast tumors associated with aggressiveness of the tumor.Indian J Cancer. 46(4): 303-310.
- 3. Roberts SG (2006) The modulation of WTI transcription function by cofactors.BiochemSoc Symp (73):191-201.
- Nishida S, Sugiyama H (2016) Immunotherapy targeting WT1: designing a protocol for WT1 peptide-based cancer vaccine. Methods Mol Biol 1467: 221-232.
- Hou X, Maheswaran S, Bennett P, Ngwu C, Re GG (1995) WT1 suppresses synthesis of the epidermal growth factor receptor and induces apoptosis. The EMBO J 14(19): 4662-4675.
- 6. Chau YY, Hastie ND (2012) The role of Wt1 in regulating mesenchyme in cancer, development, and tissue homeostasis. Trends in genetics. 515–524.
- 7. Stirzaker C, Elena Zotenko, Jenny Z (2015) Song Methylome

- sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. Nature Communications. 6:1-11.
- 8. Zamora-Avila DE (2007) RNAi silencing of the WT1 gene inhibits cell proliferation and induces apoptosis in the B16F10 murine melanoma cell line. Melanoma research. 17: 341–348.
- 9. Tatsumi N, Oji Y, Tsuji N (2008) Wilms' tumor gene WT1-shRNA as a potent apoptosis-inducing agent for solid tumors. International journal of oncology. 32:701–711.
- 10. Morrison DJ, English MA, Licht JD (2005) WT1 induces apoptosis through transcriptional regulation of the proapoptotic Bcl-2 family member Bak. Cancer research 65: 8174–8182.
- 11. Cesaro E (2010) WT1 protein is a transcriptional activator of the antiapoptotic bag3 gene. Leukemia. 24: 1204–1206.
- 12. Uesugi K (2013) Wilms' tumor 1 gene modulates Fas-related death signals and anti-apoptotic functions in hepatocellular carcinoma. Journal of gastroenterology. 48: 1069–1080.
- 13. Caldon CE, Lee CS, Sutherland RL, Musgrove, EA (2008) Wilms' tumor protein 1: an early target of progestin regulation in T-47D breast cancer cells that modulates proliferation and differentiation. Oncogene. 27: 126–138.
- 14. Wagner N, Michiels JF, Schedl A and Wagner KD (2008) The Wilms' tumour suppressor WT1 is involved in endothelial cell proliferation and migration: expression in tumour vessels *in vivo*. Oncogene 27: 3662–3672.
- 15. Dohi S, Ohno S, Yumiko S (2010) WT1 expression correlates with angiogenesis in endometrial cancer tissue. Anticancer research 30: 3187–3192.
- 16. Han Y, Yang L (2008) Wilms' tumor 1 suppressor gene mediates antiestrogen resistance via down-regulation of estrogen receptor-alpha expression in breast cancer cells. Molecular cancer research 6: 1347–1355.
- 17. Artibani M, Sims AH, Slight J, Aitken S, Thornburn A, et al. (2017) *WT1* expression in breast cancer disrupts the epithelial/mesenchymal balance of tumour cells and correlates with the metabolic response to docetaxel. Scientific Reports 7: 45255.

- 18. Raucher FJ, Morris JF, Tournay OE (1990) Binding of the Wilm's tumor locus zinc finger protein to the EGR-1 consensus sequence. Science 250:1259-1262.
- 19. Wang ZY, Qiu QQ, Huang J (1995) Products of alternatively spliced transcript of wilm's tumor suppressor gene WT1 have altered DNA binding specificity and regulate transcription in different ways. Oncogen 10: 415-422.
- 20. Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000.Nucleic Acids Res 28(1): 45–48.
- 21. Altschul SF, Madden TL, Schaffer AA (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.
- 22. Eswar N, Eramian D, Webb B, Shen M, Sali A (2008) Protein Structure Modeling With MODELLER. Methods in Molecular Biology 426(1): 145-159.
- 23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr 26:283–291.
- 24. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8(1): 52-56.
- 25. Nakagama H, Heinrich G, Pelletier J, Housman DE (1995) Sequence and structural requirements for high-affinity DNA binding by the WT1 gene product. Mol Cell Biol. 15: 1489–1498.
- Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO (1996)
   Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. Structure 4: 1171-1180.
- 27. Hejtmancik ID, Gorman MA, Freemont PS (1998) The genesis of cataract. Our visionbecome clearer. Am J Hum Gene 62: 520-
- 28. Fukasawa KM, Hirise J, Hata T, Ono Y (2010) In rat dipeptidyl peptidase III, His568 is essential for catalysis, and Glu507 or Glu512 stabilizes the coordination bond between His455 or His450 and zinc ion," Biochimica et Biophysica Acta 1804: 2063–2069.