OFFICIAL JOURNAL

ORIGINAL ARTICLE

A Molecular Analysis of Homo Sapiens' Tumor-causing TP53 gene and associated P53 protein via Bioinformatics pipeline

Syed Awais Attique[1], Qurat-Ul-Ain[2], Syed Asim Attique Shah[3], Muhammad Hassan[4,*], Sara Zahid[5], Muhammad Usman[6], Imran Shehzad[5], Rabia Kiran[7], Atif Amin Baig[4], Ghulam Mustafa[8]

[1]Research Centre for Modelling and Simulation (RCMS), National University of Sciences and Technology (NUST), Islamabad, H-12, Pakistan
[2]Centre for Applied Molecular Biology (CAMB), University of the Punjab, Lahore, 53700, Pakistan
[3]Department of Computer Science, Air University Islamabad, Pakistan.
[4]Faculty of Medicine, Universiti Sultan Zainal Abidin (UniSZA), Jalan Sultan Mahmud, 20400, Kuala Terengganu, Terengganu Darul Iman, Malaysia
[5]Centre of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad, 38000, Pakistan
[6]Department of Computer Science, University of Agriculture, Faisalabad, 38000, Pakistan
[7]Department of Zoology, Wildlife and Fisheries, University of Agriculture, Faisalabad, 38000, Pakistan
[8]Department of Biochemistry, Government College University Faisalabad-38060, Pakistan

Abstract

Objective: The current study investigates characteristics and post-translational modification (PTM) sites of *TP53* and its encoding protein (P53) through a bioinformatics pipeline. The findings of this study will provide a gateway to novel therapeutics strategies against numerous cancers followed by *TP53* in humans. **Methods:** In this study, we utilized *in-silico* strategies to evaluate the PTM site response in P53 protein leading to cancer treatment. PSIPRED online tool predicted the secondary structure of P53 and interpret its robustness to mutation. ProtParam determined the characterized physiology and NetPhos v3.1 combined with Phos3D provided possible phosphorylation sites of P53. N- and O-glycosylation site identification evaluated by NetNGlyc 1.0 and NetOGlyc 4.0 servers, which is an important factor in determining cancerous activity level. YinOYang v1.2 predicted possible Yin-Yang sites involved in the carcinogenic activity of P53. **Results**: Human P53 protein has 42 phosphorylation sites (30 Ser + 11 Thr + 1 Tyr) followed by several kinases involved in its transcriptional activity. A total of 16 O-glycans and 2 N-glycans are determined in P53 protein. However, 8 possible YinYang sites are also predicted, which are involved in the oncogenic nature of P53 protein. **Conclusions**: *TP53* genes are responsible for several cancers in humans, especially like i-Fraumeni (LFL) syndrome. Its encoding protein P53 shows carcinogenic activity in the human body. Several predicted YinYang sites as well as N- and O-glycans are involved in P53 protein's carcinogenic behavior. These findings will be helpful in the treatment of cancer patients by targeting these identified sites.

Keywords: *TP53*; Cancer-causing gene; Phosphorylation; P53 Yin Yang sites; Therapeutics targets

Introduction

In 1969, a hereditary type Li-Fraumeni Syndrome (LFS) cancer was reported by Fraumeni and Li (Birch et al., 1994). LFS causes several types of cancers including breast cancer, acute leukemia, brain cancer, adrenal cortical tumors and soft-tissue sarcoma. In addition, melanoma (skin cancer), Wilm's tumor (kidney cancer) and higher cancer risk for germ cells, pancreas, esophagus, stomach and lungs were also observed (Bhurgri, 2004). A huge percentage of LFS (30%) was

Received: 8 February 2021**; accepted revised manuscript:** 1 April 2021 **Published online:** 25 April 2021
***Corresponding author:** Muhammad Hassan, Faculty of Medicine, Universiti Sultan Zainal Abidin (UniSZA), Jalan Sultan Mahmud, 20400, Kuala Terengganu, Terengganu Darul Iman, Malaysia
Tel: +609-6275646 **Email**: m.hassan2381996@gmail.com

 | ISSN (Online): 2716-649X

reported in women at 30 years of age (Birch et al., 1994). As compared to Caucasian women, Indian and Pakistani women were more frequently diagnosed with breast cancer (16.2 vs. 6.2%) before the age of 40 years old (Bhurgri, 2004). Germline mutation of *TP53* had been reported as a leading factor of LFS in 83% of individuals (Evans et al., 2002). Therefore, the mutational *TP53* is considered a major gene for causing cancer in humans. *TP53* mutations were found to be distinctive between different histopathological types of tumors and the tumors between youngsters and adults (Rasheed et al., 1994). In breast and lung cancer patients, mutation of the TP53 gene was experimentally found in the plasma from diseased cells' DNA (Silva et al., 1999).

Consequently, P53 protein is associated with several types of cancers, so its molecular characterization can play an important role in finding potential therapeutics against cancer. Although wet lab techniques are the best way for molecular characterization but these conventional procedures are arduous, tedious, and have moral issues. Thus, we utilized bioinformatics tools, well-trained with experimental data, for molecular characterization and identification of the mutational effect of P53 protein on its function. The findings of the current study will provide a gateway to novel therapeutic strategies against numerous cancers in humans by targeting lethal *TP53* mutations.

Methodology

Characterization of Homo sapiens tumor protein P53

TP53 gene encodes a protein called P53. The primary sequence of human P53 protein (393 residues) was retrieved from UniProtKB (Boutet et al., 2007) database using accession number P04637-1 with the entry name "Cellular tumor antigen P53". Protein BLAST of (all 393 amino acids) was performed using NCBI (National Center for Biotechnology Information) with default parameters. The primary sequences of different organisms were selected based on significant alignment score, percentage identity (> 80%), query coverage (> 80%) and e-score ≤ 0. These retrieved sequences of numerous species were used for phylogenetic analysis by MEGA software (Sohpal et al., 2010). Physicochemical properties of the human P53 protein, which include molecular weight, instability index, theoretical isoelectric pH (pI) (isoelectric pH), aliphatic index, extinction coefficient and grand average of hydropathicity (GRAVY), were studied using the ProtParam server (ProtParam, 2017). Secondary structures including alpha helix, beta sheets and coils were calculated by PsiPred (McGuffin et al., 2000), while motifs and domains were identified by MEME suite server (Bailey et al., 2009) and NCBI Conserved Domain Database (Marchler-Bauer et al., 2015) respectively. Human P53 signal peptide analysis was performed using SignalP-5.0 (Armenteros et al., 2019) server. The protein's membrane topology was determined by the web-based Protter program (Omasits et al., 2014). Intermediate, exposed, buried, ordered, disordered residues and solvent accessibility along with protein binding regions were identified using PredictProtein (Rost et al., 2004).

Phosphorylation of human P53

Post-translation modification (PTM) can adjust the steadiness, subcellular area, and enzymatic movement of proteins with assorted parts in cells. Protein phosphorylation is catalyzed by chemicals known as protein kinases, while the opposite response (i.e., the expulsion of phosphates from proteins) is interceded by protein phosphatases. The Phosphorylation sites of human P53 were identified by using NetPhos 3.1. This server is based on an artificial network technique and predicts highly potential phosphorylation sites/regions on each serine, threonine and tyrosine residue, by calculating a potential value of the scoring function. This program uses 0.5 value as a threshold for each serine, threonine and tyrosine residue for potential phosphorylation site identification (Simanon et al., 2019).

Phos3D (Durek et al., 2009) gave us potential phosphorylation sites (P-sites) along with functionally potential domains from human P53 protein tertiary structure produced homology modeling by taking "PDB ID: 4mzr.1.B" as a template using SwissModel (Schwede et al., 2003). This program uses the Support Vector Machine (SVM), a type of supervised machine learning algorithm, for potential P-sites identification. However, Phos3D is capable in identifying kinase-based phosphorylation of Ser, Thr and Tyr potential P-sites using GSK3, CDC, ATM, PKA, MAPK, PKC, CKII Ser kinases and SRC Tyr kinase.

 | ISSN (Online): 2716-649X

Scansite 4.0 was used to confirm Phos3D results, as this program also predicts phosphorylation sites using different kinase activities. Scansite can predict kinase activity as well as surface accessibility (SA) for each Ser, Thr, Tyr and all other residues. The primary sequence at each target site (seven residues on respective sides of potentially modified amino-acid) was assessed according to preferences of specific-protein kinases. In comparison with all vertebrate motifs present in the SWISS-PROT database, Scansite indicates the candidate's motif value ranking in percentiles. This program uses a 0.5 SA threshold for each amino-acid. If any residue has SA ≥ 0.5, then that residue is considered to be exposed on the surface of the protein and accessible for PTM (Obenauer et al., 2003).

N- and O-glycosylation in human P53

N-glycans are linked with ER to the nitrogen atom in the asparagine side chain by making a triplet of Asparagine-X-Serine or Asparagine-X-Threonine sequence. O-glycans are present in the Golgi apparatus and assembled a sugar molecule on Ser or Thr residue. NetOGlyc 1.0 server (Gupta et al., 2004) and NetNGlyc 4.0 server (Steentoft et al., 2013) were used for the determination of GalNAc O-glycosylation and N-glycosylation sites respectively.

O-β-GlcNAc and phosphorylation interplay in Yin-Yang sites of Human P53 protein

O-linked glycosylation comprises the transmission of β-N-acetylglucosamine to the hydroxyl group of serine/threonine residues of target proteins. Increasing evidence proposes that target protein's O-GlcNAcylation is a common process and similar to phosphorylation in executing wide ranges of biological practices (Wang et al., 2011). Human P53 protein's O-β-GlcNAc modifications were determined by the YinOYang 1.2 server (Malik et al., 2007). YinOYang computes all possible potential O-β-glycosylation sites against all serine and threonine residues in a target protein. It crosschecks O-β-GlcNAc sites with NetPhos 3.1 estimations for potential phosphorylated regions to determine Yin-Yang sites with a high level of significance for both modifications.

Results

Characterization of Human tumor suppressor protein P53

Human P53 protein has 61 residues involved in alpha helices, 81 residues involved in beta-sheets, and 251 residues involved in coils formation respectively (Figure 1).

Alpha helices and beta sheets are more robust for mutations as compared to coils. Whereas, beta sheets are less robust for mutations as compared to alpha-helices (Abrusán and Marsh, 2016). As the human P53 protein's secondary structure has more beta sheets (81 amino acids residues) as compared to alpha helices (61 amino acids residues), its secondary structure is more stable to mutations. The membrane topology of the P53 protein identified by Protter determined that the whole protein (393 amino acids residues) exists outside the membrane with no signal peptide, as shown in Figure 2.

Human P53 has a 43.65 kDa molecular weight and 6.33 pI. Its instability and aliphatic index are 73.59 and 59.08. It has a GRAVY value of -0.756. The estimated half-lives are 30 hours in mammalians (*in vitro*), 20 hours in yeasts (*in vivo*) and 10 hours in Escherichia coli (*in vivo*), as shown in Table 1**.** As the instability index of human P53 protein is greater than 40.0, it is unstable in a test tube during experimentation so needs to be dealt with care and limitations. The high aliphatic index indicates its thermo-stability due to larger occurrences of alanine, valine, isoleucine, and leucine residues (Ikai, 1980). Human P53 has a negative charge on it and is present inside the membrane.

Human P53 protein has 25 motifs with different residual ranges starting from 3-8, 11-16, 19-24, 26-31, 32-37, 38-43, 44-49, 50-55, 59-64, 66-71, 74-81, 84-91, 95-100, 101-107, 108-115, 120-126, 127-132, 134-139, 143-148, 152-157, 158-163, 167-172, 175-180, 181-186, 192-197, 199-205, 209-217, 218-223, 224-229, 232-238, 242-248, 254-259, 261-266, 267-275, 277-283, 287-292, 293-298, 300-305, 308-313, 316-321, 327-332, 334-340, 341-346, 348-354, 355-360, 361-368, 369-375, 379-385, 386-391 and 388-393 respectively (Figure 1). Four domains P53_TAD, TAD2, P53, P53 and P53_tetramer with ranges from 6-30, 35-59, 109-288 and 319-355 respectively (Figure 3).

 | ISSN (Online): 2716-649X

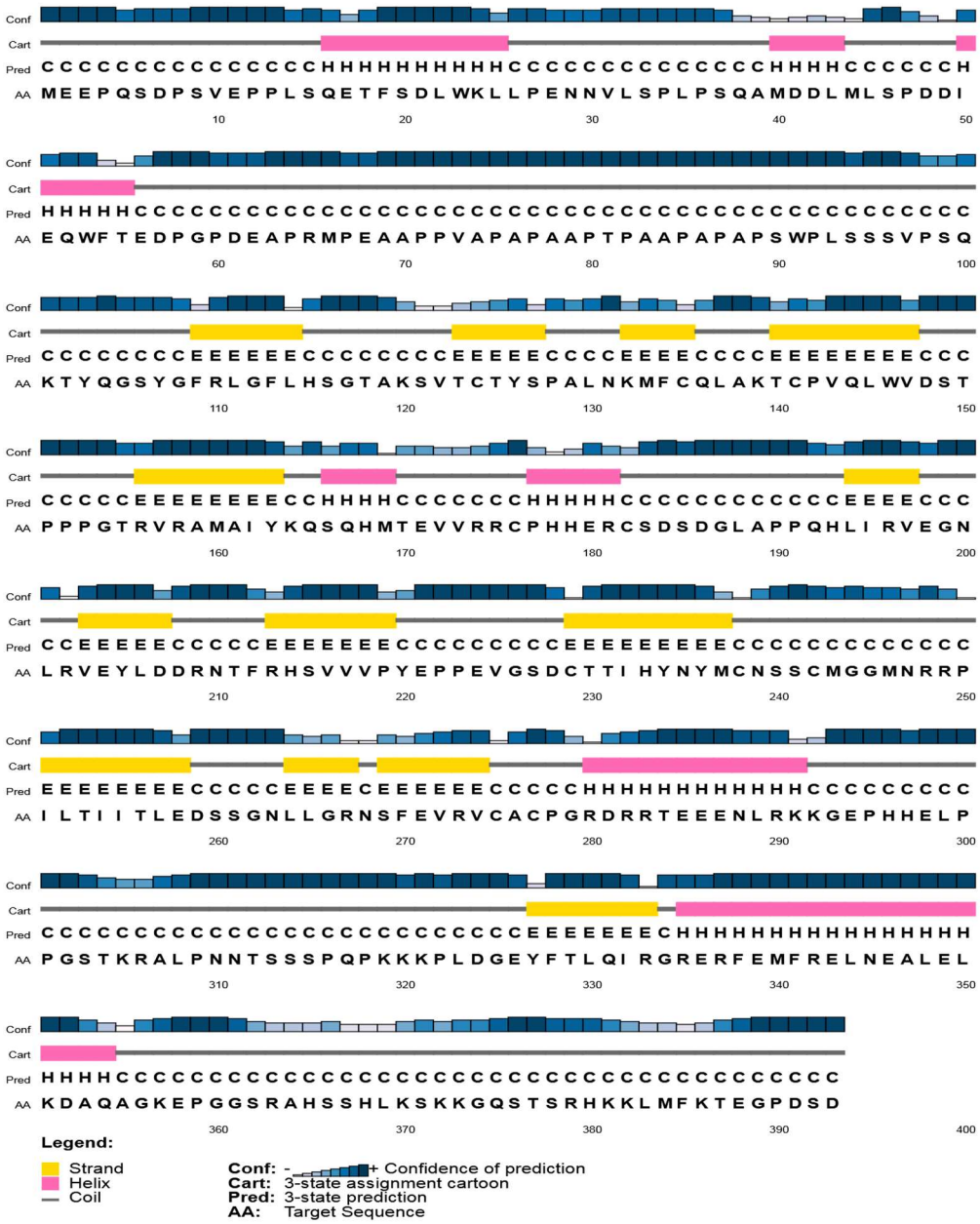

Figure 1. Human p53 protein secondary structure; alpha helices (61 residues), beta sheets (81 residues) and coils (251 residues) are presented in pink, yellow and light grey colors receptively.

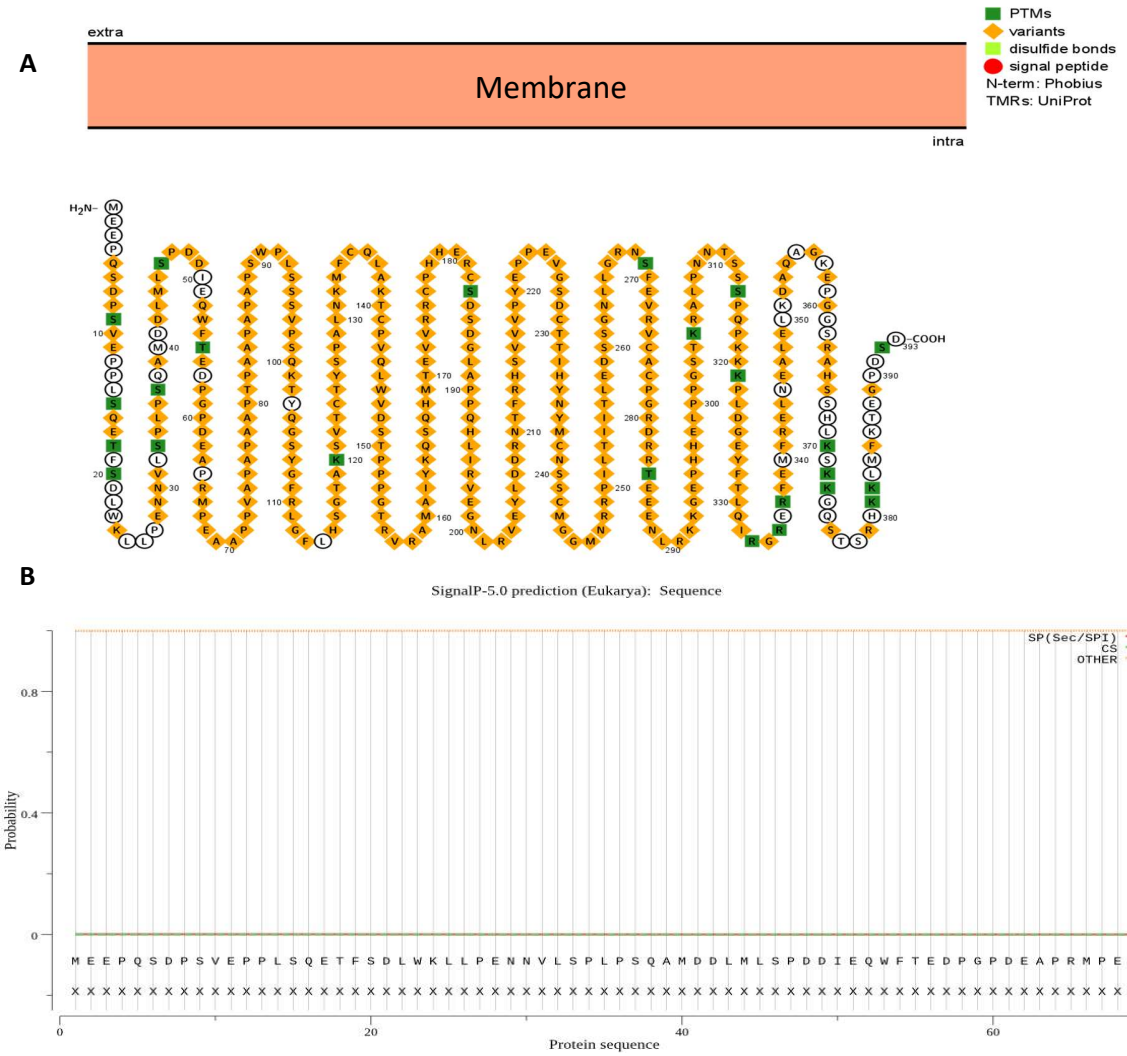

Figure 2. Representation of membrane topology and signal peptide of human P53 protein. **(A)** All residues of the protein (393 aa) are present inside the membrane, **(B)** it has no signal peptide throughout the protein**.**

Phosphorylation of human P53 protein

Identification of potential phosphorylation sites in human P53 protein was evaluated by using NetPhos v3.1 (implementing neural network algorithm) and Phos3D for kinase identification. A total of 42 significant potential sites were identified for phosphorylation. Out of these 42 phosphorylation sites, 30 were Serine (Ser), 11 were Threonine (Thr) and 1 was Tyrosine (Tyr) residues (Table 2 and Figure 4).

Table 1. Physicochemical properties of tumor suppressor protein P53 (*Homo sapiens*)

Sr. No.	*Property*	*Value*
	Molecular weight	43.65 kDa
	Theoretical pI	6.33
	Instability index	73.59
	Aliphatic index	59.08
	GRAVY	-0.756
	Estimated half-life	30 hours (mammalian, in vitro) 20 hours (yeast, in vivo) 10 hours (Escherichia coli, in vivo)

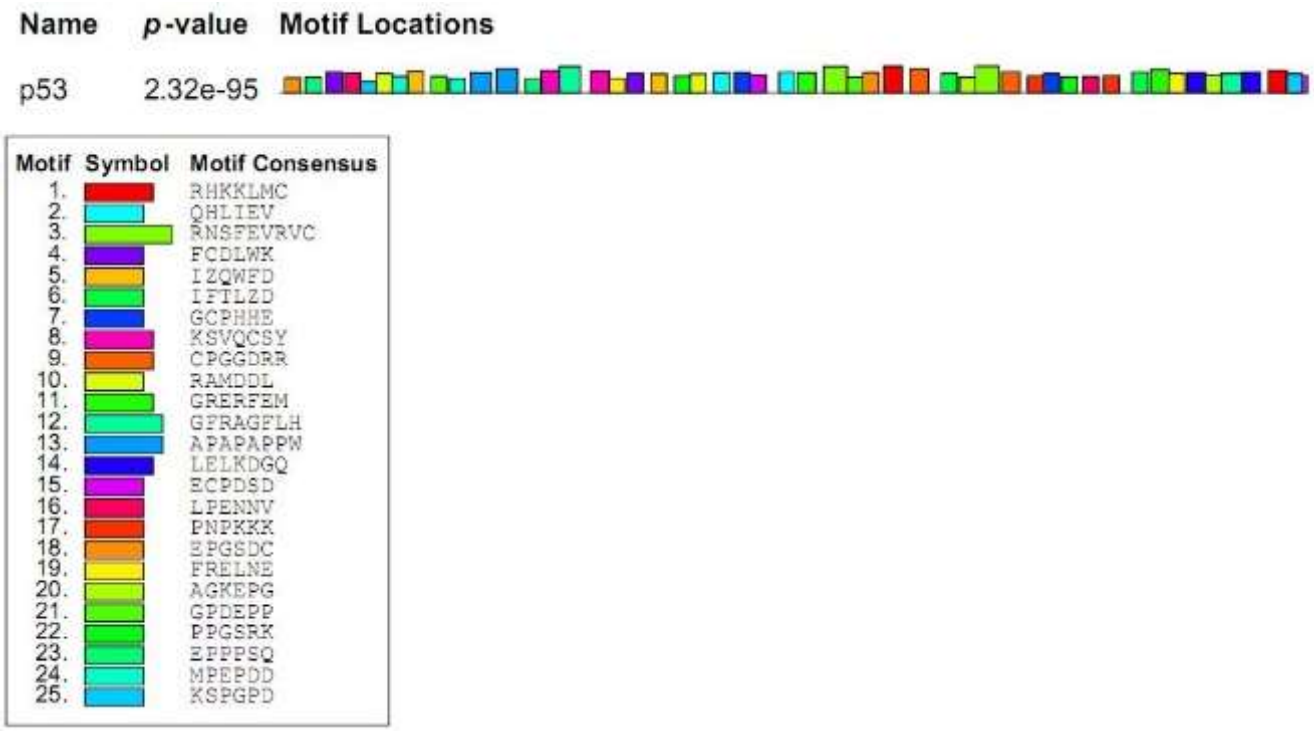

Figure 3. Identified 25 motifs with the sequences of Human P53; RHKKLMC, QHLIEV, RNSFEVRVC, FCDLWK, IZQWFD, IFTLZD, GCPHHE, KSVQCSY, CPGGDRR, RAMDDL, GRERFEM, GFRAGFLH, APAPAPPW, LELKDGQ, ECPDSD, LPENNV, PNPKKK, EPGSDC, FRELNE, AGKEPG, GPDEPP, PPGSRK, EPPPSQ, MPEPDD, and KSPGPD.

Table 2. In-silico identified and experimentally inferred phosphorylation sites, O-β-GlcNAc variations, and Yin Yang sites in human P53.

Residues		***Phosphorylation***		***Phos3D***	***Ying O Yang***	
Name	**Site**	**PI**	**CI**	**Kinases**	**O-β-GlcNAc**	**CIY**
Serine	6	-	Yes	CKII	Yes	-
=	9	(Wang and Eckhart, 1992)	Yes	-	-	-
=	15	(She et al., 2000)	Yes	DNA-PK	-	-
=	20	-	Yes	ATM	-	-
=	33	(Turenne and Price, 2001)	Yes	P38MAPK, GSK3α	-	-
=	37	(Wang and Eckhart, 1992)	Yes	DNAPK	-	-
=	46	-	Yes	ATM, CKII,GSK3	-	-
=	90	-	Yes	GSK3	Yes	-
=	94	-	-	-	Yes	-
=	95	-	-	-	Yes	-
=	96	-	Yes	PKC, CDC2	-	-
=	99	-	Yes	DNAPK, PKC	-	-
=	106	-	Yes	PKC	-	-
=	116	-	Yes	PKC	-	-
=	121	-	Yes	-	Yes	
=	127	-	Yes	P38MAPK	-	-
=	149	-	Yes	CDC2	-	-
=	166	-	Yes	ATM, CDC2, PKC	-	-
=	183	-	Yes	PKA, CDC2	-	-
=	215	-	Yes	PKA	-	-
=	240	-	Yes	CDC2	-	-
=	260	-	Yes	-	-	-
=	269	-	Yes	PKA	-	-
=	303	-	Yes	PKC	Yes	
=	313	-	-	-	Yes	-
=	314	-	Yes	PKC, CDC2	Yes	-
=	315	(Wang and Eckhart, 1992)	Yes	-	-	-
=	362	-	-	-	Yes	-
=	366	-	Yes	RSK	-	-
=	367	-	Yes	PKC	-	-
=	371	-	Yes	PKC, CDC2	-	-
=	376	-	Yes	PKC	-	-
=	378	-	Yes	PKC	-	-
=	392	-	Yes	CKII	Yes	
Threonine	18	(Wang and Eckhart, 1992)	Yes	CKI	-	-
=	55	-	Yes	CKII	-	-
=	81	-	Yes	CDK5	Yes	
=	102	-	-	-	Yes	-

=	118	-	Yes	PKC	-	-
=	123	-	-	-	Yes	-
=	125	-	Yes	PKC	-	-
=	150	-	Yes	P38MAPK	Yes	
=	155	-	Yes	-	Yes	
=	211	-	Yes	PKC	-	-
=	284	-	Yes	CKII	-	-
=	304	-	Yes	PKC	Yes	
=	312	-	-	-	Yes	
=	377	-	Yes	PKC	Yes	
Tyrosine	163	-	Yes	-	-	

PI: Previously identified; CIY: Computationally identified Yin Yang sites

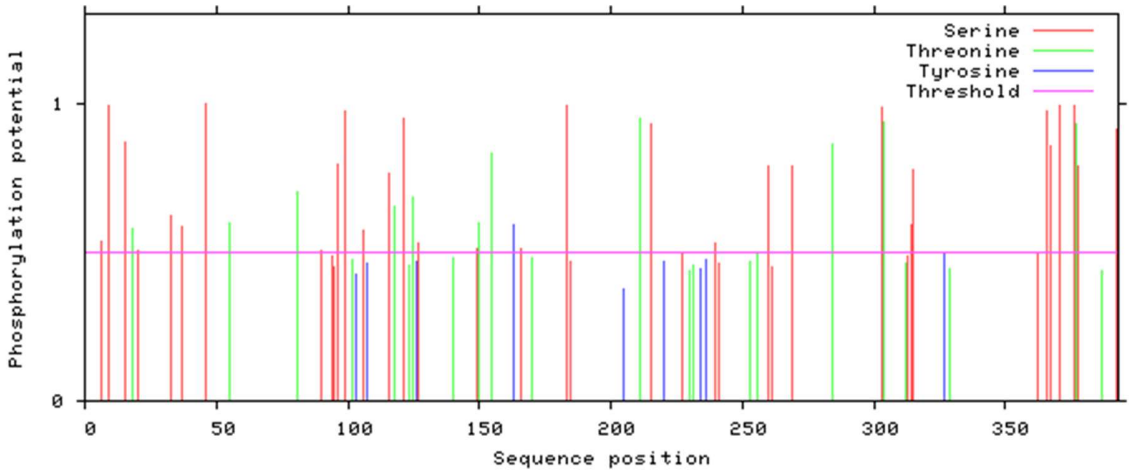

Figure 4. Representation of potential PTM sites (Serine, Threonine and Tyrosine) involved in human P53 phosphorylation. (i) The dark red vertical lines are representing Serine, (ii) light green vertical lines Threonine and (iii) blue vertical lines Tyrosine as the potential residues.

From comparative analysis, we came to know that we have identified a lot of novel phosphorylation sites in human P53 protein. However, in addition to novel phosphorylation site identification, our predicted results also included experimentally proved PTM sites, thus supporting our computational approaches.

Different kinases have been determined for P53 under multiple cellular environments against almost all phosphorylation sites (Figure 4 and Table 2). Kinase identification responsible for phosphorylation holds a major role in generating a complete functionality map of P53 as well as for the inhibitory kinases useful in LFS treatment. We used Scansite to determine the degree of environmental expose for each involved residue. Based on SA values, we identified various kinase-dependent phosphorylation sites (Table 2). Scansite and Phos3D results determined Ser33, Ser46 and Ser90 sites are phosphorylated due to GSK3 kinase. Our results elucidate that phosphorylation at site Ser33 is due to GSK3β, while GSK3α is not involved in the phosphorylation of this site. *In vivo* experiments proved that the transcriptional activity of P53 protein increased by GSK3β. Mostly DNA damage causes an increase in this protein transcriptional activity. However, in the response to DNA damage,

the kinase activity of GSK3β inhibits, concluding that this kinase is not involved in the damage of the DNA (Turenne and Price, 2001).

N- and O-glycosylation in human P53

Only two N-glycosylation sites were predicted by the NetNGlyc 1.0 server which uses neural network technology for the identification of N-C-S amino acid triplets, while NetOGlyc 4.0 server identified 16 O-glycosylation sites on different amino acid residues of human P53 as given in Table 3. Ser6, Ser7, Thr81, Ser90, Ser94, Thr304, Thr312, Ser313, Ser314, Ser315, Ser362, Ser366, Ser371, Ser376, Thr377 and Ser378 are identified as O-glycosylation sites of human P53 which play a role in O-glycan production. O-glycan reduces P53 degradation (Shental-Bechor and Levy, 2008) which increases its activity leading to cancer. On the other hand, Asn310 and Ans311 play a key role in the N-Glycosylation of human P53 protein which can influence its bioactivity and properties (Aebi, 2013, Strasser, 2014).

Discussion

The tumor suppressor protein p53 is activated due to numerous cellular responses, including genomic damage due to ionizing radiations and various genotoxic agents (Canman and Lim, 1998), the expression level of activated oncogenes (Lowe, 1999) or due to body primary cells progression to senescence (Lundberg et al., 2000). P53 protein activation by these numerous stimuli results in either apoptosis initiation or growth arrest, depending on the cellular needs (Canman and Lim, 1998, Lowe, 1999, Lundberg et al., 2000). In response to the stress P53 protein, activation results in PTMs, especially phosphorylation at serine residues (Fuchs et al., 1998).

The secondary structure of the human P53 protein (Figure 1) shows that it is more stable against mutations due to a larger number of beta sheets (81) (Abrusán and Marsh, 2016). Human P53 protein has 25 motifs and 4 domains. These identified structural motifs are involved in beta hairpins, omega loops and helix-loop-helix formation in P53 protein structure, as they provide an association between the secondary structure elements.

Here, we mainly focus on PTMs on different sites of human P53 protein, which can cause a change in the protein's functionality or cause any disorder such as cancer. In this study, we explored phosphorylation, glycosylation and yin yang sites.

We identified 42 significant phosphorylation sites. Out of these, 30 were Ser, 11 were Thr and 1 was Tyr residues (Figure 4). Phosphorylation at Ser15 is due to DNAPK kinase which causes (Table 2) reduction in binding affinity with its negative regulator (oncoprotein MDM2). Phosphorylation of P53 protein at the Ser15 site plays a crucial role in the up-regulation, stabilization and functional activation of this protein during cellular stress. Ser15 phosphorylation also plays a major role in P53 protein transactivation (Dumaz and Meek, 1999). Whereas, all the other phosphorylation sites are due to some other kinases (Manning et al., 2002). Therefore, P53 phosphorylation linked kinases are given in Table 2. Out of all identified kinases, we identified that Ser33, Ser46 and Ser90 residues are phosphorylated due to GSK3 type β kinase. This kinase plays a crucial role in P53 protein transcriptional activity. P53 is activated in response to various pathways such as activation of oncogene and genomic damages (Bønnelykke et al.). If P53 protein's activation required GSK3β by DNA damage, this kinase activity should be regulated through genomic damage. However, the activity of GSK3β kinase inhibits when cells are exposed to ionizing radiation (Turenne and Price, 2001).

 | ISSN (Online): 2716-649X

Table 3. Identified O- and N-glycosylation sites of human P53 protein

Sr. No.	***O-Glycosylation***			***N-Glycosylation***		
	Sites	**EI**	**Residues**	**Site**	**EI**	**Residues**
	6	-	Serine	310	(Seidah et al., 2003)	NNTS
	9	-	Serine	311	-	NNTS
	81	-	Threonine	-	-	-
	90	-	Serine	-	-	-
	94	-	Serine	-	-	-
	304	-	Threonine	-	-	-
	312	-	Threonine	-	-	-
	313	-	Serine	-	-	-
	314	-	Serine	-	-	-
	315	-	Serine	-	-	-
	362	-	Serine	-	-	-
	366	-	Serine	-	-	-
	371		Serine	-	-	-
	376		Serine	-	-	-
	377		Threonine	-	-	-
	378		Serine	-	-	-

O-β-GlcNAc and phosphorylation interplay in Yin-Yang sites of Human P53protein

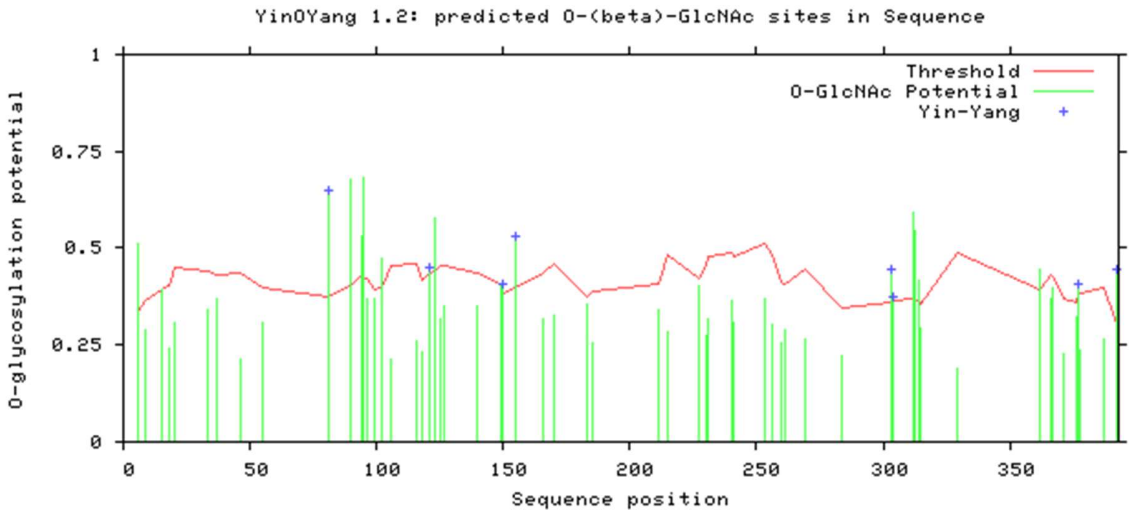

Figure 5. Pictorial representation of potential O-β-GlcNAc modification sites on Serine and Threonine residues in human P53 protein. The vertical green lines show O-β-GlcNAc modification sites against potential Thr/Ser residues, while the pinkish-red zig-zag horizontal line shows the threshold for potential modification. The blue plus signs (symbols) encircled at the top of 8 vertical lines are representing positive Yin Yang sites.

N-and O-glycosylation influence the activity, conformation and properties of any protein. So, in addition to Ser, Thr and Tyr residue phosphorylation, we also identified different N- and O-glycans in human P53. Asn310 and Asn311 play a major role in N-glycosylation. Ser6, Ser7, Thr81, Ser90, Ser94, Thr304, Thr312, Ser313, Ser314, Ser315, Ser362, Ser366, Ser371, Ser376, Thr377 and Ser378 are involved in O-glycosylation. O-glycan enhances and N-glycan reduces protein's activation, therefore, these identified O- and N-glycans can be used as therapeutics for cancer treatment specially LFS (Zandberg et al., 2011) (Table 3).

In addition to Ser, Thr and Tyr residue phosphorylation, we also identified *O*-β-glycosylation. O-β-GlcNAc and phosphorylation modifications are supposed to establish an interplay considered as Ying Yang sites (Butt et al., 2012). It has been argued that Yin-Yang sites can act as an oncogene that causes oncogenesis because of its ability to regulate cell growth control genes. While this is an attractive theory, there has been no evidence that Yin-Yang, when over-expressed, can acutely transform cells or develop tumors within animals. So, whether Yin-Yang is a "classic" oncogene is still uncertain. But Yin-Yang regulates a broad variety of different cell functions, which may provide insight into its role in oncogenesis. We identified 18 significant O-β-GlcNAc modifications. Out of these, 8 sites (Ser121, Ser303, Ser392, Thr81, Thr150, Thr155, Thr304 and Thr377) were predicted as positive Yin Yang sites (Figure 4 and Table 2). These identified 8 conserved Yin Yang sites, also showing phosphorylation, can regulate P53 protein's activity of cancer-causing. From this study, we came to know that phosphorylation as well as all the above-discussed PTMs are very necessary for P53 protein's activity, functionality and role as therapeutics.

Declaration

Acknowledgment

We thank our colleagues Sumaira Nishat and Dr. Hassan Tariq from the University of Agriculture Faisalabad, Pakistan, who guided with bioinformatics tools that greatly assisted the research. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest/Competing interests

The authors declared no conflict of interest.

References

Abrusán, G. & Marsh, J. A. 2016. Alpha helices are more robust to mutations than beta strands. *PLoS computational biology,* 12.

Armenteros, J. J. A., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., Von Heijne, G. & Nielsen, H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology,* 37**,** 420-423.

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. & Noble, W. S. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic acids research,* 37**,** W202-W208.

Bhurgri, Y. J. A. P. J. C. P. 2004. Karachi cancer registry data--implications for the national cancer control program of pakistan. 5**,** 77-82.

Birch, J. M., Hartley, A. L., Tricker, K. J., Prosser, J., Condie, A., Kelsey, A. M., Harris, M., Jones, P. H. M., Binchy, A. & Crowther, D. J. C. R. 1994. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. 54**,** 1298-1304.

Bønnelykke, K., Sleiman, P., Nielsen, K., Kreiner-Møller, E., Mercader, J. M., Belgrave, D., Den Dekker, H. T., Husby, A., Sevelsted, A., Faura-Tellez, G., Mortensen, L. J., Paternoster, L., Flaaten, R., Mølgaard, A., Smart, D. E., Thomsen, P. F., Rasmussen, M. A., Bonàs-Guarch, S., Holst, C., Nohr, E. A., Yadav, R., March, M. E., Blicher, T., Lackie, P. M., Jaddoe, V. W. V., Simpson, A., Holloway, J. W., Duijts, L., Custovic, A., Davies, D. E., Torrents, D., Gupta, R., Hollegaard, M. V., Hougaard, D. M., Hakonarson, H. & Bisgaard, H. 2014. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nature Genetics,* 46**,** 51-55.

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M. & Bairoch, A. 2007. Uniprotkb/swiss-prot. *Plant bioinformatics*. Springer.

Butt, A. M., Feng, D., Idrees, M., Tong, Y. & Lu, J. 2012. Computational identification and modeling of crosstalk between Phosphorylation, O-β-glycosylation and Methylation of FoxO3 and implications for cancer therapeutics. *International journal of molecular sciences,* 13**,** 2918-2938.

Canman, C. E. & Lim, D.-S. 1998. The role of ATM in DNA damage responses and cancer. *Oncogene,* 17**,** 3301-3308.

Dumaz, N. & Meek, D. W. 1999. Serine 15 phosphorylation stimulates p53 transactivation but does not directly influence interaction with HDM2. *The EMBO journal,* 18**,** 7002-7010.

Durek, P., Schudoma, C., Weckwerth, W., Selbig, J. & Walther, D. 2009. Detection and characterization of 3D-signature phosphorylation site motifs and their contribution towards improved phosphorylation site prediction in proteins. *BMC bioinformatics,* 10**,** 117.

Evans, D., Birch, J., Thorneycroft, M., Mcgown, G., Lalloo, F. & Varley, J. J. J. O. M. G. 2002. Low rate of TP53 germline mutations in breast cancer/sarcoma families not fulfilling classical criteria for Li-Fraumeni syndrome. 39**,** 941-944.

Fuchs, S. Y., Fried, V. A. & Ronai, Z. E. 1998. Stress-activated kinases regulate protein stability. *Oncogene,* 17**,** 1483-1490.

Gupta, R., Jung, E. & Brunak, S. 2004. Prediction of N-glycosylation sites in human proteins.

Ikai, A. 1980. Thermostability and aliphatic index of globular proteins. *The Journal of Biochemistry,* 88**,** 1895-1898.

Lowe, S. 1999. Activation of p53 by oncogenes. *Endocrine-related cancer,* 6**,** 45-48.

Lundberg, A. S., Hahn, W. C., Gupta, P. & Weinberg, R. A. 2000. Genes involved in senescence and immortalization. *Current opinion in cell biology,* 12**,** 705-709.

Malik, S. A., Ahmad, I., Khan, T. S., Shakoori, A. R. & Nasir-Ud-Din 2007. Runx1 transcription repression and stability: interplay between phosphorylation and O-GlcNAc modification. *PAKISTAN JOURNAL OF ZOOLOGY,* 39**,** 299.

Manning, G., Whyte, D., Martinez, R., Hunter, T. & Sudarsanam, S. 2002. cAMP-dependent protein kinase from Plasmodium falciparum: an update. *Trends in Biochemical Sciences,* 27**,** 514-520.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M. & Hurwitz, D. I. 2015. CDD: NCBI's conserved domain database. *Nucleic acids research,* 43**,** D222-D226.

Mcguffin, L. J., Bryson, K. & Jones, D. T. 2000. The PSIPRED protein structure prediction server. *Bioinformatics,* 16**,** 404-405.

Obenauer, J. C., Cantley, L. C. & Yaffe, M. B. 2003. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic acids research,* 31**,** 3635-3641.

Omasits, U., Ahrens, C. H., Müller, S. & Wollscheid, B. 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics,* 30**,** 884-886.

Protparam, E. 2017. ExPASy-ProtParam tool.

Rasheed, B. A., Mclendon, R. E., Herndon, J. E., Friedman, H. S., Friedman, A. H., Bigner, D. D. & Bigner, S. H. J. C. R. 1994. Alterations of the TP53 gene in human gliomas. 54**,** 1324-1330.

Rost, B., Yachdav, G. & Liu, J. 2004. The predictprotein server. *Nucleic acids research,* 32**,** W321-W326.

Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acids research,* 31**,** 3381-3385.

Seidah, N. G., Benjannet, S., Wickham, L., Marcinkiewicz, J., Jasmin, S. B., Stifani, S., Basak, A., Prat, A. & Chrétien, M. 2003. The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation. *Proceedings of the National Academy of Sciences,* 100**,** 928-933.

She, Q.-B., Chen, N. & Dong, Z. 2000. ERKs and p38 kinase phosphorylate p53 protein at serine 15 in response to UV radiation. *Journal of Biological Chemistry,* 275**,** 20444-20449.

Silva, J. M., Gonzalez, R., Dominguez, G., Garcia, J. M., España, P., Bonilla, F. J. G., Chromosomes & Cancer 1999. TP53 gene mutations in plasma DNA of cancer patients. 24**,** 160-161.

Simanon, N., Adisakwattana, P., Thiangtrongjit, T., Limpanont, Y., Chusongsang, P., Chusongsang, Y., Anuntakarun, S., Payungporn, S., Ampawong, S. & Reamtong, O. 2019. Phosphoproteomics analysis of male and female Schistosoma mekongi adult worms. *Scientific reports,* 9**,** 1-10.

Sohpal, V. K., Dey, A. & Singh, A. 2010. MEGA biocentric software for sequence and phylogenetic analysis: a review. *International journal of bioinformatics research and applications,* 6**,** 230-240.

Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T. B., Lavrsen, K., Dabelsteen, S., Pedersen, N. B. & Marcos-Silva, L. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *The EMBO journal,* 32**,** 1478-1488.

Turenne, G. A. & Price, B. D. 2001. Glycogen synthase kinase3 beta phosphorylates serine 33 of p53 and activates p53's transcriptional activity. *BMC cell biology,* 2**,** 12.

Wang, J., Torii, M., Liu, H., Hart, G. W. & Hu, Z.-Z. 2011. dbOGAP-an integrated bioinformatics resource for protein O-GlcNAcylation. *BMC bioinformatics,* 12**,** 91.

Wang, Y. & Eckhart, W. 1992. Phosphorylation sites in the amino-terminal region of mouse p53. *Proceedings of the National Academy of Sciences,* 89**,** 4231-4235.

Zandberg, W. F., Benjannet, S., Hamelin, J., Pinto, B. M. & Seidah, N. G. 2011. N-glycosylation controls trafficking, zymogen activation and substrate processing of proprotein convertases PC1/3 and subtilisin kexin isozyme-1. *Glycobiology,* 21**,** 1290-1300.

 | ISSN (Online): 2716-649X