UCP Journal of Science and Technology
Vol. 3, Issue 2 (January - June 2026)
Journal website: http://ojs.ucp.edu.pk/index.php/ucpjst/index

In Silico Insights into the Impact of Deleterious SNPs on the Functioning and Dynamics of TP53 and Its Targeted Therapy by COTI-2 and PRIMA1 in Breast Cancer Patients

Ambreen[1]*, Iltamaa Khurram[1], Tooba Khan[1], Heera Batool[1]

Abstract

Breast cancer is the most frequent malignancy in women globally and has a high death rate. In Pakistan, one in every nine women is at risk. Tumor protein 53 (TP53) is a transcription factor that suppresses tumors by regulating gene expression in response to stress. In 40-60% of breast tumors, mutations in TP53 contribute to tumor growth, medication resistance, and a poor prognosis. This study uses an in silico approach to find deleterious missense SNPs in TP53 and assess their effects on protein dynamics. Deleterious TP53 variants were screened through SNP prediction tools. Four of the 25 detected deleterious variants (Q16K, L130Q, P223R, and R273I) were chosen based on two criteria: pathogenicity threshold (>80%) and their presence in functionally significant domains. Structural stability studies using I-Mutant2.0 and MUPro revealed that L130Q is the most destabilizing variant, showing consistently lower stability across both tools. By contrast, R273I is predicted to be stabilizing in MUPro and Dynamut, whereas I-Mutant considers it destabilizing. Molecular docking analysis was performed using CB-dock to explore mutant TP53 interactions with small-molecule inhibitors COTI-2 and PRIMA. This study reveals the structural instability caused by *TP53* mutations and their impact on drug interactions. To restore TP53 activity and improve the effectiveness of TP53-targeted inhibitors in breast cancer treatment, future research should investigate targeted therapeutic approaches.

Keywords: Breast Cancer, Tumor Suppressor Gene, Missense Variant, TP53, Protein Dynamics, Molecular Docking.

1 Introduction

Breast cancer (BC) is currently regarded as a serious public health concern (Lundberg & Phoosuwan, 2022) since it is the most common type of cancer among women globally (Siegel, Miller, & Jemal, 2018). Breast cancer is the fifth leading cause of death from cancer. Around 7.8 million women were diagnosed with breast cancer in the last five years worldwide. Moreover, breast cancer is predicted to become more common in Pakistan, where one in nine women is at risk of getting the disease, which is said to be the highest rate in Asia. (Sohail & Alam, 2007)Hereditary factors and genetic abnormalities that may result in abnormal cell proliferation,

[1]Informatix Biolabs, Lahore, Pakistan
*Corresponding author's E-mail: ambreenhussain2002@gmail.com
Received: 26 April 2025; Received in revised form: 16 June 2026; Accepted: 18 June 2026.
Available online: 13 July 2026.

DOI: https://doi.org/10.24312/ucp-jst.03.02.523

suppression of apoptosis, and an increase in angiogenesis and metastasis have been the subject of much recent discussion (Lukong, 2017). According to an estimation, 20% of individuals have a family history of breast cancer, while 5–10% have hereditary BCs, which are brought on by germline mutations in tumor suppressor genes such as TP53, PTEN, BRCA1, and BRCA2 (Economopoulou, Dimitriadis, & Psyrri, 2015). Although the primary treatments for breast cancer are radiotherapy, chemotherapy, and surgery, patients with advanced-stage tumors have a poor prognosis with serious adverse consequences. Therefore, alternative screening methods are required in developing and underdeveloped countries (Nounou et al., 2015).

For many years, the main goal of genetics has been to identify genetic variations that might contribute to the onset and course of complex diseases like cancer. It could be useful to identify risk variables for both prognosis and disease prevention. However, this is costly and time-consuming (Singh et al., 2021). Fortunately, bioinformatics advancements have made it easier to screen for single-nucleotide polymorphisms (SNPs) in the genome that may be linked to diseases (Khan et al., 2022). These tools can assess SNPs from a variety of categories, including frameshift, nonsense, untranslated regions (UTRs), promoter regions, splice sites, and missense/non-synonymous. Of these mutations, non-synonymous SNPs (nsSNPs) and nonsense SNPs are considered coding variants because they directly affect a protein's folding and function. (Mooney, 2005) In silico tools for analysing the harmful effects of SNPs often use information gathered from sequences, protein structural properties, mutant and wild-type residue properties, and the evolutionary conservation of nucleotides or amino acids. (Selga, Koc, Chawade, & Ortiz, 2020) The use of functional bioinformatics has aided researchers in identifying possible disease-causing variations that may be confirmed by genotyping or sequencing-based experiments (Emadi, Akhoundi, Kalantar, & Emadi-Baygi, 2020).

Previous studies have shown that P53 codes for a transcription factor known as tumor protein 53, which contributes to the development of BC (Ahmad et al., 2023; Li et al., 2020). It is a tumor suppressor that regulates the cell-cycle checkpoint, senescence, DNA repair, apoptosis, and other processes. About 40–60% of BC cases have somatic TP53 mutations, which are prevalent in many cancer types (Pollock et al., 2022). Their TP53 status greatly influences cancer patients' dependence on a wide range of anticancer therapies that are highly associated with a poor prognosis in breast cancer (O'Connor et al., 1997) because they can potentially cause medication resistance, relapse, and disease failure. About 80% of triple-negative breast cancers have TP53 mutations because they frequently accompany clinically aggressive tumors. (Cheng, Ding, Xu, Zhu, & Jia, 2020). In addition to having a high risk of metastasis and recurrence in clinical settings, triple-negative breast cancer has significant resistance to chemotherapy and radiation therapies (Bai, Ni, Beretov, Graham, & Li, 2021). TP53 is activated through various receptors and activates the main downstream signalling pathways, which can ultimately lead to the development of cancer (Wang, Guo, Wei, & Chen, 2023). Inhibiting TP53 mutations and restoring mutant p53 function are two possible therapy approaches for p53 mutations (Liu et al., 2019). Based on the distinct functional limitations and varied structure of mutant p53, several techniques are being explored to restore p53 function (Bykov, Eriksson, Bianchi, & Wiman, 2018). Drugs like PRIMA-1 and COTI-2

have promising therapeutic strategies for cancer treatment because they reactivate P53 mutants by restoring its tumor-suppressing power and inducing cancer cell death (Lu et al., 2016). Therefore, this study aims to filter the most deleterious missense SNPS of TP53, observe the impact of variations in the DNA-binding domain and transactivation domain on protein dynamics and stability, and unravel the impact of targeted inhibition by COTI-2 and PRIMA1.

The novelty of the study stems from the fact that, while many studies have investigated TP53 mutations in cancer, the majority have either focused on the structural consequences of a small number of alterations or on the prevalence and clinical significance of frequently occurring mutations. The present study, on the other hand, employs a comprehensive in silico approach that includes pathogenicity identification, structural and functional analysis, protein stability and flexibility evaluations, gene ontology annotation, and therapeutic assessment of mutant TP53. Furthermore, this analysis goes beyond the well-known hotspot mutations and investigates four well-established and less-studied detrimental TP53 variants (Q16K, L130Q, P223R, and R273I) with the goal of expanding the knowledge about these mutations. Another unique aspect of this investigation is the comparison of two mutant p53-targeting drugs, COTI-2 and PRIMA-1, against both wild-type and mutant TP53 proteins. This study offers a more comprehensive knowledge of how particular TP53 variations may affect protein behavior and therapeutic response in breast cancer by integrating mutation prioritization with structural and pharmacological investigations.

2 Methodology

2.1 Sequence Retrieval

The nucleotide and amino acid sequences of the human cellular tumor antigen p53 (TP53) with accession number P04637 were downloaded in FASTA format from the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/).

2.2 Retrieval of Mutational Hotspots Data

SNP information for the human TP53 gene, having accession ID ENSG00000141510, genome assembly GRCh38:CM000679.2, and transcript ID ENST00000269305.9, was retrieved on February 21, 2025, using the Ensemble Genome Browser version 113 (https://asia.ensembl.org/index.html). Different types of variations were observed throughout the length of the TP53 protein**.** In this study, nonsynonymous variants, specifically missense variants of TP53, were retrieved from the database.

2.3 Selection of Deleterious Missense Variants

The pathogenicity of missense SNPs was predicted using 5 dbSNP consensus tools such as SIFT (Sim et al., 2012), PolyPhen (Adzhubei, Jordan, & Sunyaev, 2013), CADD (Rentzsch, Witten, Cooper, Shendure, & Kircher, 2019), MetaLR, and MutationAssessor (Reva, Antipin, & Sander, 2011). Data from these tools are also integrated into the Ensembl genome browser. Chromosomal coordinates, altered base and amino acid information of SNPs, and the transcript ID are among the data that the dbSNP tool offers. Based on the criteria outlined in Table 1, the pathogenicity scores of the above-mentioned SNP tools were subsequently acquired to filter the most deleterious missense SNPs of the TP53 gene (Supplementary File**).** The highly anticipated pathogenicity of the SNPs led

Table 1. Tools used to identify deleterious SNPs of the TP53 gene.

SNP Tool	Score	Interpretation	Reference
SIFT	0 or 1	Tolerated or Deleterious	(Sim et al., 2012)
PolyPhen	0 or 0.5 or 1	Benign or Possibly Damaging or Probably Damaging	(Adzhubei et al., 2013)
CADD	0 or 1	Likely Benign or Likely Deleterious	(Rentzsch et al., 2019)
MetaLR	0 or 1	Tolerated or Damaging	-
Mutation Assessor	0 or 0.5 or 1	Low, Medium, or High	(Reva et al., 2011)

to their selection. Since mutations in different TP53 domains have varied functional effects, further filtration was also carried out based on the location of the 25 most harmful variations. L130Q had an 80% pathogenicity score, whereas Q16K, P223R, and R273I all had a 90% pathogenicity score. While L130Q, P223R, and R273I were discovered from TP53's DNA-binding core domain, which is crucial for TP53's tumor suppressor activity, Q16K was detected from the transactivation domain. These variations were picked because they may impact the structure, function, and DNA interaction of TP53, which could result in cancer.

All nsSNPs in the TP53 gene were collected from publicly available databases and subjected to a rigorous filtering process to identify high-confidence detrimental alterations. Entries with inadequate annotation and synonymous variations were removed during the initial screening stage. The remaining missense variations were ranked using a consensus of many in silico functional prediction methods, including SIFT, PolyPhen-2, and PROVEAN. Variants that were reliably predicted to be detrimental or harmful across most of these networks were retained for further examination. To further refine the criteria, evolutionary conservation of affected residues was assessed; highly conserved sites were assumed to be more likely to be functionally important. Protein stability changes following mutation were evaluated using the available prediction techniques, and variants anticipated to severely destabilize the TP53 protein were given priority. Furthermore, variants were cross-referenced with the ClinVar and dbSNP databases to include clinical significance comments, and known benign polymorphisms were eliminated from the dataset. Where allele frequency information was provided, frequent variants (minor allele frequency > 0.01) were removed to focus on rare, potentially disease-associated mutations. Our multi-step integrative approach ensured that only high-confidence detrimental TP53 SNPs were selected for further structural and functional analyses.

2.4 Physicochemical Properties Analysis

The physicochemical characteristics of the TP53 protein were calculated using the ExPaSy ProtParam tool (https://web.expasy.org/protparam/), which computed metrics such as molecular weight, atomic and amino acid composition, theoretical pI, extinction coefficient, instability index, estimated half-life, aliphatic index, and the grand average of hydrophobicity (GRAVY).

2.5 Evolutionary Conservation and Phylogenetic Analysis

The progressive alignment of amino acid sequences of TP53 gene isoforms (P53α, P53β, and P53γ) was carried out using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo). The analysis revealed information on conserved domains among the variants, demarcating structurally and functionally important areas. As different TP53 isoforms play differential roles in tumor suppression, knowing their sequence conservation aids in the evaluation of how mutations could affect TP53 differently in breast cancer. To gain a better understanding of evolutionary conservation, phylogenetic analysis of all the TP53 protein isoforms was carried out using MEGA11, employing 1000 bootstrap replicates and the Maximum Likelihood methodology. The analysis identified the evolutionary divergence of TP53 isoforms along with conserved regions that would be more functionally essential. Through the identification of mutations within these conserved domains, we can better grasp their possible role in TP53 dysfunction and in the development of breast cancer.

2.6 Domain Analysis

The domain analysis of TP53 was carried out using InterPro (https://www.ebi.ac.uk/interpro/search/text/), a program that groups proteins into families and forecasts domains and significant areas to offer functional analysis of the target protein. The transactivation domain (TAD), proline-rich region, DNA binding domain, tetramerization domain, and regulatory region were among the TP53 domains. A particular range of amino acid sequences defined these locations. The tertiary structure of individual domains was then obtained using these sequences in PYMOL (https://www.pymol.org/), which was subsequently utilized to depict the TP53 exploded view. As the identified variants are in disparate domains of TP53, the exploded view representation in PyMOL was utilized to give structural insights into the respective domains containing these SNPs. This view emphasizes the spatial arrangement of the mutations in the protein structure.

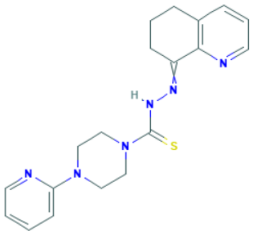
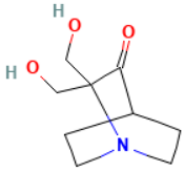
2.7 Primary and Secondary Structure Prediction

Ensembl genome browser (https://www.ensembl.org/) was used for the retrieval of the TP53 gene sequence having Ensembl ID: ENSG00000141510**.** The genomic sequence was translated into protein using the ExPaSy Translate tool (https://www.expasy.org/), and the longest open reading frame was located and used for further analysis. Secondary structure prediction of TP53 helps identify functional domains (e.g., DNA-binding regions) and structural vulnerabilities linked to mutations in cancer. PSIPRED (https://bioinf.cs.ucl.ac.uk/psipred/) was used for secondary structure prediction and functional analysis. The PSIPRED Workbench is a computational platform offering diverse tools for predicting protein structural and functional features. By analyzing amino acid sequences, it facilitates predictions such as secondary structure elements (e.g., alpha-helices, beta-strands), identification of disordered segments, and arrangement of transmembrane helices in the TP53 protein.

2.8 Tertiary Structure Prediction

The online protein modeling database AlphaFold (https://alphafold.ebi.ac.uk/), which is based on the ab initio technique, was used to predict the tertiary structure of TP53. PyMOL (https://www.pymol.org/), a molecular visualization tool, was used to visualize the 3-D structure of the 393 amino acid long TP53 protein with UniProt ID "P04637," which was retrieved in PDB format from AlphaFold.

Table 2. Physicochemical attributes of COTI-2 and PRIMA1 retrieved from the PubChem Database.

Compound Name	COTI-2	PRIMA1
Pubchem CID	74334923	322968
Molecular Weight (g/mol)	366.5	185.22
Molecular Formula	$C_{19}H_{22}N_6S$	$C_9H_{15}NO_3$
SMILES	C1CC2=C(C(=NNC(=S)N3CCN(CC3)C4=CC=CC=N4)C1)N=CC=C2	C1CN2CCC1C(=O)C2(CO)CO
Structure		

2.9 Validation and Refinement of Predicted Structure

UCLA-DOE LAB—SAVES v6.0 (https://saves.mbi.ucla.edu/) was used to validate the quality of the predicted TP53 structure. Protein structures were verified using the ERRAT server, where the error function is derived from the reported structure's non-bonded atom-to-atom interaction data. The "stereochemical quality" of the target protein structure was evaluated using PROCHECK's Ramachandran plot. Through the Galaxy web server (https://galaxy.seoklab.org/index.html), the TP53 structure was further refined.

2.10 Structural Homology Assessment

The structure of 1AIE (deep teal), which is the crystal structure of the tetramerization domain of TP53, was obtained in PDB format from the Protein Data Bank (https://www.rcsb.org/), and its superimposition with complete wild-type TP53 (blue) was carried out via PYMOL. The overlapping regions between the two sequences illustrated the structural alignment. This allowed for the validation of the structural homology of the AlphaFold-predicted TP53 structure (model).

2.11 Prediction of Subcellular Location

The subcellular location of the human TP53 protein was predicted using DeepLoc-1.0 (https://services.healthtech.dtu.dk/services/DeepLoc-2.0/), which predicts localization based on the likelihood score for each organelle. The likelihood that a protein will be found in a specific organelle increases with a higher value, and vice versa. A server called SecretomeP 1.0f (https://services.healthtech.dtu.dk/services/SecretomeP-1.0/), which generates ab initio predictions of non-classical protein secretion in eukaryotes, further validated the location and nature of the TP53 protein.

2.12 In-situ Mutagenesis

In situ mutagenesis was induced in wild-type TP53 at amino acid coordinates Q16K (rs2151047550), L130Q (rs2151033511), P223R (rs138983188), and R273I (rs121912660) via PyMOL (https://www.pymol.org/), and the structures were saved in PDB format.

2.13 Stability and Flexibility Analysis

Online tools such as MUPro (https://mupro.proteomics.ics.uci.edu/) and I-Mutant (https://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) were used to study and illustrate the effect of the

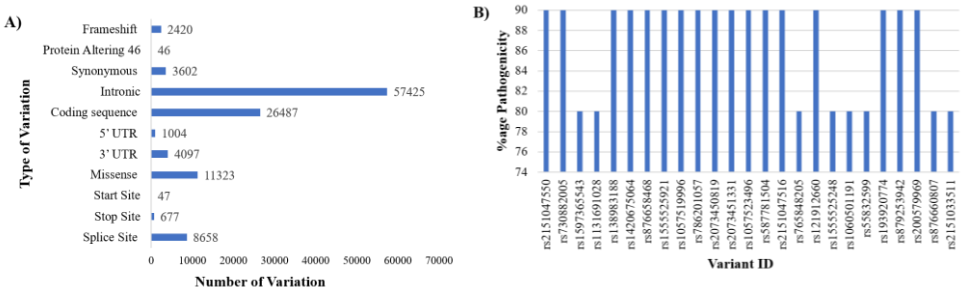


Figure 1. The distribution and pathogenicity of TP53 gene-associated SNPs retrieved from the ENSEMBLE genome browser. **(A)** The bar graph classifies many types of variations found in the TP53 gene, such as intronic, coding sequence, missense, synonymous, frameshift, and splice site variations, and their corresponding frequencies. **(B)** The histogram depicts the estimated pathogenicity of chosen deleterious variants in TP53, with the 80% and above the pathogenicity threshold, as assessed by SNP tools.

mutations on the stability of the protein. In the I-Mutant server, DDG values less than 0 indicate destabilization, while values greater than 0 indicate higher stability. In MUPro, the protein stability indicator was the Delta G value, where a decrease in the Delta G value indicates increased protein stability and vice versa. Whereas the DynaMut server (https://biosig.lab.uq.edu.au/dynamut/) was used to observe the effect of mutations on protein flexibility and stability. The results were represented in the form of vibrational entropy change (DDG) and free energy. This analysis is important because it defines the pattern of interaction of mutant proteins with others.

2.14 Gene Ontology Analysis

GO annotations of the TP53 protein are typically derived from experimental evidence, literature, and computational methods and taken from PSIPRED (https://bioinf.cs.ucl.ac.uk/psipred/). The Gene Ontology encompasses three primary domains: molecular functions, biological processes, and cellular components.

2.15 Ligand Structures Retrieval

The structural and functional characteristics of compounds, namely COTI-2 and PRIMA1, with IDs 74334923 and 322968, respectively, were obtained using the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) on February 26, 2025. To carry out molecular docking, ligands' three-dimensional structures were retrieved in SDF format. Table 2 shows the general attributes of the compounds selected for docking with wild-type and mutant TP53 proteins.

2.16 Pharmacokinetic Analysis

The pharmacokinetics of the chosen inhibitors were investigated using SwissAdme (http://www.swissadme.ch/). The filtration parameters were determined to be: Molecular mass (Mw) < 500 Daltons; hydrogen bond donor (HBD) and acceptor (HBA) parameters < 5; number of rotatable bonds < 10; and log P (octanol-water partition coefficient) parameter < 5. (Supplementary File).

2.17 Molecular Docking

The wild-type TP53 and its variations (Q16K, L130Q, P223R, and R273I) were molecularly docked with inhibitors such as COTI-2 and PRIMA1 using CBDock 2 (https://cadd.labshare.cn/cb-dock2/). The 3D complexes were then visualized in PyMol, and the 2D interactions were visualized in Ligplot. The docking tool produced five different poses for every protein-ligand combination. The best pose was picked based on the Vina score, and the complex with the lowest scoring

Table 3. TP53 deleterious variants with 80% and above percentage pathogenicity threshold.

Sr. No.	Variant ID	Percentage Pathogenicity	Residue	Residue Coordinates
1	**rs2151047550**	**90**	**Q/K**	**16**
2	rs730882005	90	C/F	238
3	rs1597365543	80	G/S	226
4	rs1131691028	80	E/V	224
5	**rs138983188**	**90**	**P/R**	**223**
6	rs1420675064	90	P/R	219
7	rs876658468	90	H/D	193
8	rs1555525921	90	A/S	189
9	rs1057519996	90	K/R	132
10	rs786201057	90	T/M	125
11	rs2073450819	90	K/R	120
12	rs2073451331	90	A/G	119
13	rs1057523496	90	F/I	109
14	rs587781504	90	G/V	105
15	rs2151047516	90	E/V	17
16	rs765848205	80	M/R	230
17	**rs121912660**	**90**	**R/I**	**273**
18	rs1555525248	80	G/R	272
19	rs1060501191	80	E/Q	264
20	rs55832599	80	R/W	260
21	rs193920774	90	G/V	259
22	rs879253942	90	L/Q	258
23	rs200579969	90	G/S	255
24	rs876660807	80	N/Y	232
25	**rs2151033511**	**80**	**L/Q**	**130**

function was selected for further analysis. The wild-type and mutant proteins were docked with PRIMA-1 using the following grid box parameters: (8, -9, and 5) along the x, y, and z axes, respectively. The grid box coordinates for the x, y, and z axes in COTI-2 were (-8, 16, 26), respectively. The interaction between residues was further validated through Ligplot.

3. Results

3.1 Variant Selection and Filtration of Deleterious Missense SNPs

115786 variants were identified in TP53 SNP data, including coding regions (e.g., start and end codons, synonymous and nonsynonymous variations) and regulatory regions (e.g., splice sites, intronic regions, and UTR regions) along the length of the TP53 protein (Figure 1A). Specifically, the impact of missense

Table 4. Analysis of physicochemical properties of TP53 Protein.

Properties	WILD TYPE	Q16K	P223R	R273I	L130Q
Amino Acids	393	393	393	393	393
Molecular Weight (Da)	43653.18	43653.22	43712.25	43610.15	43668.15
pI	6.33	6.47	6.47	6.20	6.33
GRAVY	-0.756 (slightly hydrophilic)	-0.757	-0.764	-0.733	-0.775
Instability index	73.59 (unstable)	72.61 (unstable)	72.47 (unstable)	73.38 (unstable)	73.59 (unstable)
Aliphatic index	59.08	59.08	59.08	60.08	58.09
Estimated coefficient value	36,035	36035	36035	36035	36035

variations (11323 in total) on the structure and function of the TP53 protein was observed in this study (Supplementary file). Five different consensus tools, namely SIFT, PolyPhen, CADD, MutationAssessor, and MetaLR, were utilized to filter out deleterious variants and estimate % pathogenicity. 25 missense variations were estimated to be deleterious by 80% of the aforementioned tools, based on the cumulative study of the SNP effect (Figure 1B**).** The tools evaluate the pathogenicity of variants by assigning scores. SIFT assigns scores from 0 (tolerated) to 1 (deleterious). PolyPhen

A)

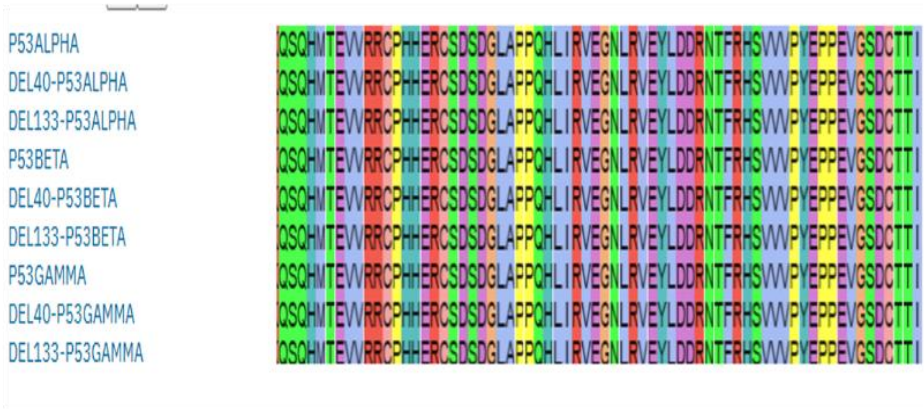


Figure 2. Evolutionary conservation and phylogenetic analysis of isoforms of TP53. **(A)** Multiple sequence alignment among highly related sequences of TP53 isoforms. The image employs the Clustal2 color scheme, with red representing positively charged residues, green polar, magenta negatively charged, yellow polar, cyan nonpolar, and blue hydrophobic residues**.**

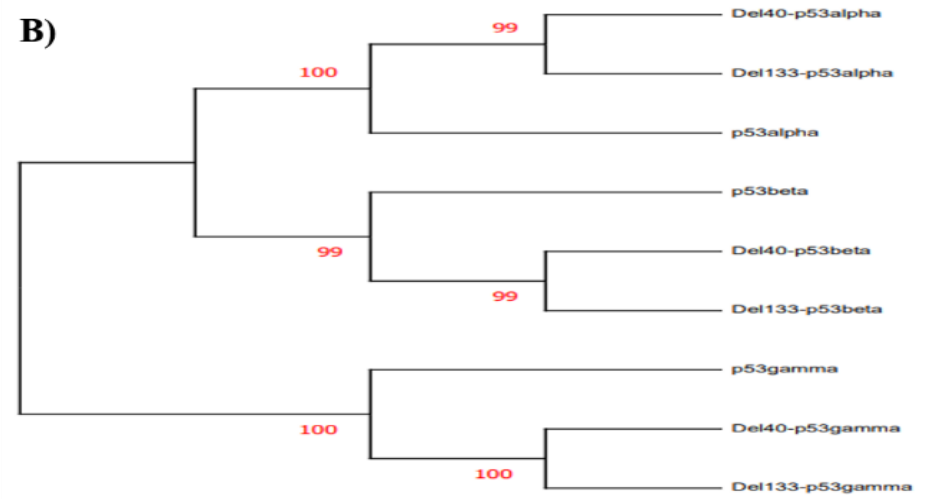


Figure 2. Evolutionary conservation and phylogenetic analysis of isoforms of TP53. **(B)** Phylogram of TP53 isoforms depicting evolutionary relationship.

classifies them as 0 (benign), 0.5 (possibly damaged), and 1 (probably damaged); CADD categorizes them as 0 (likely benign) or 1 (likely deleterious). Furthermore, MetaLR differentiated between tolerated (0) or damaging (1), and Mutation Accessor graded variants into low (0), medium (0.5), and high (1). Out of 25 filtered variants, 4 variants, namely Q16K, L130Q, P223R, and R273I, were randomly picked from the TAD and DNA-binding domain of TP53 **(Table 3).** The distribution of deleterious variants was observed to be significantly higher in the DNA-binding domain of TP53 than in the TAD or any other domain.

3.2 Physicochemical Properties Analysis

The structural and functional differences between human TP53 and its variants were assessed by analyzing their physicochemical properties. There were minor differences in the molecular weight, with R273I having the lowest (43610.15 Da) and P223R having the highest (43712.25 Da) molecular weight. The pI increased in Q16K and P223R (6.47) but decreased in R273I (6.20), suggesting potential charge variations. The change in GRAVY value is also seen, from slightly hydrophilic (-0.756) to most hydrophilic (-0.775). Across all the variants, the instability index remained unstable with minimal stability gains in Q16K and P223R. The aliphatic index changed slightly, with the most thermostable being R273I (60.08). These results indicate that TP53 stability, solubility, and expression level may be indirectly affected by mutations. (Table 4).

3.3 Evolutionary Conservation and Phylogenetic Analysis

The alignment represents amino acids of the isoforms of TP53 with differentiated colors to show conservation and functional similarity. Specific deletions were indicated in Del40-P53α and Del133-P53β variants. Most of the residues were found to be conserved in all isoforms of TP53. Figure 2A displays the sections of amino acids that are well conserved and also the level of conservation status. On the other hand, the color scheme depicts the characteristics of the amino acid residues. The Maximum Likelihood phylogenetic tree was constructed with the help of

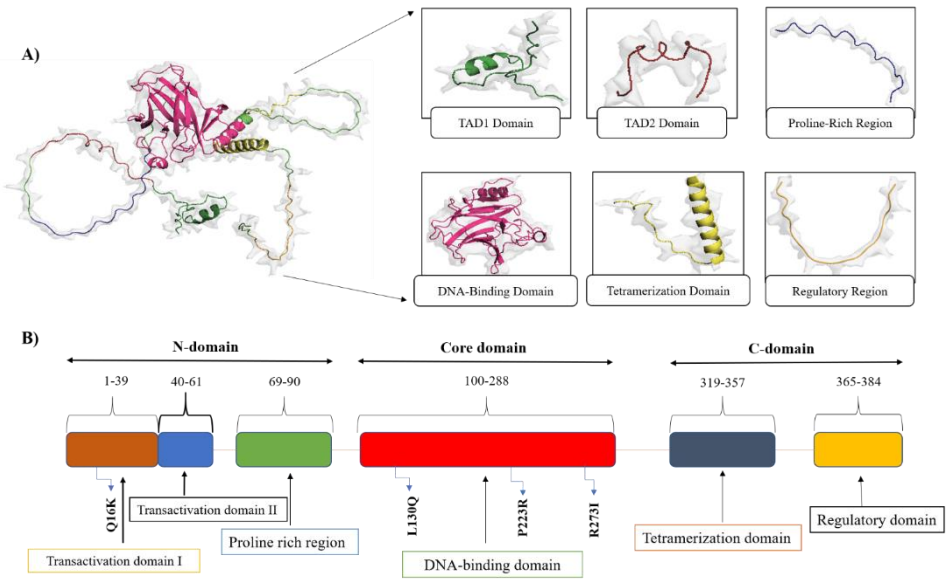


Figure 3. Domain Analysis of TP53. **(A)** Exploded view of all the domains of the predicted 3D structure of TP53. (Green color: TAD1 domain (1-39aa), red: TAD2 domain (40-61aa), blue color: Proline-rich region (69-90aa), pink color: DNA binding domain (100-288aa), yellow color: Tetramerization domain (319-357aa), orange color: Regulatory region (365-384aa)). **(B)** Sequential display of all the domains of the TP53 protein.

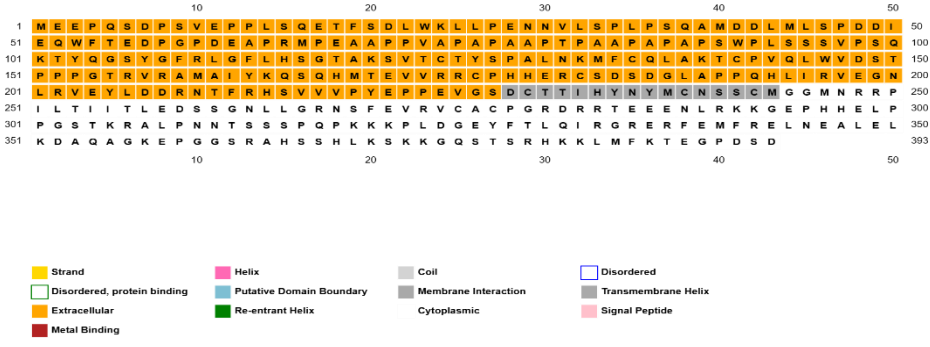


Figure 4. (A) The N-terminal region (residues 1-200) is rich in β-strands (yellow), while the C-**terminal** (251-393) appears largely unstructured. Additional features include metal-binding sites, helices, and disordered regions, indicating possible functional roles in protein stability and interactions.

MEGA11 software to depict the evolutionary relationship of the nine isoforms of P53 (P53α, P53β, P53γ, and their deletion variants). The P53 alpha, beta, and gamma form distinct clades with their deleted variant, i.e., Del40 and Del133. The bootstrap value was represented by numbers (99, 100), which depict the statistical measurement of confidence of each group. A value of 100 indicates very high confidence in the grouping, while 99 value numbers are considered highly reliable (Figure 2 B).

3.4 Domain Analysis

The transactivation domain, proline-rich area, DNA binding domain, tetramerization domain, and regulatory region are the five primary domains that

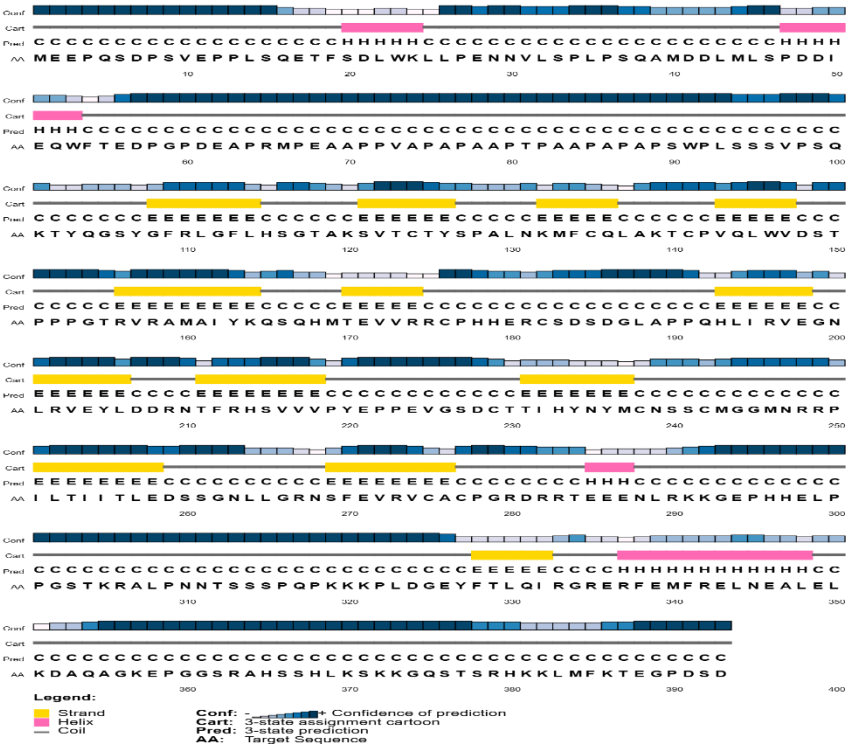


Figure 4. (B) The sequence depicts secondary structure prediction of the TP53 protein, highlighting helices (pink), β-strands (yellow), and coils (white). Confidence scores (blue bars) indicate prediction reliability.

have been identified in TP53. With a length of 1-61 residues, the transactivation domain (TAD) was separated into two domains: TAD1 (1-39 residues) and TAD2 (40-61 residues). The DNA-binding domain ranges from 100 to 288 residues, the tetramerization domain ranges from 319 to 357 residues, the regulatory area spans from 365 to 384 residues, and the proline-rich region extends from 69 to 90 residues (Figure 3B). Using PyMOL, each domain designated by a specific range of amino acids was highlighted in a different color (Figure 3A). The TAD domain of TP53 regulates gene transcription through interaction with coactivators such as p300; in contrast, the DNA-binding domain stabilizes the attachment of the target gene to TP53. Alterations in these domains affect TP53's tumor suppressor function, leading to cancer progression.

3.5 Primary and Secondary Structure Prediction

The genomic sequence of TP53 was successfully obtained from the Ensembl genome browser with the ID ENSG00000141510 and was precisely translated into a protein sequence through the ExPASy Translate tool. The longest open reading frame was selected for further structural and functional study to analyze the mutation correlated with breast cancer. The predicted secondary structure of TP53 demonstrated a composition of β-strands (yellow), α-helices (pink), and extensive coil regions (blue/gray), as determined by PSIPRED. These coil regions suggest structural flexibility or

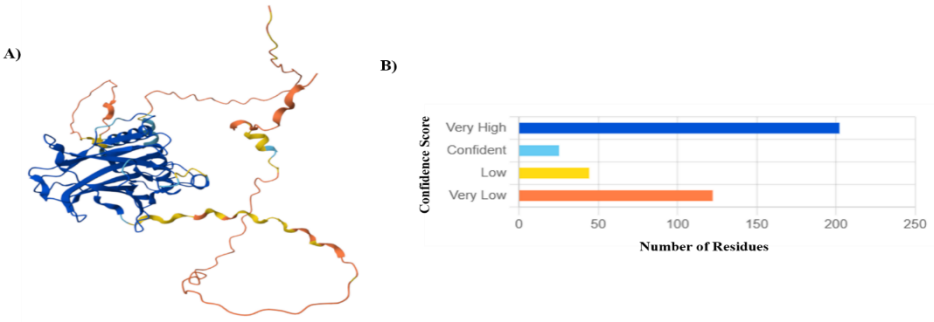


Figure 5. 3-D structural prediction of TP53. **(A)** AlphaFold predicted the tertiary structure of TP53. **(B)** Model confidence of each of the amino acid residues of TP53.

disorder, while the β-strands and α-helices likely stabilize functional domains critical for its activity. Sequence analysis further underscored extracellular motifs and potential membrane-binding sites, offering insights into TP53's structural architecture and biologically relevant regions. (Jones, 1999)

3.6 Tertiary Structure Prediction

The 3-D structure of the human TP53 protein was predicted and retrieved in PDB format from AlphaFold and was then visualized in PyMOL (Figure 5A). In addition, AlphaFold generates a per-residue confidence score (plDDT) ranging from 0 to 100 to determine the expected precision of a protein's overall 3-D structure. The more certain the tool is in the accuracy of the anticipated structure, the higher the score (Figure 5B). In terms of plDDT (global), the TP53 model confidence was 75.13.

AlphaFold showed a high confidence score for most of the amino acid residues, i.e., 227 out of 393, affirming the high accuracy of the predicted 3-D protein structure of TP53 (Table 5).

3.7 Validation and Refinement of Predicted Structure

The accuracy of the predicted protein structure is mostly dependent on the bond angles (phi and psi) of amino acids, which affect a model's quality estimate. According to the Ramachandran plot, phi and psi bond angles were present in the most favored region for 78.3% of amino acid residues (253), the additionally allowed region for 15.2% (49), the generously allowed region for 4.6% (15), and the disordered region for 1.9% (6) of amino acid residues. The predicted protein model of TP53 was verified to be accurate since the majority of amino acid residues (98.1%) are located in the quadrants that represent stable secondary structures (Figure 6A**).** Furthermore, by highlighting areas with high error values, the ERRAT plot evaluated the accuracy of the predicted TP53 structure. An ERRAT

Table 5. Confidence score per amino acid residue of TP53 based on piDDT score.

Color	Model Confidence	plDDT Score	Number of amino acids
	Very high	>90	202
	Confident	>70 but <90	25
	Low	>50 but <70	44
	Very Low	<50	122

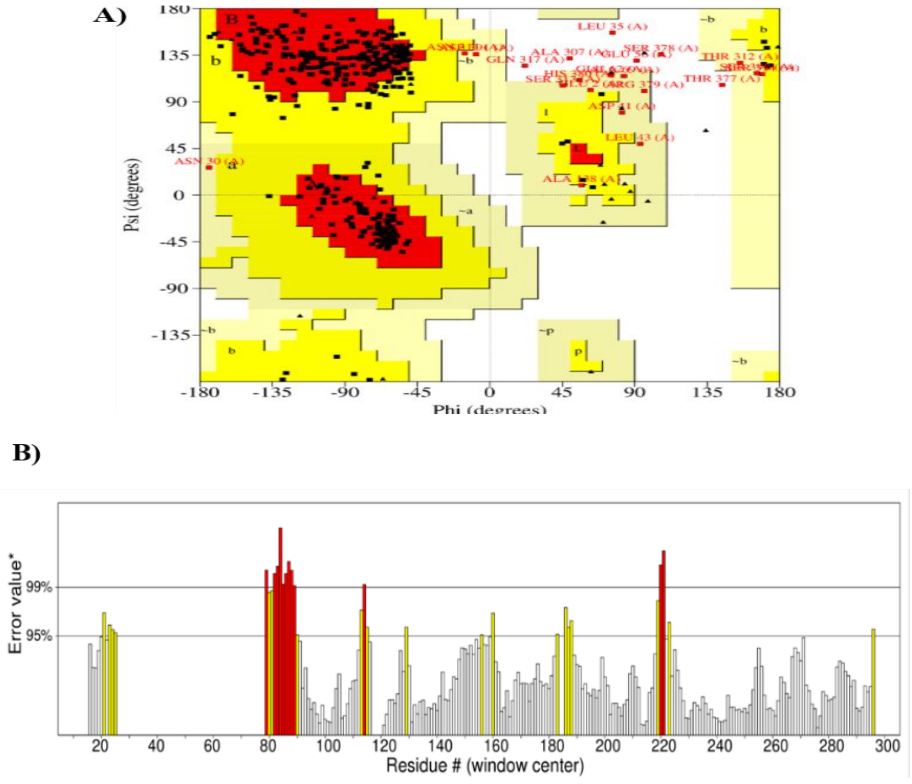


Figure 6. Validation of the predicted structure of TP53. **(A)** The Ramachandran plot is used to identify secondary structures in the TP53 protein. The Ramachandran plot's four quadrants are used to categorize amino acid residues according to the dihedral angles' (phi and psi) permitted and prohibited values. The amino acid residues with left-handed alpha-helical conformation were found in the first quadrant (Q1), those with beta-sheets were found in the second quadrant (Q2), and those with right-handed alpha-helices were found in the third quadrant. On the other hand, outliers in the disordered areas were found in the fourth quadrant. **B)** Quality estimate of TP53 by ERRAT. The x-axis refers to residue positions, and the y-axis shows the error value. The ERRAT server's Quality assessment of TP53 confirmed that the anticipated structure is correct.

score around 80% or higher suggested a well-predicted model, and TP53 attained an 86% score, indicating high accuracy. Reliable locations are typically indicated by error levels less than 95%, while problematic areas are suggested by values greater than 99%. Significant variations (red bars) are shown in some areas of this plot, especially those surrounding residues 70–90 and 220–240, which may indicate structural errors or poorly modeled areas. On the other hand, gray and yellow sections, respectively, denote places that are generally acceptable or fairly dependable. We may conclude that the model is dependable because a significant amount of protein is covered in grey and yellow regions, and the total quality factor is 86.40 (Figure 6B).

Better stereochemical precision was ensured by refining the predicted TP53 structure using the Galaxy web server. This analysis showed a large decrease in bad rotamers and an increase in Ramachandran-favored residues (up to 99%). The MolProbity and collision scores

Table 6. Refinement of the 3D structure of TP53 using Galaxy web server.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	1.793	1.0	2.6	83.6
MODEL 1	0.9109	0.608	1.307	5.7	0.3	98.7
MODEL 2	0.8957	0.650	1.337	6.1	0.6	99.0
MODEL 3	0.9065	0.632	1.275	5.2	0.0	98.7
MODEL 4	0.9008	0.623	1.240	4.7	0.3	98.5
MODEL 5	0.9052	0.639	1.416	7.6	0.9	98.5

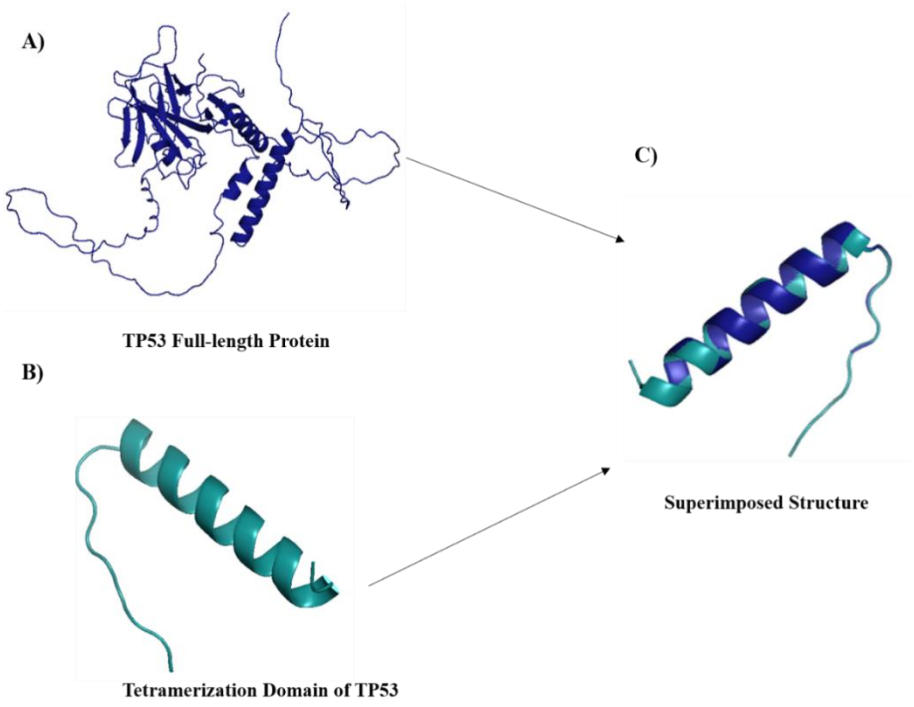


Figure 7. Illustration of structural homology of the predicted structure of TP53 (blue) by superimposing it with template 1AIE (deep teal). The common area depicts sequences shared by both proteins.

Increased slightly, suggesting small steric trade-offs, despite the RMSD values (0.608–0.650), suggesting considerable structural modifications. The improved models maintained their resemblance to the original structure, as seen by the GDT-HA values staying near 0.90. With its greatest Ramachandran favored percentage (99%) and balanced structural quality, MODEL 2 seems to be the best option overall (Table 6).

3.8 Structural Homology Assessment

To find out the structural similarities and differences between TP53 (the

Table 7. Prediction of subcellular localization of TP53 via DeepLoc.

Localization	Probability Score
Cytoplasm	0.3601
Nucleus	**0.909**
Extracellular	0.0452
Cell Membrane	0.1394
Mitochondrion	0.0802
Plastid	0.0038
Endoplasmic Reticulum	0.0539
Lysosome/Vacuole	0.0575
Golgi Apparatus	0.0401
Peroxisome	0.0037

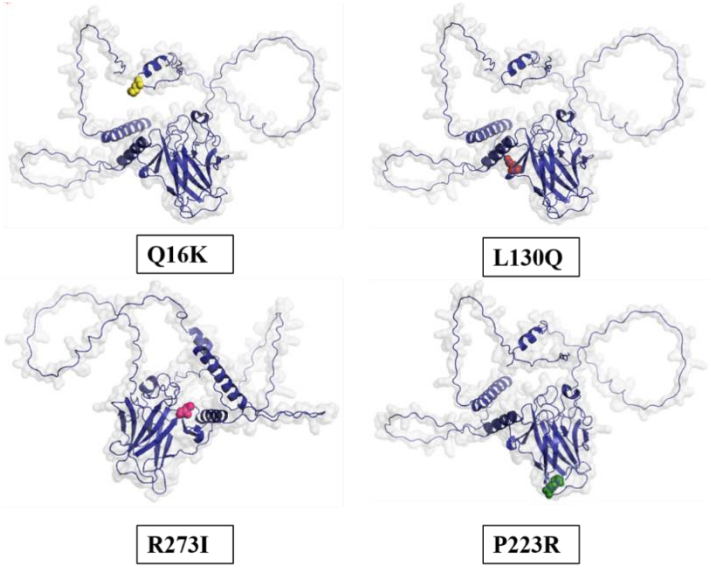


Figure 8. Mutagenesis induced in the selected variants of TP53.

predicted model) and 1AIE (the template), the 2 structures were visualized in PyMOL, and the structural homology between the 2 proteins was depicted by the superimposition of the model (blue) with the template (deep teal) (Figure 7). The superimposed areas depict the structural homology and thereby validate that the predicted structure is accurate.

3.9 Prediction of Subcellular Location

Various pathways and signals followed by all the possible subcellular compartments imply the probability or likelihood of the presence of the target protein in that compartment. The probability score for each possible location is shown in Table 7. The higher the score, the more likely a protein is present in that

Table 8. Nature of TP53 protein predicted by the Secretome server.

Name of Protein	NN-Score	Odds	Weighted
TP53	0.860	4.595	0.009

Table 9. Stability analysis of Q16K, L130Q, P223R, and R273I variants of TP53 using I-Mutant.

Position	Wild Type Residue	Mutant Residue	Stability	DDG	Reliability Index
16	Q	K	Decrease	-0.05	4
130	L	Q	Decrease	-1.45	9
223	P	R	Decrease	-1.51	8
273	R	I	Decrease	-0.22	8

Table 10. Stability analysis of Q16K, L130Q, P223R, and R273I variants of TP53 using MUPro

Position	Wild Type Residue	Mutant Residue	Stability	DDG
16	Q	K	Decreased	-0.58052438
130	L	Q	Decreased	-2.0673338
223	P	R	Decreased	-0.50987904
273	R	I	Increased	0.12806487

specific compartment. The highest score (0.909) observed in the case of the nucleus showed that TP53 is localized in the nucleus.

The nature of the protein (secretory or non-secretory) determined through the SecretomeP1.0 server is shown in Table 8. The first score is the neural network output score. For non-classically secreted proteins, the NN score should exceed the threshold of 0.6 (in the case of mammalian sequences) to indicate possible secretion. The 0.860 NN score for TP53 confirmed that it is a secretory protein.

3.10 In-situ Mutagenesis

The in situ mutagenesis was induced in wild-type TP53 at amino acid coordinates Q16K (rs2151047550), L130Q (rs2151033511), P223R (rs138983188), and R273I (rs121912660) via PYMOL (DeLano, 2002), and the structures were saved in PDB format (Figure 8).

3.11 Molecular Stability Analysis

To examine the effects of particular missense mutations (Q16K, L130Q, R273I, and P223R) on the structural stability of proteins, stability analysis was performed using I-Mutant2.0. All of the TP53 variations reduced stability as compared to the wild-type structure, according to I-Mutant v2.0 analysis. Variants Q16K, L130Q, R273I, and P223R were predicted to have free energy change (DDG) values of -0.05, -1.45, -1.51, and -0.22 Kcal/mol, respectively. The wild-type complex was destabilized following mutation, as indicated by the

Table 11. Stability and flexibility analysis of Q16K, L130Q, P223R, and R273I variants of TP53 using Dynamut.

Position	DDG (Kcal/mol)	Impact on Stability	DDS (Kcal/mol)	Impact on Flexibility
Q16K	0.020	Increased	-0.029	Decreased
L130Q	-1.513	Decreased	0.449	Increased
P223R	0.739	Increased	-1.070	Decreased
R273I	0.079	Increased	0.130	Increased

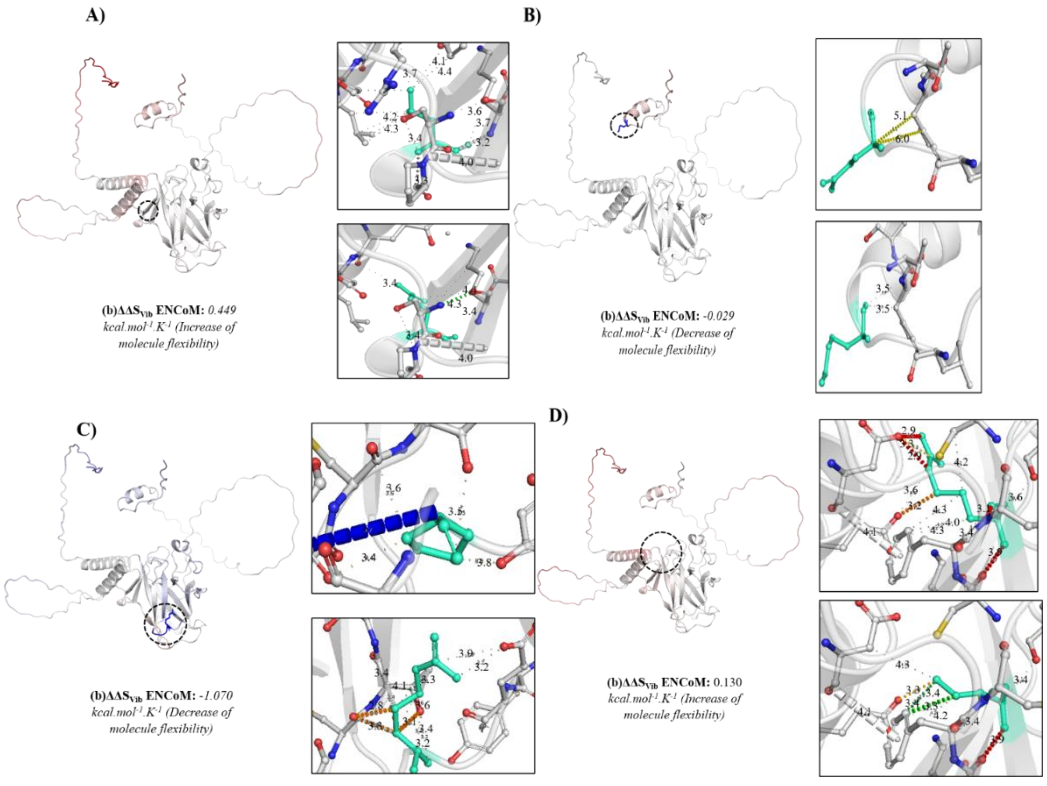


Figure 9. Flexibility analysis of (A) L130Q, (B) Q16K, (C) P223R, and (D) R2731 using Dynamut.

DDG value for each variant being less than 0 (Table 9).

However, MUPro analysis revealed that all of the TP53 variations reduced stability as compared to the wild-type structure, except R273I. Variants Q16K, L130Q, and R273I had DDG values of -0.58052438, -2.0673338, and 0.50987904Kcal/mol, respectively, which indicates that the wild-type complex was destabilized following mutation, as indicated by the DDG value being less than 0 (Table 10).

3.12 Molecular Flexibility Analysis

Additionally, Dynamut evaluated the impact of missense variants on the flexibility and stability of protein structure using vibrational entropy change (DDS) and DDG, respectively. Except for L130Q, every version had a stabilizing effect. Furthermore, protein flexibility was increased by L130Q and R273I and decreased by Q16K and P223 (Table 11 and Figure 9).

3.13 Gene Ontology Analysis

3.13.1 Biological Process Predictions

TP53 is associated with the regulation of metabolic processes, gene expression, RNA processing, and transcription (with statistical confidence exceeding 0.7). These associations underscore TP53's essential function in maintaining cellular equilibrium, directing transcriptional programs, and coordinating biosynthetic processes.

Table 12. Biological processes are summarized in the given table.

GO Term	Name	Prob	SVM Reliability
GO:0019222	regulation of metabolic process	0.943	H
GO:0010468	regulation of gene expression	0.908	H
GO:0034645	cellular macromolecule biosynthetic process	0.880	H
GO:0008380	RNA splicing	0.870	H
GO:2001141	regulation of the RNA biosynthetic process	0.862	H
GO:1903506	regulation of nucleic acid-templated transcription	0.834	H
GO:0051171	regulation of nitrogen compound metabolic process	0.822	H
GO:0009059	macromolecule biosynthetic process	0.769	H
GO:0051252	regulation of the RNA metabolic process	0.763	H
GO:0006355	regulation of transcription, DNA-templated	0.761	H
GO:0006397	mRNA processing	0.742	H
GO:0006351	transcription, DNA-templated	0.737	H
GO:0010629	negative regulation of gene expression	0.701	H
GO:0000375	RNA splicing, via transesterification reactions	0.670	H
GO:0006810	transport	0.647	H
GO:0000398	mRNA splicing, via the spliceosome	0.612	H
GO:0006396	RNA processing	0.587	H
GO:0006357	regulation of transcription from the RNA polymerase II promoter	0.584	H
GO:0009890	negative regulation of biosynthetic process	0.545	H
GO:0006366	transcription from the RNA polymerase II promoter	0.543	H
GO:0031328	positive regulation of cellular biosynthetic process	0.505	H
GO:0002376	immune system process	0.504	H

Table 13. TP53 molecular function predictions as a transcriptional regulator are summarized in the table.

GO Term	Name	Prob	SVM Reliability
GO:0003676	nucleic acid binding	0.967	H
GO:0000166	nucleotide binding	0.922	H
GO:0030554	adenyl nucleotide binding	0.906	H
GO:0001228	RNA polymerase II transcription regulatory region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription	0.889	H
GO:0003677	DNA binding	0.868	H
GO:0046914	transition metal ion binding	0.795	H
GO:0003723	RNA binding	0.727	H

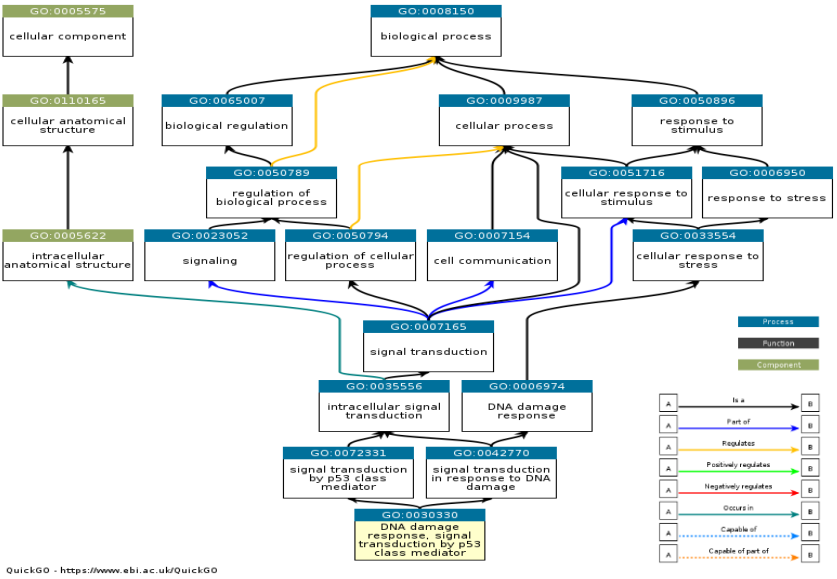


Figure 10. Blue boxes indicate Biological Processes of Tp53, Green boxes indicate Cellular components of Tp53

3.13.2 Molecular Function Predictions

TP53 exhibits associations with nucleic acid interaction, transcription factor activity, and RNA binding (probabilities surpassing 0.8), solidifying its identity as a transcriptional regulator. These capabilities enable TP53 to directly engage with DNA/RNA, influencing gene activation and safeguarding genomic integrity.

3.13.3 Cellular Component Predictions

TP53 localizes primarily within nuclear substructures, notably the nucleolus (0.819) and nuclear bodies (0.547), which are critical hubs for transcriptional control and stress-responsive signaling. This spatial distribution aligns with its role in mediating DNA damage responses and modulating transcriptional networks.

3.13.4 Ancestor Chart

TP53 governs diverse biological mechanisms, including cellular regulation and responses to environmental or internal stimuli. It is a central player in signal transduction pathways, particularly those involving intracellular communication and DNA damage recognition. TP53 predominantly also operates within subcellular compartments, particularly nuclear structures such as the nucleolus and nuclear bodies. The functions are summarized in the table. The hierarchical organization of Gene Ontology (GO) terms associated with TP53 is represented by the given chart. (https://www.ebi.ac.uk

Table 14. Cellular component predictions are given in the table.

GO Term	Name	Prob	SVM Reliability
GO:0005730	nucleolus	0.819	H
GO:0016604	nuclear body	0.547	H

(Cozzetto, Minneci, Currant, & Jones, 2016)

Table 15. Selection of the best inhibitor for protein based on physicochemical properties of compounds

Sr. No.	pdbqt	Compounds	Molecular Weight (g/mol)	HBD	HBA	logP	nRot	RO5 (Violations)
1	74334923	COTI-2	366.5	1	3	2.84	4	0
2	322968	PRIMA1	185.22	2	4	1.51	2	0

/QuickGO/GTerm?id=GO:0030330#term=history)

3.14 Virtual Screening of Best Inhibitors

The compounds' structures were obtained from PubChem and virtually screened according to their physicochemical characteristics. The COTI-2 and PRIMA1 compounds for targeted inhibition of TP53 were chosen; further information is available in **supplementary files.** The chosen chemicals underwent a pharmacokinetics analysis. **Table 15** lists the targeted proteins' chosen inhibitors. Both the

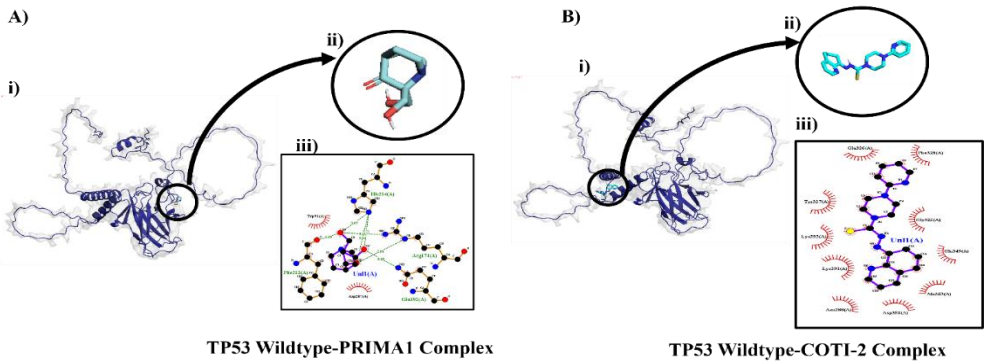


Figure 11. Molecular Docking Analysis. **(A)** Wild-type TP53-PRIMA1 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (PRIMA1) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein. **(B)** Wild-type TP53-COTI2 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (COTI2) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein.

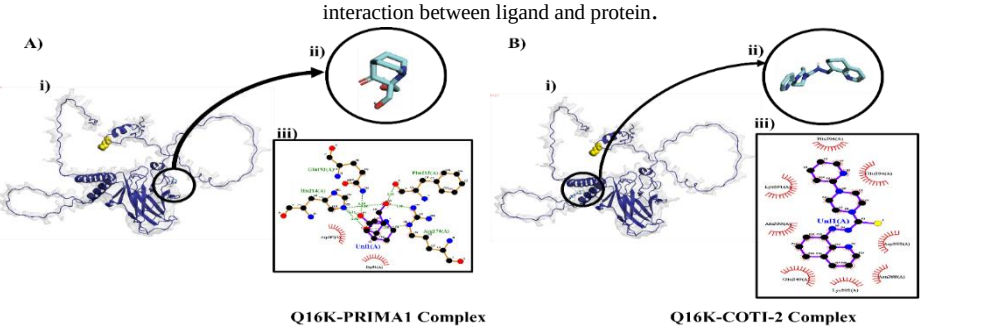


Figure 12. Molecular Docking Analysis. **(A)** Q16K variant-PRIMA1 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (PRIMA1) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein. **(B)** Q16K mutant TP53-COTI2 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (COTI2) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein

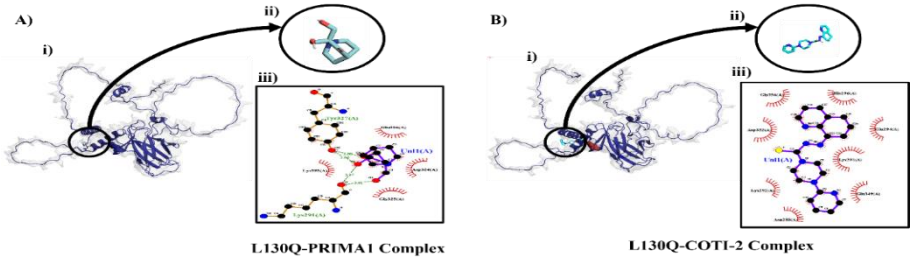


Figure 13. Molecular Docking Analysis. **(A)** L130Q variant-PRIMA1 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (PRIMA1) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein. **(B)** L130Q mutant TP53-COTI2 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (COTI2) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein

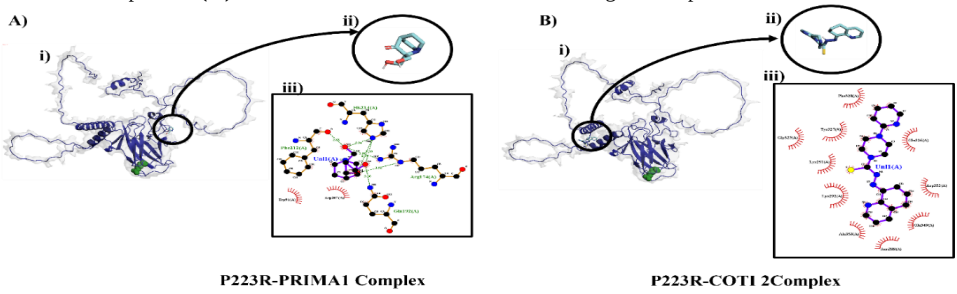


Figure 14. Molecular Docking Analysis. **(A)** P223R variant-PRIMA1 Docked Complex. **(i)** The 3D structure of molecular docking. (**ii**) The 3D structure of ligand (PRIMA1) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein. **(B)** P223R mutant TP53-COTI2 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (COTI2) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein

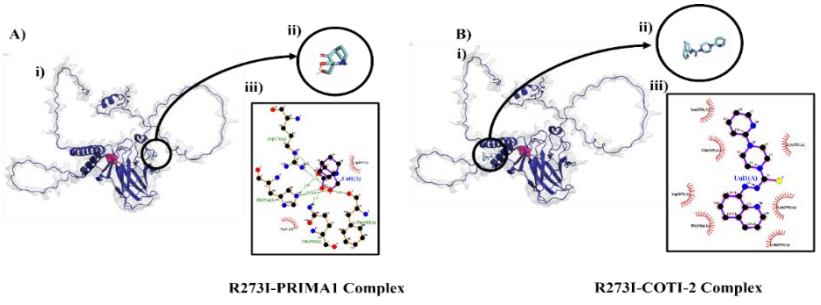


Figure 15. Molecular Docking Analysis. **(A)** R273I variant-PRIMA1 Docked Complex. **(i)** The 3D structure of molecular docking. (**ii**) The 3D structure of ligand (PRIMA1) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein. **(B)** R273I mutant TP53-COTI2 Docked Complex. **(i)** The 3D structure of molecular docking. **(ii)** The 3D structure of ligand (COTI2) docked with protein. **(iii)** The 2D residue-wise interaction between ligand and protein

compounds showed 0 Lipinski's rule of violation.

3.15 Molecular Docking

CB Dock2 was used to perform protein-ligand docking, and the 3D structure of the docked complex and the interactions between the docked molecules, residue by residue, were then shown using DiscoveryStudio. Following docking, the model with the lowest Vina

score value was selected as the best model (Table 16).

Figure 11-15 shows the 3D and 2D interaction of wild type TP53 and mutant TP53 (Q16K, L130Q, P223R, and R273I) with PRIMA1 and COTI2.

The type of interactions between the protein and ligand is summarized in Table 17.

4. Discussion

The most prevalent disease in women worldwide is breast cancer (BC) (Siegel et al., 2018), which is caused by risk factors like genetic abnormalities that trigger many cancer hallmarks (Lukong, 2017). 20% of patients have a family history of BC, according to statistical research, and 5–10% have hereditary BCs, which are brought on by germline mutations in tumor suppressor genes such as TP53 (Economopoulou et al., 2015). To find genetic variants that lead to cancer, computational biology technologies are being applied. These findings align with the studies of Andrew C.R.Martin, Angelo M.Facchiano, et al. about the Integration of mutation data and structural analysis of the TP53 tumor-suppressor protein(Martin et al., 2002). An important part of the prognosis of the disease may be the in-silico identification of SNPs and investigation of their effects on protein dynamics and stability. TP53 mutations are often seen in clinically aggressive cancers. One potential therapeutic strategy for p53 mutations is to inhibit TP53 mutations and restore mutant p53 function (Liu et al., 2019). A pathogenicity threshold of 80% or above was applied to the TP53 missense variations that were obtained from ENSEMBL to filter out the 25 most harmful SNPs that were predicted by the majority of SNP tools, including SIFT, PolyPhen, CADD, MetaLR, and MutationAssessor. Of these twenty-five variations, four variants Q16K, L130Q, P223R, and R273I were selected for additional examination. Numerous studies have examined the impact of mutations in tumor suppressor genes and proto-oncogenes to ascertain the overall impact on the structure, function, flexibility, and interaction patterns of these proteins in patients with breast cancer (Deng, Zhou, Fan, & Yuan, 2017; Lilyquist, Ruddy, Vachon, & Couch, 2018). The domain

Table 17. Types of Interactions between Ligand-Protein Complexes.

Protein	Ligand	Hydrogen Bonds	Hydrophobic Interactions
Wild Type TP53	PRIMA1	7	2
	COTI-2	0	10
Q16K	PRIMA1	7	2
	COTI2	0	8
L130Q	PRIMA1	4	4
	COTI2	0	8
P223R	PRIMA1	7	2
	COTI2	0	10
R273I	PRIMA1	7	2
	COTI2	0	7

analysis showed that L130Q, P223R, and R273I are located in the core DNA-binding domain of TP53, while the Q16K variation is located in the TAD domain. Our findings aligned with the results of Young et al.'s domain study of TP53. This study identifies TP53 mutations, particularly in DNA-binding domains (e.g., His158, Ser245), as independent predictors of poor overall survival in DLBCL patients, with Loop-L2 mutations showing no impact, and highlights their prognostic stratification in germinal center B-cell subtypes, validated alongside the International Prognostic Index (Young et al., 2008). The secondary structure prediction was done through PSIPRED to identify the structural expressions linked to cancer mutations and find functional domains. The DNA-binding domain of TP53 (residues 102–292) relies on a beta-sandwich scaffold and alpha-helical motifs to recognize specific DNA sequences. Any disruption in these structures can weaken TP53's ability to activate genes that suppress tumors. These alpha-helices and beta-sheets are essential for TP53's role as a transcription factor (GO:0003700), helping it regulate important genes like CDKN1A (p21) and BAX. Our findings align with research by Joerger and Fersht, who demonstrated that TP53's ability to regulate gene expression and to bind DNA could be disrupted due to the structural instability in the beta-sandwich domain. Also, cancer development often involves compromised p53 tumor suppressor function through genetic or regulatory disruptions, driving therapeutic innovations like structural correction agents for mutant p53 and MDM2/MDMX antagonists for cancers retaining functional p53, with these approaches now in clinical development and highlighting the pathway's evolutionary conservation while advancing personalized oncology(Joerger & Fersht, 2007). Following AlphaFold's 3D structure prediction of TP53, ERRAT and the Ramachandran plot were used to confirm it. The accuracy of the protein structure was confirmed by an overall quality score of 98.1 and 86.40 by the Ramachandran plot and errat. Our findings are consistent with earlier research that found that over 99 percent of the residues were found in the quadrants that correspond to stable secondary structures (Kasilingam & Elengoe, 2018).

To further refine the criteria, evolutionary conservation of affected residues was assessed; highly conserved sites were assumed to be more likely to be functionally important. Protein stability changes following mutation were evaluated using the available prediction techniques, and variants anticipated to severely destabilise the TP53 protein were given priority. Furthermore, variants were cross-referenced with the ClinVar and dbSNP databases to include clinical significance comments, and known benign polymorphisms were eliminated from the dataset. Where allele frequency information was provided, frequent variants (minor allele frequency > 0.01) were removed to focus on rare, potentially disease-associated mutations. Our multi-step integrative approach ensured that only high-confidence detrimental TP53 SNPs were selected for further structural and functional analyses. Since Q16K's unique location in the N-terminal transactivation domain, where somatic TP53 mutations are uncommon in patient samples, its functional influence is more likely to be caused by altered protein turnover or disrupted co-activator interactions rather than direct interference with DNA binding. By narrowing the gap between well-defined hotspot biology and less-defined mutation positions, these four variants strengthen the findings' translational significance and broaden the analytical focus of this study beyond the DNA-

binding domain, which has predominated previous TP53 structural research.

Considering their well-known and mechanistically significant functions in restoring p53 activity, COTI-2 and PRIMA-1 were identified as possible therapeutic targets. This selection is consistent with the functional impact of the TP53 mutations studied in this work. PRIMA-1 and its methylated derivative APR-246 are tiny compounds that have been extensively investigated due to their unusual ability to refold mutant p53 into a conformation comparable to wild-type p53. This step restores sequence-specific DNA-binding activity and activates apoptotic pathways that rely on p53. It is the gold standard for mutant p53 reactivation approaches because it is successful across a wide range of tumour types and TP53 mutations. Because of its dual action mechanism—reactivating mutant p53 and blocking important carcinogenic signalling pathways like AMPK regulation and PI3K/AKT/mTOR—COTI-2 was selected. COTI-2 has demonstrated anti-proliferative and pro-apoptotic effects in preclinical cancer models, including those resistant to conventional therapy, and it has demonstrated efficacy against a wider range of TP53 mutations than previous p53-targeting medications. Its ability to specifically target tumour cells while avoiding healthy cells is another therapeutic advantage [40]. We can assess the therapeutic restoration of TP53 function over a broader spectrum of detrimental mutations, both in terms of structure and function, by combining PRIMA-1 and COTI-2, two mutant p53-reactivating agents with distinct but complementary methods.

It is crucial to recognise the inherent limits of purely in silico techniques, even while the molecular docking and molecular dynamics simulation results offer insightful information about the binding affinity, stability, and interaction patterns of the chosen drugs with mutant TP53. The quality of the structural models, force fields, and scoring functions used in computational forecasts may not adequately represent the dynamic complexity of biological systems under physiological settings. Furthermore, docking algorithms offer an estimate of binding poses and energies that might not necessarily correspond with functional biological activity or actual binding affinities. In a similar vein, molecular dynamics simulations are constrained by simulation time scales and might not capture rare event dynamics or long-term conformational changes. Therefore, experimental validation through in vitro and in vivo research is necessary to demonstrate their therapeutic significance, even though the current findings provide a solid theoretical foundation for the potential efficacy of COTI-2 and PRIMA-1 against harmful TP53 mutations.

The detected harmful TP53 SNPs were compared with clinically reported TP53 mutations in breast cancer patients recorded in public variant databases and prior literature to increase the findings' translational applicability. The DNA-binding domain, which is known to contain a significant percentage of pathogenic variants linked to loss of tumour suppressor function, is one of the hotspot regions of TP53 that are commonly reported in breast cancer. Several of the high-confidence mutations found in this study correspond to these regions. However, in breast cancer cohorts, mutations at conserved residues within this domain have been linked to poor clinical outcomes, increased genomic instability, and decreased transcriptional activity. This overlap implies that the chosen SNPs may reflect functionally significant mutations contributing to the

pathophysiology of breast cancer and validates the clinical significance of the computationally prioritised variations. A portion of the discovered variations, however, have little to no direct clinical annotation in breast cancer databases, suggesting that they may be new or poorly characterised possibilities that need more research.

Because the work is solely based on in silico analysis without experimental confirmation, the therapeutic implications of the current findings should be regarded cautiously. Although favorable binding affinity and stable interactions of COTI-2 and PRIMA-1 with specific harmful TP53 mutations are suggested by the docking and molecular dynamics data, these findings are prediction computational results rather than verified biological impacts. As a result, any conclusions about direct therapeutic potential or clinical efficacy are yet conjectural. The findings mainly offer a theoretical framework for ranking potential drugs and locating structurally significant TP53 alterations that might react to reactivation techniques. To validate these anticipated interactions and ascertain their actual therapeutic significance in a biological setting, experimental research is necessary, including in vitro functional tests and in vivo validation models.

5. Conclusion

This study focused on deleterious SNPs' effect on the functioning of TP53. It includes the protein's evolutionary data and phylogenetic analysis. The AlphaFold predicted tertiary structure was validated and refined to perform further analysis. TP53 is localized in the nucleus and is secretory in nature, which supports its role in DNA damage response and transcriptional regulation. To study the impact of mutation on TP53's stability and functioning, in-situ mutagenesis was performed on TP53 at coordinates Q16K, L130Q, P223R, and R273I. Furthermore, stability analysis revealed that most of the variations result in reduced stability of the protein. GO analysis revealed that TP53's biological roles include regulating metabolism, gene expression, RNA processing, and transcription with high confidence. Functionally, it interacts with nucleic acids, acts as a transcription factor, and binds RNA (probability >0.8). For virtual screening, COTI-2 and PRIMA1 were identified as potential TP53 inhibitors based on pharmacokinetics and Lipinski's rule compliance. The molecular docking analysis of wild-type and mutated TP53 was carried out to point out the differences in the interaction pattern of the mutated protein with COTI-2 and PRIMA1 as compared to wild-type TP53.

Additional Information and Declaration

Funding

The authors received no funding for this work.

Competing Interests

The authors declare there are no competing interests.

Data Availability

The following information was supplied regarding data availability:
Raw data are available in the Supplementary Files.

References

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, *76*(1), 7-20.

Ahmad, H., Ali, A., Ali, R., Khalil, A. T., Khan, I., Khan, M. M., & Alorini, M. (2023). Mutational landscape and in-silico analysis of TP53, PIK3CA, and PTEN in patients with breast cancer from khyber Pakhtunkhwa. *ACS omega*, *8*(45), 43318-43331.

Bai, X., Ni, J., Beretov, J., Graham, P., & Li, Y. (2021). Triple-negative breast cancer therapeutic resistance: Where is the Achilles' heel? *Cancer letters*, *497*, 100-111.

Bykov, V. J. N., Eriksson, S. E., Bianchi, J., & Wiman, K. G. (2018). Targeting mutant p53 for efficient cancer therapy. *Nature Reviews Cancer*, *18*(2), 89-102.

Cheng, J. n., Ding, X., Xu, S., Zhu, B., & Jia, Q. (2020). Gene expression profiling identified TP53MutPIK3CAWild as a potential biomarker for patients with triple-negative breast cancer treated with immune checkpoint inhibitors. *Oncology letters*, *19*(4), 2817-2824.

Cozzetto, D., Minneci, F., Currant, H., & Jones, D. T. (2016). FFPred 3: feature-based function prediction for all Gene Ontology domains. *Scientific reports*, *6*(1), 31865.

DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool. *J CCP4 Newsl. Protein Crystallogr*, *40*(1), 82-92.

Deng, N. A., Zhou, H., Fan, H., & Yuan, Y. (2017). Single nucleotide polymorphisms and cancer susceptibility. *Oncotarget*, *8*(66), 110635.

Economopoulou, P., Dimitriadis, G., & Psyrri, A. (2015). Beyond BRCA: new hereditary breast cancer susceptibility genes. *Cancer treatment reviews*, *41*(1), 1-8.

Emadi, E., Akhoundi, F., Kalantar, S. M., & Emadi-Baygi, M. (2020). Predicting the most deleterious missense nsSNPs of the protein isoforms of the human HLA-G gene and in silico evaluation of their structural and functional consequences. *BMC genetics*, *21*, 1-27.

Joerger, A., & Fersht, A. R. (2007). Structure–function–rescue: the diverse nature of common p53 cancer mutants. *Oncogene*, *26*(15), 2226-2242.

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of molecular biology*, *292*(2), 195-202.

Kasilingam, T., & Elengoe, A. (2018). In silico molecular modeling and docking of apigenin against the lung cancer cell proteins. *Asian J. Pharm. Clin. Res*, *11*(9), 246-252.

Khan, K., Shah, H., Rehman, A., Badshah, Y., Ashraf, N. M., & Shabbir, M. (2022). Influence of PRKCE non-synonymous variants on protein dynamics and functionality. *Human Molecular Genetics*, *31*(13), 2236-2261.

Li, X., Chen, X., Wen, L., Wang, Y., Chen, B., Xue, Y., . . . Liao, N. (2020). Impact of TP53 mutations in breast cancer: Clinicopathological features and prognosisImpact of TP53 mutations in breast CA. *Thoracic Cancer*, *11*(7), 1861-1868.

Lilyquist, J., Ruddy, K. J., Vachon, C. M., & Couch, F. J. (2018). Common genetic variation and breast cancer risk—past, present, and future. *Cancer Epidemiology, Biomarkers & Prevention*, *27*(4), 380-394.

Liu, Y., Wang, X., Wang, G., Yang, Y., Yuan, Y., & Ouyang, L. (2019). The past, present and future of potential small-molecule drugs targeting p53-MDM2/MDMX for cancer therapy. *European journal of medicinal chemistry*, *176*, 92-104.

Lu, T., Zou, Y., Xu, G., Potter, J. A., Taylor, G. L., Duan, Q., . . . Ye, D. (2016). PRIMA-1Met suppresses colorectal cancer independent of p53 by targeting MEK. *Oncotarget*, *7*(50), 83017.

Lukong, K. E. (2017). Understanding breast cancer–The long and winding road. *BBA clinical*, *7*, 64-77.

Lundberg, P. C., & Phoosuwan, N. (2022). Life situations of Swedish women after mastectomy due to breast cancer: A

qualitative study. *European Journal of Oncology Nursing, 57*, 102116.

Martin, A. C. R., Facchiano, A. M., Cuff, A. L., Hernandez-Boussard, T., Olivier, M., Hainaut, P., & Thornton, J. M. (2002). Integrating mutation data and structural analysis of the TP53 tumor-suppressor protein. *Human mutation, 19*(2), 149-164.

Meric-Bernstam, F., Zheng, X., Shariati, M., Damodaran, S., Wathoo, C., Brusco, L., . . . Basho, R. K. (2018). Survival outcomes by TP53 mutation status in metastatic breast cancer. *JCO precision oncology, 2*, 1-15.

Mooney, S. (2005). Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis. *Briefings in bioinformatics, 6*(1), 44-56.

Nounou, M. I., ElAmrawy, F., Ahmed, N., Abdelraouf, K., Goda, S., & Syed-Sha-Qhattal, H. (2015). Breast cancer: conventional diagnosis and treatment modalities and recent patents and technologies. *Breast cancer: basic and clinical research, 9*, BCBCR-S29420.

O'Connor, P. M., Jackman, J., Bae, I., Myers, T. G., Fan, S., Mutoh, M., . . . Weinstein, J. N. (1997). Characterization of the p53 tumor suppressor pathway in cell lines of the National Cancer Institute anticancer drug screen and correlations with the growth-inhibitory potency of 123 anticancer agents. *Cancer research, 57*(19), 4285-4300.

Pollock, N. C., Ramroop, J. R., Hampel, H., Troester, M. A., Conway, K., Hu, J. J., . . . Ziv, E. (2022). Differences in somatic TP53 mutation type in breast tumors by race and receptor status. *Breast cancer research and treatment, 192*(3), 639-648.

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic acids research, 47*(D1), D886-D894.

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic acids research, 39*(17), e118-e118.

Selga, C., Koc, A., Chawade, A., & Ortiz, R. (2020). A bioinformatics pipeline to identify a subset of SNPs for genomics-assisted potato breeding. *Plants, 10*(1), 30.

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: a cancer journal for clinicians, 68*(1), 7-30.

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic acids research, 40*(W1), W452-W457.

Singh, A. K., Olsen, M. F., Lavik, L. A. S., Vold, T., Drabløs, F., & Sjursen, W. (2021). Detecting copy number variation in next generation sequencing data from diagnostic gene panels. *BMC Medical Genomics, 14*(1), 214.

Sohail, S., & Alam, S. N. (2007). Breast cancer in Pakistan-awareness and early detection.

Wang, H., Guo, M., Wei, H., & Chen, Y. (2023). Targeting p53 pathways: Mechanisms, structures and advances in therapy. *Signal transduction and targeted therapy, 8*(1), 92.

Young, K. H., Leroy, K., Møller, M. B., Colleoni, G. W. B., Sánchez-Beato, M., Kerbauy, F. R., . . . Gaulard, P. (2008). Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood, The Journal of the American Society of Hematology, 112*(8), 3088-3098.

Zhou, X., Hao, Q., & Lu, H. (2019). Mutant p53 in cancer therapy—the barrier or the path. *Journal of molecular cell biology, 11*(4), 293-305.