



Comparative Modeling and Molecular Docking Study of P53 and AKT1, Genes of Lung Cancer Pathways

Asif Mir^{*}, Syeda Naqsh e Zahra, Sobiah Rauf

Department of Bioinformatics and Biotechnology, International Islamic University, Islamabad, Pakistan

Email address:

asif.mir@iiu.edu.pk (A. Mir)

^{*}Corresponding author

To cite this article:

Asif Mir, Syeda Naqsh e Zahra, Sobiah Rauf. Comparative Modeling and Molecular Docking Study of P53 and AKT1, Genes of Lung Cancer Pathways. *International Journal of Clinical Oncology and Cancer Research*. Vol. 1, No. 1, 2016, pp. 6-14.

doi: 10.11648/j.ijcoocr.20160101.12

Received: October 15, 2016; **Accepted:** December 3, 2016; **Published:** January 9, 2017

Abstract: The fundamentals of structure-based drug designing rely on protein-ligand interactions, which play a significant role to open a gateway from identification of active residues and development of potential drugs. The endeavor behind this work is to select most susceptible genes p53 and AKT1 which plays a vital role in lung cancer pathogenicity. In a queue to deliberate the crucial role of these genes *in-silico* experimental strategy was adopted. 3-D structure of p53 generated by YASARA showed 50.9% sequence identity with 2PCX-A and Z-score of -0.276 while AKT1 showed 66.3% sequence identity with 3QKL-A and Z-score of 0.036. Mutational analysis revealed that R273L and C275Y mutations of p53 destabilize the DNA binding domain, while E17K mutation of AKT1 directly affect the binding of the ligand as this residues lines the pocket. Molecular docking was performed using ligands Staurosporine and Nutlin-3 retrieved form ZINC database. Blind docking experiment revealed that p53 involve non polar (Leu206, Leu188, Pro190), acidic (Glu204, Tyr 205) and basic (Arg202) as most interacting residues. AKT1 interactions with ligand Staurosporine revealed nonpolar (Val164, Phe438, Phe442, Phe 236, Phe 237, Phe 161), polar (Gly159, Gly157, Gly234, Gly 278), basic (Lys163, Lys158, Lys 276, Lys 179), acidic (Asp439, Glu278) as most interacting residues. It is assumed that current study will play a significant contribution to design potential drug inhibitors by utilizing most interactive residue information with Nutlin-3 and Staurosporine ligands to restrain the interaction between p53 pathways and epidermal growth pathways. Structural based receptor-ligand interactions likely to be used against anti-cancer therapy.

Keywords: P53, AKT1, Lung Cancer, Comparative Modeling, Molecular Docking, Mutational Analysis

1. Background

Lung cancer is one of the most common type of cancer worldwide [1]. The mortality rate in lung cancer is 1.3 million per annum causing it to be the most prevalent cancer type with a very poor diagnosis and the foremost cause of mortality due to cancer in both well developed and developing countries. By the end of 20th century, lung cancer had become one of the major causes of preventable death. Strong evidences showed that 85-90% is caused due to tobacco smoking resulting in bronchogenic carcinoma [2, 3]. According to epidemiologic studies, cigarette smoking is the major factor contributing towards development of lung cancer and it is reported that almost 85% cases of Non-Small cell lung cancer (NSLC) and 98% cases of Small cell lung

cancer (SCLC) develop in smokers. Classification of lung cancer is shown in Figure 1. However, about 50% of lung cancers are apparent in ex-smokers and 25% of cases in the world occur in patients who never smoke. Extent of smoking and the quantity of cigarettes smoked in a day increases the lung cancer development chances among the smokers. In developing countries, lung cancer will remain a major health problem for the near future as cigarette smoking becomes more prevalent.

Lung cancer is initiated by activation of oncogenes or tumor suppressor genes inactivation. Previously reported research studies indicated various causes of lung cancer. Mutations in genes like K-ras proto-oncogene cause 10–30% adenocarcinomas in lung. Epidermal growth factor receptor (EGFR) modulates proliferation of cell, angiogenesis, tumor invasion and apoptosis. Oncogenic mutations and EGFR

amplifications occur frequently in non-small cell lung cancer so EGFR inhibitors can be used for treatment of this common disease. Some other oncogenes involved in non-small cell lung cancer are NKX2-1, PIK3CA, c-MET and BRAF. Tumor suppressor gene inactivation has a crucial role in carcinogenesis of lung. Moreover, p53 which is most common tumor suppressor gene is mutated in about 60–75% lung cancer patients including both NSCLC and SCLC while RB is mutated in SCLC [4]. Moreover, DNA repair

capacities are one of the major factors that can affect lung cancer susceptibility. Research studies indicate that DNA repair gene polymorphisms also enhance the lung cancer risk, of which nucleotide excision repair genes (ERCC1-2 and XPA), mismatch repair genes (MLH1 and MSH2), base excision repair genes (XRCC1 and OGG1) and double-strand DNA break repair genes (XRCC3) are the most common one in causing lung cancer [5].

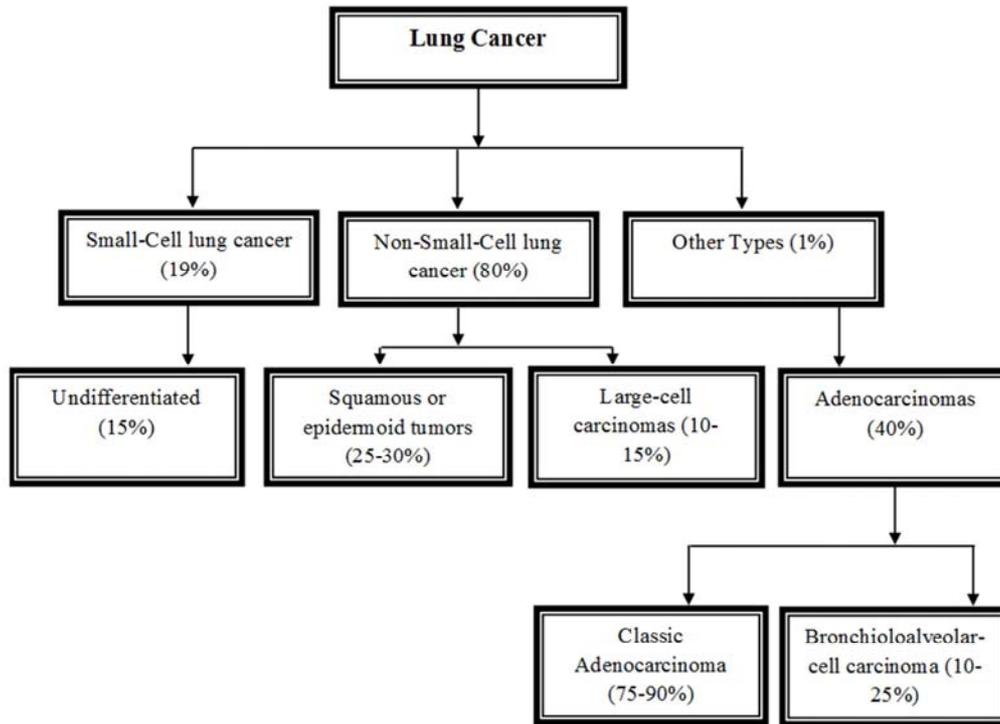


Figure 1. Classification of Lung Cancer.

Several different signaling pathways have their functions altered, therefore, the genetic anomalies associated with risk of lung cancer should be studied by considering signalling pathways instead of focusing on individual factors involved. These pathways have significant importance in context of targeted therapy. Major signalling pathways that could play significant role in lung cancer therapy include (1) Growth promoting pathways (Epidermal Growth Factor Receptor/Ras/ Phosphatidylinositol 3-Kinase) (2) Growth inhibitory pathways (p53/Rb/P14ARF, STK11) (3) Apoptotic pathways (Bcl-2/Bax/Fas/FasL).

Over expression of EGFR genes are reported in cancers of the brain, breast, endometrium, head and neck, oesophagus, stomach, colon, ovary, bladder and lung [6]. Mutations in EGFR pathway genes have been found in ~10–30% cases of NSCLC but these mutations occur very rarely in other tumors [7]. Rare somatic mutations are also reported in other genes of EGFR pathway genes, like BRAF, HER2, PIK3CA and HER4.

The p53 pathway includes several different genes which belong to various upstream and downstream sub-pathways. p53 serve as gatekeeper of the cell which protect the cell

from genetic instability and acts to sense multiple stress signals, some of which include DNA damage, activation of oncogene and hypoxia.

In current research study, Epidermal growth factor receptor (EGFR) signaling pathway and p53 pathways are selected for application of bioinformatics techniques like digital candidate gene approach (DigiCGA), expression analysis, Comparative modeling and molecular docking will be used to avoid time, expenditure, labor cost and quick analysis for prioritized candidate genes as well as to get deep understanding of the genes involved in these pathways of lung cancer.

2. Materials and Methods

The methodology used in present research involves the identification of two pathways, (i) Epidermal growth factor receptor (EGFR) signalling pathway (ii) p53 pathway already been reported in several research studies and considered to play key role in pathogenesis of lung cancer. Flow chart of methodology adopted in current study is shown in Figure 2. Summary of tools used in current research is displayed in Table 1.

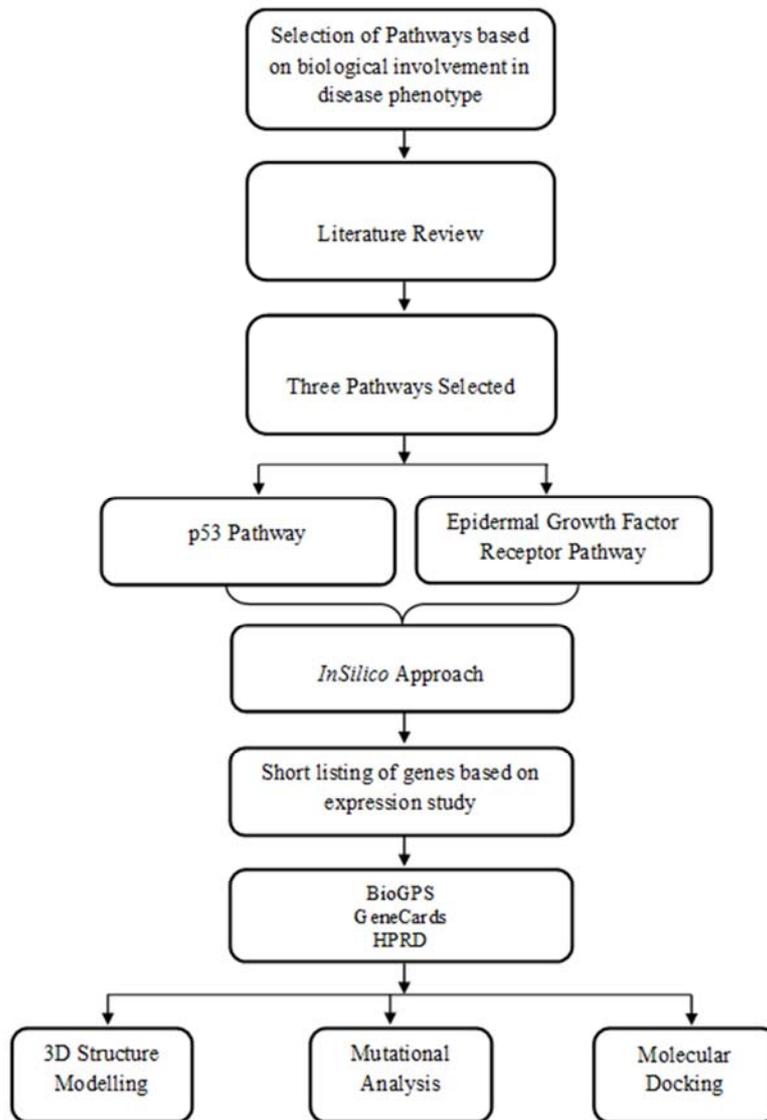


Figure 2. Protocol for the computational analysis of lung cancer.

Table 1. Summary of Tools Used In Current Research.

S/N	Tools/Database	Output
1	BioGPS	
2	HPRD	Expression profiling
3	Gene Cards	
4	YASARA	Structure Prediction, Molecular Docking
5	HOPE server	Mutational analysis
6	ZINC database	Ligand Retrieval
7	ChemDraw	Ligand Drawing
8	Chimera	Protein-Ligand interactions
9	LIGPLOT	Protein-Ligand interactions

2.1. Expression Profiling of Genes

In order to shortlist most plausible candidate genes out of total thirty two genes present in selected lung cancer signaling pathways, expression profiling was performed by using expression based databases like Human Protein

Reference Database (HPRD), BioGene Portal System (BioGPS) and GeneCards. HPRD (<http://www.hprd.org/>) consists of all curated proteomic information of human proteins like protein–protein interactions, post-translational modifications (PTMs) and tissue expression based on experimental proven data [8]. BioGPS (<http://biogps.org/>) works on the principle of gene annotation which relies on three major characteristics of gene of interest i.e. (i) molecular function, (ii) biological process, and (iii) cellular location [9] while GeneCards (<http://www.genecards.org/>) is a human gene database, which is responsible to give information about all known and predicted human genes. For the better understanding of disease pathogenicity, it retrieves biological information based on transcriptomic, genetic, proteomic and functional aspects.

2.2. Comparative Modeling

The amino acid sequences of p53 and AKT1 were

retrieved from NCBI database using accession number P04637 and P31749 respectively. These sequences of both genes were subjected to position specific iterated (PSI) BLAST [10] search against protein databank (PDB) [11] for identification of best possible template structures.

Homology/comparative modeling approach was implemented for whole protein sequences by using YASARA software, which is a tool for computational analysis with 3-D structure prediction, molecular docking, molecular visualization and molecular dynamics to predict the 3-dimensional structures as their structures were not resolved and reported yet in RCSB Protein Data Bank.

2.3. Model Evaluation

Predicted 3D structures were evaluated by PROCHECK and by ERRAT [12] servers.

2.4. Molecular Docking

Molecular docking studies were performed on set of selected genes, in which blind docking technique was adopted, in order to predict the structure of the intermolecular complex formed between prioritized proteins and their respective ligands through YASARA software. Ligands for

p53 and AKT1 were retrieved, from ZINC database [13], which have evidences of interaction with respective proteins. Nutlin-3 was found as a ligand for p53, it inhibits the interaction between p53 and MDM2 and ligand for AKT1 was Staurosporine which is ATP-competitive kinase inhibitor. Chem Draw Ultra Version 8.0 was used to draw required ligand retrieved for receptor-ligand docking [14]. Both these proteins were subjected to 50 docking runs and almost 25 complexes were generated. Complexes with the lowest docked binding energy were selected for further post-docking analysis. Post-docking Analysis was performed with LIGPLOT and CHIMERA [15] software and interactions between receptor and ligand were analyzed.

3. Results and Discussion

3.1. Expression Profiling

In order to short list the genes involved in selected pathways, expression profiling was done. p53 showed medium expression where as high expression was retrieved for AKT1 using BioGPS, HPRD and Genecards databases as depicted in Table 2.

Table 2. Expression profiles of genes of p53 pathway.

Gene	Molecular Function	Biological Process	Cellular Location	Expression in Lungs
P53	Transcription factor activity	Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism; Apoptosis	Nucleus	Medium
AKT1	Protein serine/threonine kinase activity	Cell communication; Signal transduction	Cytoplasm	High

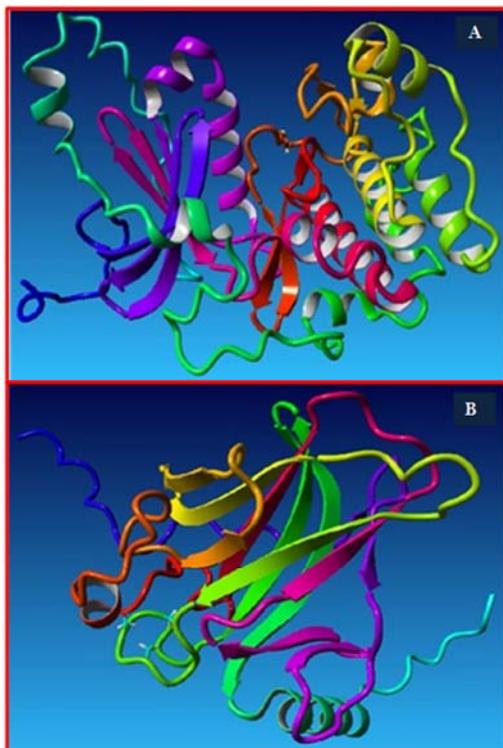


Figure 3. Predicted Structures: A) AKT1 Protein Structure using 3QKL template with Z-Score of 0.036, B) TP53 Protein Structure using 2PCX template with Z-Score of -0.276.

3.2. Comparative Modeling

YASARA software generated a hybrid structure using 2-5 templates which are ranked on the basis of alignment score (PSI_BLAST) and structural quality (Z_Score) according to WHAT CHECK [19] obtained from the PDBfinder2 database for candidate genes.

For AKT1, five templates 3QKL-A, 3CQW-A, 3D0E, 4EJN-A, 3O96-A with optimal alignment of first template and good alignment for remaining templates sorted by their overall quality Z-scores and E-values are mentioned in Table 3. Hybrid structure was generated using best aligned parts of all the five templates (Figure 3 part A). While for p53, Table 4 lists the template models sorted by their overall quality Z-scores and E-values. Hybrid structure was generated using best aligned parts of templates 3IGL-A, 2PCX-A, 4AGQ-B, 3Q01, 2RMN-A02 with good and satisfactory alignment respectively (Figure 3 Part B).

Table 3. Templates for AKT1 sorted by their overall quality Z-scores and E-values.

Model ID	Z-Score	Alignment	BLAST E-value
3QKL-A	0.036	Optimal	0
3CQW-A	-0.124	Good	0
3D0E	-0.225	Good	0
4EJN-A	-0.318	Good	0
3O96-A	-0.881	Good	0

Table 4. Templates for TP53 sorted by their overall quality Z-scores and E-values.

Model ID	Z-Score	Alignment	BLAST E-value
3IGL-A	-0.685	Good	4e-086
2PCX-A	-0.276	Good	2e-085
4AGQ-B	-0.331	Good	6e-086
3Q01	-0.446	Good	1e-095
2RMN-A02	-1.401	Satisfactory	2e-086

Table 5. Results of PROCHECK and ERRAT. MFR stands for Most Favored Region, AAR stands for Additionally Allowed Region, GAR stands for Generously Allowed Region and DAR stands for Disallowed Region.

Protein	PROCHECK				ERRAT
	Ramachandran Plot Data				
	MFR	AAR	GAR	DAR	Overall Quality Factor
AKT1	89.9%	8.8%	1.0%	0.3%	94.260
P53	92.3%	7.1%	0.5%	0.0%	90.385

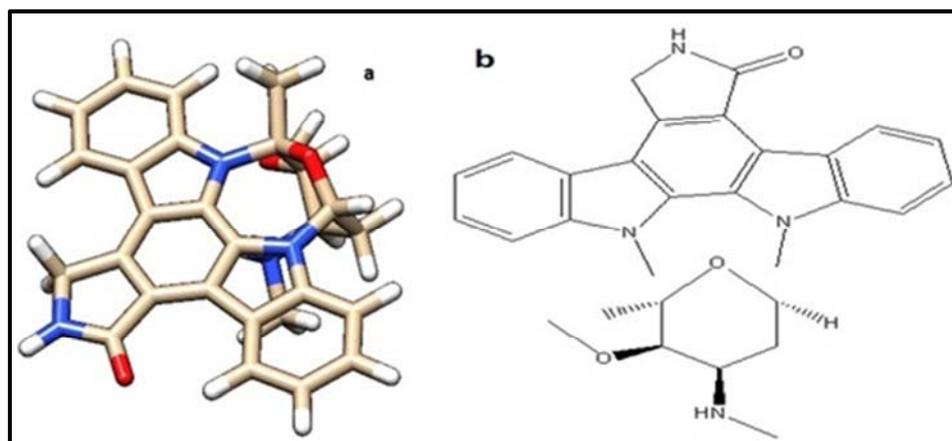
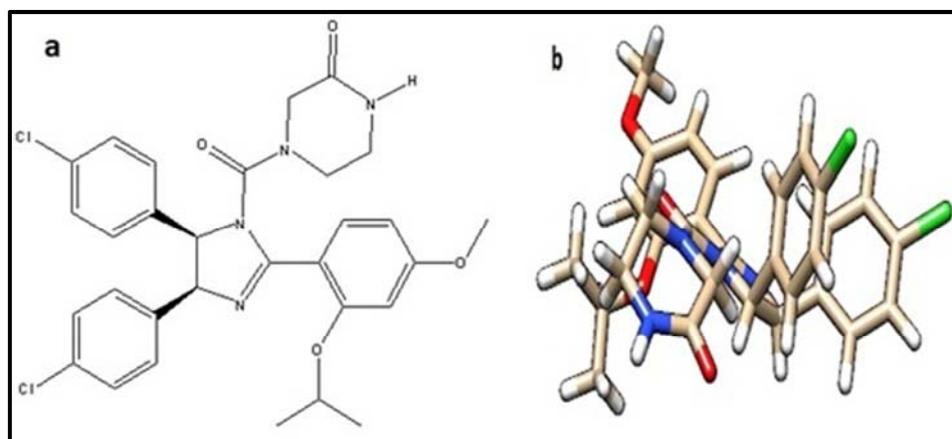
3.4. Molecular Docking

In order to obtain an optimal docking conformation we undertook 50 runs of docking for p53 and AKT1 proteins each Nutlin-3 and Staurosporine respectively. Complexes having lowest binding energies are selected for post docking

3.3. Model Evaluation

The predicted models were subjected to PROCHECK and ERRAT to find the reliability of models. Table 5 shows the Ramachandran plot characteristics and overall quality factor obtained for p53 and AKT1 predicted models which supported the quality of the modeled structures.

analysis. Structures of ligands for AKT1 and p53 are shown in Figure 4 and 5. Both ligands were found to bind with their respective receptor proteins as shown in docked complexes in Figure 6 part A and B. Table 6 displayed the binding energies of docked complexes of our proteins.

**Figure 4.** (a) Structure of AKT1 Ligand Staurosporine (2,3,10,11,12,13-hexahydro-10R-methoxy-9S-methyl-11R-methylamino-9S,13R-epoxy-1H,9H-diindolo[1,2,3-gh;3',2',1'-lm]pyrrolo[3,4-j][1,7]benzodiazonin-1-one) drawn through Chimera (b) Structure Of Staurosporine drawn through ChemDraw. Staurosporine is an ATP-competitive kinase inhibitor due to this property it binds to many kinases with high affinity.**Figure 5.** a) p53 Ligand Nutlin-3 (4-[4, 5-Bis (4-chlorophenyl)-2-(4-methoxy-2-propan-2-yloxyphenyl)-4, 5-dihydroimidazole-1-carbonyl] piperazin-2-one) drawn through ChemDraw (b) Structure of Nutlin-3 drawn through Chimera. Nutlin-3 inhibit the interaction of mdm2 and p53 and induce a growth inhibitory state in cancer cells.

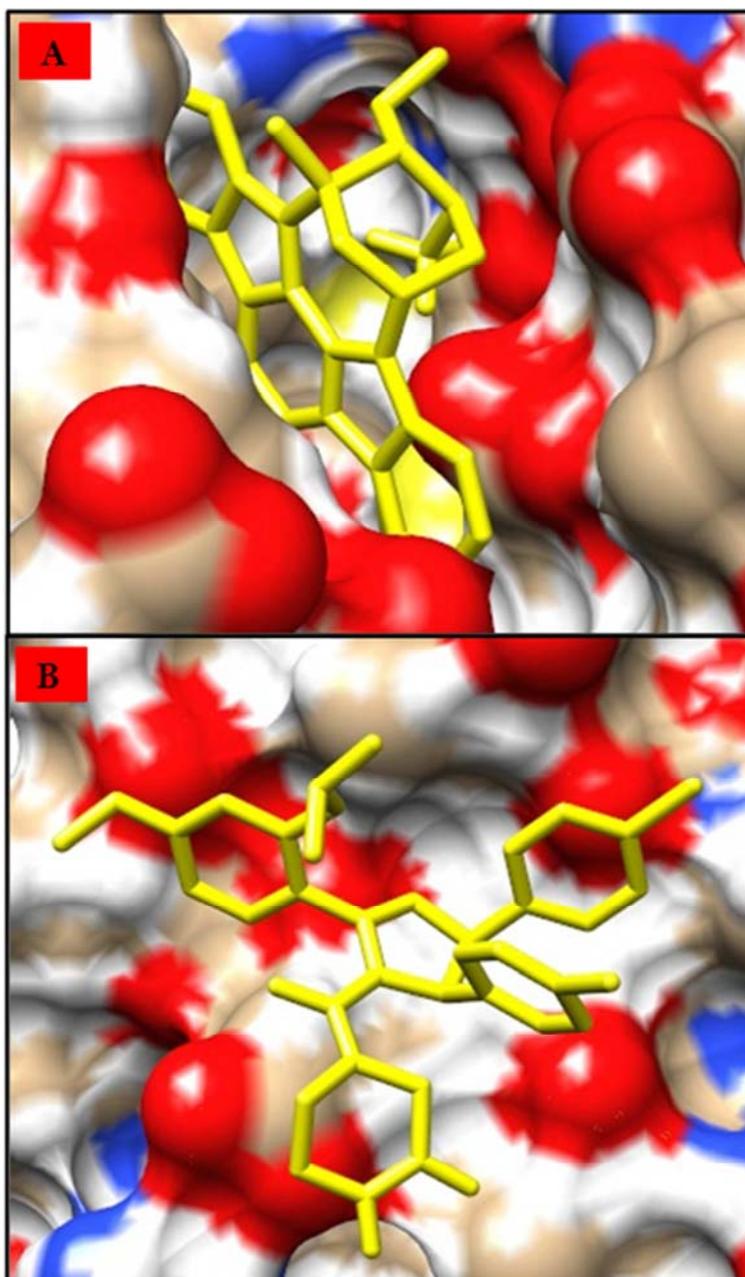


Figure 6. A Frontal View of the Docking Complexes A) Receptor AKT1 with ligand Staurosporine B) Receptor p53 with Ligandnutlin-3. Ligand is shown in "Stick" format and yellow color whereas the receptor is shown in "Surface" format.

Table 6. Binding Energies and Inhibition Constants of Docked Complexes and Interacting Receptor Protein Residues.

Protein	Ligand	Bind.energy [kcal/mol]	Inhib.constant Ki	Interacting Receptor Protein Residues
TP53	Nutlin-3	000005.1800	00000159140000.0000	GLY 187, LEU 188, PRO 190, LEU 201, ARG 202, VAL 203, GLU 204, TYR 205, LEU 206
AKT1	Staurosporine	000008.7200	00000000405620.0000	LEU156, GLY 157, LYS 158, GLY 159, THR 160, PHE 161, GLY 162, LYS 163, VAL 164, LYS 179, LEU 181, GLU 234, PHE 236, PHE 237, LYS 276, GLU 278, ASN 279, PHE 438, ASP 439, PHE 442

Results of interactions obtained through Chimera for AKT1 and p53 are shown in Figure 7 part A and B, respectively. Figure 8 and 9 obtained using LIGPLOT indicate binding of ligand within the active site for AKT1 and p53 respectively. The amino acid residues of p53 that played key role in interacting with Nutlin-3 ligand are that

are also significantly involved in the binding interactions are Leu 206, Glu 204, Tyr 205, Pro 190, Arg 202 and Leu 188 whereas AKT1 residues that are significantly involved in the binding interactions with Staurosporine are Lys 163, Gly 159, Val 164, Gly 157, Lys 276, Glu 278, Glu 234, Phe 236, Phe 237, Asp 439, Phe 438, Phe 442, Phe 161 and Lys 179.

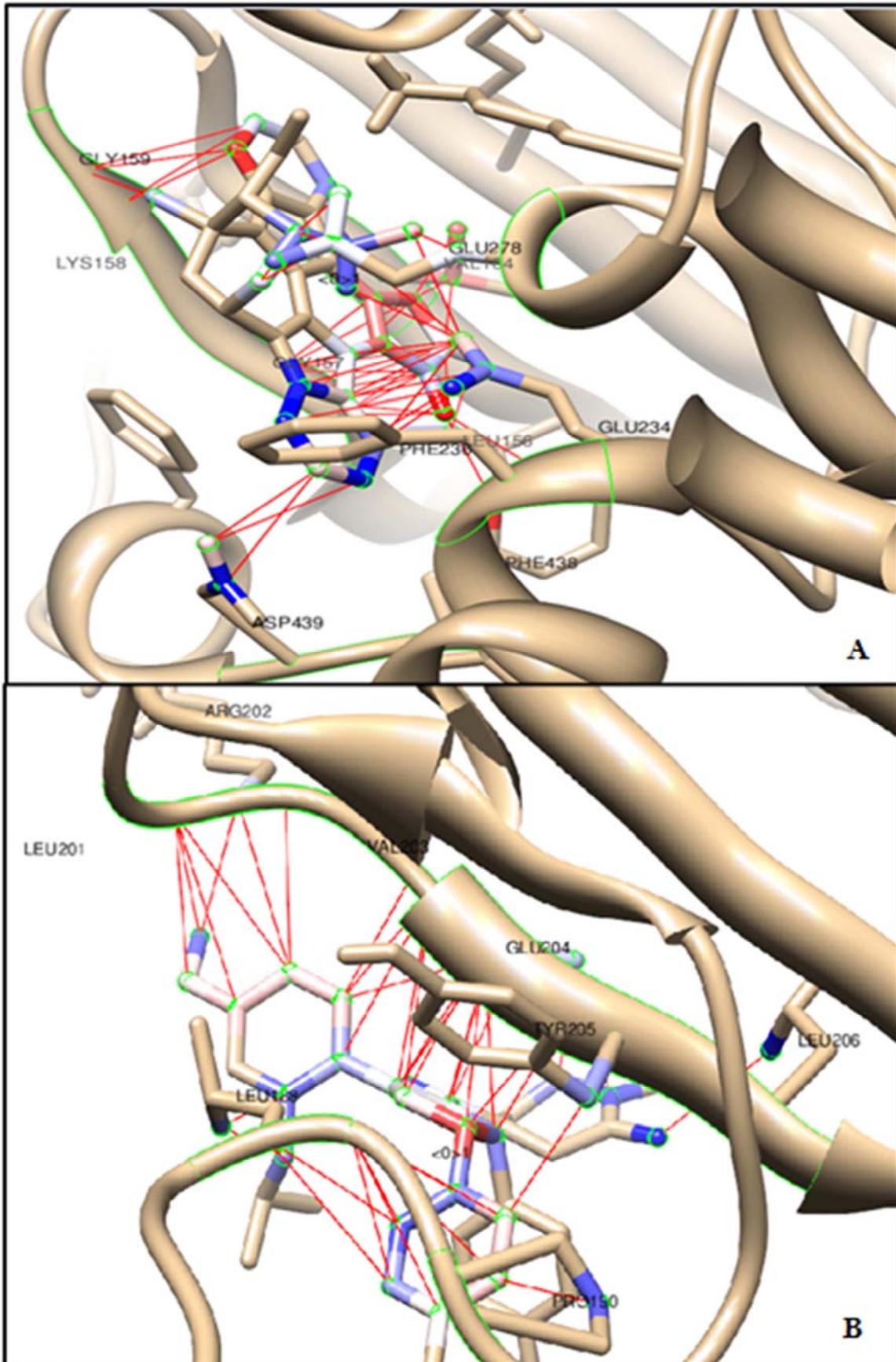


Figure 7. Interactions retrieved through CHIMERA A) important amino acid residues of AKT1 and Staurosporine. B) Amino acid residues of p53 and Nutlin-3.

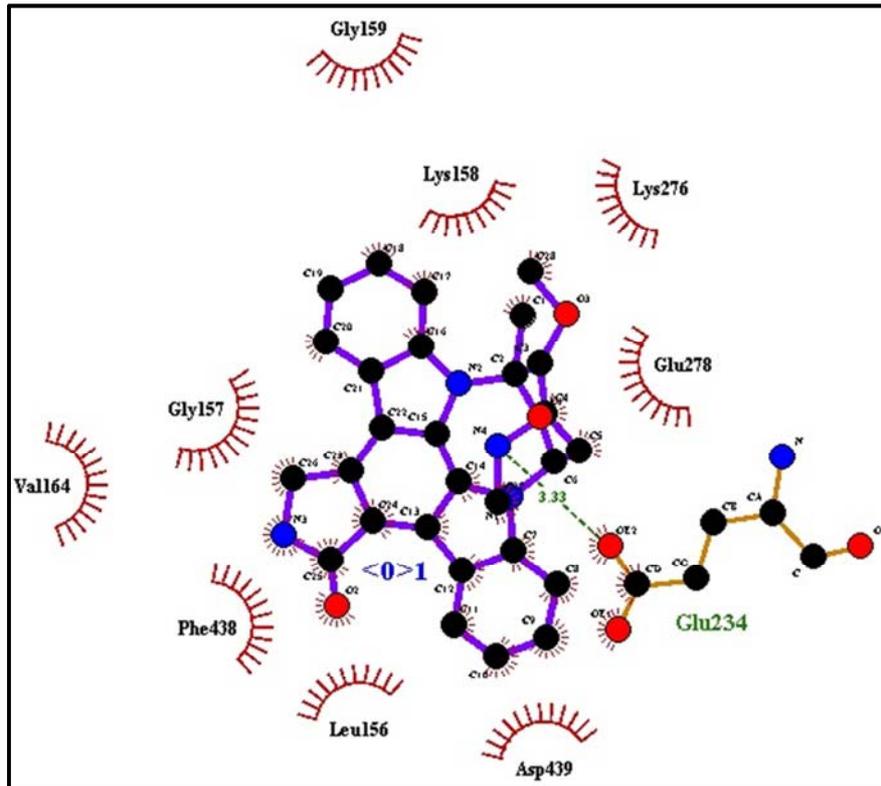


Figure 8. The interaction of Staurosporine with AKT1 revealing non polar (Val164, Phe438, Phe442, Phe 236, Phe 237, Phe 161), polar (Gly159, Gly157, Gly234, Gly 278), basic (Lys163, Lys158, Lys 276, Lys 179), acidic (Asp439, Glu278) as most interacting residues with receptor.

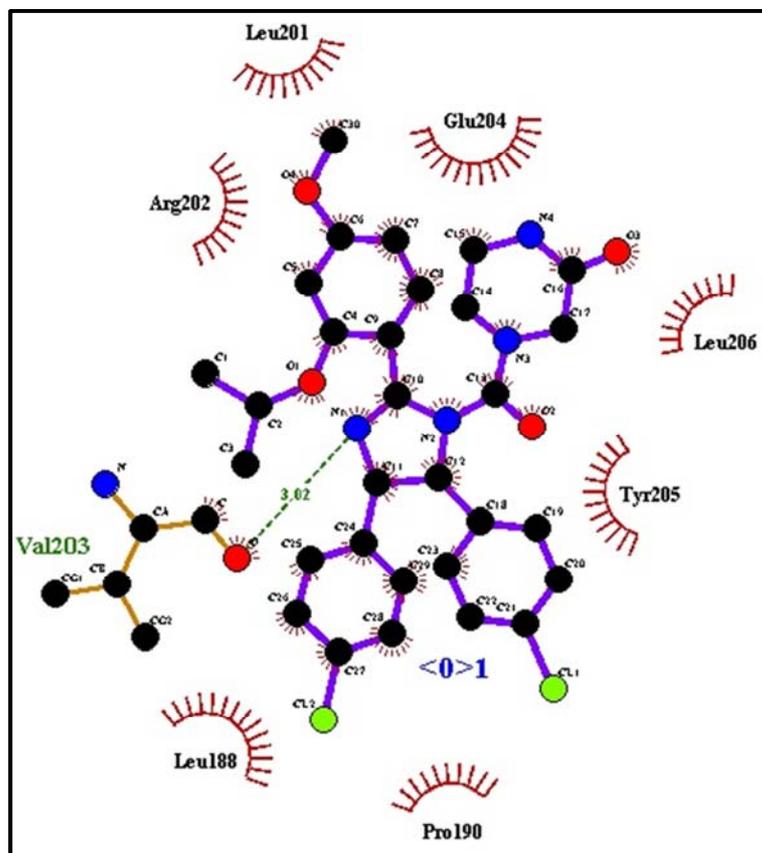


Figure 9. Interaction of Nutlin-3 with p53 involving major amino acids in interactions i.e. Non polar (Leu206, Leu188, Pro190), acidic (Glu204, Tyr 205) and basic (Arg202) as most interacting residues.

4. Conclusion

In present study, two key signaling pathways were selected and prioritized candidate genes from these signaling pathways were p53 and AKT1. p53 showed medium expression and AKT1 showed high expression in lungs. Comparative modeling with model assessment showed that generated 3D structure for most plausible candidate genes are reliable and scientists may go for wet lab experimentation by using mentioned templates. The binding features of prioritized genes with their ligands were also studied using docking approach. To inhibit the interaction between p53 and MDMD2 genes, which are closely functional partners of p53 pathway, Nutlin-3 ligand was retrieved which is commonly used against many cancer therapies by inducing a growth-inhibiting state called senescence in cancer cells, therefore thought to work best on tumors that contain normal or wild type p53. The second ligand used against AKT1 is Staurosporine, which avert the binding of ATP with kinase, alternatively inhibit protein kinases. By utilizing the core functions of ligands against mutated p53 and AKT1 proteins in docking study revealed the functionally important residues among p53 and Nutlin- 3, and AKT1 with staurosporine, which may play a significant role in disrupting the network of p53 and epidermal growth pathways, leading to make a potential drug against lung cancer.

Acknowledgements

We are thankful to Department of Bioinformatics and Biotechnology, International Islamic University, Islamabad for providing us a working platform.

References

- [1] Kiyohara C, Otsu A, Shirakawa T, Fukuda S, Hopkin MJ (2002) Genetic polymorphisms and lung cancer susceptibility: a review. *Lung Cancer* 37: 241-256.
- [2] Blot WJ, McLaughlin JK (1998) Passive smoking and lung cancer risk: what is the story now. *Journal of the National Cancer Institute* 90: 1416-1417.
- [3] Hackshaw KA, Law RM, Wald JN (1997) The accumulated evidence on lung cancer and environmental tobacco smoke. *British Medical Journal* 315: 980.
- [4] Wang Y, Yang H, Li H, Li L, Wang H, Liu C, Zheng Y (2009) Association between X-ray repair cross complementing group 1 codon 399 and 194 polymorphisms and lung cancer risk: A meta-analysis. *Cancer Letters* 285: 134-140.
- [5] Zhou W, Liu G, Miller PD, Thurston WS, Xu LL, Wain CJ, Lynch JT, Su L, Christiani CD (2003) Polymorphisms in the DNA Repair Genes XRCC1 and ERCC2, Smoking, and Lung Cancer Risk. *Cancer Epidemiology, Biomarkers & Prevention* 12: 359–365.
- [6] Nicholson IR, Gee WMJ, Harper EM (2001) EGFR and cancer prognosis. *European Journal of Cancer* 37: 9–15.
- [7] Shigematsu H, Gazdar AF (2006) Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *International Journal of Cancer* 118: 257–262.
- [8] Prasad KST, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, Balakrishnan L, Marimuthu A, Banerjee S, Somanathan SD, Sebastian A, Rani S, Ray S, Kishore HJC, Kanth S, Ahmed M, Kashyap KM, Mohmood R, Ramachandra LY, Krishna V, Rahiman AB, Mohan S, Ranganathan P, Ramabadran S, Chaerkady R, Pandey A (2009) Human Protein Reference Database—2009 update. *Nucleic Acids Research* 37: D767–D772.
- [9] Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge LC, Haase J, Janes J, WHuss J, Su IA (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biology* 10: R130.
- [10] Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK - a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26:283–291.
- [11] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Research* 28:235-242.
- [12] Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science* 2: 1511-1519.
- [13] Irwin JJ, Shoichet KB (2005) ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *Journal of Chemical Information and Modeling* 45: 177–182.
- [14] Mendelsohn LD (2004) ChemDraw 8 Ultra: Windows and Macintosh Versions. *Journal of Chemical Information and Computer Sciences* 44: 2225–2226.
- [15] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 13: 1605-12.