Impact of Single amino acid Polymorphisms in Protein-Protein interactions in tumorigenic cluster A and cluster B of VHL: Computational molecular dynamics

George Priya Doss C[*1], Chiranjib Chakraborty[2], Siddarth R[1], Nagasundaram N[1], Magesh R[3], Udhayakumar S[1], Priyanka N[1], Priyadarshini Christy J[1]

[*1]*Medical Biotechnology Division, School of Biosciences and Technology, VIT University, Vellore 632014, Tamil Nadu, India;*[2]*Department of Bio-informatics, School of Computer and Information Sciences, Galgotias University, Greater Noida, India;*

[3]*epartment of Biotechnology, Faculty of Biomedical Sciences, Technology & Research, Sri Ramachandra University, Chennai 600116, India**E-mail: georgecp77@yahoo.co.in

Abstract. The VHL tumour suppressor gene codes for von-Hippel Lindau (VHL) protein which acts by association with Elongin C and Hypoxia Inducing Factor-1 alpha (HIF). Two functional sub-domains in the protein model of VHL contain binding sites for these proteins. Alpha domain of VHL protein contains the binding site for Elongin C (tumorigenic cluster A) whereas; the beta domain contains the binding site for HIF-1 alpha (tumorigenic cluster B). The complex formation between VHL, Elongin C and HIF- 1 alpha protein is crucial for the regulation of ubiquitin dependent proteolysis of oncogenic proteins. Henceforth defects in the complex formation may lead to von-Hippel Lindau disease. A majority of VHL disease causing mutations found to occur in these binding sites are highly conserved. In this computational analysis, we predicted the functional effect of Single amino acid Polymorphisms (SAPs) in VHL-interacting proteins with the aid of pathogenicity and protein stability predictors. Furthermore, molecular dynamic behavior of the native and mutant protein complexes were analyzed by using different parameters such as Root Mean Square Deviation, Root Mean Square Fluctuation and Hydrogen bonding variations Molecular dynamics studies were performed for five variants S111R, H115Q and W117C, L158V and C162F of VHL protein that occur in the Elongin-C and HIF-1 alpha protein binding domain. Through the molecular dynamics simulations analysis, we observed that the stability and flexibility between VHL-Elongin-C and VHL-HIF-1 complexes were altered. Our strategy may be valuable for understanding SAP effects on protein-protein interactions with function and their role in human genetic diseases, linking structure function relationship and also for the development of novel pharmacological strategies.

Keywords: VHL, Mutation, nsSNP, pVHL-HIF-1, pVHL-Elongin-c, Protein-Protein interaction, Molecular dynamics.

1 Introduction

DNA variations alter the functionality of genes and consequently the structure and function of the gene product. These changes can be restricted to a single nucleotide such as Single Nucleotide Polymorphisms (SNPs). SNPs can be deletions, insertions or substitutions and found to be the most common type of variations that occur in the humans [1]. A particular class of such SNPs that occur in the coding region, called non-synonymous SNPs (nsSNPs) can cause change in the amino acid composition of the protein also known as Single Amino acid Polymorphisms (SAPs). According to the OMIM (Online Mendelian Inheritance in Man) and HGMD (Human Gene Mutation Database) database, these SAPs which result from nsSNPs were identified to be detrimental, and responsible for many disease associations in humans. Disease associated mutations that occur in protein interaction interface (active sites, allosteric binding sites etc.) may disrupt the protein interaction network either by gain or loss of interactions. These interactions are involved in metabolic process, signalling and gene regulatory networks. Therefore, better understanding of the protein–protein interactions may elucidate the molecular mechanisms leading to disease and genotype-phenotype associations. Several studies have predicted the impact of mutation in protein-protein interaction network leading to disease. Recent in depth review by Yates et al 2103 described the effect of nsSNPs on protein-protein interactions in causing various diseases [2]. Owing to increase efficiency in high-throughput sequencing technologies and other developments such as the human genome project and the 1000 genome project, the number of common mutations seen in humans stands at approximately 67,000 to 200,000. Hence, an experimental analysis of these nsSNPs for their possible impact on the protein function is almost impossible and would be time consuming. Moreover, according to Zhernakova et al. (2009), experimental approaches will have less statistical power to distinguish between disease causing nsSNPs and non disease-causing nsSNPs [3].

Von Hippel-Lindau disease is an inherited genetic syndrome that causes certain benign and malignant tumours in multiple organs of the body [4]. This syndrome is caused by detrimental mutations in the VHL tumour suppressor gene that result in Single Amino Acid Polymorphisms (SAPs). Individuals with these variations are at high risk of retinal angioma, hemangioblastomas of the central nervous system (CNSH), renal cysts and carcinoma (RCC), pancreatic cysts and tumours, pheochromocytoma, endolymphatic sac tumours and/or epididymal cystadenoma. According to Neumann and Weistler (1991), VHL disease can be classified as Type 1 (without pheochromocytoma) and Type 2 (with pheochromocytoma). Type 2 VHL disease can then be further subdivided, as claimed by Brauch et al. (1995), into Type 2A (with pheochromocytoma) and Type 2B (with pheochromocytoma and renal cell carcinoma). Type 2C was given by Hoffman et al. (2001) referring to patients with cases of

isolated pheochromocytoma in the absence of renal cell carcinoma and hemangioblastomas. The $pVHL_{30}$ contains two functional sub-domains; a beta domain comprising the residues 63-154 and residues 193-204 and a helical alpha domain comprising the residues 155-192. (16) Also, two important binding sites within the VHL protein have been identified. One binding site is responsible for Elongin-C binding in the alpha domain between the amino acid residues 157-170. The other binding site is responsible for the binding of HIF1alpha with beta domain and located between the amino acid residues 91- 113. A large proportion of the tumor causing nsSNPs are found to be located in one of these two binding sites. The formation of a complex between pVHL- HIF-1 alpha (Fig 1A) and pVHL-Elongin C (Fig 1B) and the resulting protein-protein interactions are vital for the deregulation of HIF-1 alpha under normoxia conditions to prevent the formation of tumors. The protein complex under study is between three proteins; i.e. pVHL, Elongin C and HIF-1 alpha. There are two functional sub domains within pVHL that contain separate binding sites for both Elongin-C and HIF-1 alpha. The binding site for Elongin-C lies within the alpha domain (155 to 192) of pVHL between the residues 157 to 170 in the amino acid sequence of pVHL. The binding site for HIF-1 alpha lies within the beta domain (63 to 154 and 193 to 204) of pVHL between the residues 91 to 113 in the amino acid sequence. In this study, we have taken VHL as a protein model to decipher the effect of nsSNPs occurring in protein–protein interactions and study their effects on the VHL protein's structure and function using computational tools. In addition, we have looked closer at the dynamics of the interactions involved in the complex formation between the three proteins. With the aid of molecular dynamics simulations, we were able to observe altered stability and flexibility in the VHL-HIF-1alpha and VHL-Elongin-C complexes.

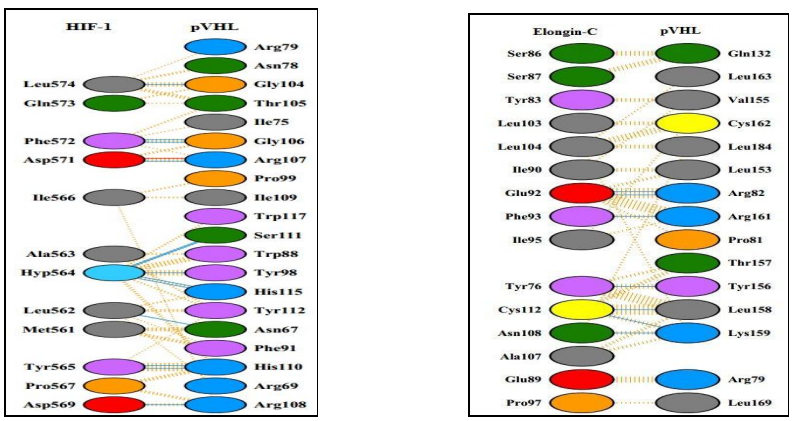
2 Material and Methods

2.1 Obtaining the datasets

For our study, we considered nsSNPs which are involved in the interaction between pVHL- Elongin C and pVHL-HIF-1 alpha. We retrieved 35 nsSNPs from the dbSNP [5] and UniProtKB [6] database that are present in the protein interaction between the two complexes. In the initial screening, 35 SNPs were subjected to various pathogenicity prediction tools. For an analysis on the effect of the variants, we obtained the amino acid sequence of the gene product form UniProtKB; UniProt ID: P40337. For structural analysis PDB structure was retrieved from PDB database with ID 1LM8 [7]. Swiss-Pdb Viewer [8] was used to mutate the pVHL protein structure. Pathogenicity testing of the nsSNPs

2.2 Pathogenicity testing of the nsSNPs

We submitted 35 nsSNPs to four different tools namely SIFT [9], PolyPhen 2.0 [10], PANTHER [11], and PHD-SNP [12] for disease pathogenicity prediction.



(A) **(B)**

Fig. 1. pVHL protein interacting amino acid residue with HIF-1 and Elongin-C. (A) Interacting amino acid residue between pVHL-HIF-1 alpha proteins. (B) Interacting amino acid residue between pVHL-Elongin-C proteins. In both the figures, orange lines show non-bonded contacts and blue color indicate hydrogen bonds.

2.3 Protein Stability Prediction

In order to analyze the impact of the deleterious variants on the stability of pVHL protein, we used a computational tool I-STABLE [13], MUPRO [14], I-Mutant2.0 [15] and SDM [16].

2.4 Molecular Dynamics Analysis

Molecular dynamics simulation was performed by using Gromacs 4.5.3 package [17]. The protein molecule was solvated in a rectangular box with water molecules (TIP3P) at 10 Å marginal radiuses. At physiological pH, the protein structures were found to be positively charged, thus in order to make the system electrically neutral, we added 6 chlorine ions (Cl^-) in the simulation box using the 'genion' tool and the genion tool replaces water molecules by negative ions at the position of the first atoms with the most favourable electrostatic potential or at random position. Emtol convergence criterion was set to 1000 kcal/mol. After this, whole molecular system was subjected to energy minimization by steepest descent algorithm implementing GROMOS96 43a1 force field. SHAKE algorithm was used to constrain bond lengths involving hydrogen, permitting a time step of 2 fs. Van der Waals and coulomb interactions were truncated at 1.0 nm. The non-bonded pair list was updated every 10 steps and conformations were stored every 0.5 ps.

3 Results

3.1 Predicted highly deleterious variants in pVHL gene

In this study, we employed four different *in silico* tools SIFT, PolyPhen 2, PANTHER and PHD-SNP to determine the pathogenic effect of mutations. SIFT predicts whether an amino acid substitution affects the protein function based on sequence homology and the physical properties of amino acid. Twenty three nsSNPs having a tolerance index score of < 0.05 were identified as deleterious. We further analyzed these ten nsSNPs using PolyPhen2 based on empirical derived rules to predict the impact of nsSNP. Thirty four nsSNPs were predicted to be probably damaging the protein structure and function. In order to verify the prediction of SIFT scores, we used HMM, based evolutionary approach PANTHER to validate its impact on protein function upon single point mutation. Out of these thirty five nsSNPs taken for our analysis, twenty four nsSNPs were designated as deleterious. PhD-SNP is a support vector machine (SVM) based computational tool that can predict whether a nsSNP can de disease-related or not using protein sequence information. Among the 35 mutations, PhD-SNP predicted 31 mutations to be deleterious. Comparing the scores of *in silico* tools used in this study, we picked out the variants that were predicted to be deleterious by all the four tools. This way the number of mutations being analysed was brought down to a mere 13 deleterious ones. The complex formation between pVHL, Elongin C and HIF-1 alpha are crucial for the suppression of tumour by the VHL tumour suppressor protein. Any mutations that are present within the binding sites of these proteins would have higher chances of altering the complex formation and are thus clinically more significant. Five nsSNPs were found to occur within the binding site of Elongin C and HIF-1 alpha which defects lead to VHL syndrome. The five nsSNPs S111R, H115Q, W117C, L158V and C162F (Table 1), occur within the binding sites of Elongin C and HIF1-alpha. The nsSNPs L158V and C162F occur within the binding site of Elongin C in the alpha domain of pVHL. This mutation hotspot is designated as "tumorigenic cluster A". The nsSNPs S111R, H115Q and W117C occur in the beta domain of pVHL and this mutation hotspot is designated "tumorigenic cluster B". In tumorigenic cluster B, the mutation S111R occurs within the binding site of pVHL. Further in order to check the stability changes of these five nsSNPs S111R, H115Q, W117C, L158V and C162F, we used I-STABLE, MUPRO, I-Mutant2 and SDM stability changes analysis prediction tools. Among the 5 variants analyzed, all the prediction tools predicted that the variant H115Q, W117C, L158V and C162F decrease the stability of VHL protein. But I-STABLE and MUPRO predicted the variant H115Q increases the stability of protein (Table 2).

Table 1. List of deleterious variants in protein-protein interacting regions related to VHL syndrome and it's SIFT, PolyPhen2, I-Mutant3, PANTHER and PhD-SNP scores.

Variant	Amino acid Position	SIFT	PolyPhen 2	PANTHER	PHD-SNP	Reference
VAR_005716	S111R	0	1	-3.72511	Disease	[18]
VAR_005723	H115Q	0	0.996	-6.61107	Disease	[19]
VAR_005725	W117C	0	0.996	-7.65404	Disease	[18]
VAR_005749	L158V	0	0.99	-3.78806	Disease	[18]
VAR_005754	C162F	0	1	-3.72391	Disease	[18,19]

3.2 Molecular Dynamics simulation analysis on pVHL, Elongin-C and HIF-1 complexes

Proteins are highly dynamic macromolecules as they constantly undergo many conformational changes that allow them to act as transporters, signaling molecules, activators and repressors. The interactions between proteins and hormones, drugs or other proteins are also dynamic in nature [20]. These conformational changes and dynamic interactions are even more important when studying a complex of proteins such as the complex between pVHL, Elongin-C and HIF-1 alpha. Hence, we performed molecular dynamic simulations to study how the dynamics of a mutant protein complex differs from that of a native complex through certain parameters such as Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF) and Hydrogen bonding.

Table 2. List of deleterious variants in protein-protein interacting regions related to VHL syndrome and it's I-STABLE, MUPRO, SDM and I-Muatant2.0 prediction results.

Position	I-Stable	Mupro	I-Muatant2.0	SDM
S111R	Increase	Increase	-0.31	Slightly destabilized- non diseased
H115Q	Decrease	Decrease	-0.34	Highly destabilized- disease associated
W117C	Decrease	Decrease	-1.41	Slightly destabilized- non diseased
L158V	Decrease	Decrease	-1.15	Slightly destabilized- non diseased
C162F	Decrease	Decrease	0.72	Highly destabilized- disease associated

3.3 Root Mean Square Deviation (RMSD) between pVHL and HIF-1 alpha and pVHL and Elongin C

Therefore, RMSD was calculated to study the difference in average distance between the backbone atoms that play a role in the complex formation between the mutant pVHL, in positions S111R, H115Q and W117C (tumorigenic cluster A), and Elongin C. Similarly, the RMSD was calculated for the backbone atoms that play a role in the complex formation between the mutant pVHL in the, positions L158V and C162F (tumorigenic cluster B), and HIF-1alpha. We also calculated the RMSD for all the atoms in the native complex between pVHL-Elongin C and between pVHL- HIF-1 alpha, which was considered as the central criterion to measure the convergence of the protein system concerned. From the calculations, we were able to make a few observations.

From Fig 2A we can observe that there is a notable difference in the RMSD values for the native and mutant complexes between pVHL and HIF-1 alpha in the initial few

picoseconds. An RMSD value of between 0.15nm to 0.30nm is maintained at the beginning. However as the simulation progresses, the complex with the mutation at one hundred and eleventh position (S111R) spikes to an RMSD value of 0.45nm when the other complexes have RMSD values at around 0.35 nm. However, the complex with mutation at one hundred and eleventh position matches with the native complex as the simulation reaches 5000 ps and maintains the same till about 7500 ps with RMSD values between 0.30 nm to 0.45 nm. At this same time, the complex with mutation at position one hundred and fifteenth positions (H115Q) and the complex with mutation at position one hundred and seventeenth positions (W117C) have characteristically lower RMSD values of around 0.20 nm to 0.35 nm. The native and all the mutant complexes have the same RMSD values of between 0.35 nm to 0.45 nm from 10000 ps until the simulation reaches 12500 ps.

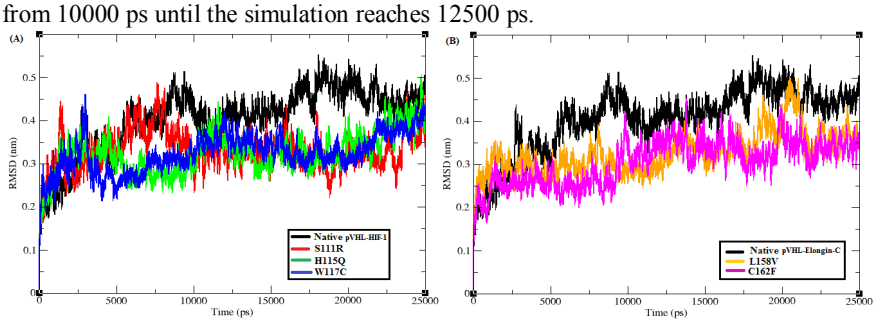


Fig. 2. Backbone RMSD of pVHL in complexes with HIF-1 alpha and Elongin-C. In both figure, the ordinate is RMSD (nm), and the abscissa is time (ps). **(A)** Black, Red, Green and Blue lines show native, S111R, H115Q and W117C structures respectively. **(B)** Black, orange, and pink lines show native, L158V and C162F structures respectively.

However, from 12500 ps, the mutant complexes show a strong deviation from the RMSD values of the native complex. From 12500 ps the mutant complexes show RMSD values between 0.20nm and 0.40nm until about 22500 ps which is deviant from the RMSD values of the native complexes that are between 0.40nm to 0.55 nm. This deviation in the RMSD values starts receding as the simulation reaches 25000 ps. The mutant complexes pVHL and Elongin C shows notable change in the Root Mean Square Deviation calculated in the initial few picoseconds of the simulation. The mutant complexes carry an altered amino acid residue at positions 158 (L158V) and 162 (C162F). As we can see from the Fig 2B, there is no observable change in the RMSD values for the native and mutant complexes. The RMSD maintains within a value between 0.15 to 0.35 nm for up to 2500 ps. Beyond 2500 ps the RMSD of the native complex starts increasing from 0.30 nm to about 0.55 nm till the simulation reaches 25000 ps where the simulation ends. The mutant complexes, on the other hand, show RMSD values which range between 0.20 nm and 0.45 nm. This difference in the RMSD values is also found between the two mutant complexes.
While the mutant complex with an altered amino acid residue at position 158 (L158V) shows a peak in the RMSD value at 20000 ps which equals to nearly 0.45 nm, this peak is not seen in the other mutant complex with altered amino acid residue at posi-

tion 162 (C162F). Such differences in the RMSD values are proof that the dynamics of the protein complex is affected in the mutant complexes when compared to the native complex. The consistent deviations reflect that the mutations could have a terrific impact on the dynamic behavior of the mutant complexes.

3.4 Root Mean Square Fluctuation (RMSF)

To explore the local conformational fluctuation changes between native and mutant protein complexes further, we calculated C-alpha atoms RMSFs (Root Mean Square Fluctuations) from trajectory files obtained from MD simulations. This shows the fluctuation of each amino acid from its time averaged position. The RMSF value of native and mutant models of complexes pVHL and HIF-1 and pVHL and Elongin C was shown in Fig 3A and 3B respectively. The RMSF profile over the simulation period of last 10ns shows fluctuations changes were found in mutant complexes S111R, H115Q and W117C of pVHL and HIF-1protein complex. Large fluctuations from averaged MD conformations were observed in all three mutant complexes S111R, H115Q and W117C in compare with the native complex of pVHL-HIF-1. The pVHL- Elongin-C mutant complexes L158V and C162F also obtained different fluctuations at each amino acid residue point in compare with native pVHL- Elongin-C complex. In particular drastic fluctuation changes were seen in the residue points ~ 50-60, ~200, ~34-350 of the mutant complexes L158V and C162F.

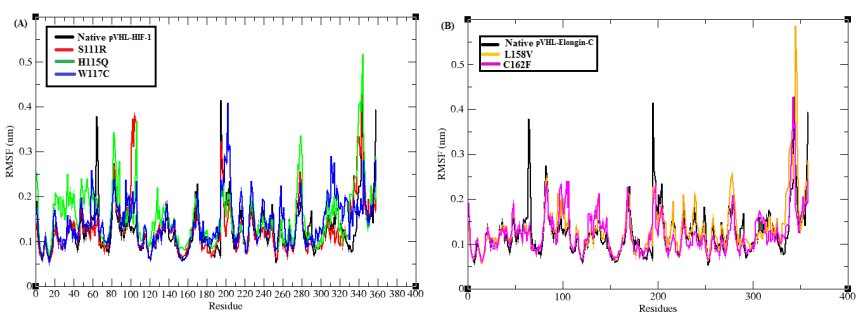


Fig. 3. C-alpha RMSF of pVHL in complexes with HIF-1 alpha and Elongin-C. The ordinate is RMSF (nm), and the abscissa is residues. **(A)** Black, Red, Green and Blue lines show native, S111R, H115Q and W117C structures respectively. **(B)** Black, orange, and pink lines show native, L158V and C162F structures respectively.

3.5 Effect of pVHL mutations in hydrogen bonding and minimum distance changes of pVHL-HIF-1 and pVHL- Elongin-C complexes

Hydrogen bonds analyses are particularly essential in determining binding specificity of protein [21] because they play the most prominent role in biological interaction processes. Fig 4 and 5 depict the number of hydrogen bonds formed between pVHL-HIF-1and pVHL- Elongin-C complexes respectively in both native and mutant form. Between the native complexes of pVHL-HIF-1 ~ 3-14 hydrogen bonds were observed, but between the mutant complexes S111R, H115Q and W117C high number

of hydrogen bonds were observed as ~5-15, ~4-13 and ~5-15 respectively. In another native complex, pVHL- Elongin-C number of hydrogen bonds observed between the molecules as ~3-13. In the mutant complex L158V increase in the number of hydrogen bonds between the molecules were observed as ~4-15 but in another mutant complex C162F decrease in the number of hydrogen bonds were observed between the molecules as ~2-12.

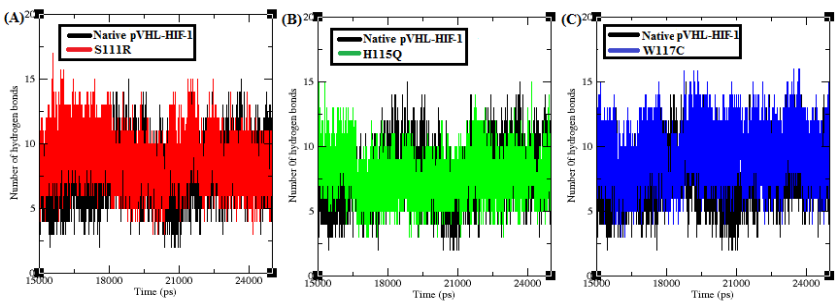


Fig 4. Total number of hydrogen bonds formed between pVHL-HIF-1 alpha complexes in native and mutant state. Black, Red, Green, and Blue lines indicate the hydrogen bonds of native, S111R, H115Q and W117C respectively.

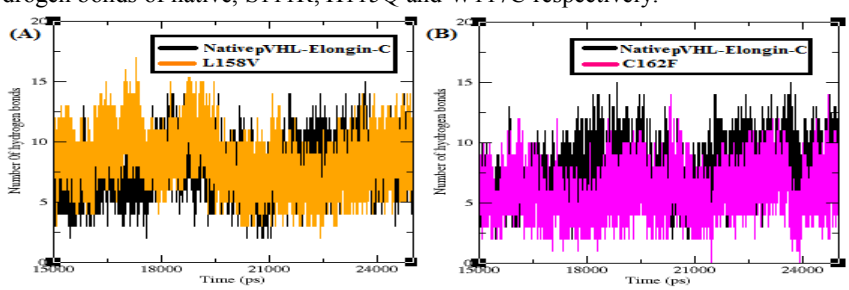


Fig. 5. Total number of hydrogen bonds formed between pVHL-Elongin-C complexes in native and mutant state**.** Black, orange, and pink lines show native, L158V and C162F structures respectively.

3 Discussion

Proteins play a major role in all the activities that transpire within a single cell. Proteins sometimes work alone but most often in concord with other proteins [22] such protein-protein interactions play a central role in most biological processes. Most regulated biological reactions are initiated or inhibited through the interactions between different proteins and their ability to form specific complexes. According to Marcotte et al. (1999), the number of interacting pairs of proteins in Escherichia coli is around 6800 and in Saccharomyces cerevisiae it is way over 45,000. This information is available as a result of completely sequenced genomes of these model organisms. It is an obvious fact that the number of such interacting pairs of proteins would be much more in humans [23]. The association between gene-phenotype and protein interactions is related intimately. The mechanism by which genes cause diseases if

understood reveal important information regarding the interactions that take place between the proteins that they code for. In the same manner, understanding the interactions between proteins and the complexes that they form will help us understand the diseases that their inaction cause [24]. Understanding the physical properties and structural characteristics that allow proteins to bind to one another is at the heart of many challenging problems that include understanding many diseases [25]. In our research on common mutations seen in the VHL gene, we analyzed nsSNPs turned out to be extremely deleterious, namely, S111R, H115Q, W117C, L158V and C162F. Using molecular dynamics simulations we conducted a thorough analysis of infinitesimal fluctuations in the dynamics of the protein complexes formed between pVHL and the mutants in the tumorigenic cluster and tumorigenic cluster B. Molecular dynamics approach was carried out to study the molecular dynamic behavior of five nsSNPs namely S111R, H115Q, W117C, L158V and C162F. In the 25 ns simulation trajectory, different parameters were applied to analyze the level of protein complexes structural changes. Molecular stability and flexibility changes were observed by RMSD and RMSF analysis. RMSD analysis results inferred that the native complex of pVHL- HIF-1-Elongin-C attained maximum deviation of ~0.55 nm, whereas mutant complexes attained less deviation in compare with the native protein complex. We noticed RMSF values for each residue of native pVHL- HIF-1-Elongin-C protein complex. Both high and low peaks were observed at some residue position in mutant complexes in compare with native complex. Fluctuation changes elucidate the impacts of deleterious amino acid substitution in the mutant complexes. Structural mutations were found to affect buried residues in the protein core, causing changes in amino acid size, amino acid charge, hydrogen bonds, salt bridges, S-S bridges [26]. Besides from the different electrostatic interactions, the hydrogen bonds across the binding interfaces and in the protein interiors act as the main contributor to maintaining protein structural stability. Furthermore, incorporation of deleterious nsSNPs may change the original electrostatic formations and distances that could be affect protein native structure. The number of hydrogen bonds formed between pVHL-HIF-1and pVHL- Elongin-C complexes in both native and mutant form was analyzed. In native complexes of pVHL-HIF-1 ~, 3-14 hydrogen bonds were observed, but between the mutant complexes high number of hydrogen bonds was observed. In another native complex, pVHL- Elongin-C number of hydrogen bonds observed between the molecules as ~3-13. In mutant complex L158V number of hydrogen bond increased but in mutant complex C162F decrease in the number of hydrogen bonds was observed. Increase in the number of hydrogen bond increase the affinity between protein and vice versa. All the substituted amino acid improves the affinity between respective proteins except substitution in the C162F complex.

5 Conclusions

Understanding the genotype-phenotype relationship through SNPs is the first and most important step in research and development. Analysing the impact of nsSNPs in protein structures and protein-protein interactions will aid in better understanding the molecular effects of pathogen-city. Experimental methods are required to validate

functional variants in disease association. Laboratory methods are expensive and time consuming and so there is a need for fast and accurate methods for predictions. Above approaches will definitely assist experimental biologists in understanding the disease association of SAPs by mapping as many structural mutations as possible analyze mutations with respect to protein function. Overall, our research pay the way for cost efficient methodology to filter the deleterious variants from neutral once, and their molecular effect in protein-protein interactions by molecular dynamic approach.

Acknowledgements

The authors also take this opportunity to thank the management of VIT and Galgotias University for providing the facilities and encouragement to carry out this work.

References

1. Cargill, M., et al.: Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet 22, 231–238 (1999)
2. Yates, C.M., Sternberg, M.J.: The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions. J Mol Biol. 425(21),3949-63(2013)
3. Zhernakova, A., van Diemen, C.C., and Wijmenga, C.: Detecting shared pathogenesis from the shared genetics of immune-related diseases. Nat Rev Genet. 10, 43–55 (2009)
4. Hes, F.J., Hoppener, J.W.M., Lips, C.J.M.: Pheochromocytoma in von Hippel-Lindau Disease. J. Clin. Endocr Metab. 88, 969–974 (2003)
5. Sherry, S.T., et al.: dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 29(1), 308-311 (2001)
6. Amos, B., Rolf, A.: The SWISS-PROT Protein Sequence Data Bank and Its New Supplement TREMBL. Nucleic Acids Res. 24(1), 21–25 (1996)
7. Min, J.H., Yang, H., Ivan, M., Gertler, F., Kaelin, W.G. Jr., Pavletich, N.P.: Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. Science.7,296(5574)1886-9 (2002)
8. Guex, N., Peitsch, M.C.: SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis. 18, 2714-2723 (1997)
9. Kumar, P., Henikoff, S., Ng, P.C.: SIFT: Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 4, 1073-81 (2009)
10. Ramensky, V., Bork, P., Sunyaev, S.: Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30, 3894-3900 (2002)
11. Mi, H., Guo, N., Kejariwal, A., Thomas, P.D.: PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. Nucleic Acids Res. 35, 247–52 (2007)
12. Capriotti, E., Calabrese, R., Casadio, R.: Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics. 22, 2729-2734 (2006)

13. Chi-Wei Chen1, Jerome Lin1, Yen-Wei Chu1.: IStable: off-the-shelf predictor integration for predicting protein stability changes. BMC Bioinformatics.14,(Suppl 2):S5(2013)
14. Jianlin Cheng, Arlo Randall, and Pierre Baldi, .Prediction of Protein Stability Changes for Single-Site Mutations Using Support Vector Machines .PROTEINS: Structure, Function, and Bioinformatics. 62, 1125–1132 (2006)
15. Emidio Capriotti, Piero Fariselli and Rita Casadio.: I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research. Vol. 33, W306–W310 (2005)
16. Catherine, L., Worth, Robert Preissner and Tom, L., Blundell.: SDM-a server for predicting effects of mutations on protein stability and malfunction. Nucleic Acids Research. Vol.39, Web server issue.W215– W222 (2011)
17. Van Gunsteren,W.F., Billeter, S.R., Eising, A.A., Hunenberger, P.H., Kruger, P.: Bio molecular Simulation: The GROMOS96 Manual and User Guide vdf Hochschulverlag AG an der ETH Zurich and BIOMOS b.v: Zurich, Groningen(1996)
18. Zbar, B., Kishida, T., Chen, F., Schmidt, L., Maher, E.R., Richards, F,M., Crossey, P.A., Webster, A.R., Affara, N.A., Ferguson-Smith, M.A., Brauch, H., Glavac, D., Neumann, H.P., Tisherman, S., Mulvihill, J.J., Gross, D.J., Shuin, T., Whaley, .J, Seizinger, B., Kley, N., Olschwang, S., Boisson, C., Richard, S., Lips, C.H., Lerman, M., et al.: Germline mutations in the Von Hippel-Lindau disease (VHL) gene in families from North America, Europe, and Japan. Hum Mutat. 8(4),348-57(1996)
19. Olschwang, S., Richard, S., Boisson, C., Giraud, S., Laurent-Puig, P., Resche, F., Thomas, G.: Germline mutation profile of the VHL gene in von Hippel-Lindau disease and in sporadic hemangioblastoma. Hum Mutat.12(6),424-30(1998)
20. Gergely Losonczy, et al.: Three novel germ-line VHL mutations in Hungarian von Hippel-Lindau patients, including a nonsense mutation in a fifteen-year-old boy with renal cell carcinoma. BMC Medical Genetics. 14:3 (2013)
21. Fersht, A.R.: Basis of biological specificity. Trends Biochem Sci. 9,145-147 (1984)
22. Shrake, A., Rupley, J.A.: Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J. Mol. Biol. 79,351-371(1997)
23. Tanja Kortemme and David Baker.: A simple physical model for binding energy hot spots in protein–protein complexes. PNAS ,vol. 99, pp.14116–14121(2002)
24. Marcotte, E.M., et al.: Detecting protein function and protein-protein interactions from genome sequences. Science. 285(5428),751-3 (1999)
25. Maricel, G.: Protein interactions and disease: computational approaches to uncover the etiology of diseases. Briefings in Bioinformatics. 5, 333-346 (2007)
26. Jiajia Chen and Bairong Shen.: Computational Analysis of Amino Acid Mutation: A Proteome Wide Perspective. Current Proteomics. 6, 4 (2009)