Biochemistry and Biophysics Reports 10 (2017) 260–266

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Prediction of a highly deleterious mutation E17K in AKT-1 gene: An *in silico* approach

Imran Khan, Irfan A. Ansari*

Department of Biosciences, Integral University, Lucknow, INDIA

ARTICLE INFO

Keywords:
AKT1
Signaling pathways
SNP
in silico
Deleterious

ABSTRACT

The AKT1 (v-akt murine thymoma viral oncogene homologue 1) kinase is a member of most frequently activated proliferation and survival signaling pathway in cancer. Recently, hyperactivation of AKT1, due to functional point mutation in the pleckstrin homology (PH) domain of AKT1 gene, has been found to be associated with human colorectal, breast and ovarian cancer. Thus, considering its crucial role in cellular signaling pathway, a functional analysis of missense mutations of AKT1 gene was undertaken in this study. Twenty nine nsSNPs (non-synonymous single nucleotide polymorphism) within coding region of AKT1 gene were selected for our investigation and six SNPs were found to be deleterious by combinatorial predictions of various computational tools. RMSD values were calculated for the mutant models which predicted four substitutions (E17K, E319G, D32E and A255T) to be highly deleterious. The insight of the structural attribute was gained through analysis of, secondary structures, solvent accessibility and intermolecular hydrogen bond analysis which confirmed one missense mutation (E17K) to be highly deleterious nsSNPs. In conclusion, the investigated gene AKT1 has twenty nine SNPs in the coding region and through progressive analysis using different bioinformatics tools one highly deleterious SNP with rs121434592 was profiled. Thus, results of this study can pave a new platform to sort nsSNPs for several important regulatory genes that can be undertaken for the confirmation of their phenotype and their correlation with diseased status in case control studies.

1. Introduction

Akt, also termed protein kinase B (PKB), is evolutionary conserved serine/threonine protein kinase and constitutes fundamental intracellular signaling systems for the regulation of multiple cellular and physiological activities, such as cell growth, proliferation, protection from apoptosis, modulation of DNA damage response and genome stability, motility, angiogenesis, and metabolism [12,19,49]. In mammals, the Akt protein kinase is expressed in three isoforms: Akt1, Akt2, and Akt3, which are also termed PKBα, PKBβ, and PKBγ, respectively [5,14]. Moreover, they are encoded on three distinct chromosomes, share a considerable homology, and contain three common structures: the N-terminal pleckstrin homology (PH) domain, the catalytic kinase domain (KD), and the C-terminal regulatory hydrophobic motif (HM) [51]. The catalytic and regulatory domains are both critical for the biological actions mediated by Akt protein kinases [23,40].

The AKT-1 (v-akt murine thymoma viral oncogene homologue 1) kinase is a member of possibly the most frequently activated proliferation and survival signaling pathway in cancer [34,53,59]. The activation of AKT1 is driven by membrane localization, which is in turn initiated by the binding of the pleckstrin homology (PH) domain to phosphatidylinositol-3,4,5-trisphosphate or phosphatidylinositol-3,4-bisphosphate, followed by phosphorylation of the regulatory amino acids serine 473 (Ser 473) and threonine 308 (Thr 308) [30,58].

Hyperactivation of AKT protein, resulting in deregulation of the PI3K/AKT pathway, has been found to be associated with several human cancers [11,21,36,38,41,56]. The AKT1 gene is composed of 14 exons and encodes 480-amino acid protein (Akt1). Recently, functional point mutation in the PH domain of AKT1 gene in human colorectal, breast and ovarian cancer has been found to be associated with AKT1 hyperactivation and tumor formation [7,35]. Thus, considering the importance of AKT1 in different cellular signaling pathways implicated in human cancers, identifying and profiling functional missense mutations could be a significant approach in designing better diagnostic and therapeutic approaches.

Non-synonymous single nucleotide polymorphisms (nsSNPs) are the simplest form of genetic variations occurring in coding region that alter the encoded amino acid [22]. With recent advances in high-throughput sequencing technologies, a huge amount of genetic variations has been identified and deposited in databases. The identification of pathological

Abbreviation: SNP, Single nucleotide polymorphism; RMSD, Root mean square deviation

* Corresponding author.
E-mail addresses: iaansari@iul.ac.in, ahmadirfan.amu@gmail.com (I.A. Ansari).

http://dx.doi.org/10.1016/j.bbrep.2017.04.013
Received 24 September 2016; Received in revised form 28 January 2017; Accepted 19 April 2017
Available online 21 April 2017

genetic variants may prove crucial in gene level studies and future target based therapies or even personalized medicine. Furthermore, efficient and effective identification of each of these functionally significant nsSNPs associated with the diseased status is a time consuming as well as an expensive task. For this purpose, various bioinformatics tools, designed on the basis of recent findings in protein structure research and evolutionary biology, may prove useful in predicting the functional importance of nsSNPs [6,13,24,37] (Conde et al., 2006). Over past few years, several in silico studies have attempted to screen missense/nsSNPs within the protein coding region of a gene and have shown these bioinformatics tools to be efficient and effective platform to prioritize SNPs for their association in disease pathology [18,20,25–28,31]. Over past few years, several in silico studies have attempted to screen missense/non-synonymous single nucleotide polymorphisms (nsSNPs) within the protein coding region of a gene and have shown these bioinformatics tools to be efficient and effective platform to prioritize SNPs for their association in disease pathology [18,20,25–28,45]. These nsSNPs within the coding region alter the encoded amino acid and further resulting in altered physiochemical properties of native protein [43,44,61,67]. Recent advancements in the field of computational biology facilitates the screening of deleterious nsSNPs through considering various aspects like sequence conservation among various species, structural features and physiochemical properties of protein [48,60,66]. So, we employed two diverse approaches for the analysis of deleterious nsSNPs through empirical based methods and support vector based approaches to profile deleterious nsSNPs of AKT-1 gene. Thus, the aim of the study is to identify nsSNPs for AKT-1 gene which are likely to alter the structural and functional aspects of the AKT-1 protein and to pave a effective method to classify the disease associated amino acid variants in several other crucial cellular proteins.

2. Materials and methods

2.1. SNP data retrival

The SNPs were retrieved from SNP database of National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/snp) using various limits of *Homo sapiens*, coding non-synonymous, stop gained, coding synonymous, mRNA UTR (5' and 3') and intronic region [52].

2.1.1. SIFT

SIFT (Sorting Intolerant from Tolerant) (http://sift.jcvi.org/www/SIFT_BLink_submit.html) uses sequence homology based algorithm to classify the effect of amino acid substitutions on protein function. We performed updated version of SIFT called SIFT-Blink by submitting query in the form of amino acid substitutions obtained and the gene identification number from NCBI. The SIFT predictions are provided as a normalized probability score chart for all 20 amino acids. SIFT scores less than 0.05 is predicted as deleterious [24,39].

2.1.2. PolyPhen 2

PolyPhen-2 (Polymorphism Phenotyping v2), (http://genetics.bwh.harvard.edu/pph2/) is an automated structure homology based tool, that predicts the structural and functional impact of amino acid substitution on the human protein. PolyPhen-2 characterizes an amino acid substitution through gathering information on sequence, structural and phylogenetic features of the query gene [1,46]. We submitted the query in the form protein FASTA sequence with mutational positions each with two amino acids variants. A position-specific independent counts (PSIC) score and their difference is generated for two variants. The PSIC score difference of 1.5 or above is characterized as damaging. The PolyPhen scores are classified as probably damaging (2.00), possibly damaging (1.50–1.99), potentially damaging (1.25–1.49), or benign (0.00–0.99).

2.1.3. I-Mutant 2.0

I-Mutant2.0 (http://folding.uib.es/cgi-bin/i-mutant2.0.cgi) is a support vector machine (SVM)-based web server for the automatic prediction of protein stability changes upon single point mutations. I-Mutant 2.0 predicts the free energy change (DDG) calculated by subtracting the free energy change of the mutant protein from the free energy change of the native protein (Kcal/mol) [4]. We provided query protein FASTA sequence along with amino acid substitutions and position and free energy change (DDG) were obtained. A zero value of DDG predicts high stability of mutant protein and more negative the value lesser the stability of the mutant protein.

2.1.4. PROVEAN

PROVEAN (Protein Variation Effect Analyzer), (http://provean.jcvi.org) is a bioinformatics tool which predicts functional effects on protein's biological function including insertions, deletions, and multiple amino acid substitutions [9]. This tool provide a score called PROVEAN score which shows difference between the deleterious and the neutral effects on protein's biological function. The approach utilizes pairwise sequence alignment scores to generate pre-computed prediction at every position of amino acid in human protein sequences. The input given is the FASTA sequence of query protein and the effects are predicted on the basis of threshold value table obtained. The analysis predicts the SNPs to be deleterious or tolerant based on their default threshold value of 0.1. If the PROVEAN score is smaller than or equal to a given threshold, the variation is predicted as deleterious.

2.2. 3D-Structure modeling and RMSD calculation

To evaluate the structural stability of the protein upon substitution, the X-ray crystallographic 3-D structure of AKT-1 protein (pdb id-3O96) was downloaded from the UniProt, further to compare and validate the results 3 well documented mutations were selected as positive controls for which we selected phosphatidylinositol 3,4,5-triphosphate 3-phosphatase (PTEN) and serine/threonine-protein kinase B-raf (BRAF) and 3D structures were downloaded having pdb id 5BUG and 5CSX respectively [64]. The structure was validated using PROCHECK (http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) [31,32]. The mutant protein models were generated using SWISS PDB Viewer [16]. Often homology modeling and mutant protein modeling results in generation of new loops and deletion of some loops in the protein structure. Thus to bring the proteins to its most favorable conformation energy minimization is required. Energy minimizations for the native and the mutant structures performed by NOMAD-Ref (http://lorentz.immstr.pasteur.fr/nomad-ref.php) which uses Gromacs algorithm by default [33]. Finally the difference in the total energy and Root Mean Square Deviation (RMSD) values were calculated via SWISS PDB viewer between the native and mutants protein structures.

2.3. Structural analysis

2.3.1. SPPIDER

SPPIDER Solvent accessibility based Protein-Protein Interface iDEntification and Recognition (http://sppider.cchmc.org/) is a web based server to predict the solvent accessibility and secondary structures in the 3D structure of proteins [42]. The server employs Polyview3D – Dictionary of Secondary Structure Protein (DSSP) for the prediction of secondary structure and solvent accessibilities (SABLE) for solvent accessibility.

2.3.2. Hydrogen Bond Analysis Tool (HBAT)

HBAT is an automated tool for the analysis of protein structure PDB files for all non-bonded interactions. The compactness of HBAT in delivering information on all interactions in a protein structure in a user friendly way is an advantage over other available tools. The program requires no additional information on PDB structures with bound

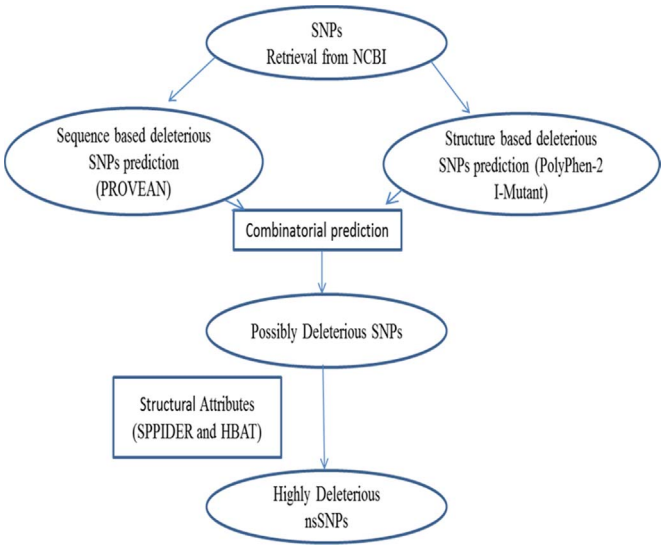

Fig. 1. Schematic representation of methodology.

ligands [57].

3. Results

3.1. DATA mining

For the current study, we have selected SNPs of AKT-1 gene for *Homo sapiens* using various limits of coding non-synonymous (nsSNPs) region, coding synonymous (sSNPs) region, stop-gain, mRNA untranslated regions UTR (5' and 3') and intronic regions. Out of 786SNPs, coding region contains 29 nsSNPs (3.74%) and 55 sSNPs (7.10%); stop-gain region contains 1 SNPs (0.13%); and non-coding regions contain 633 SNPs (81.78%) in intronic region and 68 SNPs (4.59%) in mRNA UTR region with 25 SNPs in 5' UTR and 43 SNPs in 3' UTR. Since, vast majority of SNPs were found to be in intronic region, therefore, SNPs within coding non-synonymous region or regulatory region (29 nsSNPs) of AKT-1 gene were selected for our investigation. The functional impact of these sorted nsSNPs was further assessed by various bioinformatics tools, schematic representation of methodology shown in Fig 1.

3.2. Analysis of deleterious nsSNPs using SIFT blink

On the basis of sequence homology and the physical properties of amino acids, SIFT predicts whether the amino acid substitution alter the protein function. Protein sequence with mutational positions and amino acid residues were submitted to SIFT Blink server as input. The intolerant range of SIFT is ≤ 0.05 which indicates that nsSNP is more damaging/deleterious to the protein function and score of > 0.05 predicts the tolerant range. Out of 29 nsSNPs, 11nsSNPs (37.93%) were predicted to be intolerant with 0.00scores; 7nsSNPs (24.13%) showed the score ranging from 0.01 to 0.05; and 11 nsSNPs (37.93%) were found to be tolerant with score ranging from 0.06 to 0.10. Thus, total 18nsSNPs (62.06%) were predicted to be intolerant that may alter functional properties of protein, shown in Fig. 2. These scores enable the quantitative comparison and ranking the nsSNPs according to their deleterious nature and allow researchers to decide which SNP to be targeted for further investigation.

3.3. Analysis of deleterious nsSNPs using PolyPhen-2

The PolyPhen prediction is based on the number of sequences, phylogenetic and structural features characterizing the amino acid substitution. Same protein sequence with mutational positions and amino acid substitutions, submitted to SIFT as input, were also submitted to Polyphen 2 (i.e. 29 nsSNPs). PolyPhen scores comprise a range from zero to a positive number, where zero indicates a neutral effect on amino acid substitution and higher the positivity, higher will be the detrimental effect of substitution on protein function. Out of 29 nsSNPs, 9nsSNPs (31.03%) were predicted to be probably damaging depicting score values ranging from 0.8 to 1; 3nsSNPs (10.34%) were found to be possibly damaging having score values 0.6–0.8; and 17nsSNPs (58.62%) were identified as benign (neutral) showing score value zero, shown in Fig. 2. Thus PolyPhen score is useful in quantitative characterization of the damaging effects of nsSNPs on protein function.

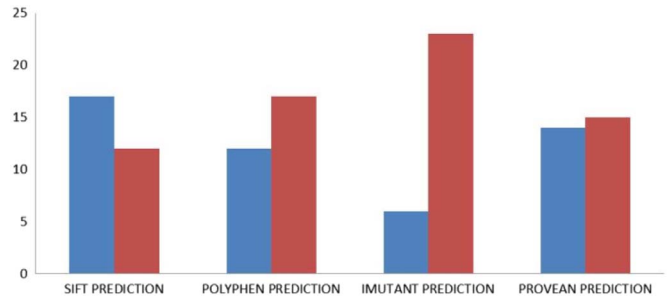

Fig. 2. Prediction of deleterious nsSNPs of AKT-1 gene by SIFT Blink, PolyPhen 2.0, I-Mutant 2.0 and PPOVEAN. The bar diagram indicates the number of deleterious and benign nsSNPs predicted by various tools, the blue bar indicates the deleterious nsSNPs and the red bar indicates the benign nsSNPs.

3.4. Analysis of deleterious nsSNPs using I-Mutant 2.0

The query protein sequence along with mutational positions and amino acid substitutions, submitted to SIFT Blink and PolyPhen 2.0, were also submitted as input to I-Mutant 2.0. I-Mutant 2.0 predicts the alterations in protein stability due to presence of single point mutations. According to I-Mutant 2.0, more negative the value of free energy change (DDG), lesser will be the stability of protein. According to these score, out of 29 nsSNPs, 6 variants P388T, E319G, A188T, A399T, R222H and T479Mshowed DDG values -2.26, -2.08, -3.01, -3.2, -2.26 and -2.81respectively which were considered to be least stable and most deleterious nsSNPs. The other 11 variants E17K, L357P, V167A, V461L, D3N, R406H, G410S, D46E, R69Q, D32E, R15Q, R367H, D262E, M363V, E375K and A255Tshowing DDG values ranging from -0.05 to -1.9 were further identified as less stable and deleterious nsSNPs. The remaining 12 variants showing DDG values ≥ 0.00 were grouped as non-deleterious. Thus, total 17nsSNPs/variants (58.62%) were predicted to be deleterious to the protein stability, shown in Fig. 2.

3.5. Analysis of deleterious nsSNPs using provean

Out of 29 nsSNPs analyzed via PROVEAN, 14 nsSNPs were predicted to be deleterious. The other 15 nsSNPs were categorized as neutral by PROVEAN. Thus, total 48.27% nsSNP/variants were predicted to be deleterious.

3.6. Combinatorial prediction of functionally deleterious nsSNPS

As each tool works on a specific algorithm which is based on variable biological aspect; therefore, through combining different computational tools, prediction of deleterious nsSNPs can be performed more accurately [15,65]. Thus to reduce the false positive predictions, we opted for a combinatorial approach involving empirical and support vector machine (SVM) based tools and selecting only those SNPs for further trajectory analysis which were commonly predicted to be deleterious by all the tools. Furthermore, to be more informative in

Table 1
Combined prediction of SIFT Blink, PolyPhen 2.0, I-Mutant 2.0 and PROVEAN for possible deleterious nsSNPs of AKT-1 gene.

S. No.	rs ID	Allele	Residue change	SIFT prediction	PolyPhen score	I-Mutant DDG	PROVEAN
1	rs121434592	G/A	E17K	0	1	−1.03	Deleterious
2	rs11555431	C/A	P388T	0	0.977	−2.26	Deleterious
3	rs12881616	G/G	E319G	0	0.993	−2.08	Deleterious
4	rs188580689	G/A	A399T	0.01	0.842	−3.2	Deleterious
5	rs201636005	T/G	D32E	0	0.96	−0.87	Deleterious
6	rs375395037	G/A	A255T	0	0.989	−0.45	Deleterious
7	PTEN	G/T	R159S	0	0.986	−3.56	Deleterious
8	BRAF	G/T	G466V	0	0.979	−3.45	Deleterious

comparing the results of all the tools, we included two documented mutations in colorectal cancer (R159S in PTEN gene and G466V in BRAF gene) as positive controls for our analysis [8,54]. By comparing the result of SIFT Blink, PolyPhen-2 and I-Mutant 2.0 tools, 6 variants E17K, P388T, E319G, A399T, D32E and A255T having SNP IDs rs121434592, rs11555431, rs12881616, rs188580689, rs201636005 and rs375395037 respectively, were predicted to be functionally deleterious by all four tools. Interestingly, all the control mutations were also predicted to be deleterious by all the tools used. Table 1 displays the predicted score of SIFT Blink, PolyPhen 2.0 and I-Mutant 2.0. These 6 nsSNPs (E17K, P388T, E319G, A399T, D32E and A255T) were selected for further trajectory analysis.

3.7. Modeling of mutant protein structures and RMSD calculation

Protein stability can be significantly altered by the presence of SNP. Thus, for better insight on single nucleotide change and its role in protein functional and structural stability, it is mandatory to gather knowledge about 3D structure. The 3D structure of AKT-1 protein (ID-3O96) was obtained from Universal Protein Resource (UniProt) having 480 amino acid sequence and 367 amino acid residues in structure. Later PROCHECK server was used to validate the native structure, the structure had resolution of 2 A° and 87% of residues were found in most favored region in the ramachandran plot. Out of six, four mutant protein models with nsSNPs (E17K, E319G, D32E and A255T) were generated using SWISSPDB Viewer, shown in Table 2. The nsSNPs P388T and A399T were not modeled due to unavailable amino acid residues in any available protein structure in UniProt. Thus for further analysis three-dimensional structures were energy minimized using NOMAD-Ref server. RMSD values were calculated for all 4 mutant models by comparing the energy minimized structures of native and mutant proteins as shown in Table 2. All mutant models showed minor deviations in RMSD values from native structure. It is reported that higher the RMSD values grater is the deviation between mutant and native structure [17]. Since the deviation was very small we also under took the total energy change in mutant models from native structure. The total energy of native model was −2181.194 KJ/mol and when compared with mutant models notable amount of variation was observed. Mutant A255T and E319G showed decreased total energy −1554.896 KJ/mol and −1905.074 KJ/mol respectively. The mutant E17K accounted very high total energy value of −6981.152 KJ/mol, further D32E mutant model showed very minor change in total energy change of −2407.920 KJ/mol. Thus, 3 nsSNPs (E17K, E319G and A255T) were selected for further trajectory analysis to investigate precisely the deleterious effect of these highly deleterious nsSNPs on the 3D AKT-1 protein structure and function.

Table 2
Root mean square deviation (RMSD) and total energy deviation of mutant proteins from native AKT-1 protein.

			Total Energy (KJ/mol)	
S. No.	Residue change	Backbone RMSD (Å)	Mutant	Native
1	Native			−2181.194
2	A255T	0.064	−1554.896	
3	D32E	0.128	−2407.920	
4	E17K	0.029	-6981.152	
5	E319G	0.107	−1905.074	

3.8. Structural investigation

3.8.1. SPPIDER

For further analysis of secondary structures and solvent accessibility, all 3 mutant proteins models (E17K, E319G and A255T) with slight deviation in RMSD values and total energy, were analyzed and compared with native models shown in Table 3. SPPIDER server predicted 1 nsSNP (E17K) with variation in the solvent accessibility from buried to exposed, thus increasing the solvent accessibility of the mutant model. Further the results of PTEN R159S and BRAF G466V used as positive control also showed variation in SPPIDER results, thus supporting the authenticity of SPPIDER tool.

3.8.2. Hydrogen Bond Analysis Tool (HBAT)

The analysis of intermolecular hydrogen bond frequencies for E17K nsSNP was calculated by HBAT. A countable change was noticed in the highly deleterious nsSNP E17K model compared to native protein structure shown in Table 4. Mutant protein model E17K showed decrease in the strong H-bonds like N-H…O, O-H…O, N-H…N and O-H…N. There was also variation in frequency of some weak bonds. These results support the previous trajectory analysis, that the predicted nsSNP (E17K) is possibly highly deleterious. Similar to SPPIDER analysis, HBAT demonstrated notable variation in both the control mutations (PTEN R159S and BRAF G466V). The investigation reveals that this nsSNP was the most possible highly deleterious and may affect the 3D-structure of the AKT-1 protein.

4. Discussion

A disease gene variant manifests in a disease state via complex

Table 3
Comparative analysis of solvent accessibility (SABLE) and secondary structures of native and mutant AKT-1 protein predicted by SPPIDER.

SPPIDER					
		Secondary Structure-DSSP		Solvent accessibility – SABLE	
S. No.	Residue change	Native	Mutant	Native	Mutant
1	A255T	H	H	0	0
2	E17K	C	C	3	4
3	E319G	H	H	1	1
4	R159S	H	C	1	1
5	G466V	C	C	0	1

Note: [H- Helix, C-Coil, E-beta Strand], [9-Fully Exposed; 0 – Fully Buried].

Table 4
Intermolecular Hydrogen Bond frequency prediction of highly deleterious nsSNPs.

HBAT										
S. No.	Residue change	N-H.. O	O-H.. O	N-H.. N	O-H.. N	C-H.. O	C-H.. N	N-H.. S	O-H.. S	C-H.. S
1	Native	401	36	194	14	60	18	0	0	2
2	E17K	384	13	194	4	56	20	0	0	3
3	Native	1306	76	590	20	232	66	7	0	13
4	R159S	1299	76	589	20	232	66	7	0	13
5	Native	289	9	167	8	27	14	0	0	0
6	G466V	273	9	163	5	17	14	0	0	0

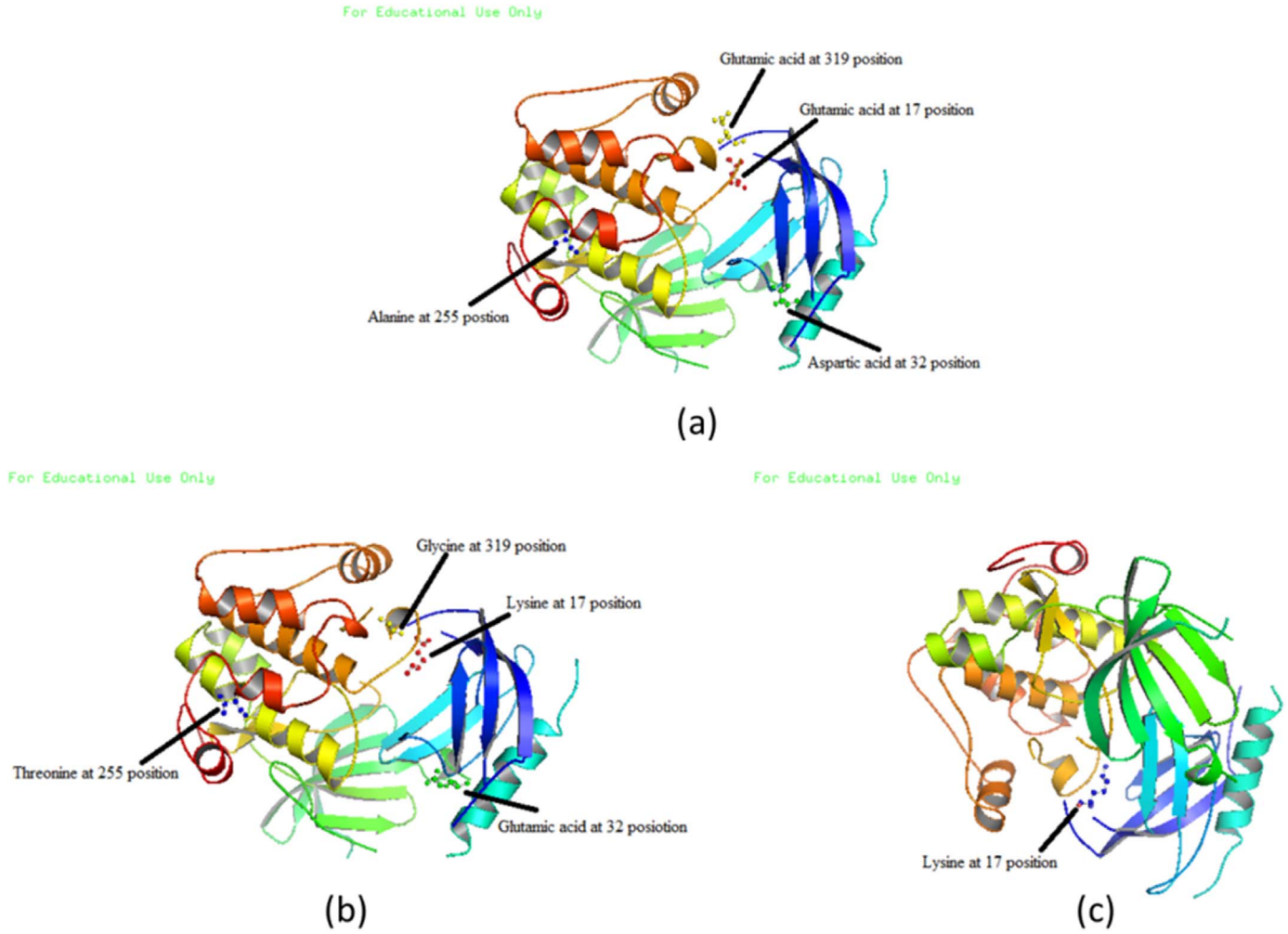

Fig. 3. 3D structure of the native and mutant AKT-1 protein models (a) Native, (b) Mutant model depicting E17K, E319G, D32E and A255T (c) Mutant E17K.

association of patient genetic background, environmental factors and stochastic influences. The clinical appearance of a mutation is also an important variable among different individuals [62]. These limitations in predicting the genotype-phenotype correlation with disease-association makes it costly and highly challenging [29]. To overcome above restrictions, recently bioinformatics tools have emerged as a valuable option for mutation analysis [18,20,25–28,45,47]. Aberrant activation of AKT-1 has been implicated in various human cancers and is also associated with drug resistance [2,55,63]. Genetic alterations in AKT1 gene have been demonstrated in many tumors including colon, breast and ovarian cancer [2,63]. Key regulatory mutations in AKT1 gene has not been widely reported in cancer. Thus the current study was focused on profiling the deleterious nsSNPs of AKT1 gene associated with various cancers. Initially, the SNP data was extracted from the NCBI database and nsSNPs were sorted for further computational analysis by different bioinformatics tools. It is well documented that using multiple tools and algorithms for prioritizing functional mutations enhances the accuracy of prediction [15,65]. Thus, we employed various tools like SIFT Blink, PolyPhen-2, I-Mutant 2.0 and PANTHER to predict the functional phenotypes of nsSNPs based on protein sequence cross species conservation, protein structure, and physicochemical properties [48,60,66]. To dismiss any possible false positive results, we incorporated a combinatorial approach by comparing results of all the tools. Thus, we selected only 6 nsSNPs (E17K, P388T, E319G, A399T, D32E, and A255T) out of all 29 nsSNPs for further investigation which were predicted as deleterious by all selected tools. To further improve our analysis and gain better understanding of functional impact of nucleotide substitutions in a gene sequence, we downloaded the template structure 3O96 of AKT1 protein from UniProt and checked for the stability via PROCHECK server and the structure was found to be a good

quality model having 90% residues in most favorable region A, B, L on ramachandran plot. To detect the degree of variation in arrangement of atoms in the native protein structure, we compared the RMSD value and total energy of native and mutant models (E17K, E319G, D32E, and A255T). Consecutively, it was also noticed that the total energy of these models showed deviation from native AKT1 structure. These energy calculations provide information on the stability of a protein structure. All 4 deleterious SNP models (E17K, E319G, D32E, and A255T) showed a very small deviation in RMSD values but when analyzed for total energy of the mutant models compared to native protein structure 3 mutant models (A255T, E17K and E319G) accounted notable variation. Thus, these 3 nsSNP models were subjected to further trajectory analysis. This involved the identification of the solvent accessibility and secondary structures prediction of selected mutant AKT1 models having highly deleterious nsSNPs (A255T, E17K and E319G). A further deeper insight of the solvent accessibility and secondary structures of these mutant AKT1models were gained from SPPIDER. This tool identified one mutant AKT1 model having E17K nsSNP with deviated accessible surface area from buried to exposed, thus increasing the solvent accessibility of the mutant model. Amino acid accessible surface area is a key functional property of a protein that also influences the DNA binding property of a protein [3,50]. We further analyzed the hydrogen bonds in the native and mutant AKT1 model with E17K deleterious SNP. The results showed a significant decrease in the number of hydrogen bonds in the structure. Thus, the trajectory analysis revealed the deleterious nature of the E17K (rs121434592) nsSNP in AKT1 gene. The 3D structures of the native and the mutant AKT1 protein were drawn through PyMol molecular graphics system providing a clear image of the residues changes in Fig. 3 [10].

Recently, the substitution of glutamic acid to lysine at amino acid 17 (E17K) in the lipid binding pocket of the PH domain of AKT1 has been reported to be associated with human colorectal cancer, breast cancer and ovarian cancers [7,35]. Thus, results of our study were found to be in accordance with the recent wet lab reports which predicted E17K as a highly deleterious nsSNP in various cancers. However, prediction of SNPs via *in silico* methods can only be substantiated by further epidemiological studies in wet lab experiments. These *in silico* studies can only facilitate the researchers to narrow down the cumbersome procedure to study each and every SNP present in a particular gene.

5. Conclusion

In conclusion, this study attempted to prioritize high risk missense mutation in the coding region of AKT-1 gene using various bioinformatics tools. The finding of the study predicted E17K, P388T, E319G, A399T, D32E and A255T SNPs as deleterious via combinatorial approach. Furthermore, systematic structural investigation leads to finding E17K as highly deleterious SNP. Although not much insight on the pathological outcomes and activitites of SNP have been gained, but the results strongly suggest that it can act as crucial variable in providing stability to structural and functional attributes of AKT-1 protein. The study overlay an alternative approach which is inexpensive and less time consuming to prioritize amino acid substitutions and short list candidate SNPs for other genes involved in regulatory cellular pathways that can be undertaken for further epidemiological studies to assess their phenotypic association and correlation with different human cancers.

Acknowledgments

The authors thank the management of the Integral University for providing the facilities to carry out this study.

References

[1] I.A. Adzhubei, S. Schmidt, L. Peshkin, V.E. Ramensky, A. Gerasimova, P. Bork, A.S. Kondrashov, S.R. Sunyaev, A method and server for predicting damaging missense mutations, Nat. Methods 7 (2010) 248–249.
[2] D.A. Altomare, J.R. Testa, Perturbations of the AKTsignaling pathway in human cancer, Oncogene 24 (2005) 7455–7764.
[3] G.J. Bartlett, C.T. Porter, N. Borkakoti, J.M. Thornton, Analysis of catalytic residues in enzyme active sites, J. Mol. Biol. 324 (2002) 105–121.
[4] K.A. Bava, M.M. Gromiha, H. Uedaira, K. Kitajima, A. Sarai, ProTherm, version 4.0: thermodynamic database for proteins and mutants, Nucl. Acids. Res. 32 (2004) D120–D121.
[5] J.M. Beaulieu, T. Del'guidice, T.D. Sotnikova, M. Lemasson, R.R. Gainetdinov, Beyond cAMP: the regulation of Akt and GSK3 by dopamine receptors, Front. Mol. Neurosci. 4 (2011) 38.
[6] Y. Bromberg, B. Rost, SNAP: predict effect of nonsynonymous polymorphisms on function, Nucleic Acids Res. 35 (2007) 3823–3835.
[7] J.D. Carpten, A.L. Faber, C. Horn, G.P. Donoho, S.L. Briggs, C.M. Robbins, G. Hostetter, S. Boguslawski, T.Y. Moses, S. Savage, M. Uhlik, A. Lin, J. Du, Y.W. Qian, D.J. Zeckner, G. Tucker-Kellogg, J. Touchman, K. Patel, S. Mousses, M. Bittner, R. Schevitz, M.H. Lai, K.L. Blanchard, J.E. Thomas, A transforming mutation in the pleckstrin homology domain of AKT1 in cancer, Nature 448 (2007) 439–444.
[8] E. Chan, PTENc. 477G > T (R159S) Mutations in colorectal cancer. My Cancer Genome ⟨https://www.mycancergenome.org/content/disease/colorectal-cancer/pten/24/⟩ (Updated March 6), 2015.
[9] Y. Choi, G.E. Sims, S. Murphy, J.R. Miller, A.P. Chan, Predicting the functional effect of amino acid substitutions and indels, PLoS One 7 (2012) e46688.
[10] W.L. DeLano, The PyMOL molecular graphics system, Version 1.3r1, DeLano Scientific LLC, South San Carlos, CA, 2010.
[11] S. Dimmeler, I. Fleming, B. Fisslthaler, C. Hermann, R. Busse, A.M. Zeiher, Activation of nitric oxide synthase in endothelial cells by Akt-dependent phosphorylation, Nature 399 (1999) 601–605.
[12] J. Downward, PI 3-kinase, Akt and cell survival, Semin. Cell Dev. Biol. 15 (2) (2004) 177–182.
[13] M.Y. Frederic, M. Lalande, C. Boileau, D. Hamroun, M. Claustres, et al., UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity -application to four genes: FBN1, FBN2, TGFBR1, and TGFBR2, Hum. Mutat. 30 (2009) 952–959.
[14] E. Gonzalez, T.E. McGraw, The Akt kinases: isoform specificity in metabolism and cancer, Cell Cycle 8 (16) (2009) 2502–2508.
[15] G. Grillo, A. Turi, F. Licciulli, F. Mignone, S. Liuni, S. Banfi, V.A. Gennarino, D.S. Horner, G. Pavesi, E. Picardi, G. Pesole, UTRdb and UTRsite (RELEASE 2010) a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs, Nucl. Acids Res. 38 (2010) D75–D80.
[16] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[17] J.H. Han, N. Kerrison, C. Chothia, S.A. Teichmann, Divergence of interdomain geometry in two-domain proteins, Structure 14 (2006) 935–945.
[18] B. Kamaraj, A. Kumar, R. Purohit, Evolutionary reconstruction and population genetics analysis of aurora kinases, Plos one 8 (9) (2013) e75763.
[19] J. Karar, A. Maity, PI3K/AKT/mTOR pathway in angiogenesis, Front. Mol. Neurosci. 4 (2011) 51.
[20] I. Khan, I.A. Ansari, P. Singh, J.F.P. Dass, Prediction of functionally significant single nucleotide polymorphisms [SNPs] in PTEN tumor suppressor gene: an in silico approach, Biotechnol. Appl. Biochem. (2016), http://dx.doi.org/10.1002/bab.1483.
[21] A. Khwaja, Akt is more than just a bad kinase, Nature 401 (1999) 33–34.
[22] M. Kimura, The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge, 1983.
[23] A. Krze´slak, L. Pomorski, A. Lipi´nska, Expression, localization, and phosphorylation of Akt1 in benign and malignant thyroid lesions, Endocr. Pathol. 22 (4) (2011) 206–211.
[24] P. Kumar, S. Henikoff, P.C. Ng, Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4 (2009) 1073–1081.
[25] A. Kumar, R. Purohit, Computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E, Mutat. Res. 738–739 (2012) 28–37.
[26] A. Kumar, V. Rajendran, R. Sethumadhavan, R. Purohit, Identifying novel oncogenes: a machine learning approach, Interdiscip. Sci. Comput. Life Sci. 5 (2013) 241–246.
[27] A. Kumar, V. Rajendran, R. Sethumadhavan, R. Purohit, Roadmap to determine the point mutations involved in cardiomyopathy disorder: A Bayesian approach, Gene 519 (2013) 34–40.
[28] A. Kumar, V. Rajendran, R. Sethumadhavan, R. Purohit, In silico prediction of a disease-associated STIL mutant and its affect on the recruitment of centromere protein J (CENPJ), FEBS Open Bio 2 (2012) 285–293.
[29] A. Kumar, V. Rajendran, R. Sethumadhavan, P. Shikla, S. Tiwari, R. Purohit, Computational SNP analysis: current approaches and future prospects, Cell Biochem. Biophys. 68 (2014) 233–239.
[30] K.E. Landgraf, C. Pilling, J.J. Falke, Molecular mechanism of an oncogenic mutation that alters membrane targeting: Glu17Lys modifies the PIP lipid specificity of the AKT1 PH domain, Biochemistry 47 (2008) 12260–12269.
[31] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK—a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283–291.
[32] R.A. Laskowski, M.W. MacArthur, J.M. Thornton, PROCHECK: validation of protein structure coordinates, in: M.G. Rossmann, E. Arnold (Eds.), International Tables of Crystallography, Vol F. Crystallography of Biological Macromolecules. Kluwer,

Dordrecht, 2001, pp. 722–725.
[33] E. Lindahl, C. Azuara, P. Koehl, M. Delarue, NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis, Nucl. Acids Res. 34 (2006) W52–W56.
[34] M.J. Lindhurst, J.C. Sapp, J.K. Teer, J.J. Johnston, E.M. Finn, K. Peters, J. Turner, J.L. Cannons, D. Bick, L. Blakemore, C. Blumhorst, K. Brockmann, P. Calder, N. Cherman, M.A. Deardorff, D.B. Everman, G. Golas, R.M. Greenstein, B.M. Kato, K.M. Keppler-Noreuil, S.A. Kuznetsov, R.T. Miyamoto, K. Newman, D. Ng, K. O'Brien, S. Rothenberg, D.J. Schwartzentruber, V. Singhal, R. Tirabosco, J. Upton, S. Wientroub, E.H. Zackai, K. Hoag, T. Whitewood-Neal, P.G. Robey, P.L. Schwartzberg, T.N. Darling, L.L. Tosi, J.C. Mullikin, L.G. Biesecker, A mosaic activating mutation in AKT1 associated with the Proteus syndrome, N. Engl. J. Med. 365 (2011) 611–619.
[35] D. Malanga, M. Scrima, C. De Marco, F. Fabiani, N. De Rosa, S. de Gisi, N. Malara, R. Savino, G. Rocco, G. Chiappetta, R. Franco, V. Tirino, G. Pirozzi, G. Viglietto, Activating E17K mutation in the gene encoding the protein kinase AKT1 in a subset of squamous cell carcinoma of the lung, Cell Cycle 7 (2008) 665–669.
[36] I. Mende, S. Malstrom, P.N. Tsichlis, P.K. Vogt, M. Aoki, Oncogenic transformation induced by membrane-targeted Akt2 and Akt3, Oncogene 20 (2001) 4419–4423.
[37] H. Mi, N. Guo, A. Kejariwal, P.D. Thomas, PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways, Nucleic Acids Res. 35 (2007) 247–252.
[38] R.C. Muise-Helmericks, H.L. Grimes, A. Bellacosa, S.E. Malstrom, P.N. Tsichlis, N. Rosen, Cyclin D expression is controlled post-transcriptionally via a phosphatidylinositol 3-kinase/Akt-dependent pathway, J. Biol. Chem. 273 (1998) 29864–29872.
[39] P.C. Ng, S. Henikoff, Predicting deleterious amino acid substitutions, Genome Res. 11 (2001) 863–874.
[40] M. Noguchi, V. Ropars, C. Roumestand, F. Suizu, Protooncogene TCL1: more than just a coactivator for Akt, FASEB J. 21 (10) (2007) 2273–2284.
[41] O.N. Ozes, L.D. Mayo, J.A. Gustin, S.R. Pfeffer, L.M. Pfeffer, D.B. Donner, NF-kappaB activation by tumour necrosis factor requires the Akt serine-threonine kinase, Nature 401 (1999) 82–85.
[42] A. Porollo, J. Meller, Prediction-based fingerprints of protein-protein interactions, Prot. Struct. Funct. Bioinform. 66 (2007) 630–645.
[43] R. Purohit, R. Rajasekaran, C. Sudandiradoss, C. George Priya Doss, K. Ramanathan, Sethumadhavan, Studies on flexibility and binding affinity of Asp25 of HIV-1 protease mutants, Int. J. Biol. Macromol. 42 (2008) 386–391.
[44] R. Purohit, R. Sethumadhavan, Structural basis for the resilience of Darunavir (TMC114) resistance major flap mutations of HIV-1 protease, Interdiscip. Sci. Comput. Life Sci. 1 (2009) 320–328.
[45] B. Rajith, G.P.C. Doss, Path to facilitate the prediction of functional amino acid substitutions in red blood cell disorders—a computational approach, PLoS One 6 (2011) e24607.
[46] V. Ramensky, P. Bork, S. Sunyaev, Human non-synonymous SNPs: server and survey, Nucl. Acids Res. 30 (2002) 3894–3900.
[47] A.S. Ramesh, R. Sethumadhavan, P. Thiagarajan, Structure–function studies on non-synonymous SNPs of chemokine receptor gene implicated in cardiovascular disease: aa computational approach, Protein J. 32 (2013) 657–665.
[48] J. Reumers, S. Maurer-Stroh, J. Schymkowitz, F. Rousseau, SNP effect v2.0: a new step in investigating the molecular phenotypic effects of human nonsynonymous SNPs, Bioinformatics 22 (2006) 2183–2185.
[49] G. Romano, The complex biology of the receptor for the insulin-like growth factor-1, Drug News Pers. 16 (8) (2003) 525–531.
[50] B. Rost, C. Sander, Conservation and prediction of solvent accessibility in protein families, Proteins 20 (1994) 216–226.
[51] S.M. Schultzes, B.A. Hemmings, M. Niessen, O. Tschopp, PI3K/AKT, MAPK and AMPK signalling: protein kinases in glucose homeostasis, Expert Rev. Mol. Med. 14 (2012) e1.
[52] S.T. Sherry, M.H. Ward, M. Kholodov, J. Baker, L. Phan, E.M. Smigielski, K. Sirotkin, dbSNP: the NCBI database of genetic variation, Nucl. Acids Res. 29 (2001) 308–311.
[53] K. Shoji, The oncogenic mutation in the pleckstrin homology domain of AKT1 in endometrial carcinomas, Br. J. Cancer 101 (2009) 145–148.
[54] D. Stover, BRAFc. 1397G > T (G466V) Mutation in colorectal cancer. My Cancer Genome ⟨https://www.mycancergenome.org/content/disease/colorectal-cancer/braf/70/⟩ (Updated February 18), 2015.
[55] M. Sun, G. Wang, J.E. Paciga, R.I. Feldman, Z.Q. Yuan, X.L. Ma, et al., AKT1/PKBalpha kinase is frequently elevatedin human cancers and its constitutive activation is requiredfor oncogenic transformation in NIH3T3 cells, Am. J. Pathol. 159 (2001) 431–437.
[56] J.R. Testa, A. Bellacosa, AKT plays a central role in tumorigenesis, Proc. Natl. Acad. Sci. USA 98 (2001) 10983–10985.
[57] A. Tiwari, S.K. Panigrahi, HBAT: a complete package for analyzing strong and weak hydrogen bonds in macromolecular crystal structures, Silico Biol. 7 (2007) 651–661.
[58] I. Vivanco, C.L. Sawyers, The phosphatidylinositol 3-kinase AKT pathway in human cancer, Nat. Rev. Cancer 2 (2002) 489–501.
[59] P. Vogiatzi, A. Giordano, Following the tracks of AKT1 gene, Cancer Biol. Ther. 6 (2007) 1521–1524.
[60] P. Wang, M. Dai, W. Xuan, R.C. McEachin, A.U. Jackson, L.J. Scott, B. Athey, S.J. Watson, F. Meng, SNP function portal: a web database for exploring the function implication of SNP alleles, Bioinformatics 22 (2006) E523–E529.
[61] Z. Wang, J. Moult, Three-dimensional structural location and molecular functional effects of missense SNPs in the T cell receptor V beta domain, Proteins 53 (2003) 748–757.
[62] H. Watkins, Hypertrophic cardiomyopathy: from molecular and genetic mechanisms to clinical management, Eur. Heart J. 3 (2001) L43–L50.
[63] K.A. West, S.S. Castillo, P.A. Dennis, Activation of the PI3K/Akt pathway and chemo-therapeutic resistance, Drug Resist. Updates 5 (2002) 234–248.
[64] W.I. Wu, W.C. Voegtli, H.L. Sturgis, F.P. Dizon, G.P. Vigers, B.J. Brandhuber, Crystal structure of human AKT1 with an allosteric inhibitor reveals a new mode of kinase inhibition, PLoS One 5 (9) (2010) e12913.
[65] H.Y. Yuan, J.J. Chiou, W.H. Tseng, C.H. Liu, C.K. Liu, Y.J. Lin, H.H. Wang, A. Yao, Y.T. Chen, C.N. Hsu, FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization, Nucl. Acids Res. 34 (2006) 635–641.
[66] P. Yue, E. Melamud, J. Moult, SNPs3D: candidate gene and SNP selection for association studies, BMC Bioinform. 7 (2006) 166.
[67] Z. Zhang, M.A. Miteva, L. Wang, E. Alexov, Analyzing effects of naturally occurring missense mutations, Comput. Math. Methods Med. 805827 (2012).