Research Article **Open Access**

Muhammad Naveed*, Sana Tehreem, Shamsa Mubeen, Fareeha Nadeem, Fatima Zafar, Muhammad Irshad

In-silico analysis of non-synonymous-SNPs of STEAP2: To provoke the progression of prostate cancer

DOI 10.1515/biol-2016-0054
Received June 21, 2016; accepted October 6, 2016

Abstract: As a novel biomarker from the STEAP family, STEAP2 encodes six transmembrane epithelial antigens to prostate cancer. The overexpression of STEAP2 is predicted as the second most common cancer in the world that is responsible for male cancer-related deaths. Non-synonymous SNPs are important group of SNPs which lead to alternations in encoded polypeptides. Changes in the amino acid sequence of gene products can lead to abnormal tissue function. The present study firstly sorted out those SNPs which exist in the coding region of STEAP2 and evaluated their impact through computational tools. Secondly, the three-dimensional structure of STEAP2 was formed through I-TASSER and validated by different software. Genomic data has been retrieved from the 1000 Genome project and Ensembl and subsequently analysed using computational tools. Out of 177 non-synonymous single nucleotide polymorphisms (nsSNPs) within the coding region, 42 mis-sense SNPs have been predicted as deleterious by all analyses. Our research shows a well-designed computational methodology to inspect the prostate cancer associated nsSNPs. It can be concluded that these nsSNPs can play their role in the up-regulation of STEAP2 which further leads to progression of prostate cancer. It can benefit scientists in the handling of cancer-associated diseases related to STEAP2 through developing novel drug therapies.

Keywords: In-silico analysis; STEAP2; nsSNPs; Prostate cancer; Deleterious Mutations; Novel drug therapies

***Corresponding author: Muhammad Naveed,** Department of Biochemistry and Molecular Biology, University of Gujrat, Pakistan 50700, E-mail: naveed.quaidian@gmail.com; dr.naveed@uog.edu.pk
Sana Tehreem, Shamsa Mubeen, Fareeha Nadeem, Fatima Zafar and Muhammad Irshad, Department of Biochemistry and Molecular Biology, University of Gujrat, Pakistan 50700

1 Introduction

Prostate cancer is a heterogeneous disease with respect to its behavior and response to hormonal therapy [1]. This cancer predominantly affects men over the age of 75 years with a range from 45 to 80 years of age. It is usually a slow-growing asymptomatic cancer in its early stages and monitors as an indolent course without the progression of metastatic disease in many men [2]. The discovery of novel diagnostic markers and therapeutic targets is of supreme prominence effective treatment of prostate cancer. The identification of novel cancer-associated genes involves detecting genes that are differentially expressed in cancer cells.

STEAP2 is presumed to encode a six-transmembrane protein whose formulation is rich in prostate proteins [3]. It belongs to the human six-transmembrane epithelial antigen of the prostate (STEAP) protein family and contains at least four homologous members. STEAP1, STEAP2, STEAP3 and STEAP4 are vastly expressed in prostate tumors. STEAP2 is positioned at the chromosomal region 7q21 and encodes 490 amino acids. It has three basic domains which are NAD(P)-binding domain, Ferric reductase transmembrane component-like domain and Pyrroline-5-carboxylate reductase, catalytic, N-terminal.

It is mostly found within the prostate epithelial cells in the Golgi complex plasma membrane that is linked with early endosomes and trans- Golgi network (TGN) that act as a shuttle between the Golgi apparatus and plasma membrane. Furthermore, STEAP2 localizes early but not late endosomes or lysosomes. Different STEAP protein family members have unique patterns of expression and functions within the cell, as they differ in their subcellular localization. These characteristics persist, despite the fact that they have identical domain structures with four to six transmembrane domains with intracellular N- and C- terminals**.** This data suggests that STEAP2 involved in endocytic and secretory pathways even though the

consequence of this is still ill-defined [4].

Real-time quantitative PCR (qPCR) revealed that in prostate cancer, the expression levels of gene are 10 times higher that within a typical prostate compared to in other tissues studied. The STEAP2 was expressed in substantial levels only in androgen-responsive LNCaP prostate cancer cell lines. The expression of STEAP2 was considerably higher in benign prostate hyperplasmias as both untreated primary and hormone-refractory prostate carcinomas, suggest that it may be convoluted in the cause of prostate cancer [5]. It was found that insane expression of STEAP2 greatly increased propagation of DU145 prostate cancer cells, as well as COS-7 cells *in vitro*. Equally, small interfering RNA-mediated knockdown of STEAP2 expression in LNCaP cells hindered cell growth and partly caused cell cycle arrest [3,6]. STEAP2 is a possible agnostic target in prostate cancer as a cell surface antigen.

Altered expression of STEAP family members STEAP2, STEAP3 and STEAP4 except STEAP1 leads to increase iron uptake by cells due to co localization with transferrin receptor 1(TfR1) during endocytic pathway. It is believed that ferric iron which is delivered by transferrin into the endosome is reduced to its divalent ferrous form by these trans-membrane proteins. This step will facilitate divalent metal transporter 1 that transfers this divalent iron from the endosome into the cytosol (Figure 1). Hence, the patterns of expression and functional activity in these specific tissues is significant in copper and iron homeostasis, and the subcellular localization of these proteins advocates that they may be significant metallo-reductases.

Single nucleotide polymorphisms (SNPs) are the most common form of genetic mutation within the human genome. One of the two main types of coding region SNP's are non-synonymous SNPs which result in single amino acid polymorphisms that may alter protein stability, charge, solubility, structure, or function potentially causing pathogenic phenotypes. Based on these properties, they are of particular concern for further experimental assessment [7]. These nsSNPs are associated with various complex diseases in humans.

To filter major diseases associated nsSNPs from nonsignificant SNPs many computational algorithms were implemented. Insilico studies have provided a well-

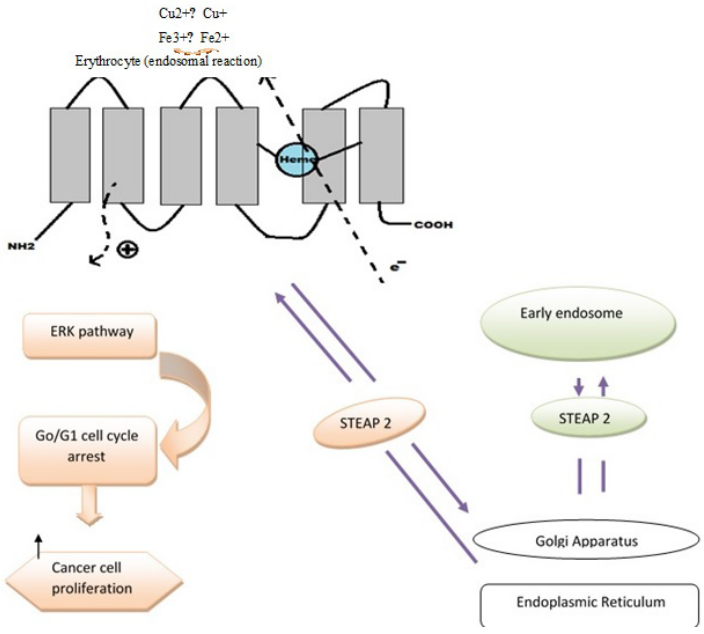

Figure 1. Schematic Diagram of STEAP2 protein structure, its localization in cells and its physiologic functions. STEAP2 functions involve in both endocytic and secretory pathways. By activating the ERK, STEAP2 induces partial arrest on the G0-G1cell cycle phases in cancer cells and therefore increases the proliferation of cell anti-apoptosis and tumor development [17]. Co-localization of STEAP 2 with Tf and TfR1 leads to Transferrin cycle of endosomal erythroid cells. STEAP2 is involved in the proper uptake of $Fe3^{+}$ and $Cu2^{+}$ and their reduction to $Fe2^{+}$ and Cu^{+}. Furthermore, recent information has indicated that the STEAP2 Protein activates ERK extracellular signal regulated kinase pathway, which in turn activates the downstream effectors molecules which are Matrix Metalloproteases (MMPs). In PCa, MMPs -2, -7, -9, -13 and -14 have been shown to be involved and their over-expression promotes PCa growth and metastasis [2]. In early endosomes and the TGN Trans Golgi network STEAP2 can also be found and it acts as a receptor for both exogenous and endogenous ligands or as regulator of protein delivery and sorting mechanisms [2,18].

organized platform for the analysis and estimations of pathological consequence of genetic mutations and in determining their basic molecular mechanism [8-13]. In this study, we have tried to find out most deleterious and disease associated nsSNPs in coding region of STEAP2 out of dataset obtained from ENSEMBL by using computational tools and then checked their effects on structure of protein [13,14]. Out of 177 non-synonymous single nucleotide polymorphisms (nsSNPs) of coding region, 42 missense SNPs have been predicted as deleterious by Polyphen2, Phd-SNP, MutPred, PROVEAN and SIFT and the effect of nsSNPs on the stability of protein have been checked by Mupro and I-mutant 3.0 [15,16]. Three-dimensional structure of STEAP2 have been predicted through I-TASSAR and after its refinement through Galaxy Refine tool. The subsequent verification of the 3D structure was completed using Verify 3D, Errat and Ramachandran Plot. For the efficient prediction of STEAP2 function, ligand binding sites and active site residue prediction were completed using COACH. The overall workflow employed in this study is shown in Figure 2.

2 Materials and methods

2.1 Dataset used for SNP annotation

ENSEMBL v7.6 was used to retrieve the variations that were present in the STEAP2. 10 transcripts of STEAP2 are present in the database but we had selected the canonical transcript. We selected non-synonymous SNPs of canonical transcript to carry out this study. Then different software was used to check out the structural and functional effects of all non-synonymous SNPs that use different algorithms to check out the effects of SNPs on proteins [12].

2.2 Softwares used for SNPs annotations

2.2.1 PolyPhen2

PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/) uses a naïve Bayesian classifier to predict the allele function using structure and function based characteristics. The

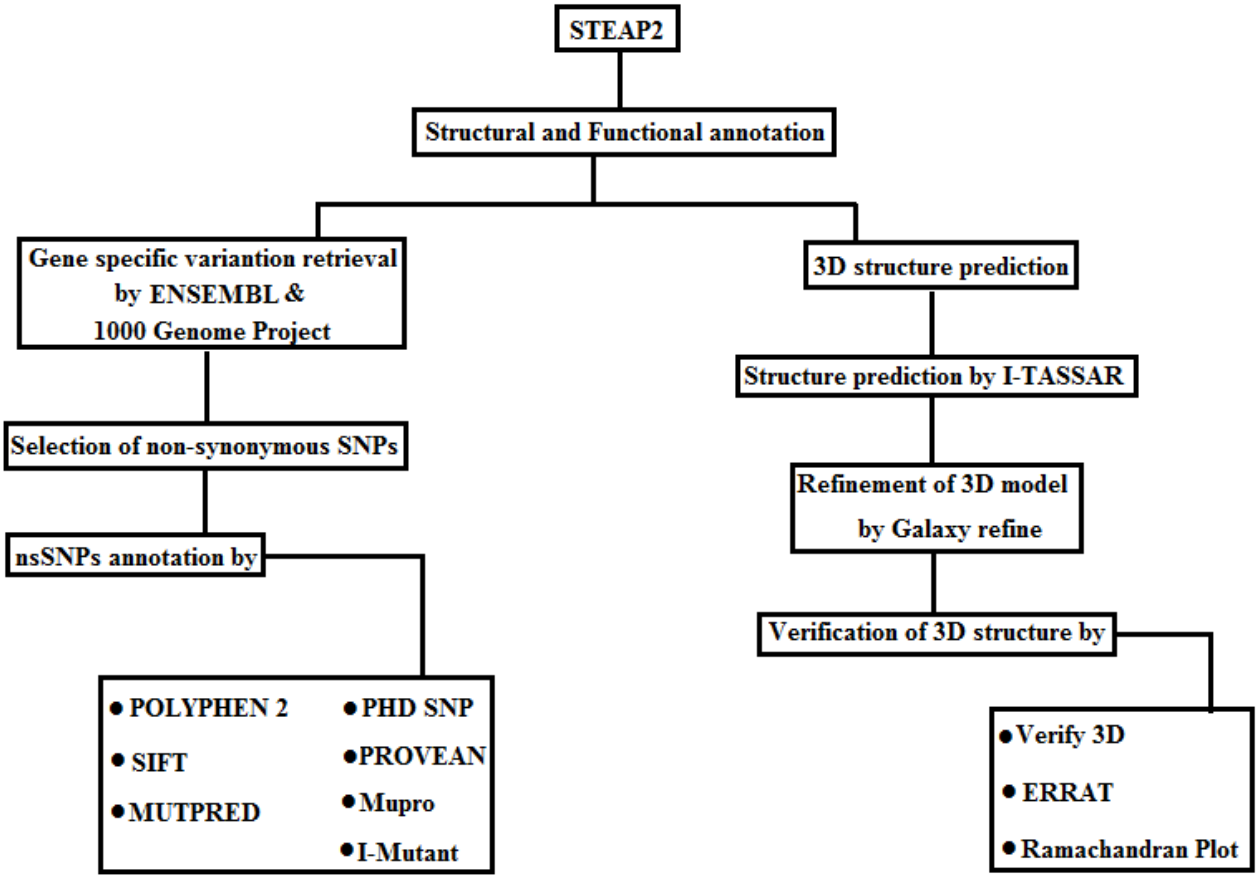

Figure 2. Flow sheet diagram of methodology.

score of polyphen2 ranges from 0-1. PolyPhen2 calculates the probability for a given mutation to be benign, possibly damaging or probably damaging. If score is nearer to 1, nsSNPs will probably be damaging [11,19].

2.2.2 PhD SNP

PhD SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms) is a web based tool that is used to annotate the non-synonymous SNPs. PhD-SNP tool (http://SNPs.biofold.org/phd-snp/phd-snp.html)is based on SVM-based classifier. The output given in the form of table which contains mutated mutation in the protein sequence. It contains both wild-type amino acid and new amino acid and tells us about the SNP is either deleterious or neutral [20,21].

2.2.3 MutPred

MutPred (http://mutpred.mutdb.org/) is based upon the SIFT algorithm. It tells us about gain or loss of 14 predicted functional and structural properties due to non-synonymous SNPs. G-value range from 0-1 in MutPred results. If the g-value is closer to 1 then amino acid substitution shows more effect on protein function [22-24].

2.2.4 PROVEAN

PROVEAN is a web based tool that is accessed at (http://provean.jcvi.org/index.php). Homologs of query sequence were searched using BLAST against the NCBI. The Cutoff value of PROVEAN is -2.5. amino acid substitutions that have value greater than cut off value will be considered as deleterious [25].

2.2.5 SIFT

SIFT (Sorting intolerance from tolerance) is a web based tool (http://sift.jcvi.org). SIFT uses the PSI-BLAST protein database and through it functionally associated protein sequences are collected. Afterwards by performing the homologous alignment of sequences, it finds out the possibility of the amino acid that exists at specific site. If score is less than 0.05 it is considered as not tolerated whereas scores greater than 0.05 are considered tolerated [21,26,27].

2.3 Predicting effect of non-synonymous SNPs on protein stability

2.3.1 Mupro

Mupro is a web based tool assessed at (http://mupro.proteomics.ics.uci.edu/). It is based upon SVM and used to predict whether a single point mutation decreases or increases the protein stability by using sequence and structure information. Protein sequences along with mutation position, the original amino acid and substituted amino acid is provided to software [21].

2.3.2 I-MUTANT 3.0

I-MUTANT 3.0 (http://gpcr2.biocomp.unibo.it/cgi/predictors/IMutant3.0/IMutant3.0.cgi) is used to determine the effect of non-synonymous SNPs on the structure and function of proteins. This software uses data from ProTherm, which is a database providing experimental proven free energy changes of protein stability due to changes in amino acids. Protein sequences are provided along with the new residue and position for obtaining the free energy change [28]. Different software use different algorithms and due to this their predictions may also differ from each other.

2.4 3D structure Prediction

Proteins 3D structures are very important to check the properties of protein like ligand binding affinities of protein in the absence and presence of mutations. The 3D structure of STEAP2 is not present in PDB. To generate an accurate and consistent tertiary structure of protein sequence of STEAP2 (ID: ENSP00000378119) retrieved from Ensembl Genome browser (http://www.ensembl.org) and then subjected to I-TASSER for three-dimensional (3D) structure prediction of STEAP2. I-TASSER I-TASSER modeling begins form a meta-server threading approach LOMETS through which structural templates are identified. LOMETS contain multiple threading programs which can generate good template alignment but I-TASSER selects the template with the highest Z-score [29-31].

2.5 Refinement of protein model

Galaxy refine tool (http://galaxy.seoklab.org/refine) is used to refine the PDB model of the protein which is

produced by I-TASSER. It is constructed on a refinement method which has been successfully verified in CASP10. The CASP10 assessment tells that this method exhibited the best presentation in educating the local structure quality. The method can improve both global and local structure quality on average for models created by protein structure prediction servers [32].

2.6 Verification of 3D structure

For verification of 3D structure verify 3D (http://services.mbi.ucla.edu/Verify_3D/), Errat (http://services.mbi.ucla.edu/ERRAT/ and Ramachandran Plot (http://services.mbi.ucla.edu/SAVES/Ramachandran/) were used. After above steps identification of ligand binding and active sites were carried out.

3 Results

3.1 SNPs annotation

The ENSEMBL database was used for the retrieval of variation of STEAP2. Out of all the variation, nsSNPs were selected for further study (Figure 3). There were 177 nsSNPs within STEAP2 in ENSEMBL. After their retrieval, different softwares were used for the structural and functional annotation of nsSNPs (Supplementary Table 1); 42 nsSNPs were predicted as deleterious by all programs and are likely to affect the structure and function of the protein resulting in different problems or diseases. Table 1 shows 42 nsSNPs which are predicted as deleterious by all of the software.

3.2 Accuracy of prediction of nsSNPs by different softwares

According to PolyPhen 2 results, 58.75% nsSNPs predicted as damaging and 41.24% were benign. SIFT predicted 50.84% as deleterious and 49.15% as benign. According to the third analysis using MutPred, 43.50% nsSNPs were predicted as deleterious and 56.49% as benign. PHD SNP and Provean were also used to make this work more specific, PHD SNP predicted 42.37% as deleterious and 57.62% as benign and PROVEAN predicted 41.80% as deleterious and 58.19% as benign. I-Mutant and Mupro predicted the effect of nsSNPs on the stability of protein structure and function. (Figure 4) (Figure 5). All nsSNPs that were predicted as deleterious by all of the softwares are: G53S, G53A, G53V, Y54C, V76A, H79D, A82V, A92G, H97P, Y98H, N118D, N129H, N129S, E131D, A161T, R163W, R163Q, A174V, R183C, L185S, L195V, F227S, D234V, V235A, H237N, Q281H, G285S, Y288C, R290I, P292A, W294G, G306R, A313T, H316R, P324L, R326G, R389G, E390G, Q395E, Q395A, S396C and C453R (Table 1).

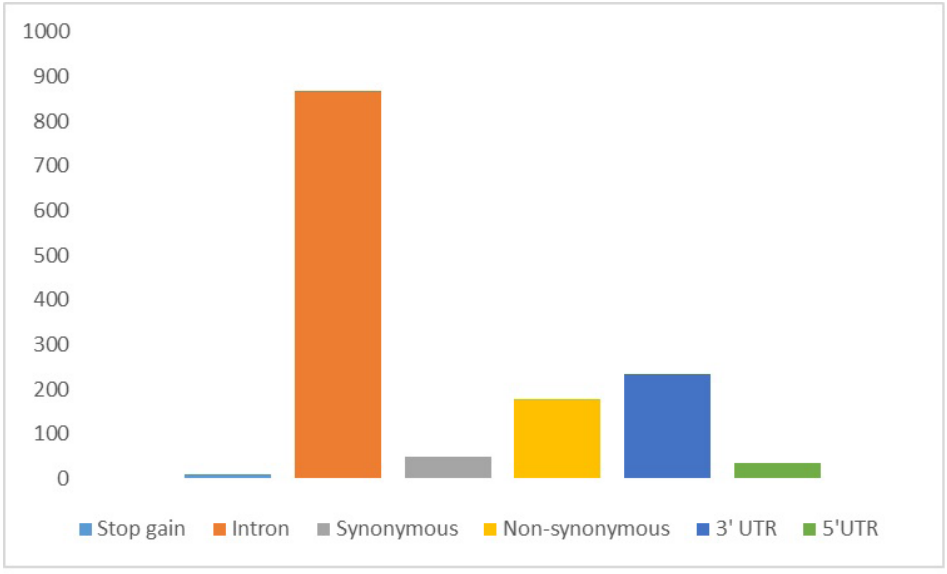

Figure 3. Chart shows distributions of SNPs present at different regions of STEAP2 gene.

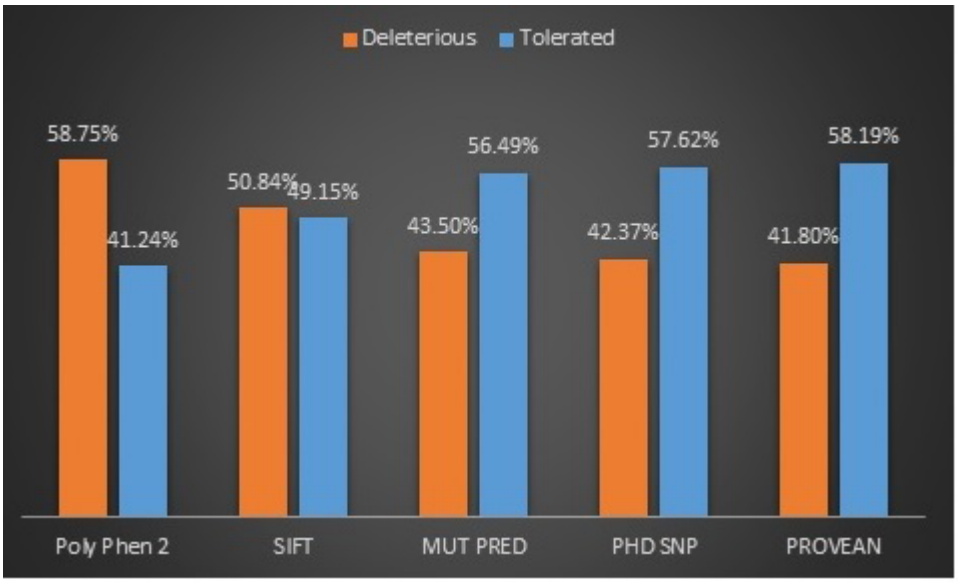

Figure 4. Percentage of deleterious and tolerated nsSNPs predicted by different softwares. These nsSNPs are most deleterious as predicted by all software which are used so these deleterious nsSNPs will affect both structure as well as function of STEAP2.

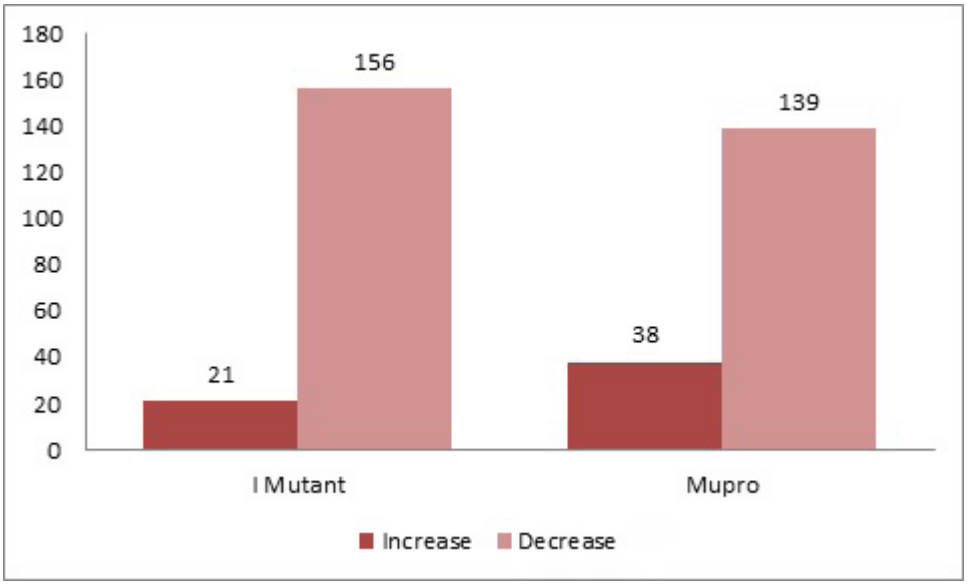

Figure 5. Effects of nsSNPs on stability of protein predicted through I- mutant and Mupro. I-Mutant predicted that 156 nsSNPs causes decrease in the stability of protein and 21 nsSNPs causes increase in the stability of protein. Mupro predicted that 139 nsSNPs causes decrease in the stability of protein and 38 causes increase in the stability of protein.

3.3 Annotation of 3D structure

3D structure of STEAP2 protein was formed by I-TASSER software. I-TASSER gives 5 different structures of this domain with their C-score (confidence score). C-score typically ranges from -5 to 2 and the higher C-score value means that model has high confidence and vice versa. The model which we have selected has a C-score of -2.08. This model has C-score higher than other models so we selected this model and further annotate this model (Figure 6).

3.4 Verification of 3D

After refinement of the PDB model, the 3D structure verification was carried using software that evaluates the 3D structure using different properties of protein. A detailed comparison of the 3D models generated by I-TASSER and Galaxy refine tool is presented in Table 2.

The following are the programs that we used to verify the STEAP2 structure:

Table 1. Identification of deleterious mutation by different softwares.

#	Deleterious Mutations	MutPred Score	PolyPhen 2 Prediction and Score	SIFT Prediction	SIFT Score	PHD SNP Prediction and Score	Provean Prediction	PROVEAN cutt off value	Mupro	I MUTANT
01	G53S	0.761	Probably damaging 1.000	Damaging	0	Disease 7	Deleterious	-4.814	Decrease	Decrease
02	G53A	0.765	Probably damaging	Damaging	0	Disease 10	Deleterious	-7.32	Decrease	Decrease
03	G53V	0.779	Probably damaging 1.000	Damaging	0	Disease 5	Deleterious	-4.829	Decrease	Decrease
04	Y54C	0.685	Probably damaging 0.996	Damaging	0	Disease 5	Deleterious	-6.288	Decrease	Decrease
05	V76A	0.686	Probably damaging 0.969	Damaging	0.01	Disease 2	Deleterious	-3.136	Decrease	Decrease
06	A82V	0.874	Probably damaging 0.998	Damaging	0	Disease 3	Deleterious	-3.509	Decrease	Increase
07	H79D	0.552	Possibly damaging 0.673	Damaging	0	Disease 2	Deleterious	-5.225	Increase	Decrease
08	A92G	0.829	Probably damaging 1.000	Damaging	0	Disease 1	Deleterious	-3.418	Decrease	Decrease
09	H97P	0.814	Probably damaging 1.000	Damaging	0	Disease 8	Deleterious	-7.649	Decrease	Increase
10	Y98H	0.847	Possibly damaging 0.980	Damaging	0.03	Disease 6	Deleterious	-3.894	Decrease	Decrease
11	N118D	0.949	Probably damaging 1.000	Damaging	0.01	Disease 7	Deleterious	-4.275	Increase	Decrease
12	N129H	0.881	Probably damaging 1.000	Damaging	0	Disease 6	Deleterious	-4.57	Increase	Decrease
13	N129S	0.82	Probably damaging	Damaging	0	Disease 4	Deleterious	-4.56	Increase	Decrease
14	E131D	0.76	Probably damaging 0.995	Damaging	0	Disease 3	Deleterious	-2.524	Decrease	Decrease
15	A161T	0.714	Possibly damaging 0.722	Damaging	0.01	Disease 3	Deleterious	-2.629	Decrease	Decrease
16	R163W	0.788	Probably damaging 0.998	Damaging	0	Disease 6	Deleterious	-6.276	Decrease	Decrease
17	R163Q	0.764	Possibly damaging	Damaging	0	Disease 5	Deleterious	-2.67	Decrease	Decrease
18	A174V	0.846	Probably damaging 1.000	Damaging	0	Disease 3	Deleterious	-3.525	Increase	Decrease
19	R183C	0.728	Probably damaging 1.000	Damaging	0.05	Disease 1	Deleterious	-4.975	Decrease	Decrease
20	L185S	0.819	Possibly damaging 0.868	Damaging	0	Disease 1	Deleterious	-4.676	Decrease	Decrease
21	L195V	0.882	Possibly damaging 0.914	Damaging	0	Disease 1	Deleterious	-2.544	Increase	Decrease
22	F227S	0.777	Possibly damaging 0.468	Damaging	0.05	Disease 5	Deleterious	-6.039	Decrease	Decrease
23	D234V	0.607	Probably damaging 0.998	Damaging	0.05	Disease 8	Deleterious	-5.392	Increase	Decrease
24	V235A	0.642	Possibly damaging 0.774	Damaging	0	Disease 1	Deleterious	-3.651	Decrease	Decrease
25	H237N	0.572	Probably damaging 0.972	Damaging	0.04	Disease 1	Deleterious	-4.353	Increase	Decrease
26	Q281H	0.712	Probably damaging 1.000	Damaging	0	Disease 7	Deleterious	-4.926	Decrease	Decrease
27	G285S	0.822	Probably damaging 1.000	Damaging	0	Disease 8	Deleterious	-5.859	Decrease	Decrease
28	Y288C	0.783	Probably damaging 1.000	Damaging	0	Disease 9	Deleterious	-8.821	Decrease	Decrease

continued **Table 1.** Identification of deleterious mutation by different softwares.

#	Deleterious Mutations	MutPred Score	PolyPhen 2 Prediction and Score	SIFT Prediction	SIFT Score	PHD SNP Prediction and Score	Provean Prediction	PROVEAN cutt off value	Mupro	I MUTANT
29	R290I	0.704	Probably damaging 0.998	Damaging	0	Disease 6	Deleterious	-7.323	Decrease	Decrease
30	P292A	0.822	Probably damaging 1.000	Damaging	0	Disease 8	Deleterious	-7.729	Decrease	Decrease
31	W294G	0.833	Probably damaging 1.000	Damaging	0	Disease 9	Deleterious	-12.56	Decrease	Decrease
32	G306R	0.834	Probably damaging 1.000	Damaging	0	Disease 9	Deleterious	-7.797	Increase	Increase
33	A313T	0.684	Probably damaging 1.000	Damaging	0.02	Disease 7	Deleterious	-3.594	Decrease	Decrease
34	P324L	0.832	Probably damaging 1.000	Damaging	0	Disease 9	Deleterious	-8.589	Decrease	Decrease
35	H316R	0.843	Probably damaging 1.000	Damaging	0	Disease 8	Deleterious	-7.762	Increase	Decrease
36	R326G	0.796	Possibly damaging 0.739	Damaging	0.01	Disease 8	Deleterious	-6.148	Decrease	Decrease
37	R389G	0.647	Probably damaging 1.000	Damaging	0	Disease 5	Deleterious	-6.468	Decrease	Decrease
38	E390G	0.677	Probably damaging 1.000	Damaging	0	Disease 4	Deleterious	-6.59	Decrease	Decrease
39	Q395E	0.688	Probably damaging 0.999	Damaging	0	Disease 7	Deleterious	-2.757	Increase	Increase
40	Q395H	0.676	Probably damaging 1.000	Damaging	0	Disease 8	Deleterious	-4.694	Decrease	Decrease
41	S396C	0.594	Probably damaging 0.998	Damaging	0	Disease 7	Deleterious	-3.815	Decrease	Decrease
42	C453R	0.873	Probably damaging 0.999	Damaging	0	Disease 8	Deleterious	-8.336	Increase	Increase

Table 2. Comparison of I-TASSER and Galaxy refine.

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Model produce by I-TASSER	**1.0000**	**0.000**	**3.364**	**17.6**	**12.2**	**75.0**
Model refine by Galaxy refine	**0.9194**	**0.501**	**2.268**	**18.6**	**1.1**	**92.8**

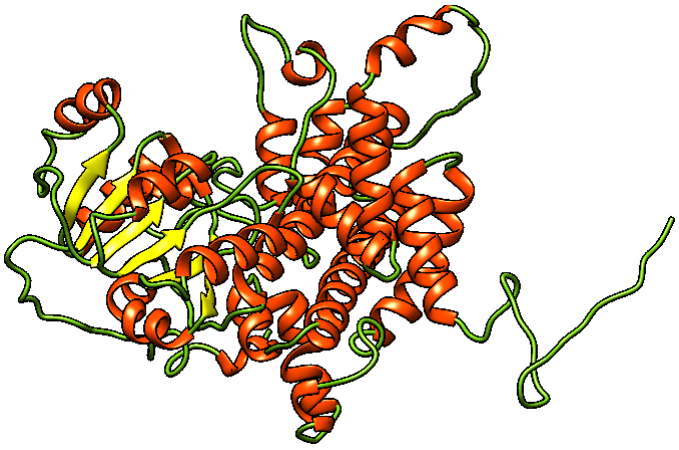

Figure 6. Visualization of 3D structure of STEAP2 by PyMOLver 1.3.

3.4.1 Verify 3D

Verify 3D determines the compatibility of amino acid sequence with the 3D structure of protein and compares the results to good structures. According to this software 65 % of amino acids should have scored greater than 0.2 to pass the structure otherwise the structure is not 3D [33,34]. In our designed 3D model 70.20% of the residues had an averaged 3D-1D score >= 0.2 (Supplementary 1).

3.4.2 Errat

This software examines the statistics of non-bonded interactions between different atom types and then plots the values of the error function versus position of residues that are calculated by comparison with statistics from highly refined structures [35] (Supplementary 2).

3.4.3 Ramachandran plot

The Ramachandran plot provides a simple view of the conformation of a protein. The distinct regions in the Ramachandran plot based on φ-ψ angles cluster reflect a particular secondary structure. The red, yellow, cream and white colors in the plot indicate most favorable, allowed, generously allowed and disallowed regions. Ideally, over 90% of the residues in these „core" regions. The proportion of residues in the „core" regions is one of the better guides to stereo chemical quality [36]. Ramachandran plot of STEAP2 protein structure shows that most of the residues are present in the core region, which indicates most favorable combination of phi-psi values (Supplementary 3).

3.5 Identification of ligand binding site

By using COACH (http://zhanglab.ccmb.med.umich.edu/COACH/) ligand binding sites in STEAP2 was analyzed. COACH is an algorithm-based web server. It was used because ligand binding sites are very important for protein function. If any mutation occurs in this region, then it can disrupt the interaction between transmembrane protein and its ligands [37].

This software uses multiple ligands to check their binding sites but ligands which have higher C-score (confidence score) indicate a more reliable prediction (Table 3). The ligand NDP have higher C-score than rest of ligands and its possible binding sites are 37, 38, 41, 60, 61, 62, 93, 94, 96, 100, 148, 149, 150, 151, 202, 205, 206, and 209 (Supplementary 4). The name or synonyms of NDP are NADPH, 2'-O-PHOSPHONOADENOSINE 5'-{3-[1-(3-Carbamoyl-1,4-Dihydropyridin-1-Yl)-1,4-Anhydro-D-Ribitol-5-Yl] DihydrogenDiphosphate}, Dihydronicotinamide-Adenine Dinuckeotide Phosphate, Reduced Nicotinamide-Adenine Dinucleotide Phosphate, Tpnh and Beta-Nicotinamide Adenine Dinucleotide Phosphate, Reduced Form.

3.6 Identification of active site residue and enzyme commission number (EC)

Active sites consist of residues that form weak interactions with the binding sites of the substrate and with the catalytic site. So, these residues are very important for a chemical reaction. By using I-TASSER software we found active site residues that are involved in catalytic reaction along with their EC number. This software predicts multiple active site residues but active site residues with highest C-score indicate reliable prediction for EC number. Active site residues with highest C-score (0.220) are 39 and 62 (Supplementary 5). The EC number 1.1.1.44

Table 3. COACH server data.

Rank	C-score	Cluster size	PDB Hit	Lig Name	Ligand Binding Site Residues
1	0.32	23	1pgoA	NDP	37, 38, 39, 41, 60, 61, 62, 93, 94, 96, 100, 118, 148, 149, 150, 151, 202, 205, 206, 209
2	0.14	11	2iz0B	ATR	39, 60, 61, 62, 65, 93, 94, 96, 100
3	0.12	10	2w8zB	6PG	118, 149, 150, 151, 202, 205, 206, 209, 210, 289, 290, 292, 317
4	0.04	5	1pgnA	POP	118, 148, 149, 150, 151, 202, 206
5	0.03	4	2iypB	5RP	49, 52, 55

is for Phosphogluconate dehydrogenase which is highly specific for NADP (+).

4 Discussion

STEAP2 (six transmembrane epithelial antigen of prostate) belongs to STEAP family which consist of four homologous members STEAP1, STEAP3 and STEAP4 and these are localized to the cell membrane. These four family members are radically distinctive and have a common characteristic of metallo reductases so these have important role in metal metabolism in mammals. In several types of human cancers such as prostate, cervix, ovary, testis, breast, colon, pancreas, iEwing sarcoma and bladder, the expression of STEAP1 and STEAP2 is significantly high. The expression of this family is different in normal and cancerous tissue making them possible candidate for the therapeutic strategies [18]. Previous studies have shown that the expression of STEAP2 is 10 times higher in prostate cancer.

STEAP2 is located at the chromosomal region 7q21 and it encodes for 490 amino acids. It has three basic domains (1) NADP binding domain with a core Rossmann-type fold which consists of 3-layers alpha/beta/alpha where the six beta strands are parallel in the order 321456 and it catalyzes the oxidation of one compound with the reduction of another (2) Ferric reductase transmembrane component like domain which not only reduces iron but also copper (3) Pyrroline-5-carboxylate reductase consists of two domains, an N-terminal catalytic domain and a C-terminal dimerization domain [38-44] .

Non-synonymous SNPs lead to alteration in the protein sequence which further may affect the protein function. In this study, we sorted out the mutations that are deleterious and might play their role in the up-regulation of STEAP2. Genomic data for variation was retrieved from ENSEMBL and 100 genome project out of which we just select 177 non-synonymous mutations. Different software based upon different algorithms were used for the annotations of nsSNPs. Only 42 nsSNPs were predicted as deleterious by all the software. Following are the details of each mutation that was selected as deleterious by most software. All 42 deleterious mutations along with their possible damaging functional and structural effects are summarized in Supplementary Table 2.

- **G53S**: Glycine is being replaced by serine at position 53. This mutation will disturb the local structure, flexibility and interaction with other molecules. Alport syndrome was caused due to homozygous mutations in the a3 and a4 genes of type IV collagen. Two different mutations were reported in a4 gene (i) a glycine to serine substitution and (ii) a serine to stop mutation [45]. **G53A:** Glycine is being replaced by alanine at position 53. The mutation will cause of lose hydrogen bonding and affect the multimeric contacts. A variation in which glycine 692 is replaced with alanine in β-amyloid precursor protein gene leads to cerebral haemorrhage and persenile dementia [46]. **G53V:** Glycine is being replaced with valine at position 53. The mutation will lose hydrogen bonding and affect the multimeric contacts. Type 2 Stickler syndrome was caused by change of glycine into valine at position 97 in COL11A1 gene [47].
- **Y54C:** Tyrosine is replaced with cysteine at position 54. The mutation will cause loss of hydrogen bonds in protein core and disturb correct folding. Replacement of tyrosine with cysteine in human **α-Synuclein enhance aggregation of protein and cellular toxicity in patient of Parkinson disease** [48]. **V76A:** Valine replaced with Alanine at position 76. Mutation will cause a possible loss of external interactions. In study, it was reported that hypophosphatasia disease was caused when valine is replaced with alanine at position 406 [49]. **A82V:** Alanine replaced with valine at position 82. Due to this mutation Rett syndrome (RTT), a neurological disorder occurs in females. A novel variation in which alanine is replaced with valine(A140V) in the MECP2 gene found in one female with minor mental retardation [50].
- **H79D:** Histidine is replaced with Aspartic acid at position 79. The mutant residue can result in protein folding problems due to a change in charge [51]. **A92G:** Alanine is being replaced with glycine at position 92. In the core of the protein the mutation will cause loss of hydrophobic interactions. Conginental long QT syndrome is a heterogeneous group of disorders due to a mutation in four genes. One mutation is that in which alanine is replaced with valine at a CpG dinucleotide [52]. **H97P:** Histidine is being replaced with proline at position 97. Mutant residue will cause a possible loss of external interactions. Cholesteryl ester storage disease (CESD) is autosomal recessive disorder that is linked to reduce activity of lysosomal acid lipase (LAL) due to a point mutation in which His108 mutates into Pro (CAT to CCT) [53].
- **Y98H:** Tyrosine is replaced by histidine at position 98. Due to this variation, hydrophobic interactions in the core of the protein will be lost. In Hemoglobin Bethesda, HC2 tyrosine is replaced with histidine and affects normal hemoglobin oxygenation function [54].
- **N118D:** Asparagine is being replaced by aspartic

acid at position 118. Due to an addition of a charge, the mutated residue causes repulsion between mutant and neighboring residues. In a study, it was reported that Hemoglobin Yoshizuka occurs due to the substitution of asparagine with aspartic acid and further it reduces oxygen affinity [55]. **N129H:** Asparagine is being replaced by histidine at position 129. Mutation disturbs the core structure and cause loss of H-bonding. **N129S:** Asparagine is being replaced by serine at position 129. Improper folding and loss of hydrogen bonds occur due to this mutation. In study, it was reported that, a mutation in lipoprotein lipase gene in which, asparagine is converted into serine is responsible for high plasma triacylglycerol along with increased BMI [56]. **E131D:** Glutamic acid is being substituted with aspartic acid at position 131. Potentially, this mutation causes loss of external interaction. In Pseudomonas aeruginosa substitution of Glutamic Acid with Aspartic Acid at position 553 severely Reduces Toxicity and activity of Enzyme [57].

- **A161T:** Alanine is being replaced by threonine at position 161. The mutation can cause loss of hydrophobic interactions with other molecules. In Plasminogen Tochigi alanine is replaced by threonine in the active site at position 600 and inactivate the plasmin [58]. **R163W:** Arginine is being replaced by tryptophan at position 163. This mutation causes loss of charge which in turn causes loss of interactions with other molecules. In the previous study, it was reported that lack of hereditary myeloperoxidase is due to a missense mutation in which arginine is replaced with tryptophan at position 569 [59]. **R163Q**: Arginine is replaced by glutamine at position 163. This mutation causes loss of charge that will lead to loss of interactions with other molecules. In bacteriorhodopsin substitution of arginine with glutamine influence formation of chromophore, proton translocation and photo cycle [60].
- **A174V:** Alanine is being replaced by valine at position 174. This mutant residue does not prefer α-helices as secondary structure**. R183C:** Arginine is replaced into cysteine at position 183. This mutation will disturb ionic interaction and result in loss of external interactions. Replacement of arginine47 with cysteine in Antithrombin III Toyama causes deficiency of heparin-binding ability [61]. **L185S:** Leucine is replaced by serine at position 185. Due to this mutation, hydrophobic interaction at the core of protein will be lost. **L195V:** Leucine is being replaced by valine at position 195.The mutation converts the wild-type residue in a residue that does not prefer α-helices as secondary structure. In Human Cyclin T1 interaction between TAR, Equine Cyclin T1 and Tat are lost due to substitution of leucine with valine [62]. **F227S:** Phenylalanine is replaced by serine at position 227. Due to this mutation, hydrophobic interactions, either on the surface or core of the protein will be lost. Striking decrease in the adhesive capacity of K88 fibrillae occurred due to replacement of phenylalanine with serine at position 150. Apparently, phenylalanine 150 plays role in the interaction of the adhesin with receptor molecule [63].
- **D234V:** Aspartic acid change into a valine at position 234. The local conformation will be destabilized to some extent because the mutant residue favors another secondary structure. **V235A:** Valine is replaced by alanine at 235 positions. Due to mutation destabilization of local conformation will occur. In patients who have retinochoroidal vascular disorder and venous thrombosis decreased activity of plasmin is reported due to replacement of alanine by valine at position 55 [64]. **H237N**: Histidine replaced by asparagine at position 237. Due to this mutation loss of interactions occur. Fumarase (or fumaratehydratase) is an enzyme that has two carboxylic acid binding sites A and B. At the A site histidine replaced by asparagine resulted in a large decrease in specific activity [65]. **Q281H:** In this mutation glutamine is replaced by histidine at position 281. Mutation will disturb the domain Ferric oxidoreductase and abolish its function. A mutant enzyme with glutamine substituted by histidine at position 373 showed a 10(4)-fold decrease in its catalytic activity [66].
- **G285S:** Glycine is being changed by serine at position 285. As serine is bigger than glycine; this might disturb torsion angles. **Y288C:** Tyrosine is being changed by cysteine at 288 position. Mutation can result in loss of hydrogen bonds and/or disturb correct folding. DAT the dopamine transporters are involved in uptake of dopamine. It consists of 60 amino acids. When tyrosine is replaced with cysteine in DAT protein it results in decrease uptake of dopamine [67]. **R290I:** Arginine is replaced with isoleucine at position 290. As isoleucine is smaller than arginine this variation might leads to loss of interaction. **P292A:** Proline is being replaced with alanine at position 292. The mutation can disturb this special conformation and abolish function. Proline mutations into alanine effect the putative transmembrane segment 6 and 10 of the glucose transporter GLUTl [68]. **W294G:** Amino acids tryptophan is replaced with glycine at position

294. Rigidity of protein and hydrophobic interactions will be lost. Lethal and severe form of Osteogenesis Imperfecta is caused by a substitution in which glycine is replaced by tryptophan in the type I collagen [69].

– **G306R:** Amino acid glycine is replaced with arginine at position 306. Hydrophobicity differences can affect the hydrophobic interactions with the membrane lipids. A case of mild osteogenesis imperfecta in a 56-year-old male. The substitution of arginine for glycine at 85 position in one of the two chains of alpha 1(I) procollagen was the main defect responsible for this disease [70]. **A313T:** Alanine is replaced with threonine at position 313. Due to this mutation, hydrophobic interactions that are present either on the surface or in the core of the protein, will be lost. Aggregation into amyloidogenic proteins or other self-aggregated amyloid fibrils in various proteins caused by the substitution of Alanine for Threonine [71]. **H316R:** Amino Acids Histidine replaced with arginine at position 316. Mutation disturbs the domain and interactions with metal ions and abolishes function. Single nucleotide polymorphism in human androgen receptor in which arginine 773 is replaced by cysteine or histidine leads towards the complete androgen insensitivity with diverse receptors phenotypes [72]. **P324L:** Proline is replaced with leucine at position 324. Size difference disturbs the interactions with metal ion which disturbs special conformation. Cone-rod dystrophy is caused by single nucleotide polymorphism in human GCAP1that is mandatory for activation of retinal guanylate cyclase-1 in which proline is replaced by leucine at 50 position [73].

– **R326G:** Arginine is replaced with glycine at position 326. This mutation can cause improper folding and loss of hydrogen bonds. Mild osteogenesis imperfecta was reported in a 56-year-old male. The basal flaw responsible for the disease was the substitution of arginine for glycine 85 (R85G) in one of the two chains of al (1) procollagen [74]. **R389G:** Amino Acids arginine is replaced with glycine at position 389.This mutation can lead towards the loss of hydrogen bonds and/or result in improper folding. **E390G:** Glutamic acid replace with glycine at position 390. This mutation will modify the charge that results in loss of interactions with other molecules. In the triple-helical domain of the type III procollagen substitution of glutamic acid for glycine at 1021 led to Ehlers-Danlos syndrome together with an enormous dissecting aortic aneurysm [75]. **Q395E:** Glutamine is replaced by glutamic acid at position 395. This mutation disturbs the domain ferric oxidoreductase and abolishes its function.

– **Q395A:** Glutamine is replaced by alanine at position 395. This mutation can disturb proper folding or/and conduces loss of hydrogen bonds. **S396C:** Serine is replaced by cysteine at position 396. Mutation disturbs Ferric reductase Trans membrane domain and abolish its function. B. licheniformisexo-small β-lactamase (ESBL) has two no sequential domains and a complex architecture. Due to mutation serine residues 126 and 265 replace with cysteine that change stability and the features of partially folded state [76]. **C453R:** Cysteine is replaced by arginine at position 453. Due to this mutation alternation of charge occur that will repel ligands or other residues with the same charge**.** Mutated recombinant von Willebrand fact prove that the type 2A-like phenotype of von Willebrand disease results due to arginine-552-cysteine (R1315C) mutation. This mutation leads to loss of function due to abnormal folding within the A1 loop of von Willebrand factor [77].

5 Conclusions

In modern genetic analysis, the characterization of disease-associated single nucleotide polymorphisms is very encouraging. Bioinformatics has reduced the cost of genotyping and helped to increase genetic association studies. Hence we conducted a bioinformatics approach to analyze the disease associated nsSNPs of STEAP2 and also reviewed the literature to check the relationship of nsSNPs with prostate cancer but it was not figured out in the literature. Out of 177 non-synonymous mutations, just 42 were predicted as deleterious by all sequence and sequence-structure based softwares and structural analysis results showed that these nsSNPs alters structure and function of STEAP2 as well as its ligand binding site and protein stability and may play their role in the up regulation of STEAP2. So, this study may be a useful tool to predict the effect of nsSNPs of STEAP2 in the upregulation of prostate cancer and may be a help tool to anticipate the consequences of mutations on the gene function.

References

[1] Edwards S., Campbell C., Flohr P., Shipley J., Giddings I., Te-Poele R., Dodson A., Foster C., Clark J., Jhavar, S., Expression analysis onto microarrays of randomly selected cDNA clones highlights HOXB13 as a marker of human prostate cancer, Br. J. Cancer, 2005, 92, 376-381.

[2] Whiteland H., Spencer-Harty S., Morgan C., Kynaston H., Thomas D. H., Bose P., Fenn N., Lewis P., Jenkins S., Doak S.

H., A role for STEAP2 in prostate cancer progression, Clin. Exp. Metastasis, 2014, 31, 909-920.
[3] Wang L., Jin Y., Arnoldussen Y. J., Jonson I., Qu, S., Mælandsmo G. M., Kristian A., Risberg B., Wæhre H., Danielsen H. E., STAMP1 is both a proliferative and an antiapoptotic factor in prostate cancer, Cancer Res., 2010, 70, 5818-5828.
[4] Grunewald T. G., Bach H., Cossarizza A., Matsumoto I., The STEAP protein family: versatile oxidoreductases and targets for cancer immunotherapy with overlapping and distinct cellular functions, Biol. Cell, 2012, 104, 641-657.
[5] Porkka K. P., Helenius M. A., Visakorpi T., Cloning and characterization of a novel six-transmembrane protein STEAP2, expressed in normal and malignant prostate, Lab. Invest., 2002, 82, 1573-1582.
[6] Ohgami R. S., Campagna D. R., McDonald A., Fleming M. D., The Steap proteins are metalloreductases, Blood, 2006, 108, 1388-1394.
[7] Purohit R., Role of ELA region in auto-activation of mutant KIT receptor: a molecular dynamics simulation insight, J. Biomol. Struct. Dyn., 2014, 32, 1033-1046.
[8] Bao L., Cui Y., Prediction of the phenotypic effects of non-synonymous single nucleotide polymorphisms using structural and evolutionary information, Bioinformatics, 2005, 21, 2185-2190.
[9] Calabrese R., Capriotti E., Fariselli P., Martelli P. L., Casadio R., Functional annotations improve the predictive score of human disease-related mutations in proteins, Hum Mutat, 2009, 30, 1237-1244.
[10] Balu K., Purohit R., Mutational analysis of TYR gene and its structural consequences in OCA1A, Gene, 2013, 513, 184-195.
[11] Kamaraj B., Purohit R., In silico screening and molecular dynamics simulation of disease-associated nsSNP in TYRP1 gene and its structural consequences in OCA3, BioMed Res. Int., 2013, http://dx.doi.org/10.1155/2013/697051.
[12] Kamaraj B., Purohit R., Mutational Analysis on Membrane Associated Transporter Protein (MATP) and Their Structural Consequences in Oculocutaeous Albinism Type 4 (OCA4)–A Molecular Dynamics Approach, J. Cell. Biochem., 2016, 117 (11), 2608-2619.
[13] Kumar A., Purohit R., Use of long term molecular dynamics simulation in predicting cancer associated SNPs, PLoS Comput. Biol., 2014, 10, e1003318.
[14] Alwi Z. B., The use of SNPs in pharmacogenomics studies, Malays. J. Med. Sci., 2005, 12(2): 4–12.
[15] Nagamani S., Singh K. D., Muthusamy K., Combined sequence and sequence-structure based methods for analyzing FGF23, CYP24A1 and VDR genes, Meta Gene, 2016, 9, 26-36.
[16] Qian X., Qin L., Xing G., Cao X., Phenotype prediction of pathogenic nonsynonymous single nucleotide polymorphisms in WFS1, Sci. Rep, 2015, 5:14731.
[17] Gioeli D., Mandell J. W., Petroni G. R., Frierson H. F., Weber M. J., Activation of mitogen-activated protein kinase associated with prostate cancer progression, Cancer Res., 1999, 59, 279-284.
[18] Gomes I. M., Maia C. J., Santos C. R., STEAP proteins: from structure to applications in cancer therapy, Mol. Cancer Res., 2012, 10, 573-587.
[19] Adzhubei I., Jordan D.M. and Sunyaev S.R., Predicting functional effect of human missense mutations using PolyPhen-2, Curr. Protoc. Hum. Genet., 2013, 7-20.
[20] Capriotti E F. P., Calabrese R., Casadio R., Predicting protein stability changes from sequences using support vector machines, Bioinformatics, 2005a, 21, doi:10.1093/bioinformatics/bti1109.
[21] Capriotti R. C. R. C., Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information, Bioinformatics, 2006, 22, 2729-2734, doi:10.1093/bioinformatics/btl423.
[22] Jianlin Cheng A. R., Pierre Baldi., Prediction of Protein Stability Changes for Single-Site Mutations Using Support Vector Machines, Proteins, 2006, 62, 1125–1132.
[23] Ng PC, Henikoff S., SIFT: Predicting amino acid changes that affect protein function, Nucleic Acids Res., 2003, 31, 3812-4, doi:10.1093/nar/gkg509.
[24] Kamaraj B., Rajendran V., Sethumadhavan R., Purohit R., In-silico screening of cancer associated mutation on PLK1 protein and its structural consequences, J. Mol. Model., 2013, 19, 5587-5599.
[25] Choi Y., Chan A. P., PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels, Bioinformatics, 2015, 31(16):2745-7..
[26] Ng PC, Henikoff S., Predicting deleterious amino acid substitutions, Genome Res., 2001, 11, 863-74, doi:10.1101/gr.176601.
[27] Kumar A., Purohit R., Computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E, Mutat. Res-Fund. Mol. M., 2012, 738, 28-37.
[28] Naveed M., Kazmi S. K., Anwar F., Arshad F., Dar T. Z., Zafar M., Computational Analysis and Polymorphism study of Tumor Suppressor Candidate Gene-3 for Non Syndromic Autosomal Recessive Mental Retardation, J. Appl. Bioinform. Comput. Biol., 2016, 5, 2. DOI:10.4172/2329-9533.1000127.
[29] Jianyi Yang, W. Z., Baoji He., Sara Elizabeth Walker., Hongjiu Zhang., Brandon Govindarajoo., Jouko Virtanen., Zhidong Xue., Hong-Bin Shen., Yang Zhang., Template-based protein structure prediction in CASP11 and retrospect of I-TASSER in the last decade, Proteins, 2015, doi:10.1002/prot.24918.
[30] Yang J., Yan, R., Roy A., Xu D., Poisson J., Zhang Y., The I-TASSER Suite: protein structure and function prediction, Nat. Methods, 2015, 12, 7-8.
[31] Naveed M., Ahmed I., Khalid N., Mumtaz A. S., Bioinformatics based structural characterization of glucose dehydrogenase (gdh) gene and growth promoting activity of Leclercia sp. QAU-66, Braz. J. Microbiol, 2014, 45, 603-611.
[32] Heo L P. H., Seok C., GalaxyRefine: Protein structure refinement driven by side-chain repacking, Nucleic Acids Res., 2013, W384-8, doi:10.1093/nar/gkt458.
[33] Eisenberg D, Lüthy R., Bowie J. U., VERIFY3D: assessment of protein models with three-dimensional profiles, Method Enzymol, 1997, 277, 396–404, doi:10.1016/S0076-6879(97)77022-8.
[34] Bowie J. U., Lüthy R., Eisenberg D., A method to identify protein sequences that fold into a known three-dimensional structure, Science, 1991, 253, 164-170, doi:10.1126/science.1853201.
[35] Kiran Bhargava R. N., Prahlad Kumar Seth., Kamlesh Kumar Pant., Rakesh Kumar Dixit., Molecular Docking studies of D2 Dopamine receptor with Risperidone derivatives, Bioinformation, 2014, 10, 008-012.

[36] Bosco K HoEmail R. B., The Ramachandran plots of glycine and pre-proline, BMC Struct. Biol., 2005, 5, doi:10.1186/1472-6807-5-14.
[37] Yang J., Roy, A., Zhang Y., BioLiP: a semi-manually curated database for biologically relevant ligand–protein interactions, Nucleic Acids Res., 2013, 41, D1096-D1103.
[38] Gibbons B. J., Hurley T. D., Structure of three class I human alcohol dehydrogenases complexed with isoenzyme specific formamide inhibitors, Biochemistry-US, 2004, 43, 12555-12562.
[39] Lamzin V. S., Dauter Z., Popov V. O., Harutyunyan E. H., Wilson K. S., High resolution structures of holo and apo formate dehydrogenase, J. Mol. Biol., 1994, 236, 759-785.
[40] Nocek B., Chang C., Li H., Lezondra L., Holzle D., Collart F., Joachimiak A., Crystal structures of δ 1-pyrroline-5-carboxylate reductase from human pathogens Neisseria meningitides and Streptococcus pyogenes, J. Mol. Biol., 2005, 354, 91-106.
[41] Read J., Winter V., Eszes C., Sessions R., Brady R., Structural basis for altered activity of M-and H-isozyme forms of human lactate dehydrogenase, Proteins, 2001, 43, 175-185.
[42] Roman D. G., Dancis A., Anderson G. J., Klausner R. D., The fission yeast ferric reductase gene frp1+ is required for ferric iron uptake and encodes a protein that is homologous to the gp91-phox subunit of the human NADPH phagocyte oxidoreductase, Mol. Cell. Biol, 1993, 13, 4342-4350.
[43] Rotrosen D., Yeung C. L., Leto T. L., Malech H. L., Kwong C., Cytochrome b558: the flavin-binding component of the phagocyte NADPH oxidase, Science, 1992, 256, 1459-1462.
[44] Warkentin E., Mamat B., Sordel-Klippert M., Wicke M., Thauer R. K., Iwata M., Iwata S., Ermler U., Shima S., Structures of F420H2: NADP+ oxidoreductase with and without its substrates bound, EMBO J., 2001, 20, 6561-6569.
[45] Lemmink H. H., Nillesen W. N., Mochizuki T., Schröder C. H., Brunner H. G., van Oost B. A., Monnens L., Smeets H., Benign familial hematuria due to mutation of the type IV collagen alpha4 gene, J. Clin. Invest., 1996, 98, 1114.
[46] Hendriks L., van Duijn C., Cras P., Cruts M., Van Hul W., van Harskamp F., Warren A., McInnis M. G., Antonarakis S. E., Martin J.-J., 1992. Presenile dementia and cerebral haemorrhage linked to a mutation at codon 692 of the β-amyloid precursor protein gene, Nat. Genet., 1992, 1, 218-221.
[47] Richards A. J., Yates J. R., Williams R., Payne S. J., Pope F. M., Scott J. D., Snead M. P., A family with Stickler syndrome type 2 has a mutation in the COL11A1 gene resulting in the substitution of glycine 97 by valine in a1 (XI) collagen, Hum. Mol. Genet., 1996, 5, 1339-1343.
[48] Zhou W., Freed C. R., Tyrosine-to-cysteine modification of human α-synuclein enhances protein aggregation and cellular toxicity, J. Biol. Chem., 2004, 279, 10128-10135.
[49] Numa N., Ishida Y., Nasu M., Sohda M., Misumi Y., Noda T., Oda K., Molecular basis of perinatal hypophosphatasia with tissue-nonspecific alkaline phosphatase bearing a conservative replacement of valine by alanine at position 406, FEBS J., 2008, 275, 2727-2737.
[50] Orrico A., Lam C.-W., Galli L., Dotti M. T., Hayek G., Tong S.-F., Poon P. M., Zappella M., Federico A., Sorrentino V., MECP2 mutation in male patients with non-specific X-linked mental retardation, FEBS Lett., 2000, 481, 285-288.
[51] Ellard F. M., Drew J., Blakemore W. E., Stuart D. I., King A. M., Evidence for the role of His-142 of protein 1C in the acid-induced disassembly of foot-and-mouth disease virus capsids, J. Gen. Virol., 1999, 80, 1911-1918.
[52] Russell M. W., Dick M., Collins F. S., Brody L. C., KVLQT1 mutations in three families with familial or sporadic long QT syndrome, Hum. Mol. Genet., 1996, 5, 1319-1324.
[53] Ries S., Büchler C., Schindler G., Aslanidis C., Ameis D., Gasche C., Jung N., Schambach A., Fehringer P., Vanier M. T., Different missense mutations in histidine-108 of lysosomal acid lipase cause cholesteryl ester storage disease in unrelated compound heterozygous and hemizygous individuals, Hum. Mutat., 1998, 12(1):44-51.
[54] Adamson J. W., Hayashi A., Stamatoyannopoulos G., Burger W. F., Erythrocyte function and marrow regulation in hemoglobin Bethesda (β145 histidine), J. Clin. Inv., 1972, 51 (11), 2883-2888.
[55] Imamura T., Fujita S., Ohta Y., Hanada M., Yanase T., Hemoglobin Yoshizuka (G10 (108) β asparagine→ aspartic acid): a new variant with a reduced oxygen affinity from a Japanese family, J. Clin. Inv., 1969, 48(12): 2341–2348.
[56] Fisher R., Mailly F., Peacock R., Hamsten A., Seed M., Yudkin J., Beisiegel U., Feussner G., Miller G., Humphries S., Interaction of the lipoprotein lipase asparagine 291--> serine mutation with body mass index determines elevated plasma triacylglycerol concentrations: a study in hyperlipidemic subjects, myocardial infarction survivors, and healthy adults, J. Lipid Res., 1995, 36, 2104-2112.
[57] Douglas C. M., Collier R., Exotoxin A of Pseudomonas aeruginosa: substitution of glutamic acid 553 with aspartic acid drastically reduces toxicity and enzymatic activity, J. Bacteriol., 1987, 169, 4967-4971.
[58] Miyata T., Iwanaga S., Sakata Y., Aoki N., Plasminogen Tochigi: inactive plasmin resulting from replacement of alanine-600 by threonine in the active site, Proc. Natl. A. Sci., 1982, 79, 6132-6136.
[59] Nauseef W. M., Brigham S., Cogley M., Hereditary myeloperoxidase deficiency due to a missense mutation of arginine 569 to tryptophan, J. Biol. Chem., 1994, 269, 1212-1216.
[60] Stern L. J., Khorana H., Structure-function studies on bacteriorhodopsin. X. Individual substitutions of arginine residues by glutamine affect chromophore formation, photocycle, and proton translocation, J. Biol. Chem., 1989, 264, 14202-14208.
[61] Koide T., Odani S., Takahashi K., Ono T., Sakuragawa N., Antithrombin III Toyama: replacement of arginine-47 by cysteine in hereditary abnormal antithrombin III that lacks heparin-binding ability, Proc. Natl. A. Sci., 1984, 81, 289-293.
[62] Taube R., Fujinaga K., Irwin D., Wimmer J., Geyer M., Peterlin B. M., Interactions between equine cyclin T1, Tat, and TAR are disrupted by a leucine-to-valine substitution found in human cyclin T1, J. Virol., 2000, 74, 892-898.
[63] Jacobs A., Roosendaal B., van Breemen J., De Graaf F., Role of phenylalanine 150 in the receptor-binding domain of the K88 fibrillar subunit, J. Bacteriol., 1987, 169, 4907-4911.
[64] Takeda-Shitaka M., Umeyama H., Effect of exceptional valine replacement for highly conserved alanine-55 on the catalytic site structure of chymotrypsin-like serine protease, Chem. Pharm. Bull., 1998, 46, 1343-1348.

[65] Weaver T., Lees M., Banaszak L., Mutations of fumarase that distinguish between the active site and a nearby dicarboxylic acid binding site, Protein Sci., 1997, 6, 834-842.
[66] Gaume B., Sharp R., Manson F., Chapman S., Reid G., Lederer F., Mutation to glutamine of histidine 373, the catalytic base of flavocytochrome b 2 (L-lactate dehydrogenase), Biochimie, 1995, 77, 621-630.
[67] Zhu J., Reith M., Role of dopamine transporter in the action of psychostimulants, nicotine, and other drugs of abuse, CNS. Neurol. Disord. Drug Targets, 2008, 7, 393.
[68] Wellner M., Monden I., Mueckler M. M., Keller K., Functional consequences of proline mutations in the putative transmembrane segments 6 and 10 of the glucose transporter GLUT1, Eur. J. Biochem., 1995, 227, 454-458.
[69] Nuytinck L., Tükel T., Kayserili H., Apak M. Y., De Paepe A., Glycine to tryptophan substitution in type I collagen in a patient with OI type III: a unique collagen mutation, J. Med. Genet., 2000, 37, 371-375.
[70] Deak S., Scholz P. M., Amenta P., Constantinou C., Levi-Minzi S., Gonzalez-Lavin L., Mackenzie J., The substitution of arginine for glycine 85 of the alpha 1 (I) procollagen chain results in mild osteogenesis imperfecta. The mutation provides direct evidence for three discrete domains of cooperative melting of intact type I collagen, J. Biol. Chem., 1991, 266, 21827-21832.
[71] Podoly E., Hanin G., Soreq H., Alanine-to-threonine substitutions and amyloid diseases: Butyrylcholinesterase as a case study, Chem-Biol. Interact., 2010, 187, 64-71.
[72] Prior L., Bordet S., Trifiro M. A., Mhatre A., Kaufman M., Pinsky L., Wrogeman K., Belsham D. D., Pereira F., Greenberg C., Replacement of arginine 773 by cysteine or histidine in the human androgen receptor causes complete androgen insensitivity with different receptor phenotypes, Am. J. Hum. Genet., 1992, 51(1):143-55.
[73] Newbold R. J., Deery E. C., Walker C. E., Wilkie S. E., Srinivasan N., Hunt D. M., Bhattacharya S. S., Warren M. J., The destabilization of human GCAP1 by a proline to leucine mutation might cause cone-rod dystrophy, Hum. Mol. Genet., 2001, 10, 47-54.
[74] Deak S., Scholz P. M., Amenta P., Constantinou C., Levi-Minzi S., Gonzalez-Lavin L., Mackenzie J., The substitution of arginine for glycine 85 of the alpha 1 (I) procollagen chain results in mild osteogenesis imperfecta. The mutation provides direct evidence for three discrete domains of cooperative melting of intact type I collagen, J. Biol. Chem., 1991, 266, 21827-21832.
[75] Narcisi P., Wu Y., Tromp G., Earley J. J., Richards A. J., Pope F. M., Kuivaniemi H., Single base mutation that substitutes glutamic acid for glycine 1021 in the COL3A1 gene and causes Ehlers–Danlos syndrome type IV, Am. J. Hum. Genet, 1993, 46, 278-283.
[76] Santos J., Risso V. A., Sica M. P., Ermácora M. R., Effects of serine-to-cysteine mutations on β-lactamase folding, Biophys. J., 2007, 93, 1707-1718.
[77] Ribba A.-S., Hilbert L., Lavergne J.-M., Fressinaud E., Boyer-Neumann C., Ternisien C., Juhan-Vague I., Goudemand J., Girma J.-P., Mazurier C., The arginine-552-cysteine (R1315C) mutation within the A1 loop of von Willebrand factor induces an abnormal folding with a loss of function resulting in type 2A–like phenotype of von Willebrand disease: study of 10 patients and mutated recombinant von Willebrand factor, Blood, 2001, 97, 952-959.

Supplemental Material: The online version of this article (DOI: 10.1515/biol-2016-0054) offers supplementary material.