In silico Analysis of the Functional and Structural Impacts of Non-synonymous Single Nucleotide Polymorphisms in the Human Paraxonase 1 Gene

Sudip Paul*, Md. Solayman, Moumoni Saha, Md. Sabir Hossain

Department of Biochemistry and Molecular Biology
Jahangirnagar University, Savar
Dhaka-1342, Bangladesh
E-mails: sudip@juniv.edu, naeem40thju@gmail.com, moumonisaha@gmail.com, sabiriu@gmail.com

Corresponding author

Received: May 9, 2015 **Accepted: September 23, 2015**

Published: September 30, 2015

Abstract: *Computational approaches could help in identifying deleterious non-synonymous single nucleotide polymorphisms (nsSNPs) in a disease related gene which is a difficult and laborious task through laboratory experiments. In the present study, we analyzed the impacts of nsSNPs on structure and function of Paraxonase 1 (PON1) using different bioinformatics tools. The human PON1 protein sequence and its corresponding gene's SNP information were collected from UniProt and dbSNP databases, respectively. We utilized SIFT, Polyphen, I-Mutant 2.0, MutPred, SNP & GO, PhD-SNP and PANTHER tools in order to examine the total 39 nsSNPs occurring in the PON1 coding region. We filtered the most pathological mutations by combining the scores of the aforementioned servers and found 8 SNPs (G344C, S302L, W281C, D279Y, H134R, F120S, L90P, C42R) as deleterious and disease causing. The PDB structure of PON1 protein was obtained from RCSB Protein Data Bank (PDB ID: 1V04). The deleterious SNPs in native PON1 were introduced using Swiss-PDB Viewer package and changes in free energy were observed for six out of eight mutant structures. Two SNPs, S302L (substitution of serine to leucine at 302 position in amino acid sequence) and L90P (substitution of leucine to proline at 90 position in amino acid sequence) caused the highest energy increase amongst all. The findings implicate that these nsSNPs would be analyzed further in detail to enumerate their possible association with the protein deteriorating and disease causal potentialities.*

Keywords: *SNP, PON1, Missense, Bioinformatics, Energy change.*

Introduction

Single nucleotide polymorphisms (SNPs) or single nucleotide DNA sequence variations are recently being studied exclusively because of their possible association with human complex diseases. SNPs make up about 90% of all human genetic variation, occurring every 100-300 bases along the 3-billion-base human genome, although their density vary between regions [14]. There are various types of SNPs, among these non-synonymous SNPs (nsSNPs) are one of the most important ones because they cause an amino acid alteration on the protein sequences and thereby can have an intense impact on protein structures and functions [12]. These nsSNPs affect gene regulation by shifting DNA and transcriptional binding factors and the preservation of the structural integrity of cells and tissues [3, 29]. Also, nsSNPs affect the functional roles of proteins in the signal transduction of visual, hormonal and other stimulants [8, 27].

There are a number of SNPs identified till date. Among these, identifying SNPs of likely functional importance still remains as a difficult task as requiring multiple testing of hundreds

or thousands of SNPs in candidate genes [22]. To overcome these limitations and serve as a complementary category of these traditional statistical methods, computational approaches that rely on properties of variants instead of experimental data of patients have been designed for the detection of deleterious variants with the growing functional annotations of the human genome sequence. Although, such methods may never be accurate enough to replace wet-lab experiments, they might be help in identifying and prioritizing a small number of susceptible and tractable candidate nsSNPs from pools of available data [19]. Recent studies have shown that computational methods are capable of well estimating the functional effects of nsSNPs [35].

Numbers of genes have been studied for SNP analysis to explore their plausible association with various diseases, PON1 gene is one of them. PON1 is one of three paraxonase gene family members, located in a gene cluster on chromosome 7q21.3-22.1 [10]. All of the paraxonases have antioxidant activities [24]. PON1 and PON3 share similar functions in association with high density lipoprotein (HDL) as described previously; however, PON3 has lower expression levels [22]. PON1 is the most abundant form and hence extensively investigated. Human PON1 (HuPON1) consists of 355 amino acids exclusively associated with HDL in association with human phosphate binding protein (HPBP). ApoA1 is major protein in HDL which stabilizes PON1 and binds it with very high affinity [18]. HuPON1 plays a major role in the prevention of atherosclerosis by protecting HDL and low density lipoprotein (LDL) against oxidative stress mediated through the uptake of oxidized-LDL by macrophages, inhibition of macrophage cholesterol biosynthesis and stimulation of HDL mediated cholesterol efflux from macrophages [2]. The low serum paraoxonase activity in type 2 diabetes mellitus was recently shown to be correlated with the levels of oxidized LDL and vascular complications [31]. Polymorphisms in the PON1 gene have been investigated with respect to their association with various human diseases linked to oxidative stress such as coronary heart disease, Parkinson's disease, type 2 diabetes mellitus and inflammatory bowel disease [16], but the findings are inconsistent. However, a polymorphism of the PON1 gene that causes reduction in enzymatic activity, Q192R was found to be significantly associated with increased risk of heart diseases [34].

Although there are presently several published articles describing the association of SNPs in the HuPON1 gene with different types of diseases, computational analysis has not yet been undertaken on the functional and structural consequences of nsSNPs in this gene. In the current study, we employed different publicly available bioinformatics tools and databases for a comprehensive analysis of nsSNPs in PON1 gene. As the majority of disease mutations affect protein stability [30, 33], we also proposed modeled protein structures for the mutant proteins and compared them with the native protein in order to evaluate stability changes.

Materials and methods

Collection of PON1 SNP dataset

The information about SNPs of PON1 gene of *Homo sapiens* was obtained from the db-SNP (http://www.ncbi.nlm.nih.gov/SNP/) [26] for further computational analysis.

Assessment of the functional impacts of deleterious nsSNPs using a sequence homology-based method (SIFT)

The functional impacts of the nsSNPs were further analyzed using SIFT (http://sift.jcvi.org) [20]. The SIFT program envisages deleterious or non-tolerated SNPs on the principle that some amino acids have a propensity to be conserved in a protein family and any substitution at these positions would influence protein function and thus have a phenotypic effect.

SIFT calculates the normalized probability in terms of SIFT score or tolerance index (TI) score for each mutation. The substitutions with normalized probabilities ≤ 0.05 are predicted to be non-tolerated or deleterious amino acids substitutions, whereas those > 0.05 are considered to be tolerated.

Investigation of the functional consequences of coding nsSNPs using structure homology-based method (PolyPhen)

To search the possible effect of an amino acid substitution on the structure and function of PON1 protein, PolyPhen V2 (http://genetics.bwh.harvard.edu/pph2) [1] server was used. The protein sequence with mutational position and two amino acid variants were submitted to the server. PolyPhen generates multiple sequence alignment of homologous protein structures, calculates the position-specific independent counts (PSIC) scores for each of the two variants, and then calculates the PSIC score difference between both the allelic variants. The higher the PSIC score difference, the higher the functional impacts a particular amino acid substitution is likely to have or the more likely it is to be damaging. The PolyPhen server classifies nsSNPs into three main categories, benign, possibly damaging, or probably damaging, and provides the corresponding specificity and sensitivity values.

Analysis of the nature of non-synonymous mutations by MutPred

The MutPred server [15] was employed to classify an amino acid substitution (AAS) as disease-associated or neutral. In addition, it predicts molecular cause of disease/deleterious AAS. MutPred is based upon SIFT and a gain/loss of 14 different structural and functional properties. The output of MutPred contains a general score (g), i.e., the probability that the amino acid substitution is deleterious/disease-associated, and top 5 property scores (p), where p is the P-value that certain structural and functional properties are impacted.

Analysis of the effects of nsSNPs on the protein stability by I-Mutant 2.0

I-Mutant 2.0 is a SVM based tools i.e., support vector machine based tool that leads to automatic protein stability change prediction which is caused by single point mutation [6]. The initiations were done either by using protein structure or more precisely from the protein sequence. The output is a free energy change value ($\Delta\Delta G$). Positive $\Delta\Delta G$ value infers that the protein being mutated is of higher stability and vice versa is also true.

Prediction of disease related nsSNPs by SNPs & GO

SNPs & GO [5] is also a support vector machine (SVM) based on the method to accurately predict the mutation related to disease from protein sequence. The input is the FASTA sequence of the whole protein, the output is based on the difference among the neutral and disease related variations of the protein sequence. The RI (reliability index) with value of greater than 5 depicts the disease related effect caused by mutation on the function of parent protein. The PHD SNP [7] and PANTHER [28] algorithms were also used in the display of output.

Modeling of nsSNPs on protein structures and calculation of their RMSD difference

Structural analysis was executed to evaluate and compare the stability of native and mutant structures. The highest resolution (2.20 Å) native structure of the HuPON1 protein available in the Protein Data Bank (PDB) [4] has an ID of 1V04 [11]. The amino acid residue substitutions were carried out using the Swiss-PDB viewer [9], followed by energy minimization of the modeled 3D structures using a version of the GROMOS 43B1 force field in GROMOS96 software package embraced in the Swiss-PDB viewer [32]. The comparison

between the resulting native and modeled structures was made by the calculation of the potential energy and RMSD values using UCFS Chimera 1.8.1 [21].

Results and discussion

Retrieval of SNPs

The dbSNP was utilized for retrieving the SNPs in the human PON1 gene using the gene ID: 5444. A total of 65 SNPs were found in the coding region, among them 21 were synonymous, 39 non-synonymous and missense, 4 non-synonymous and nonsense, and 1 frame-shift mutations (Fig. 1). Only missense non-synonymous coding SNPs were chosen for further analysis.

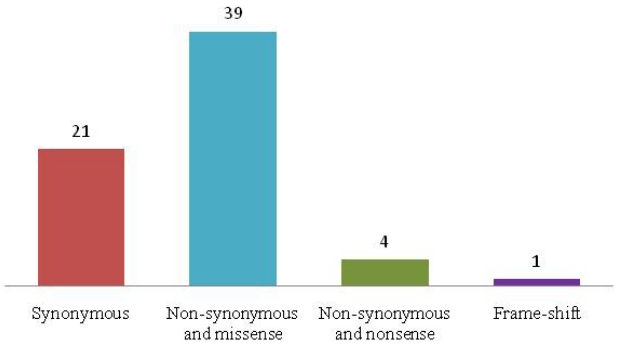

Fig. 1 Single nucleotide polymorphisms in the coding region of the PON1 gene retrieved from the dbSNP database

Prediction of tolerated and deleterious SNPs

When the SNPs were submitted to the SIFT program for predicting their effect on protein function, out of the 39 SNPs screened, 17 variants were found to be damaging and others as tolerated. Among the SNPs analyzed, SIFT did not predict the effect of one SNP (rs 149100710; E49K) on the function of PON1. The detailed result has been depicted in Table 1.

Damaged nsSNPs by PolyPhen server

All the 39 missense nsSNPs submitted to SIFT were also submitted to the PolyPhen server. 14 out of 39 SNPs were considered to be probably damaging and exhibited a range of PSIC score difference between 0.76 and 1.00 (Table 2). Six of them were found as possibly damaging and others as benign. It can be seen from Table 2 that there was significant correlation between the results obtained from the evolutionary-based approach SIFT and the structural based approach PolyPhen. Out of the total 14 SNPs predicted as probably damaging by PolyPhen also detected as damaging by SIFT suggesting that these nsSNPs may disrupt both the protein function and structure. The SNPs were also analyzed by the MutPred server and found a strong correlation between the results obtained from the PolyPhen and MutPred servers.

Damaging nsSNPs found by I-Mutant 2.0

I-Mutant 2.0 is an online server used to predict stability of the induced mutations in protein structure. The results for the inputs of all 39 missense SNPs are given in Table 2. The results are predicted to be either increase or decrease of the free energy change upon mutation. 35 out of 39 SNPs screened were found cause a decrease in the free energy.

Table 1. List of non-synonymous SNPs of the human PON1 gene analyzed by SIFT

rs ID	SNPs		SIFT	
	SNP	Amino acid change	Prediction	SIFT score
rs 368206333	G/T	G344C	DAMAGING	0
rs 141598837	A/G	K340R	TOLERATED	0.09
rs 145997673	G/A	G330S	DAMAGING	0
rs 372449149	C/T	T318I	TOLERATED	0.06
rs 185623242	C/T	S302L	DAMAGING	0
rs 199693212	T/C	F292S	TOLERATED	0.24
rs 148911901	T/A	M289K	TOLERATED	0.29
rs 369422555	G/C	W281C	DAMAGING	0
rs 72552786	G/T	D279Y	DAMAGING	0
rs 368248410	A/G	I271V	TOLERATED	0.13
rs 371803280	G/A	V268M	DAMAGING	0
rs 548299742	G/A	H246R	TOLERATED	0.17
rs 564064745	A/G	D231N	TOLERATED	0.11
rs 370355032	C/T	P210S	TOLERATED	0.13
rs 80019660	C/T	A201V	TOLERATED	0.55
rs 13306698	A/G	R160G	DAMAGING	0
rs 112078575	A/G	K151R	TOLERATED	0.68
rs 536888659	G/A	H134R	DAMAGING	0
rs 202062288	G/T	M127I	TOLERATED	0.29
rs 144390653	T/G	M127R	DAMAGING	0
rs 148785172	G/A	A126T	TOLERATED	1
rs 189946844	A/T	E123V	DAMAGING	0.02
rs 147867887	C/T	T121I	TOLERATED	0.91
rs 368620674	T/C	F120S	DAMAGING	0
rs 72552787	A/G	I102V	TOLERATED	0.19
rs 532844853	C/G	L100F	DAMAGING	0
rs 72552788	T/C	L90P	DAMAGING	0
rs 367566813	T/C	M88T	DAMAGING	0
rs 371338407	C/G	P79R	DAMAGING	0.02
rs 199616322	C/T	P59S	TOLERATED	0.22
rs 149100710	G/A	E49K	NOT PREDICTED	-
rs 144612002	A/G	I48V	TOLERATED	0.44
rs 138512790	T/C	C42R	DAMAGING	0
rs 141665531	C/T	P40L	TOLERATED	0.11
rs 551653548	A/G	R27Q	TOLERATED	1
rs 146211440	T/G	S23A	DAMAGING	0.03
rs 141948033	A/G	N19D	TOLERATED	0.16
rs 201783178	A/G	R18G	TOLERATED	1
rs 150657027	C/T	A6V	TOLERATED	1

Damaging nsSNPs found by SNPs&GO, PHD-SNP and PANTHER

SNPs&GO is a server for the prediction of single point protein mutations likely to be involved in the insurgence of diseases in humans. The results of SNPs & GO for the inputs of all 39 nsSNPs are given in Table 3. The results are displayed in terms of neutral or disease causing mutation. It was found that out of 39, 9 mutations were having disease causing abilities while the rests were neutral.

PHD-SNP is a SVM-based classifier in the newer version of which a predictor was developed based on a single SVM trained and tested on protein sequence and profile information. The results acquired from this server for all nsSNPs are given in Table 3. In this server, the results are also given in the form of neutral or disease causing mutations. It was found that 15 (out of 39) of the mutations were deleterious while the rests as neutral (Table 3).

Table 2. PolyPhen, MutPred and I-Mutant predictions for non-synonymous SNPs of the human PON1 gene

		PolyPhen				MutPred Probability of Deleterious Mutation	I-Mutant	
rs ID	AA change	PSID Score	Sensitivity	Specificity	Prediction		RI	Stability
rs 368206333	G344C	1.000	0.00	1.00	PD	0.903	6	Decrease
rs 141598837	K340R	0.107	0.93	0.86	Benign	0.509	5	Decrease
rs 145997673	G330S	0.997	0.41	0.98	PD	0.631	3	Decrease
rs 372449149	T318I	0.591	0.87	0.91	PD	0.443	5	Decrease
rs 185623242	S302L	1.000	0.00	1.00	PD	0.674	1	Decrease
rs 199693212	F292S	0.044	0.94	0.83	Benign	0.735	9	Decrease
rs 148911901	M289K	0.0000	1.00	0.00	Benign	0.511	7	Decrease
rs 369422555	W281C	1.000	0.00	1.00	PD	0.684	8	Decrease
rs 72552786	D279Y	1.000	0.00	1.00	PD	0.727	1	Increase
rs 368248410	I271V	0.033	0.95	0.82	Benign	0.525	5	Decrease
rs 371803280	V268M	0.990	0.72	0.97	PD	0.616	6	Decrease
rs 548299742	H246R	0.083	0.93	0.85	Benign	0.725	4	Decrease
rs 564064745	D231N	0.924	0.81	0.94	PD	0.799	8	Decrease
rs 370355032	P210S	0.628	0.87	0.91	PD	0.816	5	Decrease
rs 80019660	A201V	0.916	0.81	0.94	PD	0.322	3	Decrease
rs 13306698	R160G	0.916	0.81	0.94	PD	0.494	8	Decrease
rs 112078575	K151R	0.169	0.97	0.87	Benign	0.294	9	Decrease
rs 536888659	H134R	1.000	0.00	1.00	PD	0.719	4	Decrease
rs 202062288	M127I	0.0000	1.00	0.00	Benign	0.438	5	Decrease
rs 144390653	M127R	0.0000	1.00	0.00	Benign	0.622	4	Decrease
rs 148785172	A126T	0.0000	1.00	0.00	Benign	0.436	3	Decrease
rs 189946844	E123V	0.036	0.94	0.82	Benign	0.324	7	Decrease
rs 147867887	T121I	0.0000	1.00	0.00	Benign	0.344	4	Decrease
rs 368620674	F120S	0.997	0.41	0.98	PD	0.738	7	Decrease
rs 72552787	I102V	0.001	0.99	0.15	Benign	0.480	3	Increase
rs 532844853	L100F	1.000	0.00	1.00	PD	0.656	8	Decrease
rs 72552788	L90P	1.000	0.00	1.00	PD	0.718	3	Decrease
rs 367566813	M88T	0.508	0.88	0.90	PD	0.617	9	Decrease
rs 371338407	P79R	0.760	0.85	0.92	PD	0.464	6	Decrease
rs 199616322	P59S	0.004	0.97	0.59	Benign	0.611	4	Decrease
rs 149100710	E49K	0.995	0.68	0.97	PD	0.479	6	Decrease
rs 144612002	I48V	0.0000	1.00	0.00	Benign	0.381	7	Decrease
rs 138512790	C42R	1.000	0.00	1.00	PD	0.889	8	Decrease
rs 141665531	P40L	0.004	0.97	0.59	Benign	0.456	5	Decrease
rs 551653548	R27Q	1.000	0.00	1.00	PD	0.665	7	Decrease
rs 146211440	S23A	0.0000	1.00	0.00	Benign	0.234	1	Decrease
rs 141948033	N19D	0.0000	1.00	0.00	Benign	0.377	2	Increase
rs 201783178	R18G	0.0000	1.00	0.00	Benign	0.472	6	Decrease
rs 150657027	A6V	0.0000	1.00	0.00	Benign	0.338	4	Increase

AA – amino acid, PD – probably damaging

PANTHER server was also utilized in the present study. Out of total 39 SNPs, the server did not give any prediction about the disease causing ability of 10. Among the 29 predicted, 9 were found to be disease causing and the others as neutral (Table 3).

Table 3. List of non-synonymous SNPs analyzed for disease association by SNP&GO, PHD-SNP and PANTHER

AA change	PHD-SNP			PANTHER			SNPs & GO		
	Prediction	RI	Probability	Prediction	RI	Probability	Prediction	RI	Probability
G344C	Disease	8	0.877	Disease	7	0.850	Disease	5	0.766
K340R	Neutral	4	0.284	Neutral	6	0.200	Neutral	9	0.064
G330S	Disease	1	0.540	Neutral	3	0.352	Neutral	3	0.359
T318I	Neutral	5	0.264	Neutral	1	0.437	Neutral	8	0.118
S302L	Disease	3	0.641	Disease	7	0.865	Disease	1	0.539
F292S	Neutral	2	0.410	Neutral	3	0.371	Neutral	4	0.279
M289K	Disease	2	0.623	Neutral	7	0.145	Neutral	5	0.270
W281C	Disease	6	0.800	Disease	7	0.830	Disease	4	0.714
D279Y	Disease	8	0.886	Disease	6	0.801	Disease	5	0.759
I271V	Neutral	8	0.108	Neutral	7	0.139	Neutral	9	0.030
V268M	Neutral	4	0.276	Neutral	1	0.465	Neutral	8	0.087
H246R	Neutral	3	0.361	Neutral	2	0.383	Neutral	5	0.273
D231N	Disease	2	0.579	Disease	1	0.536	Neutral	2	0.410
P210S	Disease	3	0.666	Disease	0	0.513	Disease	0	0.512
A201V	Neutral	2	0.376	Neutral	5	0.250	Neutral	6	0.189
R160G	Neutral	8	0.123	Neutral	1	0.453	Neutral	9	0.054
K151R	Neutral	9	0.036	Neutral	7	0.147	Neutral	10	0.005
H134R	Disease	5	0.764	Disease	6	0.793	Disease	5	0.734
M127I	Neutral	8	0.098	Neutral	7	0.175	Neutral	9	0.044
M127R	Disease	1	0.531	Neutral	4	0.289	Neutral	5	0.260
A126T	Neutral	9	0.045	Neutral	2	0.390	Neutral	10	0.019
E123V	Neutral	2	0.412	Neutral	1	0.435	Neutral	5	0.274
T121I	Neutral	8	0.084	Neutral	4	0.294	Neutral	9	0.026
F120S	Disease	2	0.597	Disease	3	0.632	Disease	2	0.591
I102V	Neutral	8	0.092	Neutral	6	0.204	Neutral	9	0.030
L100F	Neutral	3	0.330	Disease	3	0.654	Neutral	6	0.193
L90P	Disease	3	0.626	Disease	8	0.917	Disease	1	0.531
M88T	Neutral	4	0.278	Neutral	2	0.391	Neutral	7	0.154
P79R	Neutral	1	0.459	Neutral	1	0.455	Neutral	5	0.234
P59S	Neutral	8	0.106	NP	-	-	Neutral	10	0.020
E49K	Disease	1	0.533	NP	-	-	Neutral	7	0.142
I48V	Neutral	9	0.071	NP	-	-	Neutral	10	0.020
C42R	Disease	7	0.859	NP	-	-	Disease	1	0.573
P40L	Neutral	5	0.230	NP	-	-	Neutral	9	0.052
R27Q	Disease	3	0.627	NP	-	-	Neutral	6	0.224
S23A	Neutral	9	0.059	NP	-	-	Neutral	10	0.017
N19D	Neutral	7	0.133	NP	-	-	Neutral	9	0.066
R18G	Neutral	8	0.083	NP	-	-	Neutral	9	0.035
A6V	Neutral	9	0.028	NP	-	-	Neutral	10	0.009

AA – amino acid, NP – not predicted

By combining the predictions of SIFT, PolyPhen-2, I-Mutant 2.0, MutPred, SNPs&GO, PANTHER and PHD-SNP servers, eight SNPs (G344C, S302L, W281C, D279Y, H134R, F120S, L90P and C42R) were found to be more deleterious and disease associated (Table 4). These functionally significant variants were further superimposed with the native protein structure.

Structural analysis of mutant structures

The eight predicted deleterious and disease causal SNPs (rs 368206333, rs 185623242, rs 369422555, rs 72552786, rs 536888659, rs 368620674, rs 72552788 and rs 138512790) were mapped to the PDB ID 1VO4 native structure. The amino acid residue substitutions

were performed by Swiss-Pdb Viewer independently to get eight mutant modeled structures (1VO4 G344C, 1VO4 S302L, 1VO4 W281C, 1VO4 D279Y, 1VO4 H134R, 1VO4 F120S, 1VO4 L90P and 1VO4 C42R, respectively). Then, energy minimizations were performed for the native structure (1VO4) and the mutant modeled structures.

Table 4. Total energy and RMSD of native and mutant modeled structures of PON1 protein

rs ID	Amino acid change	Total energy after minimization, (KJ/mol)	RMSD between native and mutant structures, (Å)
rs 368206333	G344C	-6290.605	0.000
rs 185623242	S302L	-3747.879	0.011
rs 369422555	W281C	-6967.503	0.000
rs 72552786	D279Y	-7043.176	0.000
rs 536888659	H134R	-7338.226	0.000
rs 368620674	F120S	-7089.467	0.000
rs 72552788	L90P	-5486.519	0.044
rs 138512790	C42R	-6888.083	0.000

Total energy of model structure (1V04) after energy minimization: -7082.146 KJ/mol.

The total free energy for the native structure (1VO4) and the eight mutant modeled structures were given in the Table 4. Six out of eight mutant modeled structures showed an increase in free energy (less favorable change) in comparison with the native structure. The mutant models 1VO4 S302L and 1V04 L90P showed the greatest increase in energy, which may be explained by the energetically unfavorable substitution of serine to leucine and leucine to proline, respectively. Among these, substitution of leucine to proline in a protein structure have been found to cause significant reduction in protein stability associated with different diseases such as neuroblastoma, Parkinson's disease, etc. [13, 17]. The remaining amino acid substitutions may not cause significant destabilization of the protein structure and hence show less energy change compared to others. Figs. 1-3 represent the successful creation of the mutations S302L and L90P in the PON1 native protein structure using Swiss-PDB viewer. The images were captured by the PyMOL Molecular Graphics System Version 1.3 [25].

It can be seen from Table 3 that the RMSD values between the native structure and the mutant modeled structures are all similar, ranging from 0.000-0.0444 Å. Because these values are low, we can suggest that these mutations do not cause a significant change in the mutant structures with respect to the native protein structure.

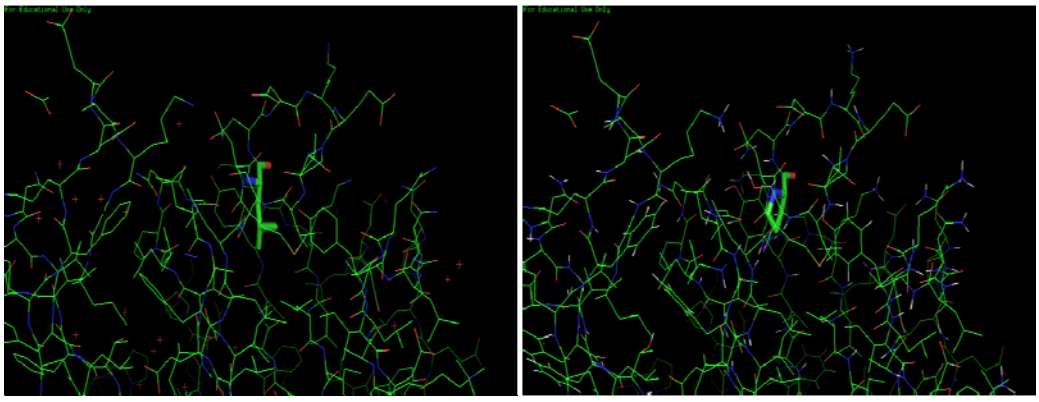

Fig. 2 The native protein structure with Leucine (90) and mutant protein structure with Proline (90) for the SNP L90P (rs 72552788) of human Paraxonase 1 (PDB ID: 1V04)

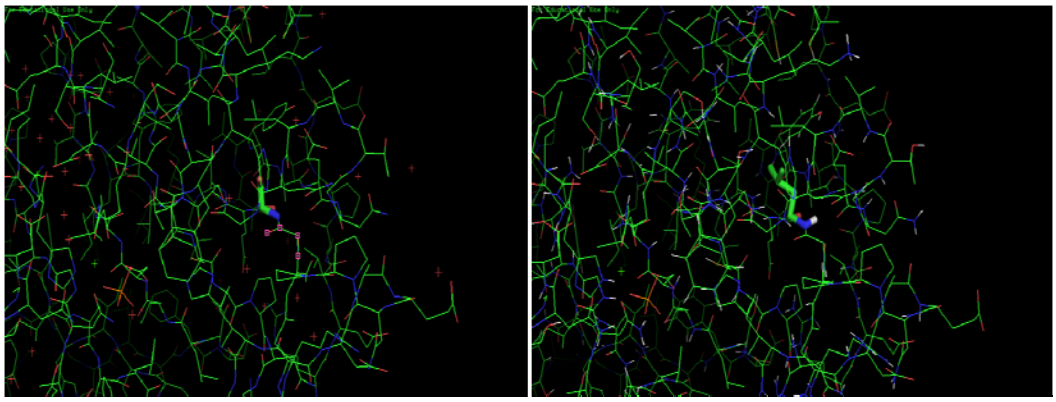

Fig. 3 The native protein structure with Serine (302) and mutant protein structure with Leucine (302) for the SNP S302L (rs 185623242) of human Paraxonase 1 (PDB ID: 1V04)

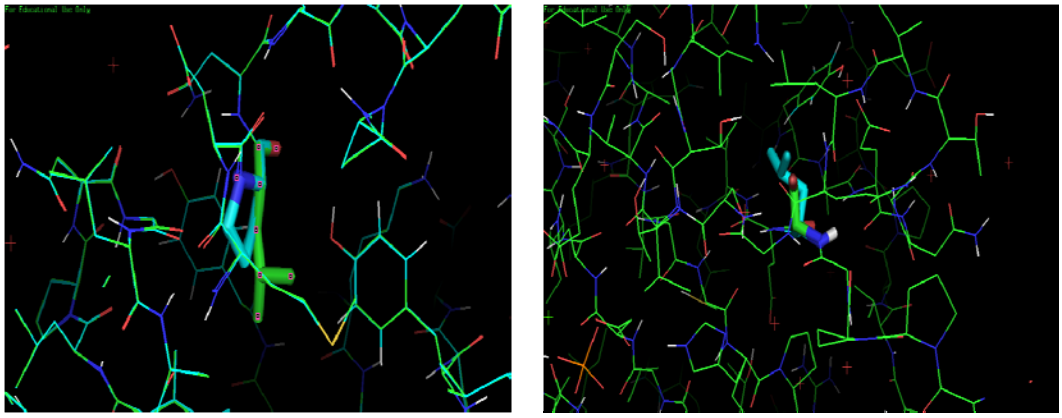

Fig. 4 The superimposed structures of the native protein with the mutant protein with the SNP L90P (rs 72552788) and the mutant protein with the SNP S302L (rs 185623242) of Paraxonase 1 (PDB ID: 1V04)

Conclusion

In this study, we investigated the functional and structural impacts of SNPs in the PON1gene using computational prediction tools. We found 39 nsSNPs in the protein coding region of PON1 gene from the dbSNP database. Out of these missense nsSNPs, eight SNPs were found to be deleterious and disease causing by SIFT, PolyPhen, MutPred, I-mutant, PHD-SNP, PANTHER, SNP&GO. Furthermore, the structural analysis results showed that the amino acid residue substitutions which had the greatest impact on the stability of the PON1 protein were mutations 1V04 L90P (rs 72552788) and 1V04 S302L (rs 185623242). Substitution of leucine by proline has been found to be associated with various diseases and particularly it may cause a significant decline in protein stability. Based on our findings, we can conclude that these SNPs should be considered as important candidates in causing diseases related to PON1 malfunction.

Abbreviations of amino acids

A: Alanine, **C**: Cysteine, **D**: Aspartic Acid, **E**: Glutamic Acid, **F**: Phenylalanine, **G**: Glycine, **H**: Histidine, **I**: Isoleucine, **K**: Lysine, **L**: Leucine, **M**: Methionine, **N**: Asparagine, **P**: Proline, **Q**: Glutamine, **R**: Arginine, **S**: Serine, **T**: Threonine, **V**: Valine, **W**: Tryptophan, **Y**: Tyrosine.

Acknowledgements

Authors greatly acknowledge the support provided by the Department of Biochemistry and Molecular Biology of Jahangirnagar University, Savar, Dhaka-1342, Bangladesh.

References

1. Adzhubei I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov, S. R. Sunyaev (2010). A Method and Server for Predicting Damaging Missense Mutations, Nature Methods, 7(4), 248-249.
2. Aviram M., M. Rosenblat (2004). Paraoxonases 1, 2, and 3, Oxidative Stress, and Macrophage Foam Cell Formation during Atherosclerosis Development, Free Radical Biology and Medicine, 37(9), 1304-1316.
3. Barroso I., M. Gurnell, V. Crowley, M. Agostini, J. Schwabe, M. Soos, G. L. Maslen, T. Williams, H. Lewis, A. Schafer (1999). Dominant Negative Mutations in Human PPARγ Associated with Severe Insulin Resistance, Diabetes Mellitus and Hypertension, Nature, 402(6764), 880-883.
4. Berman H. M., J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne (2000). The Protein Data Bank, Nucleic Acids Research, 28(1), 235-242.
5. Calabrese R., E. Capriotti, P. Fariselli, P. L. Martelli, R. Casadio (2009). Functional Annotations Improve the Predictive Score of Human Disease-related Mutations in Proteins, Human Mutation, 30(8), 1237-1244.
6. Capriotti E., P. Fariselli, R. Casadio (2005). I-Mutant2.0: Predicting Stability Changes upon Mutation from the Protein Sequence or Structure, Nucleic Acids Res, 33(2), 306-310.
7. Capriotti E., R. Calabrese, R. Casadio (2006). Predicting the Insurgence of Human Genetic Diseases Associated to Single Point Protein Mutations with Support Vector Machines and Evolutionary Information, Bioinformatics, 22(22), 2729-2734.
8. Dryja T. P., T. L. McGee, L. B. Hahn, G. S. Cowley, J. E. Olsson, E. Reichel, M. A. Sandberg, E. L. Berson (1990). Mutations within the Rhodopsin Gene in Patients with Autosomal Dominant Retinitis Pigmentosa, N Eng J Med, 323(19), 1302-1307.
9. Guex N., A. Diemand, M. C. Peitsch (1999). Protein Modelling for All, Trends in Biochemical Sciences, 24(9), 364-367.
10. Gupta N., K. Gill, S. Singh (2009). Paraoxonases: Structure, Gene Polymorphism & Role in Coronary Artery Disease, Indian Journal of Medical Research, 130, 361-8.
11. Harel M., A. Aharoni, L. Gaidukov, B. Brumshtein, O. Khersonsky, R. Meged, H. Dvir, R. B. Ravelli, A. McCarthy, L. Toker (2004). Structure and Evolution of the Serum Paraoxonase Family of Detoxifying and Anti-atherosclerotic Enzymes, Nature Structural & Molecular Biology, 11(5), 412-419.
12. Krawczak M., E. V. Ball, I. Fenton, P. D. Stenson, S. Abeysinghe, N. Thomas, D. N. Cooper (2000). Human Gene Mutation Database – A Biomedical Information and Research Resource, Human Mutation, 15(1), 45-51.
13. Kundu A., S. Bag, S. Ramaiah, A. Anbarasu (2013). Leucine to Proline Substitution by SNP at Position 197 in Caspase-9 Gene Expression Leads to Neuroblastoma: A Bioinformatics Analysis, 3 Biotech, 3(3), 225-234.
14. Lee J.-E., J. H. Choi, J. H. Lee, M. G. Lee (2005). Gene SNPs and Mutations in Clinical Genetic Testing: Haplotype-based Testing and Analysis, Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 573(1), 195-204.
15. Li B., V. G. Krishnan, M. E. Mort, F. Xin, K. K. Kamati, D. N. Cooper, S. D. Mooney, P. Radivojac (2009). Automated Inference of Molecular Mechanisms of Disease from Amino Acid Substitutions, Bioinformatics, 25(21), 2744-2750.
16. Li H.-L., D.-P. Liu, C.-C. Liang (2003). Paraoxonase Gene Polymorphisms, Oxidative Stress, and Diseases, Journal of Molecular Medicine, 81(12), 766-779.
17. Moore D. J., L. Zhang, T. M. Dawson, V. L. Dawson (2003). A Missense Mutation (L166P) in DJ-1, Linked to Familial Parkinson's Disease, Confers Reduced Protein Stability and Impairs Homo-oligomerization, J of Neurochemistry, 87(6), 1558-1567.

18. Morales R., A. Berna, P. Carpentier, C. Contreras-Martel, F. Renault, M. Nicodeme, M.-L. Chesne-Seck, F. Bernier, J. Dupuy, C. Schaeffer (2006). Serendipitous Discovery and X-ray Structure of a Human Phosphate Binding Apolipoprotein, Structure, 14(3), 601-609.
19. Nakken S., I. Alseth, T. Rognes (2007). Computational Prediction of the Effects of Non-synonymous Single Nucleotide Polymorphisms in Human DNA Repair Genes, Neuroscience, 145(4), 1273-1279.
20. Ng P. C., S. Henikoff (2006). Predicting the Effects of Amino Acid Substitutions on Protein Function, Annu Rev Genomics Hum Genet, 7, 61-80.
21. Pettersen E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin (2004). UCSF Chimera – A Visualization System for Exploratory Research and Analysis, Journal of Computational Chemistry, 25(13), 1605-1612.
22. Ramensky V., P. Bork, S. Sunyaev (2002). Human Non-synonymous SNPs: Server and Survey, Nucleic Acids Res, 30(17), 3894-3900.
23. Reddy S. T., D. J. Wadleigh, V. Grijalva, C. Ng, S. Hama, A. Gangopadhyay, D. M. Shih, A. J. Lusis, M. Navab, A. M. Fogelman (2001). Human Paraoxonase-3 is an HDL-Associated Enzyme with Biological Activity Similar to Paraoxonase-1 Protein but is not Regulated by Oxidized Lipids, Arterioscler, Thromb, and Vasc Biol, 21(4), 542-547.
24. Rosenblat M., D. Draganov, C. E. Watson, C. L. Bisgaier, B. N. La Du, M. Aviram (2003). Mouse Macrophage Paraoxonase 2 Activity is Increased Whereas Cellular Paraoxonase 3 Activity is Decreased Under Oxidative Stress, Arterioscler, Thromb, and Vasc Biol, 23(3), 468-474.
25. Schrödinger L. L. C (2010). The PyMOL Molecular Graphics System, Version 1.3.
26. Sherry S. T., M.-H. Ward, M. Kholodov, J. Baker, L. Phan, E. M. Smigielski, K. Sirotkin (2001). dbSNP: The NCBI Database of Genetic Variation, Nucleic Acids Res, 29(1), 308-311.
27. Smith E. P., J. Boyd, G. R. Frank, H. Takahashi, R. M. Cohen, B. Specker, T. C. Williams, D. B. Lubahn, K. S. Korach (1994). Estrogen Resistance Caused by a Mutation in the Estrogen-receptor Gene in a Man, N Eng J Med, 331(16), 1056-1061.
28. Thomas P. D., M. J. Campbell, A. Kejariwal, H. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan, A. Narechania (2003). PANTHER: A Library of Protein Families and Subfamilies Indexed by Function, Genome Research, 13(9), 2129-2141.
29. Thomas R., R. McConnell, J. Whittacker, P. Kirkpatrick, J. Bradley, R. Sandford (1999). Identification of Mutations in the Repeated Part of the Autosomal Dominant Polycystic Kidney Disease Type 1 Gene, PKD1, by Long-range PCR, Am J Hum Genet, 65(1), 39-49.
30. Tokuriki N., F. Stricher, L. Serrano, D. S. Tawfik (2008). How Protein Stability and New Functions Trade off, PLoS Computational Biology, 4(2), e1000002.
31. Tsuzura S., Y. Ikeda, T. Suehiro, K. Ota, F. Osaki, K. Arii, Y. Kumon, K. Hashimoto (2004). Correlation of Plasma Oxidized Low-density Lipoprotein Levels to Vascular Complications and Human Serum Paraoxonase in Patients with Type 2 Diabetes, Metabolism, 53(3), 297-302.
32. Van Gunsteren W. F., S. Billeter, A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. Scott, I. G. Tironi (1996). Biomolecular Simulation: The {GROMOS96} Manual and User Guide, Vdf Hochschulverlag AG an der ETH Zuerich, Zurich, Switzerland.
33. Wang Z., J. Moult (2001). SNPs, Protein Structure, and Disease, Human Mutation, 17(4), 263-270.
34. Wheeler J. G., B. D. Keavney, H. Watkins, R. Collins, J. Danesh (2004). Four Paraoxonase Gene Polymorphisms in 11 212 Cases of Coronary Heart Disease and 12 786 Controls: Meta-analysis of 43 Studies, The Lancet, 363(9410), 689-695.

35. Wu J., R. Jiang (2013). Prediction of Deleterious Nonsynonymous Single-nucleotide Polymorphism for Human Diseases, The Scientific World Journal, 2013:675851, doi: 10.1155/2013/675851.

Sudip Paul, M.Sc.
E-mail: sudip@juniv.edu

Sudip Paul is working as a Lecturer in the Department of Biochemistry and Molecular Biology, Jahangirnagar University, Savar, Dhaka, Bangladesh. He completed B.Sc. (Hons.) and M.Sc. in Biochemistry and Molecular Biology from Jahangirnagar University, Bangladesh. His areas of research are bioinformatics, medicinal chemistry, clinical biochemistry, molecular pharmacology and public health. He has awarded with several grants for his research projects.

Md. Solayman
E-mail: naeem40thju@gmail.com

Md. Solayman graduated from the Department of Biochemistry and Molecular Biology, Jahangirnagr University, Bangladesh. Currently, he is working on several research projects. He has excellent command over different bioinformatics tools. He is also working on drug discovery and computer aided drug designing in bangladesh institute of computational chemistry and biochemistry. His other research interests are clinical biochemistry, medicinal chemistry, medical genetics, etc.

Moumoni Saha, M.Sc.
E-mail: moumonisaha@gmail.com

Moumoni Saha completed her M.Sc. in Biochemistry and Molecular Biology from Jahangirnagar University, Bangladesh. Her scientific interests are molecular pharmacology, clinical biochemistry, medicinal chemistry and bioinformatics.

Assoc. Prof. Md. Sabir Hossain, Ph.D.
E-mail: sabiriu@gmail.com

Md. Sabir Hossain is now working as an Associate Professor in the Department of Biochemistry and Molecular Biology, Jahangirnagar University, Bangladesh. He completed his Ph.D. degree from Dhaka University. Apart from teaching different courses, he has conducted several research works on bioinformatics, nutrition and public health, clinical biochemistry etc.