

## IMPACT OF DELETERIOUS NON-SYNONYMOUS SINGLE NUCLEOTIDE POLYMORPHISMS OF CYTOKINE GENES ON NON-CLASSICAL HYDROGEN BONDS PREDISPOSING TO CARDIOVASCULAR DISEASE: AN *IN SILICO* APPROACH

SAI RAMESH A<sup>1</sup>, RAO SETHUMADHAVAN<sup>2</sup>, PADMA THIAGARAJAN<sup>2\*</sup>

<sup>1</sup>Department of Biotechnology, Sree Sastha Institute of Engineering and Technology, Chembarambakkam, Chennai - 600 123, Tamil Nadu, India. <sup>2</sup>Department of Biomedical Sciences, School of Biosciences and Technology, VIT University, Vellore - 632 014, Tamil Nadu, India. Email: padmadk4@gmail.com

Received: 01 May 2017, Revised and Accepted: 26 July 2017

### ABSTRACT

**Objective:** Cardiovascular disease (CVD) is a leading cause of death worldwide. Malfunctioning of genes that are responsible for several inflammatory processes is the major cause for its initiation. Cytokine genes are one such group of genes that are involved in the development of CVD. Hence, the prediction of potential point mutations in these genes is important for diagnostic purposes. Such mutations result in altered protein structure and function when compared to neutral ones.

**Methods:** In this study, interleukin factor 6 (IL6), tumor necrosis factor  $\alpha$  (TNF- $\alpha$ ), interleukin factor 4 (IL4), and interferon gamma have been analyzed using sorting intolerant from tolerant and PolyPhen 2.0 tools.

**Results:** Several single nucleotide polymorphisms (SNPs), in IL6, TNF- $\alpha$ , and IL4, are found to be potentially deleterious. In addition, bond analysis has also been performed on these SNPs. It has been predicted that L119P and R196H of IL6 as well as K87T and T181N of TNF- $\alpha$  are potential ns-SNP's that may cause structural and functional variations in the corresponding proteins. The hydrogen and Cation-Pi bond analysis performed in this study provides molecular-based evidence that support the predicted deleterious potential of such SNPs for these CVD candidate genes along with other conventional *in silico* tools.

**Conclusion:** The study testifies the importance of adopting a computational approach to narrow down potential point mutants for disease predictions.

**Keywords:** Cardiovascular disease, Interleukin 6, Tumor necrosis factor  $\alpha$ , Non-synonymous single nucleotide polymorphisms, Cation-Pi analysis.

© 2017 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2017.v10i11.19531>

### INTRODUCTION

Globally, cardiovascular diseases (CVDs) have the highest mortality rate as compared to any other non-communicable diseases [1]. Their primary cause is the development of atherosclerosis that leads to its clinical precipitation as myocardial ischemia and infraction [2]. The pathophysiology of atherosclerosis implicates it partially to be an inflammatory disorder, resulting in elevated levels of specific circulatory cytokines. These, in turn, may be cautiously referred to as biomarkers for predicting the prevalence and progression of the disease in symptomatic as well as non-symptomatic individuals [3]. Cytokines are glycoproteins that play a pivotal role in mediating immune responses by activation of signal transduction pathways. They broadly include interferons, chemokines, members of tumor necrosis factor (TNF) family, and interleukins. Elevated levels of interleukin 6 (IL6) and TNF- $\alpha$  along with cholesterol levels in serum have been detected consistently in CVD patients and are now recommended for use as biomarkers [4,5]. In addition, their respective genes are reported to be highly polymorphic and closely associated with immune or autoimmune disorders [6]. Non-synonymous single nucleotide polymorphisms (ns-SNPs) in their coding regions may alter some physiochemical properties of their native proteins [7]. These SNPs are of particular interest to both human geneticists and computational biologists, as one half of such known variations are related to human-inherited diseases [8]. Computational biology facilitates the screening and prediction of diseases associated with ns-SNPs through multiple tools that depend mainly on protein sequences and/or their structure [9-11]. Recent advances have now paved a way to understand the molecular level variations to a better extent. For example, changes in distances between hydrogen and

Cation-Pi bonds help the researchers to explore causes behind such mutations, thus validating their predictions [12,13].

The present work focuses on the identification of deleterious SNPs in four cytokines, namely, IL6, IL4 (both isoforms 1 and 2), TNF- $\alpha$ , and interferon gamma (IFNg). An attempt has been made to understand the levels of structural and functional variations between the wild and mutant genes, using conventional computational tools. The results are validated using Salt Bridge, Hydrogen, and Cation-Pi bond analyses.

### METHODS

#### SNP retrieval and prediction of deleterious effects

SNPs of the selected cytokines genes, namely, IL6, IL4, TNF- $\alpha$ , and IFNg were retrieved from the National Center for Biotechnology Information (NCBI) - dbSNP (<http://www.ncbi.nlm.nih.gov/snp/>) database [14,15], by limiting the studies to Homo Sapiens, SNPs, coding non-synonymous and synonymous regions, 5' and 3' untranslated region (UTR) and intron function classes. The ns-SNPs were investigated further for analyzing their deleterious effects using standard computational tools like SIFT-Blink ([http://sift.jcvi.org/www/SIFT\\_BLink\\_submit.html](http://sift.jcvi.org/www/SIFT_BLink_submit.html)) [16,17] and polymorphism phenotype 2.0 (PolyPhen 2.0) (<http://genetics.bwh.harvard.edu/ph2/>) [17].

#### Structure modeling and root mean square deviation (RMSD) calculations

These studies were carried out on ns-SNPs, identified to be deleterious, by both prediction tools. For mutant modeling, SWISS-MODEL (<http://swissmodel.expasy.org/>) was used to obtain 3D structures [18]

and validated by PROCHECK (<http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>) [19,20]. Validated structures were considered to generate mutant structures using Deepview/Swiss PdbViewer [21] and were energy minimized through normal mode analysis, deformation, and refinement (<http://lorenz.immstr.pasteur.fr/index1.php>) [22], before calculation of RMSD values.

### Trajectory analysis

Deleterious ns-SNPs, predicted by both SIFT and PolyPhen 2.0, were subjected to further investigations such as Solvent Accessibility using absolute surface area (ASA) View (<http://gibk26.bio.kyutech.ac.jp/jouhou/shandar/netasa/asaview>) [23] and Evolutionary Conservation Position Analysis by ConSurf [24]. Identification of stabilizing residues was carried out by SRide (<http://sride.enzim.hu/>) [25,26].

### Bond analysis

Salt Bridge Analysis was performed by evaluation of salt bridges in protein (ESBRI) (<http://bioinformatica.isa.cnr.it/ESBRI/input.html>) [27,28]. Hydrogen Bond and Cation-Pi Bond Analysis were carried out through hydrogen bond analysis tool 1.1 (HBAT) and CAPTURE (<http://capture.caltech.edu/>) [29], respectively.

## RESULTS

### SNP retrieval and prediction of deleterious effects

A sum of 538 SNPs were obtained from NCBI database with functions limited to Homo Sapiens, SNPs, coding non-synonymous and synonymous, introns, and UTR 3' and 5', for IL6 (139), TNF- $\alpha$  (109), IL4 (202) and IFN $\gamma$  (88). Total number of active SNPs is shown in Table 1.

### Structure modeling and RMSD calculations

The SWISS-MODEL ExPASy used in modeling the native 3D structures for each of the cytokines provided the template id and modeled residue range (Table 2). The sequence identity percentage of modeled structure to native sequence and its QMEAN Z-score are given in Table 2. The genes IL6, TNF- $\alpha$ , and IL4 had 100% sequence identity and promising QMean Z-Score. The calculated RMSD values, obtained by superimposing the energy minimized native, over the mutant, are provided in Table 3. Only L119P and R196H of IL6 and K87T and T181N of TNF- $\alpha$  had significant RMSD values, and this makes them as highly deleterious ns-SNPs.

### Trajectory analysis

Table 4 displays the ASA view values for the highly deleterious ns-SNPs of IL6 and TNF- $\alpha$ , along with the conserved regions and details on their secondary structure. The CONSURF value of ns-SNPs L119P and R196H in IL6 showed a significant change in their value. In case of TNF- $\alpha$ , a significant change in ASA view is noticed only for K87T. Similarly, Table 5 indicates the stabilizing residues identified through SRide for the same genes. R196H of IL6 and K87T and T181N of TNF- $\alpha$  showed the presence of one or more additional stabilizing residues.

### Bond analysis

ESBRI predicted the formation of seven new salt bridges at various residue positions, in each of the IL6 mutants and the loss of six and five salt bridges in L119P and R196H, respectively, when compared with that of the native protein. In case of TNF- $\alpha$ , there were two new formations and no loss of salt bridges in K87T and five formations

Table 1: Total number of active SNP's

Cytokines	Functions		Introns	UTR		Total active SNPs	
	Codon			3'	5'		
	Non-synonymous			Synonymous			
	Total	Highly deleterious by SIFT and PolyPhen					
IL6	15	2	5	83	8	4	113
TNF- $\alpha$	14	3	12	25	26	4	81
IL4 1 and 2	12	2, 1 (isoforms)	6	147	4	0	169
IFN $\gamma$	4	0	2	59	13	1	79
Total	45	8	25	314	51	9	442

TNF: Tumor necrosis factor, IL: Interleukin, IFN $\gamma$ : Interferon gamma, SNPs: Single nucleotide polymorphisms, UT: 5' and 3' untranslated region, SIFT: Sorting intolerant from tolerant

Table 2: SWISS model report for cytokines

Cytokines	Total residues	Modeled residue range	Template id	Sequence identity %	QMEAN Z-score
IL6	212	47-212	1il6	100	-1.586
TNF- $\alpha$	233	85-233	319j	100	1.313
IL4 (isoform 1)	153	25-153	2b8u	100	-0.556
IL4 (isoform 2)	137	25-137	2b8u	87.6	-2.466

TNF: Tumor necrosis factor, IL: Interleukin

Table 3: RMSD value of cytokines

Cytokines	Residue change (NCBI rs-id)	RMSD value (Å)	Total energy	
			Mutant (KJ/mol)	Native (KJ/mol)
IL6	L119P (rs11544633)	0.81	-11.654	-11.005
	R196H (rs143432552)	0.83	-11.475	
TNF- $\alpha$	K87T (rs190788828)	1.60	-8.989	-8.978
	T181N (rs140654183)	2.30	-9.275	
	I194N (rs147831616)	0.20	-9.209	
IL4 (isoform 1)	V53A (rs139863211)	0.64	-9.290	-9,266
	K150M (rs149147538)	0.33	-9.279	
IL4 (isoform 2)	K134M (rs149147538)	0.34	-8.370	-8.360

TNF: Tumor necrosis factor, IL: Interleukin, RMSD: Root mean square deviation, NCBI: National Center for Biotechnology Information

and seven deletions in T181N. Considerable variations in salt bridge distances were observed between other residue pairs in mutants and native proteins. The graphical representations are shown in Figs. 1 and 2.

The hydrogen bonds analyzed through HBA and Cation-Pi bond Capture, for the same set of genes with highly deleterious ns-SNPs are given in Tables 6, 7a, and b, respectively.

## DISCUSSION

### SNP retrieval and prediction of deleterious effects

The ns-SNPs, with ID rs11544633 and rs14343252 for IL6, rs190788828, rs140654183, and rs147831616 for TNF- $\alpha$ , rs139863211, and rs149147538 for IL4 (isoform 1) and rs149147538 for IL4 (isoform 2) were predicted, both by SIFT and PolyPhen 2.0, to be deleterious. Only deleterious ns-SNPs, predicted by SIFT and PolyPhen 2.0, were further analyzed, as a combined prediction may increase the accuracy [30].

### Structure modeling and RMSD calculations

Point mutations may have deleterious effects on structure-function profile of proteins and may also affect their structural

stabilities [31]. Functional changes can be understood by knowledge of their 3D structures. The total energy given in Table 3 depicts significant variations between the natives and mutants. Only deleterious ns-SNPs, with RMSD >0.8Å were subjected to further analysis, as higher RMSD values indicate greater deviations between native and mutant structures, since they may result in functional changes of proteins [32]. Accordingly, two deleterious ns-SNPs, each of IL6 and TNF- $\alpha$ , were predicted to have RMSD >0.8Å. These are henceforth referred to as highly deleterious SNPs.

### Trajectory analysis

Lee-Richard molecular surface area or ASA is a simple prediction tool that investigates the surface tension of an atom, in a protein molecule that is accessible to the solvent [33,23]. This analysis aids in understanding the background reasons and strengthens choice of predicted ns-SNPs as “deleterious” or “highly deleterious.” Accordingly, this study identified two ns-SNPs of IL6 (L119P and R196H) and TNF- $\alpha$  (K87T and T181N), each to be different from their respective natives, in any one of the prediction analysis. A similar trend was observed when checked for stabilizing residues in all of the highly deleterious ns-SNPs as seen from Table 5. To support and strengthen the trajectory tool findings, bond analyses were carried out for the highly deleterious ns-SNPs of IL6 and TNF- $\alpha$ , to comprehend changes that may have occurred at the molecular level.

### Bond analysis

Oppositely charged residues in a folded protein often form salt bridge bonds by Van der Waals interactions. Any significant variations observed in such residue pairs, between the native and mutants, may result in substantial changes at either structural or functional levels. Hydrogen bond analysis, performed using HBA, evaluates the native and mutant proteins based on variations observed in any one or both of geometric criteria, namely, the distance between acceptor atom (X=C, N, O) - Hydrogen (H) (XH usually 3Å) and the angle formed between the donor-hydrogen-acceptor (angle XHA usually 120°). Table 6 briefs the variations obtained for IL6 and TNF- $\alpha$ . In IL6, the existence of two hydrogen bonds at 119<sup>th</sup> position was lost when leucine is mutated to proline. In TNF- $\alpha$ , four hydrogen bonds were reduced to two when

Table 4: Trajectory analysis data of investigated cytokines

Cytokines	Solvent accessibility (SA) (ASA view)		Conserved region color (CONSURF)	
	Native (Å <sup>2</sup> )	Mutant (Å <sup>2</sup> )	Native	Mutant
IL6				
L119P	0.000	0.014	9 (HC)	8 (C)
R196H	0.424	0.462	6 (A)	5 (V)
TNF- $\alpha$				
K87T	0.520	0.396	9 (HC)	9 (HC)
T181N	0.267	0.225	3 (V)	3 (V)

HC: Highly conserved, C: Conserved, A: Average, V: Variable, TNF: Tumor necrosis factor, IL: Interleukin, ASA: Absolute surface area

Table 5: Stabilizing residue analysis of investigated cytokines using SRide

Cytokines	Residue change	Stabilizing residues (SRide)
IL6	Native	LEU112, THR147
	L119P	LEU112
	R196H	<b>ILE70</b>
TNF- $\alpha$	Native	VAL89, TYR132, LEU133, ILE134, VAL138, ASP206, LEU218, GLY229, ILE230, ILE231, and ALA232
	K87T	TYR132, LEU133, ILE134, VAL138, <b>LEU139, PHE140, GLY142</b> , ASP206, LEU218, GLY229, ILE230, and ALA232
	T181N	VAL89, VAL93, TYR132, LEU133, ILE134, <b>LEU139</b> , ASP206, LEU218, GLY229, ILE230, and ALA232

Destabilising residues in bold. TNF: Tumor necrosis factor, IL: Interleukin

Table 6: Hydrogen bond analysis of IL6 and TNF- $\alpha$

Cytokines	Residue change		Bond type	Donor-CH_No.	Acceptor CH_No.	disXH	disHA	disXA	angXHA
IL6	L119P	Native	N-H...O	LEU-119	ILE-115	1.006	1.851	2.854	173.4
			N-H...O	LEU-119	ILE-116	1.006	2.845	3.163	99.01
	R196H	Mutant	-	-	-	-	-	-	-
			Native	N-H...O	ARG-196	HIS-192	1.007	1.796	2.803
		N-H...N		ARG-196	LEU-195	1.007	2.654	2.879	92.43
		Mutant	N-H...O	HIS-196	HIS-192	1.007	1.806	2.813	178
N-H...N	HIS-196		LEU-195	1.007	2.721	2.933	91.83		
TNF- $\alpha$	K87T	Native	N-H...O	LYS-87	SER-85	1.003	1.995	2.966	162
			N-H...N	LYS-87	ASP-86	1.003	2.336	2.735	102.4
			N-H...N	LYS-87	ASN-115	1.003	2.869	3.101	93.74
			N-H...O	LYS-87	ALA-232	1.003	2.186	3.077	146.9
			N-H...O	THR-87	SER-85	1.003	2.031	3.021	168.3
			N-H...N	THR-87	ASP-86	1.003	2.323	2.749	104.3
	T181N	Native	N-H...O	THR-181	CYS-145	1.005	1.814	2.792	163.3
			N-H...O	ASN-181	ARG-179	1.001	2.104	2.894	134.2
			N-H...O	ASN-181	ARG-179	1.001	2.104	2.894	134.2
			N-H...O	ASN-181	ARG-179	1.001	2.104	2.894	134.2
			N-H...O	ASN-181	ARG-179	1.001	2.104	2.894	134.2
			N-H...O	ASN-181	ARG-179	1.001	2.104	2.894	134.2

TNF: Tumor necrosis factor, IL: Interleukin

lysine at 87 is mutated to threonine. On combining the above two bond analyses, it is seen that L119P of IL6 and K87T of TNF- $\alpha$  scored a greater risk.

To increase the accuracy of prediction, Cation-Pi Bond analysis was also performed for these predicted highly deleterious mutants (Table 7a and b). Cation-Pi bonds present a strong non-covalent binding force that contributes highly to the secondary structures of proteins.

Among them, the aromatic acid interactions, Phe/Trp/Tyr binding to Lys/Arg are common. It has also been reported that in structural biology, a protein structure includes one Cation-Pi interaction for every seventy seven amino acids and by default, arginine will be preferred over lysine [34]. In the present study, CAPTURE observed one Cation-Pi bond for every 166 amino acids in case of native IL6 and 83 amino acids, for both mutants L119P and R196H of IL6. The tool also predicted that the protein had 166 amino acids. Out of this, the native IL6, with an

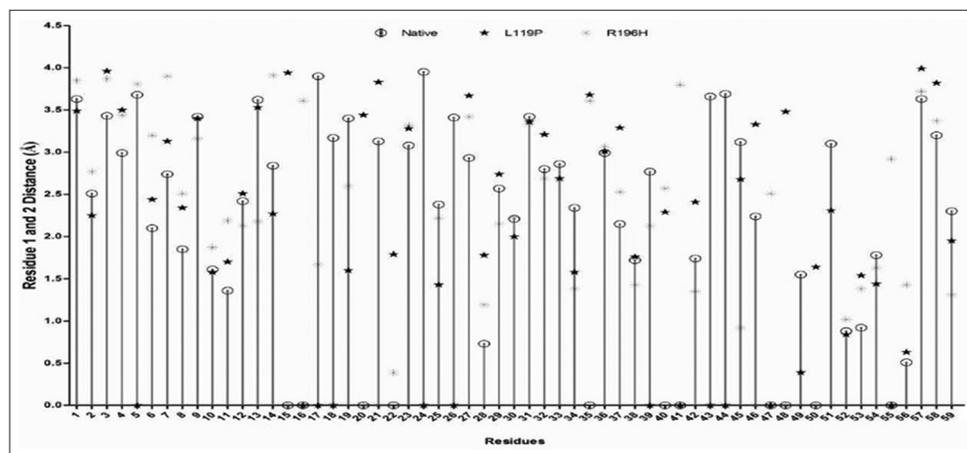


Fig. 1: Salt bridge analysis of interleukin 6

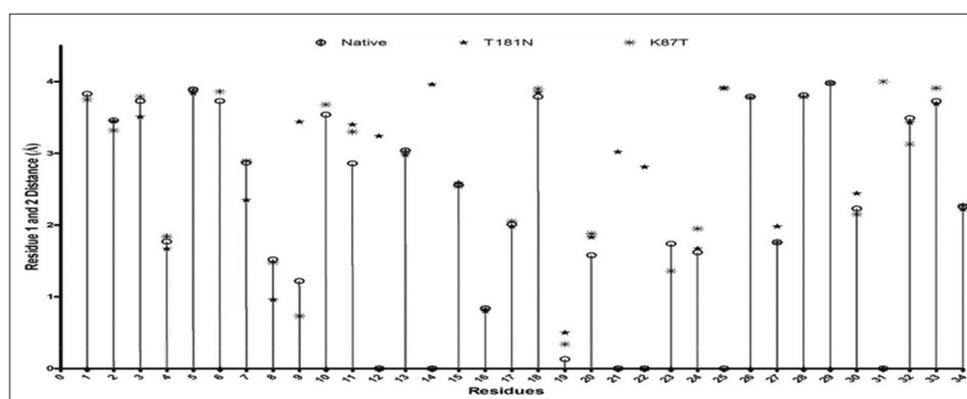


Fig. 2: Salt bridge analysis of tumor necrosis factor  $\alpha$

Table 7a: Cation-Pi interaction analysis of IL6

Residue change	Interaction R1/R2	Interacting pairs #	Energetically significant	R1/R2 #	E (total) = [E (es) + E (vdw)] kcal/mol
Native	ARG/PHE	0	0		
	ARG/TYR	1	0		
	ARG/TRP	0	0		
	LYS/PHE	1	1	156/153	-7.34 [(-6.47)+(-0.87)]
	LYS/TYR	0	0		
L119P	LYS/TRP	0	0		
	ARG/PHE	1	0		
	ARG/TYR	1	1	132/128	-3.52 [(-1.61)+(-1.91)]
	ARG/TRP	0	0		
	LYS/PHE	2	1	156/153	-7.38 [(-6.5)+(-0.88)]
R196H	LYS/TYR	1	0		
	LYS/TRP	0	0		
	ARG/PHE	1	0		
	ARG/TYR	1	1	132/128	-4.31 [(-1.9)+(-2.41)]
	ARG/TRP	0	0		
	LYS/PHE	2	1	156/153	-7.33 [(-6.46)+(-0.87)]
	LYS/TYR	1	0		
	LYS/TRP	0	0		

IL: Interleukin

Table 7b: Cation-Pi interaction analysis of TNF- $\alpha$ 

Residue change	Interaction R1/R2	Interacting pairs #	Energetically significant	R1/R2 #	E (total) = [E (es) +E (vdw)] (kcal/mol)
Native and K87T	ARG/PHE	1	0		
	ARG/TYR	1	0		
	ARG/TRP	0	0		
	LYS/PHE	0	0		
	LYS/TYR	1	0		
	LYS/TRP	0	0		
T181N	ARG/PHE	1	0		
	ARG/TYR	1	0		
	ARG/TRP	0	0		
	LYS/PHE	1	0		
	LYS/TYR	1	1	141/191	-2.22 [(-1.13)+(-1.09)]
	LYS/TRP	0	0		

TNF: Tumor necrosis factor

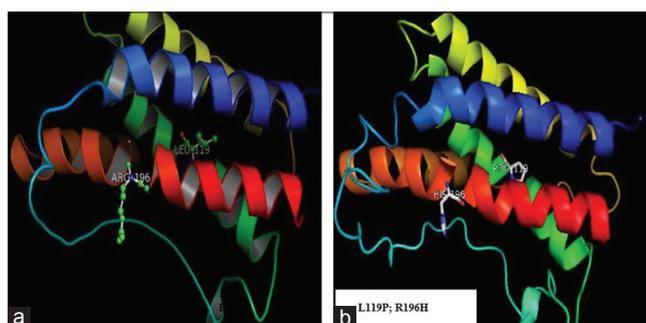


Fig. 3: (a) 3D structure of interleukin 6 (IL6) native. (b) 3D structure of IL6 mutants

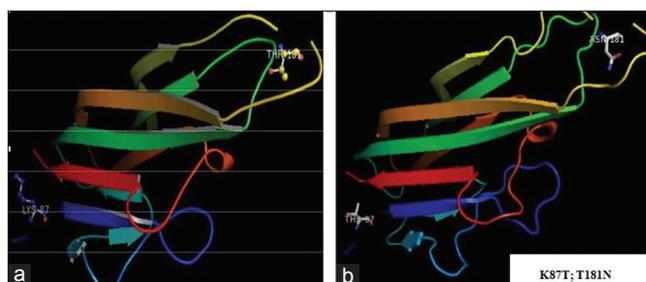


Fig. 4: (a) 3D structure of tumor necrosis factor  $\alpha$  (TNF- $\alpha$ ) native. (b) 3D structure of TNF- $\alpha$  mutants

apparent molecular weight of 19086 D, contained 13 lysine residues (7.8%), eight arginine residues (4.8%), seven phenyl alanine residues (4.2%), three tyrosine residues (1.8%), and one tryptophan residue (0.6%). For the mutant L119P of IL6, the molecular weight reduced to 19070D. However, the percentage of amino acids remained the same as in the native protein. However, in R196H, two variations were observed. Arginine residues were reduced to seven and its molecular weight dropped down to 19063D.

With TNF- $\alpha$ , both native and mutant K87T proteins do not have any Cation-Pi bond formation. However, the mutant T181N, on an average, has one Cation-Pi interaction for every 146 residues. In addition, variations in amino acid percentages, between native and mutants were also observed. Native TNF- $\alpha$  and T181N had 149 amino acid residues, out of which six lysine residues (4.0%), seven arginine residues (4.7%), four phenylalanine residues (2.7%), seven tyrosine residues (4.7%), and two tryptophan residues (1.3%) were present. Minor changes in molecular weight from 16635 D and 16648 D were also seen. For the mutant K87T, one lysine residue was less and a molecular weight of 16608 D was obtained.

Trajectory analysis and bond analysis support and confirm the fact that the highly deleterious ns-SNPs L119P and R196H of IL6 and K87T and T181N of TNF- $\alpha$  are expected to cause major structural and functional variations in the respective proteins. For these proteins, 3D structures of both mutants and native forms were designed using PyMol. These are represented in Figs. 3a and b, 4a and b [35].

## CONCLUSION

This study primarily focused on prioritizing functional SNPs of CVD candidate genes, IL6, TNF- $\alpha$ , IL4 (isoform 1 and 2), and IFN $\gamma$ . These genes play a major role in the onset and progression of CVD. Accordingly, four ns-SNPs were predicted as disease-related mutations and were strongly supported by various computational tools. The prioritization analyses performed in this study will help the molecular biologist narrow down the search from 538 SNPs to four highly deleterious ns-SNPs with added benefits such as reduction in time and cost spent for this prediction.

## ACKNOWLEDGMENTS

The authors thank VIT University for supporting this study.

## REFERENCES

- World Health Organization. Global Status Report on Non-Communicable Diseases 2010. Geneva: World Health Organization; 2011.
- Mukesh K, Vicky D, Shruti M, Dinesh S, Neha M, Mangla L. cardiovascular disease prevalence and drug utilization patterns at a tertiary care hospital in north eastern India. *Int J Pharm Pharm Sci* 2016;8:116-9.
- Tamariz L, Hare JM. Inflammatory cytokines in heart failure: Roles in aetiology and utility as biomarkers. *Eur Heart J* 2010;31(7):768-70.
- Ridker PM, Rifai N, Stampfer MJ, Hennekens CH. Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation* 2000;101(15):1767-72.
- Venu P, Rema R, Hariprasad MG. Evaluation of the antihyperlipidemic, cardioprotective activity of a polyherbal formulation. *Int J Pharm Pharm Sci* 2010;2:86-91.
- Kofler S, Nickel T, Weis M. Role of cytokines in cardiovascular diseases: A focus on endothelial responses to inflammation. *Clin Sci (Lond)* 2005;108(3):205-13.
- Zhang Z, Miteva MA, Wang L, Alexov E. Analyzing effects of naturally occurring missense mutations. *Comput Math Methods Med* 2012;2012:805827.
- Zhang Z, Norris J, Schwartz C, Alexov E. *In silico* and *in vitro* investigations of the mutability of disease-causing missense mutation sites in spermine synthase. *PLoS One* 2011;6(5):e20373.
- Hicks S, Wheeler DA, Plon SE, Kimmel M. Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. *Hum Mutat* 2011;32(6):661-8.
- Bao L, Cui Y. Prediction of the phenotypic effects of non-synonymous single nucleotide polymorphisms using structural and evolutionary information. *Bioinformatics* 2005;21(10):2185-90.

11. Ye ZQ, Zhao SQ, Gao G, Liu XQ, Langlois RE, Lu H, *et al.* Finding new structural and sequence attributes to predict possible disease association of single amino acid polymorphism (SAP). *Bioinformatics* 2007;23(12):1444-50.
12. Weiss MS, Brandl M, Sühnel J, Pal D, Hilgenfeld R. More hydrogen bonds for the (structural) biologist. *Trends Biochem Sci* 2001;26(9):521-3.
13. Dougherty DA. Cation-pi interactions in chemistry and biology: A new view of benzene, Phe, Tyr, and Trp. *Science* 1996;271(5246):163-8.
14. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, *et al.* dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res* 2001;29(1):308-11.
15. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4(7):1073-81.
16. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res* 2001;11:863-74.
17. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7(4):248-9.
18. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T. The SWISS-MODEL repository and associated resources. *Nucleic Acids Res* 2009;37:D387-92.
19. Laskowski RA, MacArthur MW, Thornton JM. PROCHECK: Validation of protein structure coordinates. *International Tables of Crystallography, Crystallography of Biological Macromolecules*. Vol. F. Netherlands: Kluwer Academic Publishers; 2001. p. 722-5.
20. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK-a program to check the stereo chemical quality of protein structures. *J Appl Crystallogr* 1993;26:283-91.
21. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 1997;18(15):2714-23.
22. Lindahl E, Azuara C, Koehl P, Delarue M. NOMAD-Ref: Visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Res* 2006;34:W52-6.
23. Ahmad S, Gromiha M, Fawareh H, Sarai A. ASAView: Database and tool for solvent accessibility representation in proteins. *BMC Bioinformatics* 2004;5:51.
24. Armon A, Graur D, Ben-Tal N. ConSurf: An algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J Mol Biol* 2001;307(1):447-63.
25. Gromiha MM, Pujadas G, Magyar C, Selvaraj S, Simon I. Locating the stabilizing residues in (alpha/beta) 8 barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation. *Proteins* 2004;55(2):316-29.
26. Magyar C, Gromiha MM, Pujadas G, Tusnady GE, Simon I. SRide: A server for identifying stabilizing residues in proteins. *Nucleic Acids Res* 2005;33:W303-5.
27. Kumar S, Nussinov R. Salt bridge stability in monomeric proteins. *J Mol Biol* 1999;293(5):1241-55.
28. Sarakatsannis JN, Duan Y. Statistical characterization of salt bridges in proteins. *Proteins* 2005;60(4):732-9.
29. Gallivan JP, Dougherty DA. Cation-pi interactions in structural biology. *Proc Natl Acad Sci U S A* 1999;96(17):9459-64.
30. Rajith B, Doss GP. Path to facilitate the prediction of functional amino acid substitutions in red blood cell disorders-a computational approach. *PLoS One* 2011;6(9):e24607.
31. Chasman D, Adams RM. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: Structure-based assessment of amino acid variation. *J Mol Biol* 2001;307(2):683-706.
32. Ramesh AS, Khan I, Farhan M, Thiagarajan P. Profiling deleterious non-synonymous SNPs of smoker's gene CYP1A1. *Cell Biochem Biophys* 2013;67(3):1391-6.
33. Lee B, Richards FM. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* 1971;55(3):379-400.
34. Dougherty DA. Cation-pi interactions involving aromatic amino acids. *J Nutr* 2007;137 6 Suppl 1:1504S-8.
35. de Lano WL. The PyMOL Molecular Graphics System, Version 1.3r1. South San Carlos, CA: DeLano Scientific LLC; 2010.