International Journal of Drug Discovery
ISSN: 0975-4423 & E-ISSN: 0975-914X, Volume 4, Issue 1, 2012, pp.-145-152.
Available online at http://www.bioinfo.in/contents.php?id=24

PHARMACOGENOINFORMATICS: MODELING OF P-GLYCOPROTEIN AND NOVEL APPROACH OF *In Silico* DRUG DESIGNING BASED ON GENETIC VARIATION OF MDR1 GENE INVOLVED IN STATIN RESISTANCE

SAI BABU M.[1,2], VENKATESWARA RAO A.[1], LAKSHMAN TEJA G.[3], SUBASH KAUL[4], JYOTHY A.[1] AND ANJANA MUNSHI[1]*

[1]Institute of Genetics and Hospital for Genetic Diseases, Osmania University, Begumpet, Hyderabad-500016, AP, India.
[2]Dr. NTR University of Health Sciences, Vijayawada-520008, AP, India.
[3]GenBios, Bangalore - 560078, Karnataka, India.
[4]Nizams Institute of Medical Sciences, Punjagutta, Hyderabad- 500082, AP, India..
*Corresponding Author: Email- anjanadurani@yahoo.co.in

Received: June 16, 2012; Accepted: June 28, 2012

Abstract- Statins are the most prescribed drugs, highly effective in reducing the risk of cardiovascular and cerebrovascular events, primarily by lowering low density lipoprotein (LDL) cholesterol. Although large clinical trials found a 27% average relative risk reduction of major coronary events, there is large variability in benefits from statin therapy. Researchers have found three SNPs (C3435T, G2677T/A, C1236T) of MDR1gene, which codes for P-Glycoprotein (P-gp) (a drug efflux transporter), responsible for the reduced bioavailability of statins. We aimed to design a new drug molecule based on synonymous and nonsynonymous SNPs of MDR1 gene, which is not a substrate to P-gp and acts directly on ßhydroxy methylglutaryl coenzyme A reductase (HMG-CoA), a target site for statins, using Insilico tools. Structural changes in mRNA due to synonymous and nonsynonymous SNPs were evaluated by SNPfold. The 3D structures of normal and mutant proteins of P-gp and HMG-CoA reductase were modeled by Molecular Operating Environment (MOE). A new lead molecule was designed from native structure by VegaZZ and parameters of drug were validated with HyperChem and Pharmacophore mapping was done using LigandScout. We docked the lead molecule with normal and mutant P-gp and found no interactions with P-gp showing that it is not a substrate for P-gp. However, it forms clear hydrogen bond interactions with HMG-CoA reductase. This is a novel approach in the field of bioinformatics and pharmacogenomics (pharmacogenoinformatics) for the development of new drug molecules based on the SNPs of genes involved in drug metabolism in a particular population.
Keywords- Statin resistance, MDR1 gene, Polymorphism, P-gp, HMG Co-A, MOE, Modeling, drug designing, docking

Citation: Sai Babu M., et al. (2012) Pharmacogenoinformatics: Modeling of P-Glycoprotein and Novel Approach of *Insilico* Drug Designing Based On Genetic Variation of MDR1 Gene Involved In Statin Resistance. International Journal of Drug Discovery, ISSN: 0975-4423 & E-ISSN: 0975-914X, Volume 4, Issue 1, pp.-145-152.

Introduction

ATP-binding cassette (ABC) superfamily of proteins includes P-glycoprotein (MDR1, ABCB1), multi drug resistance associated protein 1 (MRP1, ABCC1), canalicular mulitorganic anion transporter (cMOAT, ABCC2) and breast cancer resistance protein/ mitoxantrone resistance protein (BCRP- MXR, ABCG2). These transport proteins transport diverse array of substrates, including drugs, amino acids, toxins, lipids, sterols, bile salts, peptides, etc. These proteins are expressed in lower intestinal tract, liver and kidney. The ABC proteins are classified into seven subfamilies based on phylogenetic analysis [1]. Out of 49 ABC proteins, only a few have been characterized in terms of their biochemistry and function [1-2]. These proteins form four domains, two membrane bound domains involved in the pathway of transport of substrates and two nucleotide binding domains (NBD), which hydrolyze ATP to boost this pathway.

P-glycoprotein (P-gp) is the most important member of this family which plays a central role in safe guarding the organisms from toxicity of various endogenous and exogenous molecules [3]. P-gp encoded by MDR1 gene, is a large (170Kda) transmembrane protein consisting of 1280 amino acids [4-7]. Over expression of this protein has often been associated with the multidrug resistance (MDR1) phenotype which involves the removal of structurally unrelated compounds from and within the cells. Lipid bilayer has a main role to play in efflux function of P-gp. It acts as a 'vacuum cleaner' for hydrophobic molecules present within the membrane [8].

The clinical use of statins constantly reduces the risk of cardiovascular and cerebrovascular diseases. Genetic polymorphisms of the MDR1 gene have been reported to be associated with altera-

tion in disposition kinetics and interaction profiles of various drugs including statins (Atravostatin, Stevastatin, Simvastatin, Mevastatin..etc) [9]. More than 50 SNPs and insertion/deletion polymorphisms have been reported in MDR1 gene [10-12]. Pharmacogenomic studies have reported 3 main SNPs C3435T (exon 26, synonymous), G2677T/A (exon 21, non-synonymous) and C1236T (exon 16, synonymous), responsible for the reduced or decreased bioavailability of statins [13].

Since three dimensional structure of P-gp is not available, we aimed to study the structure using computational tools. We compared the normal structure of P-gp with mutated structure induced by nonsynonymous SNP (G2677T/A). The synonymous SNPs (C3435T and C1236T) were analyzed for change in RNA structure. The binding sites of P-gp were compared in the normal protein and built mutants. Based on the structure and function of mutated P-gp (considering above mentioned SNPs), we designed a new drug molecule (NDA) which is not a substrate to P-gp and restricts its efflux.

Methodology

SNPfold

The SNPfold, developed by Laederach Lab, was used to compute potential effect of SNP on RNA structure [14]. It computes wild-type and SNP containing sequences and calculates the partition function of RNA sequence, which is the probability of base pairing of every possible base pair. It provides better representation for non-structural RNAs like untranslated regions (UTR), non-coding RNAs or regulatory RNAs. By calculating the partition function for queried sequence it captures all its possible confirmations.

SNPs with a large error with this calculation considerably change the thermodynamic parameters of the sequence. mFold attempts to predict the structure of an RNA, SNPfold rather computes the ensemble and probability of all possible structures of an RNA.

If the correlation coefficient value is closer to 1, lesser are the chances of change in RNA structure. In case the sequence length is smaller, this is normalized by calculating p-value for the given correlation coefficient. Rank of SNP determines the change in RNA structure when compared to all possible mutations. The p-value is simply the rank divided by total number of possible mutations.

Homology Modeling

We built structures of P-gp and HMG-CoA using homology modeling based on the template structure. The modeling procedures were carried out on OCTANE R12000 silicon graphics workstation.

The structurally conserved regions (SCRs) were determined by pairwise sequence alignment. The X-ray structure of *Mus Musculus* P-gp was used as a template (PDB ID 3G60).The structure Human HMG-CoA was taken from PDB (3CCW). Five comparative models of the target sequence were built by Molecular Operating Environment (MOE). This procedure is advantageous because one can select best model from the given models. Furthermore, the more variability among the models can be used to evaluate the reliability of the modeling. Energy minimization was done by taking gradient 0.05, force field MMFF94X + Solvation and threshold RMSD values 0.05 as parameters. The initial and final energy of the protein was calculated by GizMOE using MMFF94X force field with constant gradient method. Stereochemical calculation of modeled protein, an important part of the comparative modeling process, was validated by Rampage and PROCHECK.

Lead Molecule Design

Using ChemSketch the native structure of the drug was generated by its simplified molecular-input line-entry system (SMILES) notation obtained from drug bank and structure analog of the drug was sketched. Lead molecule development was done with VegaZZ tool kit and geometrical optimization of the lead molecule was performed using HyperChem Professional. Pharmacophore mapping of the native drug molecule was carried out with LigandScout and lead drug structure was validated by Lipinski filters.

Docking

Developed lead molecule was docked with MOE to P-gp and HMG -CoA structures determined by comparative homology modeling. The steps for docking included receptor molecule preparation, preparation of ligand molecule, binding sites prediction, running the docking process and visualization of interactions.

Results

SNPfold results revealed that correlation coefficients of both synonymous and nonsynonymous SNPs are close to 1 (table 1), indicating that there is less chance of change in RNA structure (Fig.1-3).

Table 1- SNPfold scoring for synonymous and nonsynonymous SNPs.

SNP	P-value	Rank	Co-relation Coefficient
C3435T	0.6606	991/1500	0.9823
G2677T	0.3142	330/1050	0.9393
G2677A	0.8190	860/1050	0.9883
C1236T	0.2000	480/2400	0.9532

P-values of both synonymous and nonsynonymous SNPs are >0.05 and correlation coefficients are closer to 1 indicating less change in RNA structure.

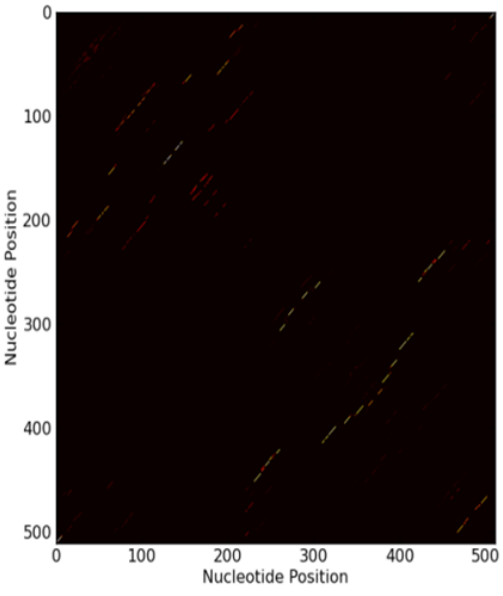


Fig. 1a- Normal sequence For C3435T polymorphism

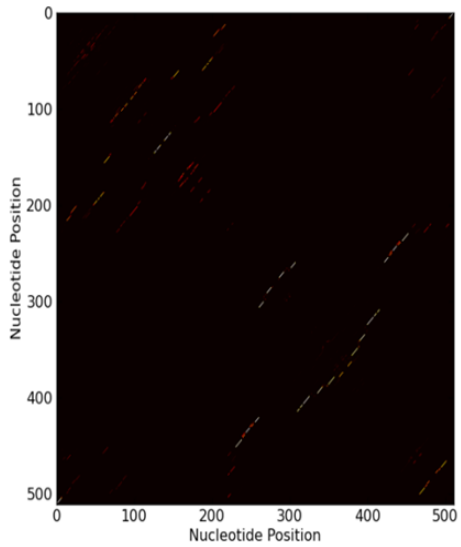


Fig. 1b- Mutant sequence showing no change in RNA structure For C3435T polymorphism

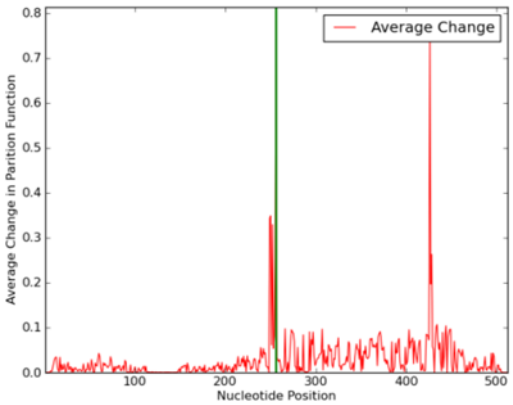


Fig. 1c- Average change is considerably low, mutant has been highlighted green For C3435T polymorphism

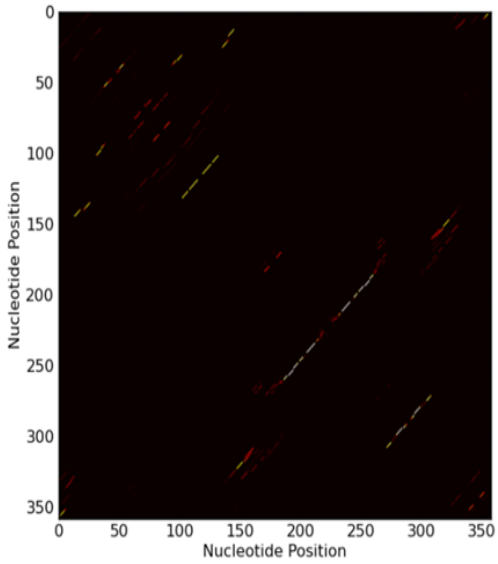


Fig. 2a- Normal For G2677T/A Polymorphism

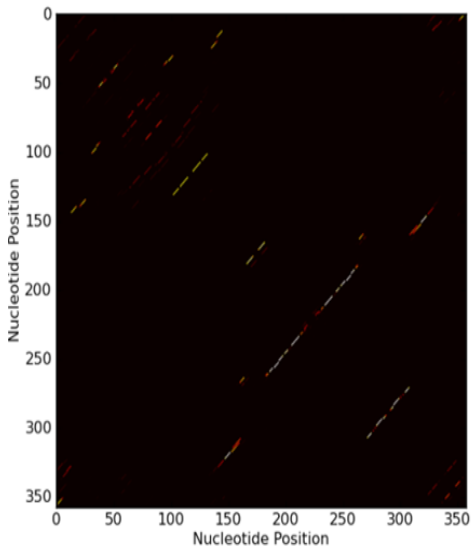


Fig. 2b- Mutants showing no change in RNA structure For G2677T/A Polymorphism

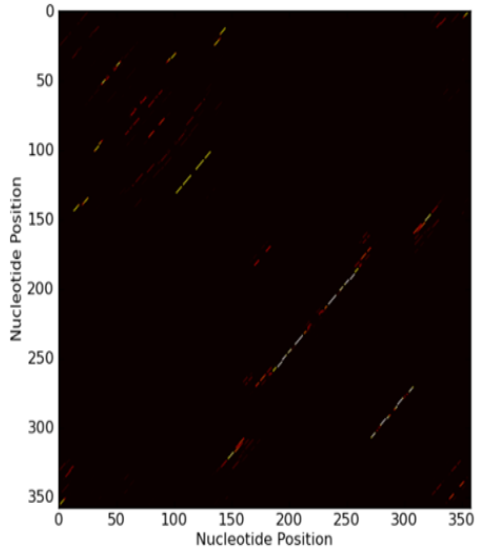


Fig. 2c- Mutants showing no change in RNA structure For G2677T/A Polymorphism

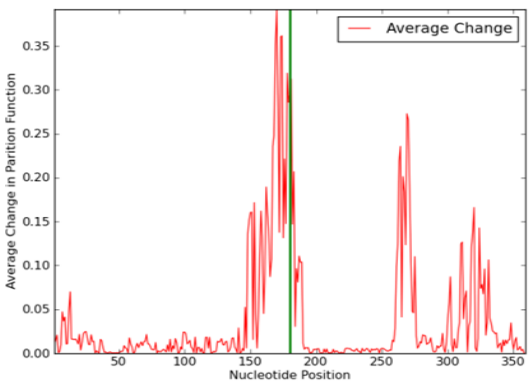


Fig. 2d- Average change is considerably low, mutants have been highlighted green For G2677T/A Polymorphism

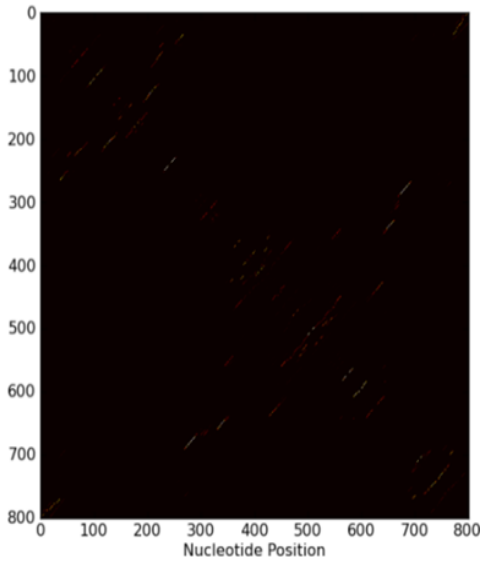


Fig. 3a- Normal sequence For C1236T polymorphism

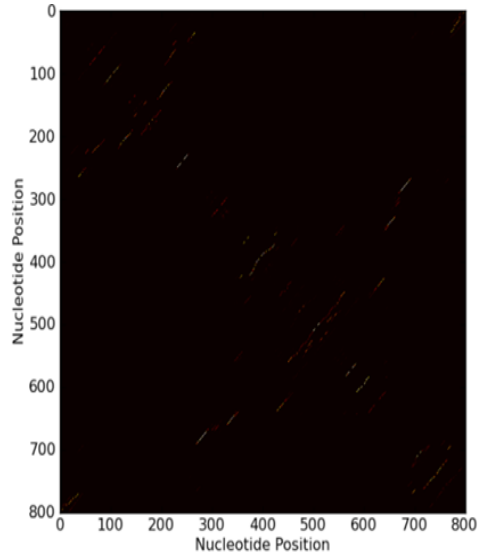


Fig. 3b- Mutant sequence showing no change in RNA structure For C1236T polymorphism

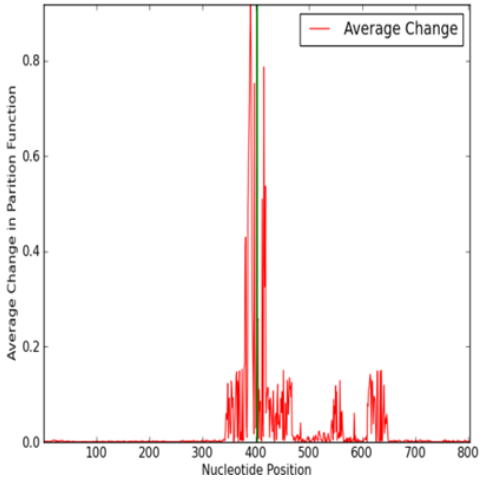


Fig. 3c- Average change is considerably low, mutant has been highlighted green For C1236T polymorphism

Five models of P-gp and HMG-CoA were generated by MOE. Energy and Root Mean Square Deviation (RMSD) values were calculated. The 3D structures of structurally conserved regions (SCRs) were similar for all the five models of P-gp and HMG-CoA. Among these, lowest energy structure was selected (Fig. 4&5). The structure validity of P-gp and HMG-CoA was checked by Rampage (Fig. 6&7). In case of P-gp Rampage analysis revealed that 86.5% residues were in favoured region,9.8% in allowed region and 3.5% in outlier region. For HMG-CoA, 96.5% of the residues were in favoured region, 3.5% residues in allowed region and 0.5% in outlier region.

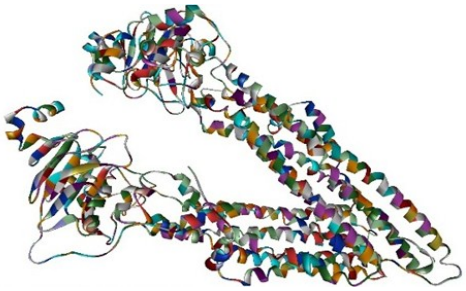


Fig. 4a- Modeled Human P-glycoprotein

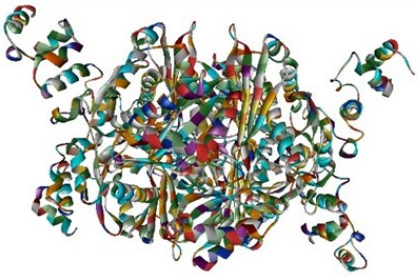


Fig. 4b- Modeled Human HMG Co-A

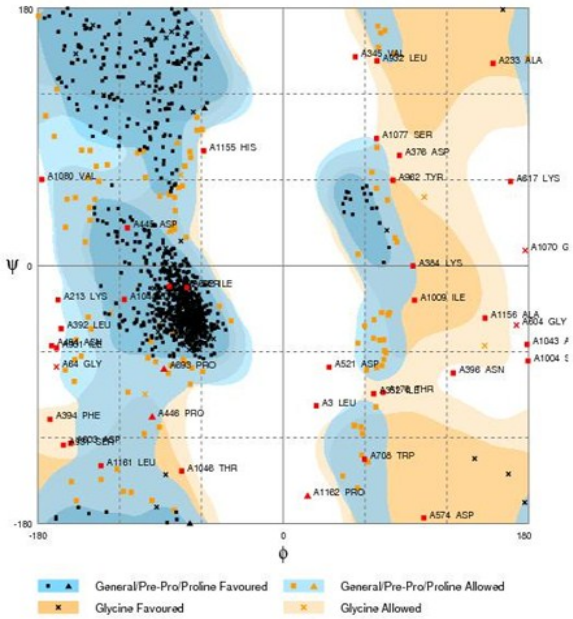


Fig. 5a- Rampage for P-Glycoprotein

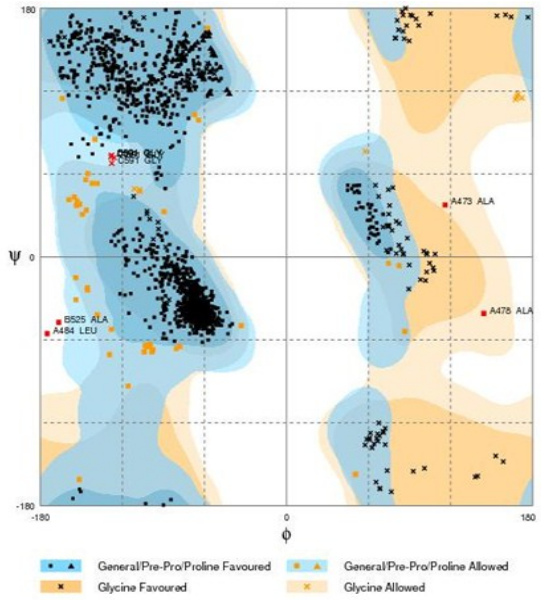


Fig. 5b- Rampage for HMG Co-A

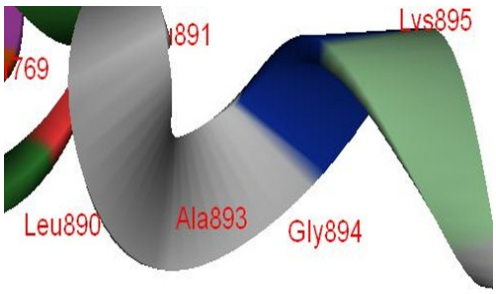


Fig. 6a- Normal region (Ala 893)

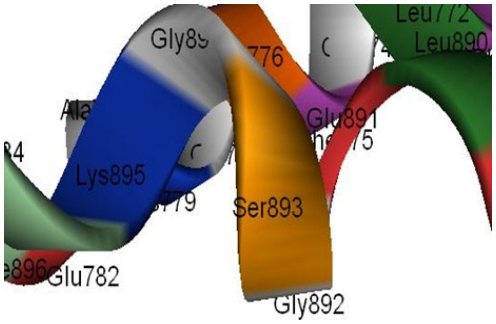


Fig. 6b- Mutated region (Ala893 → Ser893)

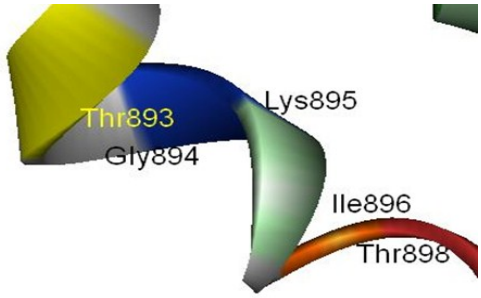


Fig. 6c- Mutated region(Ala893 → Thr893)
Fig. 6- Built Normal and Mutant structures of P-gp based on non-synomous SNP

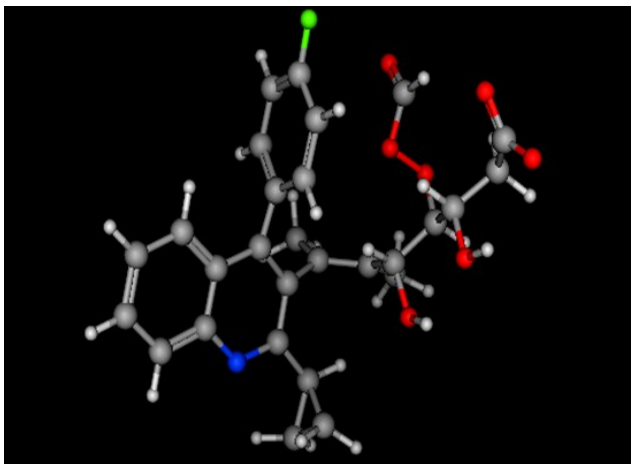


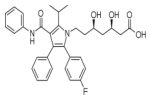
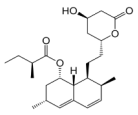
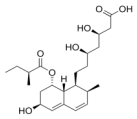
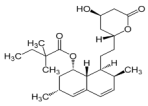
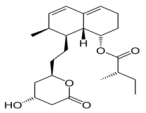
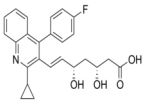
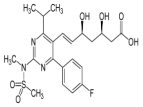
Fig. 7- New Lead drug molecule generated by adding functional group at Pharmacophore region to native statin

Mutated structures of P-gp were generated using Insight II (Accelrys). Based on nonsynonymous amino acid, changes in the P-gp structure were induced. A short deviation was observed in the helix at residues S893, T893. RMSD calculations by server SuperPose showed change of 0.002 and 0.01 for S893 and T893 respectively for global and local RMSD calculations [15].

Lead Molecule Development

Properties of various statins have been represented in table 2.

Table 2- Properties of various statin drugs based on data obtained from drug bank

Drug	Molecular formula	Structure	Mass (g/mol)	LogP	Pka
Atarvostatin	$C_{33}H_{35}F\ N_2O_5$		558.64	4.24	11.824
Lovastatin	$C_{24}H_{36}O_5$		404.54	4.11	NA
Pravastatin	$C_{23}H_{36}O_7$		424.528	2.23	14.533
Simvastatin	$C_{25}H_{38}O_5$		418.566	4.51	NA
Mevastatin	$C_{23}H_{34}O_5$		390.513	2.950	14.890
Pitavastatin	$C_{25}H_{24}F\ N\ O_4$		421.461	2.950	14.890
Rusavastatin	$C_{22}H_{28}FN_3O_6S$		481.539	1.47	14.648

The new Lead molecule was designed by VegaZZ tool. Carboxylic group was added at active pharmacoregions and Quantitative structure activity relationship (QSAR) properties of the drug were calculated with ALOGPS (Chemical formula: $C_{28}H_{30}F$ N O_7, Energy = 19.264 Kcal/mol, LogP = 4.21, Pka = 9.60, Mass = 511.59). The lead molecule satisfied Lipinski rules.

Ligand Docking with P-gp and HMG-CoA

New drug molecule was docked with P-gp as well as HMG-CoA reductase using MOE. The docking results of the Ligand with HMG-CoA are shown in table 3.

Table 3- Docking interaction results with HMG Co-A

Ligand	Receptor	Residue	Type	Score	Distance
O 43183	OG 42119	SER 784	H-donor	39.30%	1.95 A°
O 43218	O 42128	GLY 785	H-donor	20.10%	2.14 A°
N 43201	OD 42166	ASN 788	H-donor	57.70%	1.72 A°
O 43183	NH 20930	ARG 598	H-acceptor	22.90%	2.74 A°
O 43183	OG 42119	SER 784	H-acceptor	39.30%	1.95 A°
O 43216	ND 42167	ASN 788	H-acceptor	26.20%	2.71 A°
O 43218	ND 42167	ASN 788	H-acceptor	31.30%	2.58 A°
O 43219	ND 42167	ASN 788	H-acceptor	42.30%	1.58 A°

Docking scores of a lead molecule with residues of HMG Co-A reductase

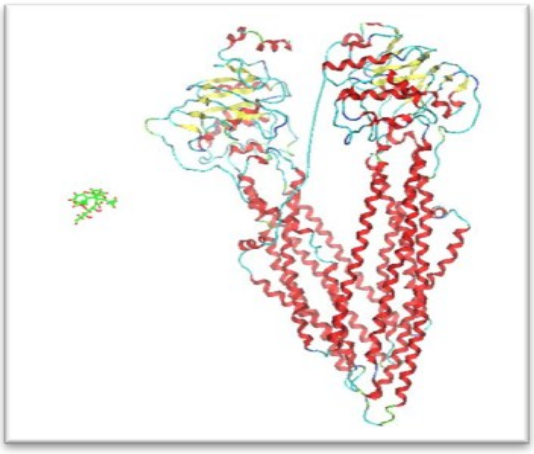


Fig. 8a- Normal P-gp

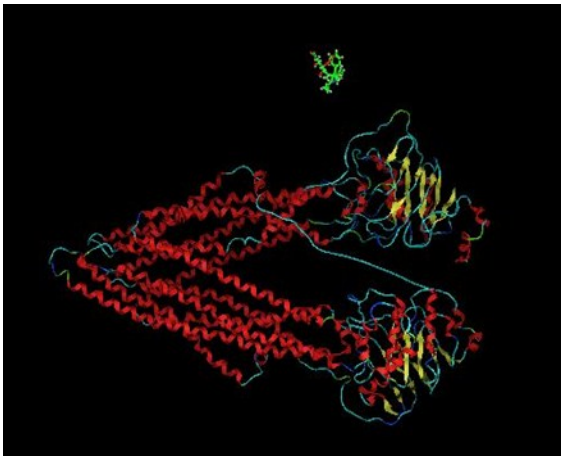


Fig. 8b- Mutant Protein (Ala to Ser)

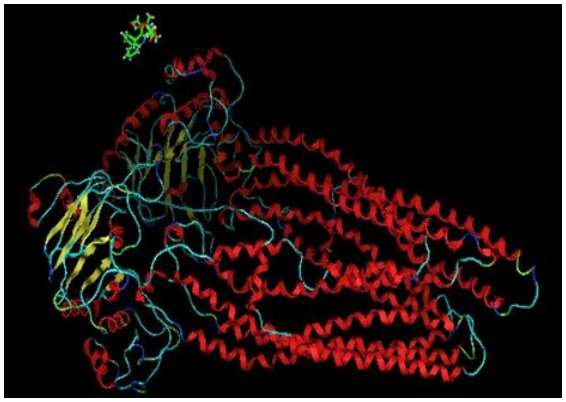


Fig. 8c- Mutant Protein (Ala to Thr)

Fig. 8- Normal and mutants of P-gp not showing any interactions with new lead drug candidate.

The binding pockets of normal P-gp and mutant P-gp did not show any interactions with lead molecule (Fig. 8). Based on the homology, eight residues were conserved in inhibitor binding site of HMG -CoA as shown in Fig. 9. It formed clear hydrogen bonds with residues SER784, GLY785, ASN 788(hydrogen donors) and lowest docking score was found to be is 20.1% with GLY785. Clear hydrogen bonds were also found with residues ARG598, SER784 and ASN788 (hydrogen acceptors) and lowest docking score was found to be 26.2%, with ASN788. Best Ligand pose energy was -6.53911 Kcal/mol. The interactions and Pharmacophore mapping carried out with LigandScout have been shown in Fig. 10.

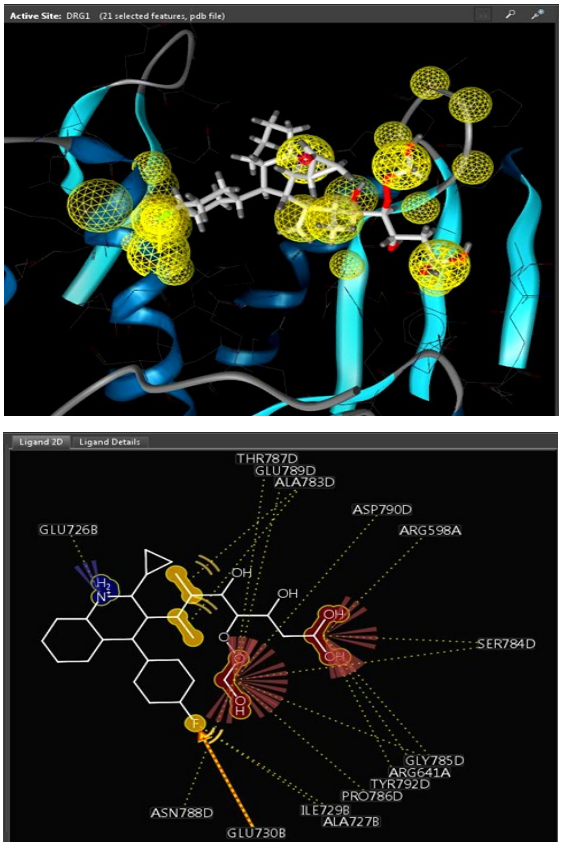


Fig. 9- Pharmacophore mapping of new drug showing interaction profile with HMG Co-A

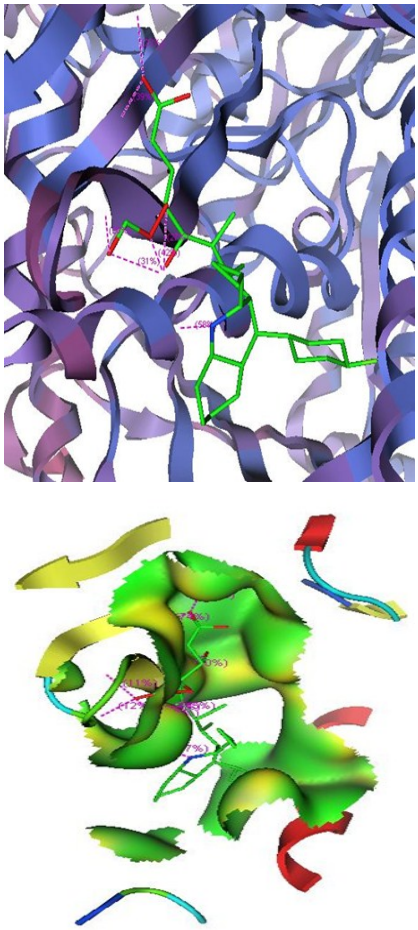


Fig. 10- Docked New drug molecule at active sites of HMG-CoA at residues ARG 598, ASN788, GLY785,SER784

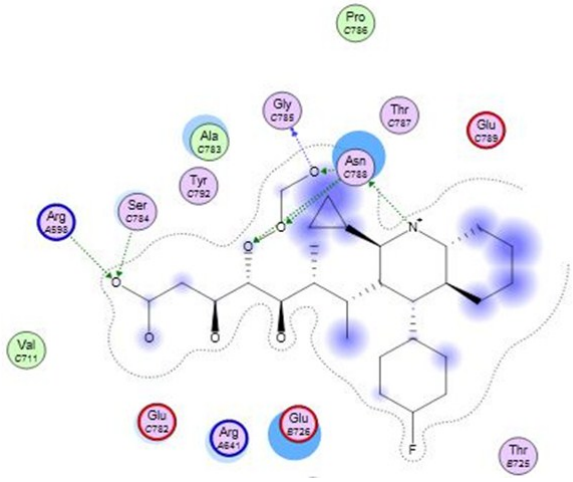


Fig. 11- Ligand interaction diagram obtained from MOE

Discussion

The clinical use of statins consistently decreases the risk of cardiovascular and cerebrovascular diseases by lowering blood cholesterol through inhibition of HMG Co-A reductase. In addition to identification of susceptible genetic markers for these diseases it is necessary to investigate how different genetic variants change the response to the currently used drugs. The Pharmacogenetics is one of the most promising fields which aims at early detection of individual characteristics of the patients that can identify them as good responders or bad responders to each pharmacological treatment.

Several studies have demonstrated contribution of genetic polymorphisms to inter-individual variation to statin therapy [16-18]. Intensive statin therapy has decreased the risk and reversed the progression of stroke and other related disorders. The reduction of LDL in response to statin therapy has been reported to vary from 10 to 70% from person to person and many individuals do not reach the target.

Statins are the substrates to P-gp which transports drugs from the cell through cell membrane with the help of ATP [19]. Polymorphisms in MDR1 gene reduce the bioavailability of statins to target molecules due to over expression of P-gp. Meta-analysis of the Pharmacogenomics data has shown that three polymorphisms C3435T, G2677T/A, C1236T of MDR1 gene are mainly involved in altered expression and structural changes in P-gp [20]. C3435T polymorphism of this gene has been associated with modified response to statin therapy. We also found that the atorvastatin could be more useful for the prevention of future cerebrovascular events in stroke patients with CC genotype than individuals with TT genotype (Unpublished data).

To overcome the decreased bioavailability of statins due to SNPs in MDR1 gene, we designed a new drug molecule which is not a substrate to P-gp, acts on HMG Co-A reductase and is transported through other efflux drug transporters and. First of all we wanted to establish if there are any structural changes in RNA due to synonymous (C3435T, C1236T) and nonsynonymous (G2677T/A) SNPs of MDR1 gene by SNPfold. Output of SNPfold confirmed that correlation coefficient values of synonymous and nonsynonymous SNPs were nearer to 1, indicating that possible change of RNA structure is low. Since the 3D structure of human P-gp is not available till date, therefore, we modeled P-gp based on normal and mutant sequence of MDR1 gene. The normal and mutant structures of modeled P-gp were compared and a short deviation in helix at position 893, where alanine is replaced by serine/ threonine, was observed. Docking studies performed on modeled P-gp and HMG Co-A with new drug molecule showed clear hydrogen bond interactions with HMG Co-A whereas there was no interaction with normal and built mutants of P-gp.

Drug influx into cytoplasm takes place through organic anion transporter protein (OATP), that modulates the cellular intake of various amphipathic compounds. These transmembrane proteins (OATP1B1, OATP2B2, OATP3B3) belong to OATP family and transport a large number of therapeutic drugs including statins. The drug molecules are effluxed out by efflux transporters including P-gp (Fig. 12).

Based on interaction studies, the new drug molecule does not seem to be substrate to P-gp. Therefore, it is expected to be effluxed through other efflux transporters. The existing statins are mainly effluxed by P-gp and over expression due to genetic variation, increases its efflux action. Our new statin molecule is not effluxed out by P-gp, therefore, the individuals with any P-gp genotype will respond equally to this new drug. In addition to this, our new drug molecule is a better inhibitor of HMG-CoA in comparison with other statins (Fig. 12). However, Experimental evidence on

live cells, pre clinical and clinical studies need to be taken up to determine the efficacy of this novel drug molecule.

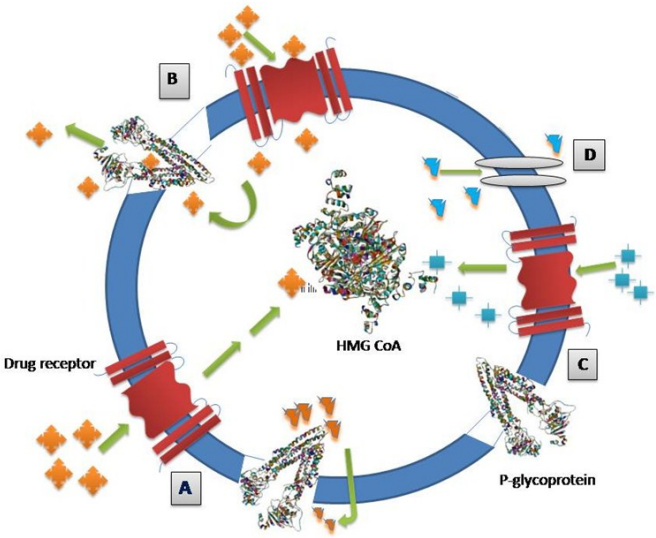


Fig 12: Drug transport mechanisms of normal and mutant P-gp with existing statins (orange) and designed drug (blue).

A- Normal Influx and efflux for a stain drug; drug influx through receptor (OATP), reaches target, xenotoxic substances effluxed through P-gp.

B- Mutant P-gp effluxing the drug without reaching the target.

C- Designed drug is not a substrate to P-gp (both Normal and Mutants) reaches the target.

D- Designed drug effluxed through other efflux transporters.

Conclusion

In conclusion this novel drug is not a substrate for P-gp and acts efficiently on HMG Co-A reductase and can be used by all the individuals irrespective of their P-gp genotypes. There will no need to determine the MDR1 genotypes before prescribing this novel drug. Therefore, it might change the concept of personalized medicine" to "One drug fits all (Populace medicine)". Lot of research has to be carried out in the field of bioinformatics to design novel drug molecules by considering the influence of genetic polymorphisms based on the pharmacogenomic data output which will be major breakthrough in the field of pharmacogenomics and bioinformatics (pharmacogenoinformatics).

Acknowledgement

The financial assistance by ICMR, India is acknowledged with thanks.

References

[1] Dean M., Rzhetsky A. and Alliments R. (2001) *Genome Res*., 11,1156-66.

[2] Steps J.A., Ralph S., Zhao Z., Baillie D.L., *et al*. (2004). *Genome Biol*., 5, R15.

[3] Lesile E.M., Deeley R.G., Cole S.P.C. (2005) *Toxicol. Appl. Pharmacol*., 204, 216-237.

[4] Gottesman M.M., Pastan I. (1993) *Ann. Rev. Biochem*., 62, 385-427.

[5] Ambudkar S.V., Dey S., Hrycyna C.A., *et al*. (1999) *Annu. Rev. Pharmacol. Toxicol*., 39, 361-98.

[6] Troutman M.D., Luo G., Gan L.S., Thakker D.R., David Rodrigues A., *Drugs and the Pharmaceutical Sciences*, 116, 295-357.

[7] Kim R.B. (2002) *Drug Metab. Rev*., 34, 47-54.

[8] Higgins C.F., linton K.J. (2004) *Nature Struct. Biol*., 11, 918-926.

[9] Kajinami K., Brousseau M.E., Ordovas J.M., *et al*. (2004) *The American journal of cardiol*., 93(8), 1046-1050.

[10] Choudhuri S., Klassen C.D. (2006) *Int. J. Toxicol*., 25, 231-259.

[11] Leschiner G.D., Andrew T., Pirmoahmed M., Leschziner G.D., Andrew T., Pirmohamed M., *et al*. (2007) *Pharmacogenomics J*., 7, 154-179.

[12] Schwab M., Eichelbaum M., Fromm M.F. (2003) *Ann. Rev. Pharmacol. Toxicol*., 43, 285-307.

[13] Bercovich D., Friedlander Y., Korem S., Houminer A., Hoffman A., Kleinberg L, *et al*. (2006) *Atherosclerosis*., 185(1), 97-107.

[14] Halvorsen M., Martin J.S., Broadaway S. and Alain Laederach (2010) *PLoS. Genet*., 6(8).

[15] Maiti R., Van Domselaar G.H., Zhang H., Wishart D.S. (2004) *Nucleic Acids Res*., 32, 590-594.

[16] Chasman D.I., Posada D., Subrahmanyan L., Cook N.R., Stanton V.P. and Ridker P.M. (2004) *JAMA*., 291, 2821-2827.

[17] Krauss R.M., Mangravite L.M., Smith J.D., Marisa W.M., Dai Wang., Xiuqing Guo *et al*. (2008) *Circulation*., 117, 1537-1544.

[18] Thompson J.F., Man M., Johnson K.J., Wood L.S., Lira M.E., Lloyd D.B. *et al*. (2005) *Pharmacogenomics J*., 5, 352-358.

[19] Carol W Holtzman., Barbara S Wiggins., Sarah A Spinler (2006) *Pharmacotherapy*., 26(11), 1601-1607.

[20] Komar A.A. (2007) *Pharmacogenomics*., 8(8), 1075-1080.