

ANALYZING THE EFFECT OF NSSNPS IN CYP1A1 TOWARDS BENZOTHAZOLES BINDING

JEMMY CHRISTY H., PREETHI B.

Department of Bioinformatics, Sathyabama University, Chennai 119, Tamilnadu, India.
Email: jemmyjacob@gmail.com

Received: 26 May 2014 Revised and Accepted: 08 Jul 2014

ABSTRACT

Objective: CYP1A1 involved in biotransformation of carcinogenic polycyclic aromatic hydrocarbons (PAHs) and heterocyclic aromatic amines/amides (HAAs) and results in electrophilic reactive intermediates that leads to toxicity and cancer, thus influencing the fields of cancer research. Benzothiazole and its analogs are known for their anti-tumor activity because they act as potent aryl hydrocarbon receptor (AhR) agonist thus binds AhR and results in induction of CYP1A1 which forms DNA adduct and leads to cell death by activation of apoptotic mechanism. The main aim of this study is to extrapolate the relationship between nsSNPs of CYP1A1 and their effects in Benzothiazoles binding capability.

Methods: Computational analysis of deleterious mutations in CYP1A1 and their impact on its structure were as well as altered drug response to Benzothiazoles based drug DF 203, NSC 674495 were studied. Furthermore molecular dynamics simulation (MDS) approach was conducted to investigate conformational changes in the mutant protein structure with respect to its native conformation.

Results: Our studies revealed that 6 deleterious nsSNPs CYP1A1 have the impact on structural stability based on secondary structural patterns and molecular dynamics and altered drug response was seen in nsSNP rs2229150 (R93W) for the drug 2-(4-amino-3-methylphenyl) benzothiazole (DF 203, NSC 674495).

Conclusion: Our study would be helpful to understand the nsSNP effect on CYP1A1 which in turn leads to carcinogenesis as well as Benzothiazole (DF-203) binding affinity and designing individualized therapeutic treatments.

Keywords: CYP1A1, nsSNPs, Molecular dynamics, Cancer, 2-(4-Amino-3-methylphenyl) benzothiazole.

INTRODUCTION

Single nucleotide polymorphism in Cytochrome P450 (CYP) enzymes play a vital role in xenobiotic metabolism can modulate the individual susceptibility to environmental contaminants exposure and the associated risk for cancer development [1]. CYP1A1 is one of the major gene among 57 functional genes of CYP gene family, is involved in biotransformation of procarcinogenic compounds as polyaromatic hydrocarbons (PAH's) and aromatic amines found in cigarette smoke to carcinogen by monooxygenation [2, 3]. CYP1A1 encodes aryl hydrocarbon hydrolase (AHH), an enzyme involved in the production of reactive epoxide intermediates from environmental contaminants such as polycyclic aromatic hydrocarbons [PAHs], polyhalogenated aromatic hydrocarbons (PHAHs) and steroid hormones that might increase the risk of oxidative stress and cancer. Several SNPs have been identified in CYP1A1, some of which lead to a more highly inducible AHH activity [4, 5]. Nonsynonymous single nucleotide polymorphisms (nsSNPs) in the gene coding regions generally considered as genetic variations and were linked to various human inherited diseases [6]. SNP at 4,887 position of exon 7 results in a amino acid change of isoleucine to valine (Ile to Val) [7]. This amino acid variation is associated with CYP1A1 induction and increased AHH enzyme activity that might cause higher rates of carcinogen activation. So individuals with the variant CYP1A1 gene (Val) may be more susceptible to xenobiotic carcinogens and health risk [8, 9]. Among the Caucasians, the CYP1A1 variant was associated with a higher risk of breast cancer [10], whereas in Chinese and Japanese this polymorphism was associated with other types of cancer, such as lung cancer [11]. The frequency of the variant Val allele differs between Caucasian and Asians and is about 0.052 and 0.228, respectively [12]. These amino acid substitutions in CYP1A1 are supposed to be the pathogenetic basis of increased susceptibility to certain diseases and altered drug metabolism [13]. There were 68 coding ns SNP's in CYP1A1 were characterized their functional impacts on disease susceptibility were also reported by means of Insilico studies [14]. So we have taken those ns SNP's as key factor for further studies on drug binding ability. 2-(4-amino-3-methylphenyl) benzothiazole (DF 203, NSC 674495) is metabolized in sensitive cancer cells can binds covalently

to CYP1A1 induce expression and antiproliferative activity through DNA adduct formation [15]. Deleterious amino acid substitution in CYP1A1 protein may affect the binding ability of CYP1A1 with DF 203 and have ability to interfere drug metabolism. So insights of these deleterious SNP's would be beneficial to identify the basis of genetic variation and their response to therapy.

MATERIALS AND METHODS

Native and mutant models preparation and analysis

Deleterious ns SNP's were retrieved from dbSNP [16] based on literature studies [14] and then CYP1A1 native model was prepared using Homology modeling server Modeller 9v7 [17] for homology modeling of CYP1A1 sequence was retrieved from Uniprot database [http://www.uniprot.org/], Modeller 9v7 implements homology modeling of proteins by satisfying spatial restraints, here we used 4I8V as template a PDB structure resolved by X ray crystallography with 2.60 resolution. ClustalW was used to align the target and template sequences and the resultant alignment was stored as PIR format [18]. The alignment and the template atom files were given as input to MODELLER 9v7, to generate the 3D structure of CYP1A1. The scripts "align-ligand.py" and "model-ligand.py" was used to generate ten rough 3D models. Modeller 9v7 automatically derives restraints from known related structures. The restraints include distances, angles, dihedral angles, pairs of dihedral angles, and some other spatial restraints. Bond and angle values are taken from CHARMM-22 [19] force field and then modeled structures were validated based on the backbone conformation of the ten models was inspected using the Phi/Psi Ramachandran plot given by PROCHECK server (http://nihserver.mbi.ucla.edu/SAVS/) and the results indicate that only one model out of ten generated models were perfectly fit with no residues in the disallowed region of Ramachandran plot [20]. Then deleterious mutant models of CYP1A1 were prepared using Pymol mutagenesis tool [21] and both native and mutant models were minimized using SWISSPDB server [22] then RMSD values were recorded for native and mutant models. Substitution of an amino acid may produce changes at the structural level. Changes in the secondary structure with respect to the substituted amino acid were analyzed using PDBsum [23].

Drug binding affinity on native and mutant models

Drug 2-(4-Amino-3-methylphenyl) benzothiazole (DF 203, NSC 674495) binding affinity towards native and mutant models of CYP1A1 was analysed using autodock4 suite, and molecular docking tool [24]. Drug structure was retrieved from PubChem compound database [25]. Autodock uses Lamarckian Genetic Algorithm (LGA) to search the best conformers. Minimum docked free energy were calculated by AutoDock for each GA run and cluster ranking for total clusters were also reported by AutoDock. Docking modes were generally considered based on two factors: The ligand association with the key residues of the receptor and docked complex thermodynamic stability. Docking mode with lowest energy with satisfied above said parameters were selected from over 10 GA runs. The grid boxes were centered on the root of macromolecule.

The intermolecular and intramolecular energy between the protein and the ligand calculated in docking simulations. To evaluate the candidate conformation, the grids were used as lookup tables which store the values used in the calculation, thus making the overall docking simulation exceptionally fast. The Graphical User Interface program 'Auto Dock Tools' was used to prepare, run, and analyze the docking simulations and Kollman united atom charges, solvation parameters, and polar hydrogens were added into the receptor PDB information for the preparation of protein in docking simulation finally Gasteiger charges were added in the ligand PDB file. Finally, the protein-ligand complexes were analyzed using DS Visualizer [26] and ligplot visualization tool [27].

Molecular Dynamics Simulation

rs2229150 (R93W) considered for further dynamic simulation studies as they were showing, we then computed the comparative analysis of structural deviations in native and mutant structure. RMSD, RMSF, SAS and Rg analysis were carried out Molecular Dynamics simulation for the native CYP1A1 and mutant structures of nsSNP rs2229150 are done using Schrödinger11 Maestro-Desmond package from the D. E. Shaw Research laboratory [28] free license is given to academic users for accessing Desmond and its source code. Desmond applies numerical techniques and parallel algorithms to attain specific and accurate performance on a platform that containing a number of processors, and also can be executed on a single-processor computer.

Query ligand-protein complex immersed in TIP3P waterbox extending 10 Å marginal radius beyond any of the complex's atom by Desmond's system builder. This step adds the counter ions to neutralize the simulation box and 0.15 M sodium and chloride ions were used to approximate physiological conditions. Then this complex was minimized to a convergence gradient threshold of 1.0 kcal/ (mol.Å). Desmond dynamics program utilizes the OPLS2005 force field and the NPT ensemble (constant number of particles, pressure and temperature) at 300 K, with default periodic boundary conditions. 5 ps were fixed for the production run. Once the simulation is Completed files saved as standard PDB files.

Then these files were analyzed using the component interactions script of Maestro [29] which does compute the native and mutant CYP1A1 models Molecular Dynamic Trajectory of RMSD, RMSF and Radius of gyration analysis and changes in energy, pressure and volume using simulation Quality Analyst [29].

RESULTS AND DISCUSSION

Based on literature studies it was clear that six nsSNPs are having the ability to change protein stability, so here we have investigated their consequences on CYP1A1 protein.

Modeling of Native and mutant models of CYP1A1 and Binding Site Prediction

Experimentally resolved structure for CYP1A1 does not exist, so native CYP1A1 MODEL was generated using Modeller 9v7, using 4I8V a crystallographic structure which has 512 amino acids and 85% identity with the CYP1A1 sequence. The modeled structure is validated using SAVES PROCHECK, it generates the Ramachandran Plot. In that 89.50% residues are in most favored region and there is

no residue present in disallowed region, this modeled native models. binding sites Arg106, Met121, Ser122, Trp131, Arg135, Leu142, Ile198, Asp313, Leu314, Phe315, Ala317, Gly318, Thr321, Val322, Ala325, Phe376, Phe381, Val382, Phe384, Thr385, Ile386, His388, Gln411 And Phe450 were predicted using the tool 3D Ligand site. Mutant of models building were carried out using PYMOL mutagenesis tool and minimized using Swiss PDB viewer saved in PDB format and visualized using by Pymol visualizer.

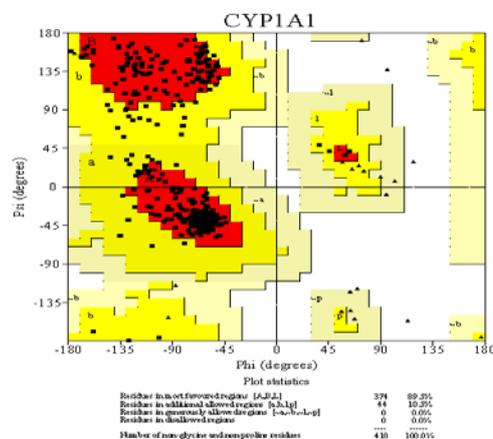
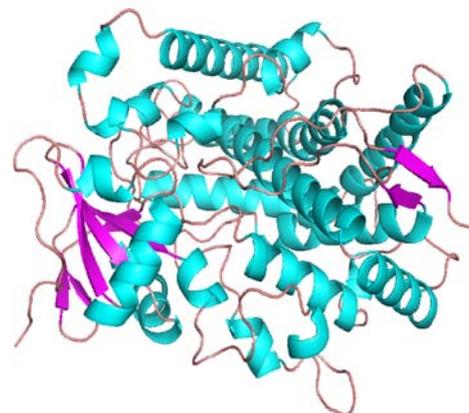


Fig. 1: Modeled CYP1A structure validated using Ramachandran Plot

Secondary Structure Analysis on Native and mutant models

Deleterious nsSNPs of CYP1A1 changes its secondary structural organization and this alteration were analyzed in PDBsum and tabulated (Table 1). It has to be noted that the observed numbers of secondary structural elements are equal in both native and mutant models except the turns in all the six mutant models.

Benzothiazoles (DF-203) binding affinity studies on native and mutant models of CYP1A1 using docking

Based on in vitro studies the drug 2-(4-Amino-3-methylphenyl) benzothiazole (DF203, NSC 674495) active towards CYP1A1 protein and has the ability to lead the programmed cell death in cancerous cells. Substitution of deleterious amino acid in CYP1A1 protein may affect the binding ability of CYP1A1 with the drug molecules. This has to be analyzed to improve the potentiality of the drugs activity on CYP1A1 protein. Hence, we analyzed the binding ability of drugs with native and mutant models of CYP1A1 protein using Autodock4, and drug DF-203 and CYP1A1 interactions were depicted in LIGPLOT. The number of hydrogen bonds formed between protein and ligand were tabulated (Table 3). Van der Waals interacting energies and electrostatic interacting energies, binding energies, torsion free energy, internal energy between CYP1A1 protein (native and mutant) and DF-203 molecule were computed and tabulated (Table 2).

Table 1: Secondary Structure changes of Native and Mutant molecule studies

	Sheets	Beta Hairpins	Psi Loop	Beta Bulge	Strands	Helices	Helix-Helix Interactions	Beta Turns	Gamma Turns
Native	3	3	1	1	8	25	45	34	4
rs2229150 (R93W)	3	3	1	1	8	25	45	35	4
rs17861094(I78T)	3	3	1	1	8	25	45	35	4
rs34260157(R279G/W)	3	3	1	1	8	25	45	35	4
rs45442501(R135W)	3	3	1	1	8	25	45	35	4
rs1048943 (I462F)	3	3	1	1	8	25	44	35	4
rs36121583(F470V)	3	3	1	1	8	25	45	35	4

Table 2: Energy changes of Native and Mutant with the Drug DF203

	Binding Energy (kcal/mol)	Intermolecular Energy (kcal/mol)	Electrostatic Energy (kcal/mol)	vander Wal's energy (kcal/mol)	Internal energy (kcal/mol)	Torsional Free Energy (kcal/mol)
Native	-6.36	-6.95	-0.04	-6.79	-0.18	0.6
rs2229150 (R93W)	-6.12	-6.51	-0.01	-6.34	-0.17	0.55
rs45442501 (R135W)	-6.08	-6.46	-0.03	-6.31	-0.17	0.55
rs34260157 (R279G)	-6.13	-6.51	-0.02	-6.34	-0.17	0.55
rs34260157 (R279W)	-6.08	-6.46	-0.02	-6.28	-0.17	0.55
rs17861094 (I78T)	-5.94	-6.32	-0.03	-6.24	-0.17	0.55
rs1048943 (I462F)	-6.26	-6.00	-0.02	-6.20	-0.17	0.55
rs36121583 (F470V)	-5.94	-6.51	-0.04	-6.31	-0.17	0.55

Table 3: DF203-CYP1A1 native and mutant H-Bond interactions

Protein	No of H bond	Donor	Acceptor	H Bond distance
Native	1	N2	O	3.28
rs2229150 (R93W)	0	0	0	0
rs45442501 (R135W)	1	SG	S	3.28
rs34260157 (R279G)	1	N2	O	3.23
rs34260157 (R279W)	1	N2	O	3.08
rs17861094 (I78T)	2	N2,SG	O,S	3.28,3.20
rs1048943 (I462F)	1	N2	O	3.12
rs36121583 (F470V)	2	N2,SG	O,S	3.24,3.20

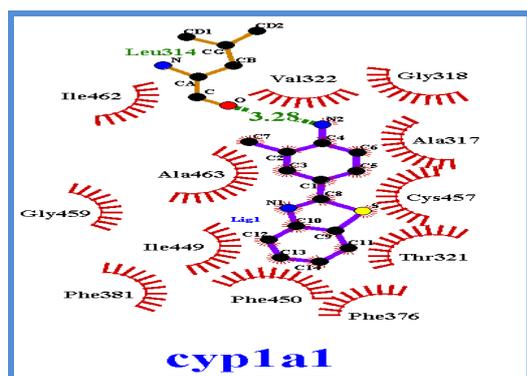


Fig. 2: CYP1A1 with DF203

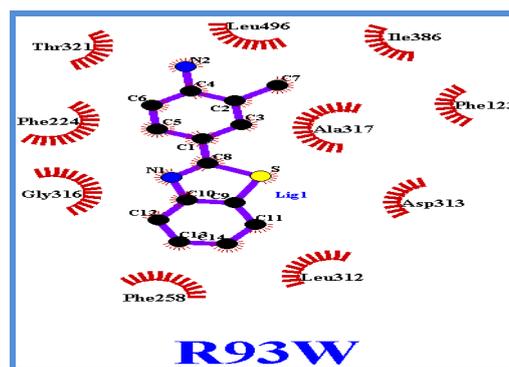


Fig. 3: R93W Mutant with DF203

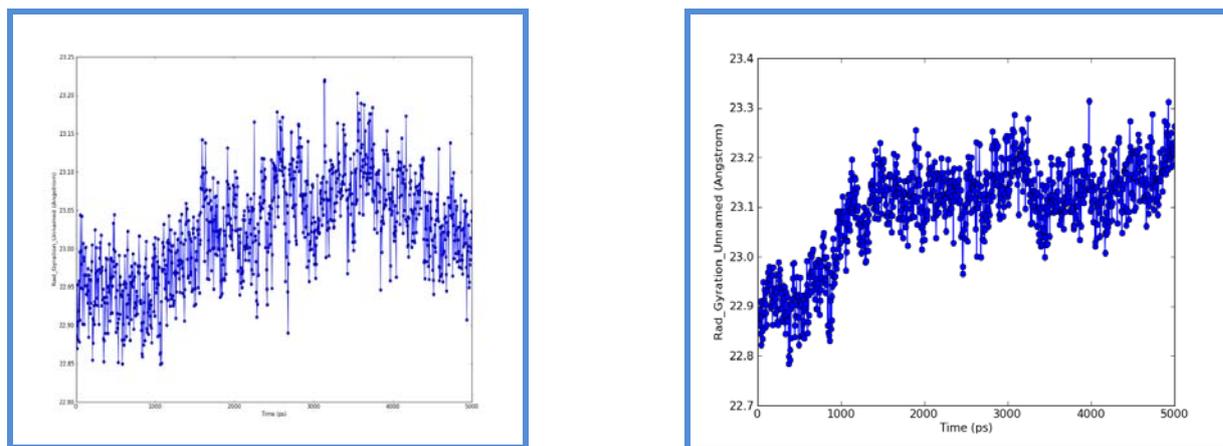


Fig. 10: Radius of gyration of α atoms of native and mutant CYP1A1 protein

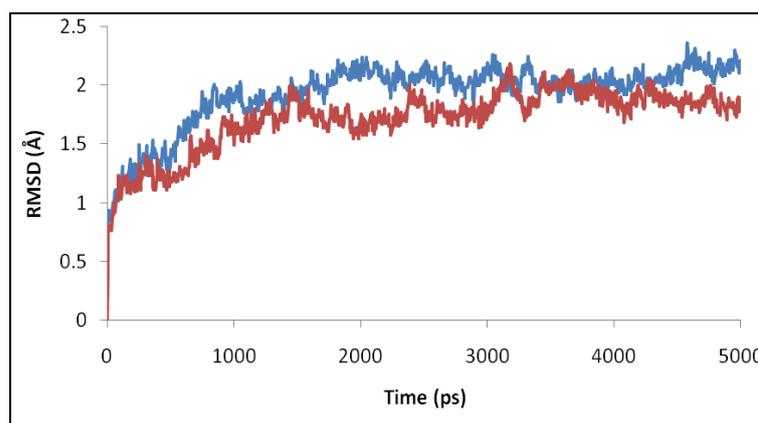


Fig. 11: Backbone RMSDs are shown as a function of time for native and mutant CYP1A1 (Native is shown in red and mutant in blue).

Molecular Dynamics Simulation

Molecular simulation dynamics approach was used to compute RMSF, RMSD and Rg values of the deleterious mutant model (R93W) and native CYP1A1 and the results provided a diverged fashion of variation between native and mutant model (R93W) of CYP1A1. Generally RMSD for all the Ca atoms were calculated from the starting structure which was described as the central origin to compute the protein system and RMSF were calculated to plot the degree of flexibility in mutant due to that mutation curve differ from native CYP1A1 during the simulation.

We observed a major fluctuation in both native and mutant between a time periods of 1000 to 5000 ps. Mutant CYP1A1 (R93W) structure showed deviation till 500 ps from their starting structure and results in a backbone RMSD of ~ 0.74 to 0.27 \AA during the simulations (Fig.6-11).

The radius of gyration (Rg) is the mass-weighted root mean square distance of group of atoms from their common centre of mass so that it provided an observation into global dimension of protein. Radius of gyration graph for alpha-carbon atoms of protein vs time depicted in (Fig.11). We observed a major fluctuation in both native and mutant between a time periods of 1000 to 5000 ns (Fig.11) observation on Rg graph, native structure was found stable than mutant.

Genomic variants especially nsSNPs in the human genome needs insight knowledge because of their impact on cancer related studies and altered drug response of anticancer therapeutic agents. These deleterious nsSNP lead to alterations in protein function especially the geometric constraints, hydrophobicity and hydrogen bond formation [30, 31]. Studies on nsSNPs effects towards protein stability revealed that approximately 25% of are deleterious and

have altered protein function [32]. Wang and Moulton reported that, the vast majority of the disease associated nsSNPs in their dataset (up to 80%) resulted in protein destabilization [33]. Molecular dynamics approach was used to gain the knowledge on native and mutant protein structure variation. Here the point mutation (R93W) which got altered drug response with 2-(4-Amino-3-methylphenyl) benzothiazole (DF203, NSC 674495) was considered for further MD studies and the Structural changes were analyzed using different parameters in 5ns simulation trajectory method. RMSD and RMSF data's observed with fluctuation revealed the changes in molecular stability of native and R93W mutant CYP1A1 models. Thus, from the RMSD and RMSF analysis, it is confirmed reduced stability R93W would affect the CYP1A1 protein structure. Protein ligand interactions analysis using docking revealed that mutant models of CYP1A1 binding affinity is reduced with the drug DF-203, thus this study would be useful in studying the drug targets in CYP1A1 for DF-203 (2-(4-Amino-3-methylphenyl) benzothiazole)

CONCLUSION

Non synonymous single nucleotide polymorphisms (nsSNPs) are the known biomarkers to disease susceptibility. In conclusion, our results showed that the analysis of six different SNPs on the CYP1A1 protein structure using Insilco. These point mutations can disturb CYP1A1 interactions with other molecules and the drugs. Current studies propose that a mutant R93W (rs2229150) out of 6 nsSNPs like rs2229150, rs45442501, rs34260157, rs17861094, rs1048943, rs36121583 identified in this study constitute altered drug response and structure instability. Based molecular docking studies altered drug response was predicted for R93W mutant with DF-203 (2-(4-Amino-3-methylphenyl) benzothiazole) Structural consequences of the deleterious predicted point mutations have been extensively

using Molecular dynamics simulation approaches and based on various parameters like potential energy, root-mean-square deviation, and root-mean-square fluctuation and radiation gyration, it is observed that deleterious nsSNP at position R93W would play a significant role in causing disease by the CYP1A1 protein. Finally this theoretical approach is entirely based on Insilco approach would be helpful in reducing the cost in experimental depiction of pathological nsSNPs and CYP1A1 based drug design.

CONFLICT OF INTERESTS

Declared None

REFERENCES

- Bozina N, Bradamante V, Lovric M. Genetic polymorphism of metabolic enzymes P450 [CYP] as a susceptibility factor for drug response, toxicity, and cancer risk. *Arh Hig Rada Toksikol*. 2009;60:21742.
- Nebert DW, Russell DW. Clinical importance of the cytochromes P450.2002; *Lancet* 360:1155-1162.
- Gonzalez FJ. Molecular genetics of the P-450 superfamily. *Pharmacol Ther* 1990;45:1-38.
- Nebert DW, Dalton TP. The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. *Nat Rev Cancer*. 2006;6:94760.
- Crofts F, Taioli E, Trachman J, Cosma GN, Currie D, Toniolo P, et al. Functional significance of different human CYP1A1 genotypes. *Carcinogenesis*. 1994;15:2961-3.
- Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA. Human Gene Mutation Database (HGMD):2003 update. *Hum Mutat*. 2003;7:577-581.
- Cascorbi I, Brockmüller J, Roots I. A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. *Cancer Res*. 1996;Nov 1;56 [21]:4965-9.
- Cosma G, Crofts F, Taioli E, Toniolo P, Garte S. Relationship between genotype and function of the human CYP1A1 gene. *J Toxicol Environ Health*. 1993;40:309-16.
- Garte S, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, et al. Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*. 2001;10:1239-48.
- Sergentanis TN, Economopoulos KP. Review Four polymorphisms in cytochrome P450 1A1 [CYP1A1] gene and breast cancer risk: a meta-analysis. *Breast Cancer Res Treat*. 2010 Jul;122[2]:459-69.
- Chen S, Xue K, Xu L, Ma G, Wu J. Polymorphisms of the CYP1A1 and GSTM1 genes in relation to individual susceptibility to lung carcinoma in Chinese population. *Mutat Res*. 2001;458:41-7.
- Garte S, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, et al. Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*. 2001;10:1239-48.
- Lin-Lin Wang, Yong Li, and Shu-Feng Zhou, A Bioinformatics Approach for the Phenotype Prediction of Nonsynonymous Single Nucleotide Polymorphisms in Human Cytochromes P450. *Drug Metab Dispos* May 2009;37:977-991.
- Sai Ramesh A., Imran Khan, Md. Farhan, Padma Thiagarajan. Profiling Deleterious Non-synonymous SNPs of Smoker's Gene CYP1A1. *Cell Biochemistry and Biophysics*. 2013;Volume 67, Issue 3, pp 1391-1396 15.
- Eileen Brantley, Valentina Trapani, Michael C. Alley, Curtis D. Hose, Tracey D. Bradshaw, Malcolm F. G. Stevens, Edward A. Sausville, and Sherman F. Stinson. Fluorinated 2-[4-Amino-3-Methylphenyl]Benzothiazoles Induce Cyp1a1 Expression, Become Metabolized, and Bind To Macromolecules In Sensitive Human Cancer Cells. *Drug Metab Dispos* 2004;32:1392-1401.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;Jan 1;29[1]:308-11.
- Fiser A, Sali A. Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol*. 2003;374:461-491.
- Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, et al. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*. 2003;31:3497-3500. doi:10.1093/nar/gkg500.
- Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234:779-815.
- Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 1996;8:477-486.
- DeLano WL. *The PyMOL Molecular Graphics System, Version 0.99*. San Carlos, CA: DeLano Scientific; 2002.
- Guex, N. and Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 1997;18:2714-2723.
- Laskowski. R.A. PDBsum: Summaries and analyses of PDB structures. *Nucleic Acids Research*, 29:221-222, 2001
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput Chem* 2009, 30:2785-2791.
- Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res*. 2009;37:W623-33.
- Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc., 2013.
- Wallace, A. C., R. A. Laskowski, and J. M. Thornton. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*. 1995;8:127-134.
- Desmond Molecular Dynamics System, version 3.1; D. E. Shaw Research: New York, NY, USA, 2012.
- Maestro-Desmond Interoperability Tools, version 3.1; Schrödinger: New York, NY, USA, 2012.
- Gromiha MM, Ponnuswamy PK. Prediction of protein secondary structures from their hydrophobic characteristics. *Int J Pept Protein Res*. 1995;45:225-240.
- Rose GD, Wolfenden R. Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu Rev Biophys Biomol Struct* 1993;22:381-415.
- Yue P, Moulton J. Identification and analysis of deleterious human SNPs. *J Mol Biol*. 2006;356:1263-1274.
- Wang Z, Moulton J. SNPs, protein structure, and disease. *Hum Mutat* 2001;17:263-270.