Computational Biology and Chemistry 78 (2019) 37–52

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac

Research Article

Computational insights into factor affecting the potency of diaryl sulfone analogs as *Escherichia coli* dihydropteroate synthase inhibitors

Bratin Kumar Das, Pushyaraga PV, Debashree Chakraborty*

Department of Chemistry, National Institute of Technology Karnataka, Surathkal, Mangalore, 575025, India

ARTICLE INFO

Keywords:
Dihydropteroate synthase
3D-QSAR
Molecular docking
Density functional theory
Free energy perturbation
Molecular dynamics simulation
Diaryl sulfone

ABSTRACT

Dihydropteroate synthase (DHPS) is an alluring target for designing novel drug candidates to prevent infections caused by pathogenic *Escherichia coli* strains. Diaryl Sulfone (SO) compounds are found to inhibit DHPS competitively with respect to the substrate *p*ABA (*p*-aminobenzoate). The extra aromatic ring of diaryl sulfone compounds found to stabilize them in highly flexible *p*ABA binding loops. In this present study, a statistically significant 3D-QSAR model was developed using a data set of diaryl sulfone compounds. The favourable and unfavourable contributions of substitutions in sulfone compounds were illustrated by contour plot obtained from the developed 3D-QSAR model. Molecular docking calculations were performed to investigate the putative binding mode of diaryl sulfone compounds at the catalytic pocket. DFT calculations were carried out using SCF approach, B3LYP- 6-31 G (d) basis set to compute the HOMO, LUMO energies and their respective location at *p*ABA binding pocket. Further, the developed model was validated by FEP (Free Energy Perturbation) calculations. The calculated relative free energy of binding between the highly potent and less potent sulfone compound was found to be −3.78 kcal/ mol which is comparable to the experimental value of −5.85 kcal/mol. A 10 ns molecular dynamics simulation of inhibitor and DHPS confirmed its stability at *p*ABA catalytic site. Outcomes of the present work provide deeper insight in designing novel drug candidates for pathogenic *Escherichia coli* strains.

1. Introduction

Dihydropteroate synthase (DHPS) is known to be a validated drug target to block folate production in bacterial cells (Domagk, 1935). This particular enzyme catalyses the condensation between 7,8-dihydropterine pyrophosphate (DHPP) and *P*-aminobenzoic acid (*p*ABA) to produce 7,8-dihydropteroate, a precursor of tetrahydrofolate (Alun and Derrick Jeremy, 2002; Griffin and Brown, 1964). Folate analogs like tetrahydrofolate reported to be a crucial co-factor for synthesizing essential amino acids, nitrogen bases and play an important role in reaction associated with one carbon transfer (Bourne, 2014; Rossi et al., 2011). Higher eukaryotic cells consume folate from dietary sources due to the presence of membrane-associated folate transport protein (Matherly and Goldman, 2003; Whetstine et al., 2002). On contrary, all bacterial cells and some lower pathogenic eukaryotes lack the folate transport protein hence compelled to synthesize folate *de novo* (Levin et al., 2004). Therefore, the growth of bacterial cells can be prohibited by inhibiting DHPS competitively.

Diaryl sulfone class of compounds are proved to be, successful in preventing broad range protozoal, microbial infections due to it's structural similarity with *p*ABA (Roland et al., 1979, Woods, 1940). They act as, alternative substrates of DHPS which subsequently form dead-end sulfa-pterin conjugates. The sulfa-pterin conjugates are also reported to block enzymes present later in folic biosynthetic pathway (Chakraborty et al., 2013; Roland et al., 1979; Yun et al., 2012). Hence sulfone class of antimicrobials plays an important role in depleting folate pool in the prokaryotic, lower pathogenic eukaryotic cells (THEN, ANGEHRN, 1973). Sulfone derivatives have been reported to be, a

Abbreviations: B3LYP-6-31G (d), Becke, three parameter Lee-Yang-Parr functional of 6-31 G basis set; 3D-QSAR, 3 dimensional quantitative structure-activity relationship; DHPS, dihydropteroate synthase; DHP^+, cationic intermediate of 7,8-dihydropterine; *E. coli*, *Escherichia coli*; FEP, free energy perturbation; F, ratio of model variance to the observed activity variance (variance ratio); HOMO, highest occupied molecular orbital; LUMO, lowest unoccupied molecular orbital; P, significance level of variance ratio; Pearson–r, linear correlation between test set structures and their activity; Q^2, test set correlation coefficient; R^2, regression coefficient; RMSE, root mean square error; RMSD, root mean square deviation; SD, standard deviation; SO, sulfone; SCF, self consistent field; UTI, urinary tract infection

* Corresponding author.
E-mail address: debashree@nitk.edu.in (D. Chakraborty).

https://doi.org/10.1016/j.compbiolchem.2018.11.005
Received 16 July 2018; Received in revised form 8 October 2018; Accepted 13 November 2018
Available online 17 November 2018
1476-9271/

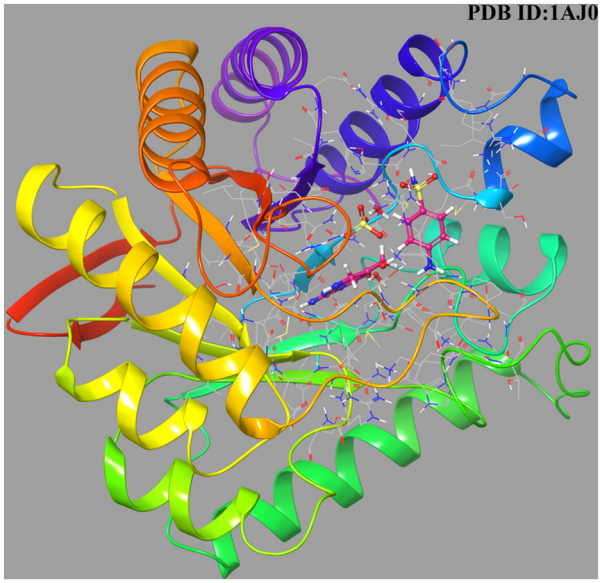


Fig. 1. Crystal structure of dihydropteroate synthase (PDB ID: 1 AJ0, resolution: 2.0 Å).

successful, widely used agent for preventing pneumonia, malaria, tuberculosis, toxoplasmosis etc. by inhibiting folate biosynthetic pathway (Campbell et al., 2012; Forgacs et al., 2009; Manyando et al., 2013; Nzila, 2006; Walzer et al., 1988). Pathogenic *Escherichia coli* strains are reported as the causative agent of various health hazards including diarrhoea, colitis, infant mortality, UTI etc. (Russo and Johnson, 2003). Therefore understanding the key non-bonded interaction of diaryl sulfone molecules at EcDHPS catalytic pocket is of great importance for designing new drug candidates of higher affinity. For greater insight into sulfonamide binding at *p*ABA, pocket of DHPS, structural and theoretical studies have been reported on the co-crystal structure of bacterial DHPS (Hampele et al., 1997; Lawrence et al., 2005; Levin et al., 2004; Morgan et al., 2011; Yun et al., 2012). One co-crystal structure of EcDHPS with sulfonamide has been reported to date (Achari et al., 1997) and presented in Fig. 1. This structure provides a valuable insight to revisit diaryl sulfone interaction at *p*ABA binding pocket.

EcDHPS structure is highly conserved and consists of two identical monomers of a classical triosephosphate isomerase (TIM) type barrel. This TIM barrel consists of eight parallel beta strands encircled by eight alpha helices. The active site of DHPS monomer is located at the C-terminal end of the β-barrel which comprises of three conserved subsite: the pterin binding site, the *p*ABA binding site and the anion binding site. The pterin binding pocket is situated in a deep cleft of the C-terminal end of β-barrel, whereas the pABA binding pocket is located at the surface and comprises of two flexible loops (loop1 and loop2) (Achari et al., 1997; Hampele et al., 1997). Initially, DHPP binds with DHPS active site and it catalyses the slow release of pyrophosphate moiety from DHPP. The released pyrophosphate moiety plays an important role in stabilizing the conformation of flexible *p*ABA binding loops (Yun et al., 2012). After *p*ABA binding, the condensation of DHP^+ and *p*ABA occurs through S_N1 reaction to form dihydropteroate. The mutation in loop 1 and loop 2 allow the binding of *p*ABA but disrupt the binding of sulfa compounds (Yun et al., 2012).

Recently, targeting DHPS (Qi et al., 2011; Zhao et al., 2016) and its mutational studies (Yun et al., 2012) have become an interesting topic for the development of antibacterial drugs candidates to prevent antibiotic-resistant pathogenic bacterial infections. Therefore, a deeper insight regarding diaryl sulfone binding between two flexible loops of EcDHPS will be helpful for synthesizing novel type of sulfa drugs with improved affinity to EcDHPS catalytic pocket. Computer simulation of biomolecules proved to be an important tool for better understanding of protein-ligand interactions and their stabilization patterns (Srivastava and Tiwari, 2017) at the molecular level. Recently, DFT studies of protein-ligand complexes are reported to be effective for understanding them at their electronic level (Tao et al., 2009). The HOMO, LUMO orbital energies and their location at the protein-ligand complex are found to be important for predicting the type of non-bonded interactions which are crucial to achieve favourable recognition of ligands by protein molecules ultimately resulting to their stability in catalytic pocket (Correa-Basurto et al., 2012). DFT calculations also facilitate the mechanistic investigation of product formation in the catalytic pocket of enzymes (Malkhasian and Howlin, 2016). Another important technique to measure the stability of the protein ligand complex is FEP (Free energy perturbation). The FEP/REST(Free energy perturbation/ replica exchange with solute tempering) are reported (Lenselink et al., 2016; Wang et al., 2015) to provide a rigorous algorithm to compute the difference in binding affinity due to presence or absence of specific functional groups in inhibitors. Several FEP studies were carried out to explain the effect of in-silico mutagenesis and SAR (structure-activity relationship) of different antagonist bound to enzymes(Chen et al., 2013; Goldfeld et al., 2015; Keränen et al., 2014). FEP calculations are also reported to be efficient in reducing the computational cost compared to absolute free energy calculations (Bash et al., 1987; Jorgensen and Ravimohan, 1998). Therefore, the utility of these tools is found to be crucial to determine the factors affecting the stabilization of diaryl sulfone derivatives at DHPS catalytic pocket. The relative stability of the protein-ligand complex calculated by FEP can be compared with the experimental data.

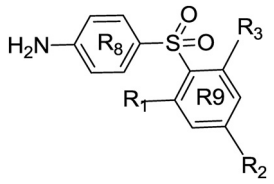
In this present study, we selected a dataset of sulfone compounds to develop a model which clearly describes the effect of substitutions, affecting the potency of sulfone compounds. Initially, 3D-QSAR model was developed to understand the structure-activity relationship (SAR) of diaryl sulfone molecules. The effect of the spatial arrangement of structural features on DHPS inhibition was explained by the contour plot obtained from 3D-QSAR model. Next molecular docking of all the sulfone molecules was carried out at the binding pocket of *p*ABA. The result of molecular docking was correlated with the experimental activity value ($pEII_{50}$). Molecular docking helped in predicting the probable binding pose of sulfone molecule with respect to *p*ABA. Further, we employ DFT (Density Functional Theorem) calculations to clarify the ligand-protein stability and interactions at electronic level. FEP/REST calculations were performed to compute the change in binding free energy due to functional group mutation. Lastly, a 10 ns molecular dynamics was perform to confirm the stability of one of the top scoring compounds. The developed model provides insightful information regarding the mode of binding and important substitution required to stabilize diaryl sulfone compounds at EcDHPS catalytic pocket. This study will help in developing a better generation of Sulfa drugs or pterin-sulfa compounds with improved affinity and therapeutic activity to prevent the infection caused by pathogenic *E.coli*.

2. Methods and computational details

2.1. Ligand preparation

A biological data set consisting of fifty diaryl sulfone(SO) derivatives(Table 1) was chosen from literature(De Benedetti et al., 1989, 1987; Hevener et al., 2010; Lopez de Compadre et al., 1987) and used in the present study. The selected molecules from the biological dataset shared the same assay procedure (Richey and Brown, 1969) with variation in substitution and potency profiles. The biological activity of the dataset was represented by the EII_{50} (Enzyme Inhibition $Index_{50}$) values and reported to have inhibitory activity spanned from 67.3μM to 0.21μM. In spite of having higher potency, **compound 1** was not considered for 3D-QSAR calculations due to the absence of molecules having potency in the same activity range (Golbraikh et al., 2003). EII_{50}

Table 1
structural detail, experimental activities, predicted activity and docking energy of sulfone compounds (1–50).



Compound No.	R1	R2	R3	Experimental activity ($pEII_{50/EXP}$)	Predicted activity ($pEII_{50/PRED}$)	**Residual activity	Docked Energy (Kcal/mol)
1	ONa	ONa	H	8.302	–	–	−6.09
2	CH_3	O^-	H	6.63	6.55	−0.08	−5.68
3	CH_3	O^-	CH_3	6.68	6.60	0.08	−5.88
4	Cl	OH	H	6.58	6.59	0.01	−5.7
5[t]	Cl	O^-	H	6.58	6.57	−0.01	−5.66
6	CH_3	OH	H	6.46	6.46	0.00	−5.86
7	CH_3	CH_3	O^-	6.3	6.02	−0.28	−5.01
8[t]	ONa	NH_2	H	6.29	6.31	0.02	−4.97
9[t]	CH_3	OH	CH_3	6.29	6.29	0.00	−5.89
10	OH	OH	H	6.14	6.17	0.03	−5.22
11	ONa	OH	OH	5.83	5.80	−0.03	−5.42
12[t]	OH	O^-	OH	5.83	5.60	−0.23	−4.74
13	H	ONa	H	5.82	5.83	0.01	−5.14
14[t]	NH_2	NH_2	H	5.79	5.82	0.03	−4.43
15[t]	CH_3	CH_3	CH_3	5.77	5.85	0.08	−5.67
16	CH_3	OCH_3	H	5.75	5.77	0.02	−5.89
17[t]	CH_3	CH_3	OH	5.75	5.72	−0.03	−5.06
18[t]	CH_3	OCH_3	CH_3	5.75	6.03	0.28	−5.76
19	CH_3	CH_3	H	5.72	5.68	−0.04	−5.50
20	OH	OH	OH	5.71	5.76	0.05	−4.74
21[t]	Cl	Cl	H	5.65	5.86	0.21	−5.66
22	NH_2	H	H	5.64	5.62	−0.02	−5.13
23	H	OH	H	5.61	5.64	−0.03	−5.09
24	H	NH_2	H	5.57	5.57	0.00	−4.49
25	H	$N(C_2H_5)_2$	H	5.57	5.53	−0.04	−4.80
26	CH_3	CH_3	OCH_3	5.52	5.53	0.01	−5.22
27	OCH_3	OCH_3	H	5.49	5.58	0.09	−5.15
28	Cl	Cl	Cl	5.48	5.81	0.33	−5.39
29	H	$N(CH_3)_2$	H	5.44	5.37	−0.07	−4.83
30[t]	H	NHOH	H	5.34	5.44	−0.10	−5.16
31	H	NHC_2H_5	H	5.34	5.33	−0.01	−4.77
32	H	COOH	H	5.29	5.29	0.00	−4.81
33	H	COONa	H	5.29	5.34	0.05	−5.31
34[t]	NH_2	NHCHO	H	5.23	4.91	−0.32	−5.47
35	H	$NHCOCH_3$	H	5.23	5.26	0.03	−4.63
36[*]	*m*-NH_2			5.22	5.24	0.02	−5.05
37	H	CH_3	H	5.21	5..18	−0.03	−4.97
38	H	OCH_3	H	5.18	5.12	−0.06	−5.28
39[t]	H	H	H	5.02	5.19	0.17	−4.88
40	H	Br	H	4.96	4.97	0.01	−5.01
41	OCH_3	OCH_3	OCH_3	4.89	4.87	−0.02	−4.91
42	H	F	H	4.74	4.73	−0.01	−4.71
43[t]	H	Cl	H	4.68	4.76	0.08	−4.76
44	H	$CONH_2$	H	4.64	4.73	0.09	−5.41
45	H	$COOCH_3$	H	4.46	4.40	−0.06	−4.62
46	NO_2	NO_2	H	4.44	4.46	0.02	−4.06
47[t]	H	$COCH_3$	H	4.36	4.37	0.18	−4.05
48	H	NO_2	H	4.34	4.35	0.01	−5.72
49	H	$CON(C_2H_5)_2$	H	4.30	4.46	0.16	−4.22
50	H	CN	H	4.17	4.07	−0.1	−4.03

1) [t]Defines that the compound is in the test-set list for 3D-QSAR analysis.
2) [*]Defines the functional group is in the *meta* position of the R9 pharmacophoric ring only.
3) **Residual activity defines the difference between phase predicted activity and experimental activity of diaryl sulfone compounds.

values were imported in Maestro (Schrödinger Release 2017-2: Maestro, Schrödinger, LLC, New York, NY, 2017) project table panel and converted to $pEII_{50}$. The 3D structures of the ligand were constructed using the builder panel in Maestro. The geometry of the ligands was optimized after generating the structure using the Ligprep module (v4.9, schrodinger2017-2). Partial atomic charges were assigned and the possible ionization states were generated at pH of 8.2 ± 0.1. The OPLS_2005 (Kaminski et al., 2001) force field was used for energy minimization of each ligand until it reached a root mean square deviation (RMSD) cut off of 0.01Å. Then the resulting structures were used for the modelling studies.

2.2. 3D-QSAR modelling

Phase (v 4.9) (Dixon et al., 2006) was used to generate 3D-QSAR models for the diaryl sulfone-based inhibitors of DHPS. The prepared structures of SO derivatives were selected in workspace navigator panel in Maestro (Schrödinger Release 2017-2: Maestro, Schrödinger, LLC,

New York, NY, 2017) interface and imported to the atom based 3D-QSAR model development panel with their respective biological activity value ($pIIE_{50}$). Before building the 3D-QSAR model, all prepared ligands were aligned using common scaffold based alignment, a type of flexible ligand alignment tool in Phase (v 4.9) and Largest Common Bemis-Murcao scaffold(Bemis and Murcko, 1996) was selected as a parameter. The atom based 3D-QSAR model considers all the atoms into account; whereas pharmacophore based QSAR model does not consider the ligand feature beyond pharmacophore model. Consequently, atom based 3D-QSAR proved to be more useful for predicting the true structure-activity relationship (Shah et al., 2010; Verma et al., 2010) of diaryl sulfone compound at the *p*ABA binding pocket. The dataset was randomly divided, 75% structures (to maintain the standard 3:1 ratio) were kept as training set and rest molecules were automatically considered as test-set by using "Automated Random Selection" option present in the Phase (v4.9) module. Care was taken in order to include the most active and inactive molecules in training set (Dixon et al., 2006; Golbraikh et al., 2003). The maximum number of Partial Least Square (PLS) factor was N/5 (Where N in the number of training set molecules). Use of higher PLS factor could cause over fitting of data, hence optimal six PLS factor was used (Polański et al., 2002). The 3D-QSAR model was generated by keeping a grid spacing of 1Å. Further, contour plot analysis was performed to interpret and recognize the important pharmacophoric requirements at spatial sites of the structures by the cubic 3D grid. The biological activity of the training set molecules was evaluated by the generated 3D-QSAR model in order to estimate the quality of the 3D-QSAR model. Lastly the stability and predictivity of the developed 3D-QSAR model was examined with an external test set of ten diaryl sulfone compounds(Lopez de Compadre et al., 1987).

2.3. Molecular docking

The co-crystal structure of DHPS with sulfonamide (Fig. 1) (PDB ID: 1AJ0, resolution 2.0Å) was obtained from Protein Data Bank (Achari et al., 1997) and the docking of the fifty drug molecules were carried out using an automated docking program namely Autodock (v4.2.6) which considers the protein to be rigid during docking. Prior to docking, DHPS structure was prepared (Sastry et al., 2013) using Autodock graphical user interface (Morris et al., 2009). The co-crystallized ligand, water molecules were removed and Gasteiger charges (Gasteiger and Marsili, 1980), polar hydrogen were added to the protein structure. The energy minimization of protein was performed till RMSD of 0.30 Å conjugate gradient steps using OPLS-2005 force-field (Sastry et al., 2013). The minimized ligands were taken and the bond associated with the sulphone groups were set as rotatable bonds (Lopez de Compadre et al., 1987). A 3D grid was created at the binding site of the protein having a size of 30Å × 24Å ×22Å with a default spacing of 0.375Å, using an Auto-Grid algorithm to quantify the binding affinity of ligand at the catalytic site of DHPS. The grid includes all eight active site residues namely Thr62, Arg63, Phe190, Ser219, Arg220, Lys221, His257 and SO_4 284 (Achari et al., 1997). The molecular docking simulations were performed using Lamarckian Genetic Algorithm (https://doi.org/10.1002/(SICI)1096-987X(19981115)19:14 < 1639::AID-JCC10 > 3.0.CO;2-B). The initial population and the number of energy evaluations were fixed to 150 and 2.5×10^6 respectively for all molecular docking simulations. The co-crystallized sulphanilamide was removed and re-docked into the same position and RMSD between them was calculated to assess the reliability of docking results (Tripuraneni and Azam, 2016). Several clusters were obtained for the docking orientation within RMS (Root Mean Square) deviation of 0.5Å and the lowest energy cluster obtained for each sulfone compounds were used for further analysis.

2.4. DFT calculation setup

Molecular orbital calculations were performed to explore the binding site of DHPS (i.e. the *p*ABA binding pocket) to analyse the protein-ligand interaction at electronic level. It was evident from the frontier molecular orbital theory that HOMO of the ligands interacts with LUMO positioned in the binding pocket of the protein (Correa-Basurto et al., 2012). In order to determine the location of HOMO, LUMO orbitals in ligands single point energy calculations were performed by Gaussian09 software package Frisch et al., 2009), using SCF (Self Consistent Field) approach (Tomasi et al., 2005). B3LYP(Lee et al., 1988) functional and 6-31G(d) (Otsuka Takao et al., 2015) basis set was used to compute single point energy of sulfone compounds. The recognition ability of the protein was explained by evaluating the HOMO and LUMO energies of the ligands (Correa-Basurto et al., 2012). The HOMO, LUMO calculations of the binding site residues was also performed to explain the ligand-binding mechanism at the catalytic pocket of DHPS. Since the calculation of the whole protein molecule with ligand is computationally expensive so we took only the interacting amino acids and ligands for this study. The generated polypeptide cluster was capped with N-actyl group at the N-terminal end and N-methyl amide group at C-terminal end (Duan et al., 2007). The interacting amino acid residues at *p*ABA binding pocket (Achari et al., 1997) were considered for single point energy calculation at B3LYP/6-31G(d) level. The resulting amino acid cluster was found to have 176 atoms. The cubegen utility in Gaussian09 software package was used to generate the HOMO, LUMO orbitals. The docked poses of highly potent and less potent compounds along with the interacting amino acids were chosen for calculating HOMO, LUMO energies using aforementioned basis set and functional. The location of HOMO, LUMO orbitals on the ligand-protein complex was visualized by Gabedit interface (v2.4.7).

2.5. Relative binding free energy calculation

FEP/REST (Free Energy Perturbation/ Replica exchange with solute tempering) calculations were performed to compute the binding free energy difference between the less potent and highly potent sulfone compounds (these compounds share the same congeneric scaffold) using Desmond MD (v4.8) (Bowers et al., 2006) suite distributed by Schrödinger. The incorporation of REST method in FEP calculations enhances sampling in phase space with an efficient hopping protocol. The efficient sampling of relevant conformations of ligands helps to measure relative binding affinity accurately within an easily accessible simulation time period (Wang et al., 2012). In FEP/REST calculation a small region of interest especially the localized region surrounding the binding pocket including the ligand is heated up by keeping rest of the system cold(Liu et al., 2005). FEP/REST captures the free energy difference of two similar systems through an alchemical transformation pathway. This pathway involves an array of discrete steps starting from initial lambda window **($\lambda = 0$)** to final lambda window **($\lambda = 1$)** (Wang et al., 2011). The free energy difference was calculated applying Bennett Acceptance Ratio (BAR) method (Bennett, 1976) and the error for each free energy calculation was estimated by bootstrapping (Paliwal and Shirts, 2011; Pohorille et al., 2010). Desmond (v 4.8) programmes are reported to provide a good single node, parallel performance compare to other MD packages (Bowers et al., 2006). The OPLS-2005 force field was employed to generate the essential force field parameters required for performing the FEP simulations due to its accuracy in predicting the free energy of solvation of drug-like molecules (Shivakumar et al., 2010). The pose view file of minimized 1AJ0 and sulfone compounds was imported in the FEP panel of Desmond (v4.8), the core RMSD difference between two compounds was 0.004 Å. The complexes of protein and ligand molecules were solvated in SPC(Simple point-charge) water model (Ferguson, 1995) in an orthorhombic box with 5Å buffer size and minimized with the limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method("Practical Methods of

Optimization, 2nd Edition," n.d.). Initially a 100 ps MD simulation was carried out under NVT ensemble applying a force constant of 50 kcal/mol/Å^2 for minimizing the protein and ligand heavy atoms. 1 fs time step was maintained for aforementioned 100 ps simulation. The solute heavy atoms were restrained to their initial position and 10 K temperature was maintained during the simulation. Next, two MD simulations of 12 ps, 36 ps were run under NVT and NPT ensemble respectively maintaining the same force constant to equilibrate the system with 1 fs time step. The system was relaxed with 240 ps MD simulation under NPT ensemble without applying any restrains on heavy atoms keeping time step as 1 fs. Ultimately, 5 ns production simulations were continued under NPT ensemble for both solvent and protein complex using Nose- Hoover thermostats(Martyna et al., 1992), Martyana-Tobias-Klein barostat(Martyna et al., 1994) at effective temperature of 300K and 1 atm pressure. For production run 1fs time step was maintained throughout the simulation time. FEP/REST simulations were carried out with 12 lambda windows and 1.2 ps interval was kept to exchange replicas between two neighbouring lambda windows (Fukunishi et al., 2002). The acceptance ratio of replica exchange was maintained at 0.3. The convergence analysis of free energy during the course of simulation was also performed to check whether the simulation time was sufficient for free energy calculation. The relative binding energy at various thermodynamic states was calculated from the trajectories using the *fepmapper_sid2pdf.py* python script for further analysis (Bowers et al., 2006).

2.6. ADME/toxicity prediction procedure

The drug-like nature of diaryl sulfone compounds were predicted by evaluating their pharmacokinetic (PK) profiles. The ADME/Tox properties of the top ten dock scored SO compounds were calculated using QikProp (v5.7)(Jorgensen and Duffy, 2000) (Schrodinger-2018/3). The structures were already prepared with Ligprep module and imported to QikProp panel for ADME calculation. The molecular weight (MW), octanol/water partition co-efficient ($logP_{o/w}$), dipole moment, H-bond donor and H-bond acceptor etc. properties were calculated for top scored diaryl sulfone compounds and evaluated in accordance with "Lipinski's rule of 5" (Lipinski et al., 1997) to access their drug-likeness.

2.7. System setup for molecular dynamics study

The dynamic behaviour of diaryl sulfone bound EcDHPS (PDB ID: 1AJ0) was observed through all atom molecular dynamics simulation study. The docking predicted binding mode of highly active sulfone compounds were evaluated by observing their dynamic nature in EcDHPS active site. Desmond (v5.4) software package(Bowers et al., 2006; Shivakumar et al., 2010) was employed to perform the molecular dynamics of 1AJ0 and **compound 4** in its bound state. OPLS_2005 all atom force field (Kaminski et al., 2001; Shivakumar et al., 2010, 2009; Jorgensen et al., 1996) was applied to generate the necessary topology and parameter files required for molecular dynamics simulation. The entire system was inserted in the centre of an orthorhombic periodic box with 10Å buffer region from the protein surface and filled with water. The periodic box volume of 1AJ0/**compound 4** complex was 313617 Å^3. The whole system was solvated with SPC (simple point-charge) water molecules (Jorgensen and Jenson, 1998; Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K - The Journal of Physic et al., 2018) and neutralized by adding thirteen Cl^- ions. The solvated protein structure in the periodic box contained 29,076 atoms including 4381 atoms of 1AJ0 and **compound 4.** Next, the system was minimized applying steepest descent algorithm and gradient threshold was kept at 25 kcal/ mol/ Å. The maximum number of iterations in the course of minimization was kept 2000 steps until a convergence threshold of 1.0 kcal/mol/Å was attained. Subsequently, NVT (constant number of atoms N, volume V and temperature T) equilibration was performed at 310K for 5 ns using Nose-Hoover thermostat (Martyna et al., 1992)(thermostat relaxation time = 200 ps) with a time step of 2 fs. Next to NVT, NPT (constant number of atoms N, pressure P and temperature T) equilibration was performed using Nose-Hoover thermostat and Martyna-Tobias-Klein barostat (Martyna et al., 1994)(P = 1 atm, barostat relaxation time = 200 ps and thermostat relaxation time = 200 ps) at 310K with a time step of 2 fs for 5 ns. The solute (protein and ligand) heavy atoms were restrained during the equilibrations, applying force constant of 100 Kcal/mol/Å^2. Lastly, the equilibrated system was taken for production run. The restrains on solute heavy atoms were removed and 10 ns production MD was carried out in NPT ensemble using Nose-Hoover thermostat and Martyna-Tobias-Klein barostat with a time step of 2 fs. A multiple time step RESPA integrator algorithm was implemented for all simulation steps with a time step of 2fs for bonded, 2fs for 'near' non-bonded and 6fs for 'far' non-bonded interactions. Simulation data were retrieved at each 10 ps and the visual inspection of three-dimensional structures, trajectories were done using Maestro graphical user interface.

3. Results and discussion

3.1. 3D-QSAR model construction and visualization of contour maps

The atom-based 3D-QSAR analysis generated a statistically significant model along with good predicted activities of test set compounds when PLS factor was 6. The Biological activity predicted by the generated 3D-QSAR model is documented in Table 1. The PLS regression summary for generated 3D-QSAR is represented in Table 2. The low value of standard deviation (SD) (0.1171) and root-mean-square deviation obtained from the present study define the significance and reliability of the model. Moreover, high regression coefficient (0.9737) for training set, in addition with the stability ranges from 0.862 to 0.58 on maximum scale of 1, F value (Variance ratio) of 166.5 with smaller P value (< 0.005) and Pearson-r of 0.9183 reflected the relevance and confidence of the model respectively (Table 2). The stability for predicting unknown compounds in the test set for the generated 3D-QSAR model was indicated by the low value of RMSE (Root Mean Square Error), P (Significance level of variance ratio) and SD (Table 2).

The 3D-QSAR was validated by its closeness in predicting the activity of test set ligands (Table 1). Fig. 2 illustrates the scatter plots which describe that the experimental and predicted activity of sulfone derivatives showed moderate difference and good linear correlation between experimental and Phase predicted biological activity values.

The developed 3D-QSAR model was further validated by predicting the experimental activity of compounds not included in model development (external test set). The predicted $pEII_{50}$ values of the compounds in external test set are documented in **Supplementary** Table 2. A scatter plot of experimental vs. predicted $pEII_{50}$ values of external test set is illustrated in **Supplementary** Fig. 1**A**. The plot of predicted activity vs. the residual activity is also shown in **Supplementary** Fig. 1**B**, which was used to identify the outlier of the developed 3D-QSAR model. It was found that, the QSAR model was able to predict the experimental activity with a correlation coefficient (R^2) of 0.74 and cross-validation coefficient (Q^2) of 0.61. The R^2 value and Q^2 above 0.5 indicate the generated 3D-QSAR model to be good and able to endorse the experimental inhibitory activity($pEII_{50}$) of compounds included in

Table 2
Regression summary of generated 3D-QSAR model.

PLS	SD	R^2	F	P	Stability	RMSE	Q^2	Pearson-r
1.	0.3744	0.6811	68.3	1.92e-09	0.862	0.39	0.6446	0.8572
2.	0.2992	0.8027	67.1	1.19e-11	0.828	0.36	0.6936	0.8709
3.	0.2160	0.9005	90.5	3.94e-15	0.702	0.31	0.7720	0.8925
4.	0.1719	0.9391	111.8	3.48e-17	0.647	0.34	0.7312	0.8713
5.	0.1418	0.9600	134.2	1.16e-18	0.632	0.29	0.8067	0.9092
6.	0.1171	0.9737	166.5	5.02e-20	0.58	0.30	0.7948	0.9183

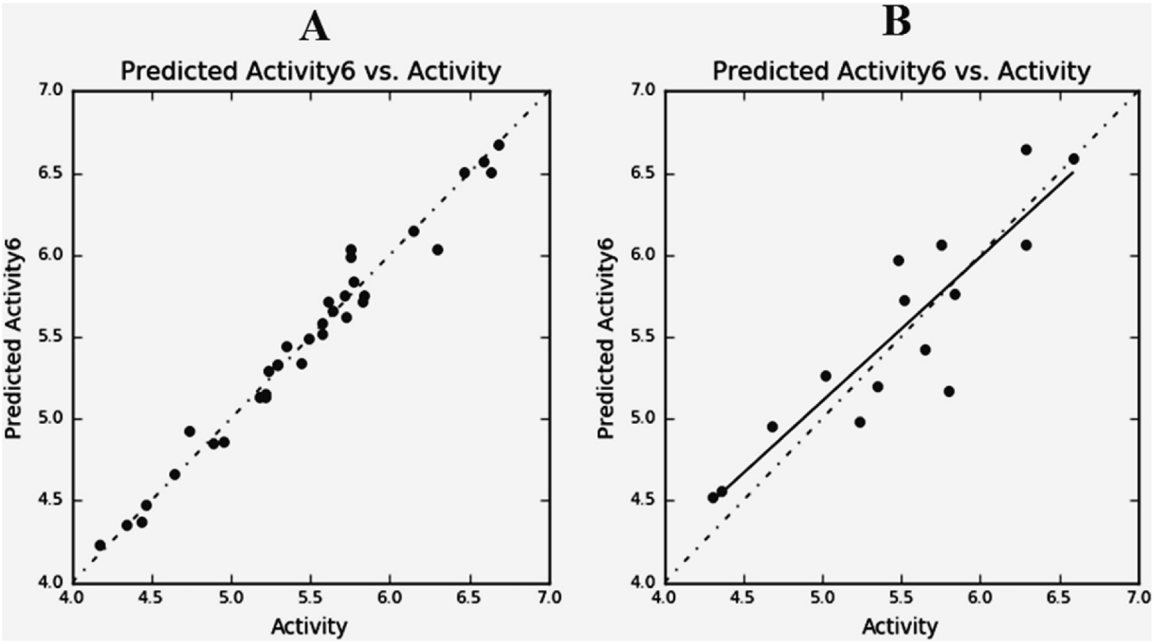


Fig. 2. Fitness graph between observed activity versus phase-predicted activity for (a) training set (b) test set compounds with the best fit line [y = 0.88x + 0.71 (R^2 = 0.81)].

external test set (Golbraikh and Tropsha, 2000). Moreover, **Supplementary** Fig. 1**B** illustrates no outlier in this study. Hence, the model can be considered as stable.

In order to interpret the effect of the spatial arrangement of structural features on biological activity, contour plot analysis was performed. The structural features were the presence of H-bond donor group, hydrophobic group, electron withdrawing group and an anionic group. Their individual positive contribution was represented by blue cubes and negative contribution was shown by red cubes. Comparisons of most significant favourable and unfavourable interactions are represented in Fig. 3. Since the substitutions were mainly situated at R9 ring, the contour plot appeared at 2', 4' and 6' position of this region. The list of functional groups present at 2', 4', 6' site of the R9 ring of all 50 compounds is given in **Supplementary** Table 1 to correlate the contour plot with the substitution. The generated 3D-QSAR model was implemented on **compound 39** ($pEII_{50}$= 5.02) for better visual inspection of the model. The effects of different substituents are discussed as follows.

3.1.1. Presence of anionic group

It can be identified from Table 1 that highly active **compound 2**($pEII_{50}$= 6.63), **compound 3**($pEII_{50}$= 6.68), **compound 4** ($pEII_{50}$=6.58) and moderately active **compound 11** ($pEII_{50}$=5.83) have phenoxide group in 2', 4' and 6' position of R9 pharmacophoric ring, which is anionic or electron donating in nature. The blue cube region at 2', 4' and 6' position (Fig. 3A**)** are representing the favourable and unfavourable contribution of anionic groups which completely agrees with the aforementioned compounds. Although **compound 1**($pEII_{50}$= 8.302) was not considered in the 3D-QSAR calculation, it is having phenoxide ions at 2' and 4' position. Presence of electron donating groups at R9 ring increases the electron density in R9 ring and SO_2 unit of SO compounds hence increase the potency of ligands (Lopez de Compadre et al., 1987). A remarkable decrease in biological activity was found for **compound 50** ($pEII_{50}$= 4.17), **compound 46** ($pEII_{50}$= 4.44), **compound 44** ($pEII_{50}$= *4.64*) due to lack of anionic or electron donating group in the blue cube region (Fig. 3**-A**) of R9 ring. However, a too bulky group at 6' position of the R9 ring may give some unfavourable anionic contribution in ligands.

3.1.2. Presence of hydrogen bond donor group

For hydrogen bond donor aspect, the blue cubes at 2' and 4' indicate the preference of hydrogen bond donor (Fig. 3B) groups at those positions. Small groups like NH_2, OH at 2' and 4' position of R9 ring generally contributes to favourable hydrogen bonding interactions. It is found from Table 1 that most of the highly active compounds like **compound 4**, **compound 6** ($pEII_{50}$= 6.46), **compound 8** ($pEII_{50}$= 6.29), **compound 9** ($pEII_{50}$= 6.29) and **compound 10**($pEII_{50}$= 6.14) have OH group in 2' position and 4' position, hence increment in activity was found for those compounds. Whereas the presence of bulky substituents like $-CONH_2$ groups and absence of hydrogen bond donor groups in those positions result in a decrease in biological activity of **compound 44**($pEII_{50}$=4.64).

3.1.3. Presence of hydrophobic group

The other important component that impact on biological activity is a hydrophobic character, as displayed in Fig. 3**C** which represents the presence of blue cubes at 2', 4', 6' positions of the R9 pharmacophoric site due to the presence of small hydrophobic groups for favourable hydrophobic interaction with 1AJ0 catalytic pocket. This assumption is supported by highly active **compound 3**, **compound 7**($pEII_{50}$= 6.30), **compound 9** and moderately active **compound 15** ($pEII_{50}$=5.77), **compound 17**($pEII_{50}$= 5.75) and **compound 19**($pEII_{50}$=5.72) because $-CH_3$ is substituted at 2' and 4' positions (Table 1) in aforementioned ligands. The presence of hydrophobic groups can induce stiffness in the covalent bond which is attached with sulphonyl group, hence decreases the entropy of ligands and increase the activity (Lopez de Compadre et al., 1987). **Compound 50**, **compound 48** ($pEII_{50}$=4.34)**, compound 44** lack hydrophobic groups in 2' and 4' which is responsible for the low 1AJ0 inhibitory activity.

3.1.4. Presence of electron-withdrawing group

The existence of red cubes at 2' and 4' position of the R9 pharmacophoric feature indicates the unfavourable position of electron withdrawing group (Fig. 3D). The presence of electron withdrawing group in R9 pharmacophoric ring reduces the electron density in this ring and sulphonyl unit, hence decreases the inhibitory activity of sulfone compounds. **Compound 50 (**$pEII_{50}$=4.17**), compound 46** ($pEII_{50}$=4.44), **compound 48** ($pEII_{50}$=4.34), **compound 47**

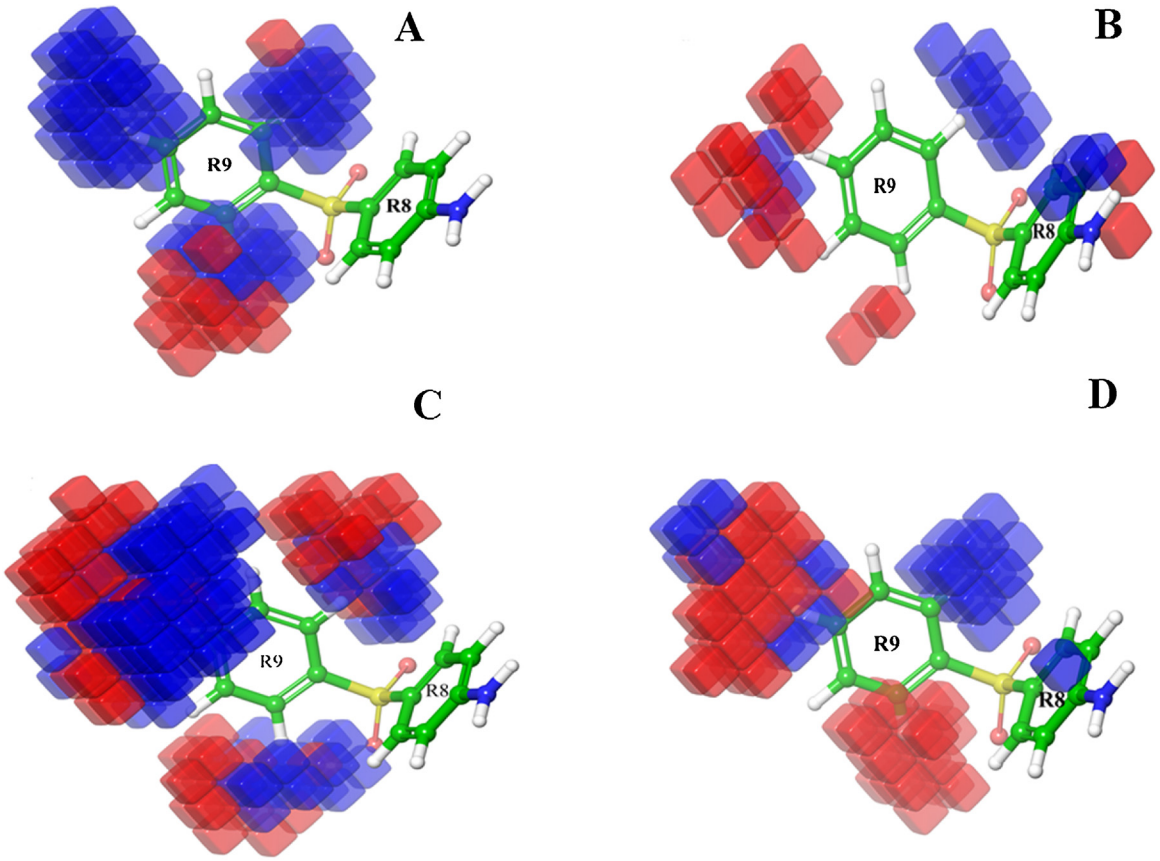


Fig. 3. Visualization of 3D -QSAR contour plot on **compound 39**(pEII50 = 5.02) Effect of A) presence of the anionic group, B) hydrogen bond donor group, C) hydrophobic group D) electron withdrawing group predicted by generated 3D-QSAR model.

($pEII_{50}$ = 4.36) and **compound 49(**$pEII_{50}$ = 4.3**)** have electron withdrawing group at 4' and 2' position resulting decrease in biological activity (Table 1)

3.2. Molecular docking

Molecular docking suggested that the interactions were mainly influenced by hydrogen bonding and π-π stacking interactions due to the presence of two aromatic rings in sulfone compounds. It is found from Table 1 that as we move from compounds substituted with electron donating groups to electron withdrawing groups, the inhibitory potency of sulfone drug molecules decreases. The substitution at ortho-position is attributed to the stability of complex as it helped the molecule to fit into the binding site more rigidly (Lopez de Compadre et al., 1987). The interactions were mainly resided in the region of Thr62 to His 257 due to the presence of catalytic site in this portion. It was found that Thr 62, Arg220 and Ser222 involved in hydrogen bonding network with most of the sulfone(SO) derivatives (**Compound 1–50)** and Phe190, His257 exhibited π-π stacking interaction with the R8, R9 aromatic rings present in the SO molecules. The binding energies of sulfone derivatives were found to span from -6.09 Kcal/mol to -4.03 Kcal/mole at the *p*ABA binding site of DHPS and showed good correlation with experimental activity value.

The ligand interaction diagram of docked compounds was illustrated in Fig. 4. The inhibitory activity of highest potent drug molecule (**Compound 1**) can be attributed due to the presence of two electron donating groups -ONa (which easily dissociates as O^- and Na^+) at *ortho* and *para* position in the ring R9 which enhance the electron density in common moiety (4-NH_2-$C_6H_4SO_2$) (ring R8) (De Benedetti et al., 1989) supporting for stronger π- electron interactions. The anionic functional group of the ligand was found to form salt bridges with Arg63 and His257 residues, resulting in extra stabilization of **compound 1** at the *p*ABA binding pocket (Fig. 4A). Similar observations were also implied by the generated 3D-QSAR model that the presence of negative ionic groups at R1, R2, and R3 are desirable for biological activity. The protein-ligand complex of **compound 1** was found to be stabilized by three hydrogen bonding interaction, π-π interaction, π-cation interaction and salt bridge formation (Fig. 4A**).** The carbonyl groups of Thr62 and Pro145 residues were contributing hydrogen bonding interactions with amine hydrogen of **compound 1** at a distance of 1.98Å (Thr62-C═O — H—N) and 1.92Å (Pro145- C═O — H—N) respectively. Sulphonyl oxygen of **compound 1** showed hydrogen bonding with the amide hydrogen of Ser222 residue (S═O —H-N- Ser222). Phe190 and Lys221 exhibited π- π stacking and π- cation interactions respectively with the ring R8 of the most potent drug. In addition to these interactions, extra stabilization was found between the phenoxide ions in ligand and the amino acid residues Arg63 and His257 due to salt bridge formation. The binding energy of highly potent **compound 3** and **compound 4** also showed good agreement in accordance with the 3D-QSAR results. The methyl groups present at 2', 6' position in the R9 ring of **compound 3** was found to involve in hydrophobic interactions with Pro232, Arg220 and Lys221 residues respectively and presented in **Supplementary** Fig. 2. Both the **compounds 3** and **4** exhibited one hydrogen bonding interaction in the complex (O^- — H-N- Arg220 and H–O — H-N- Arg220 respectively) and also showed π- π stacking interaction with His257 (Fig. 4**B & C**). Arg63 and Lys221 were found to form π-cation interaction with the ring R9 and the ring R8 respectively. The lowest potent drug molecule (**compound 50**) showed hydrogen bonding interactions with the amino acid residues Thr62 (Thr62C═O—HN), Ser222 (Ser222NH—O═S) and π-cation interaction with Lys221. The formation of the salt bridge was absent in the **compound 50** due to the absence of anionic group

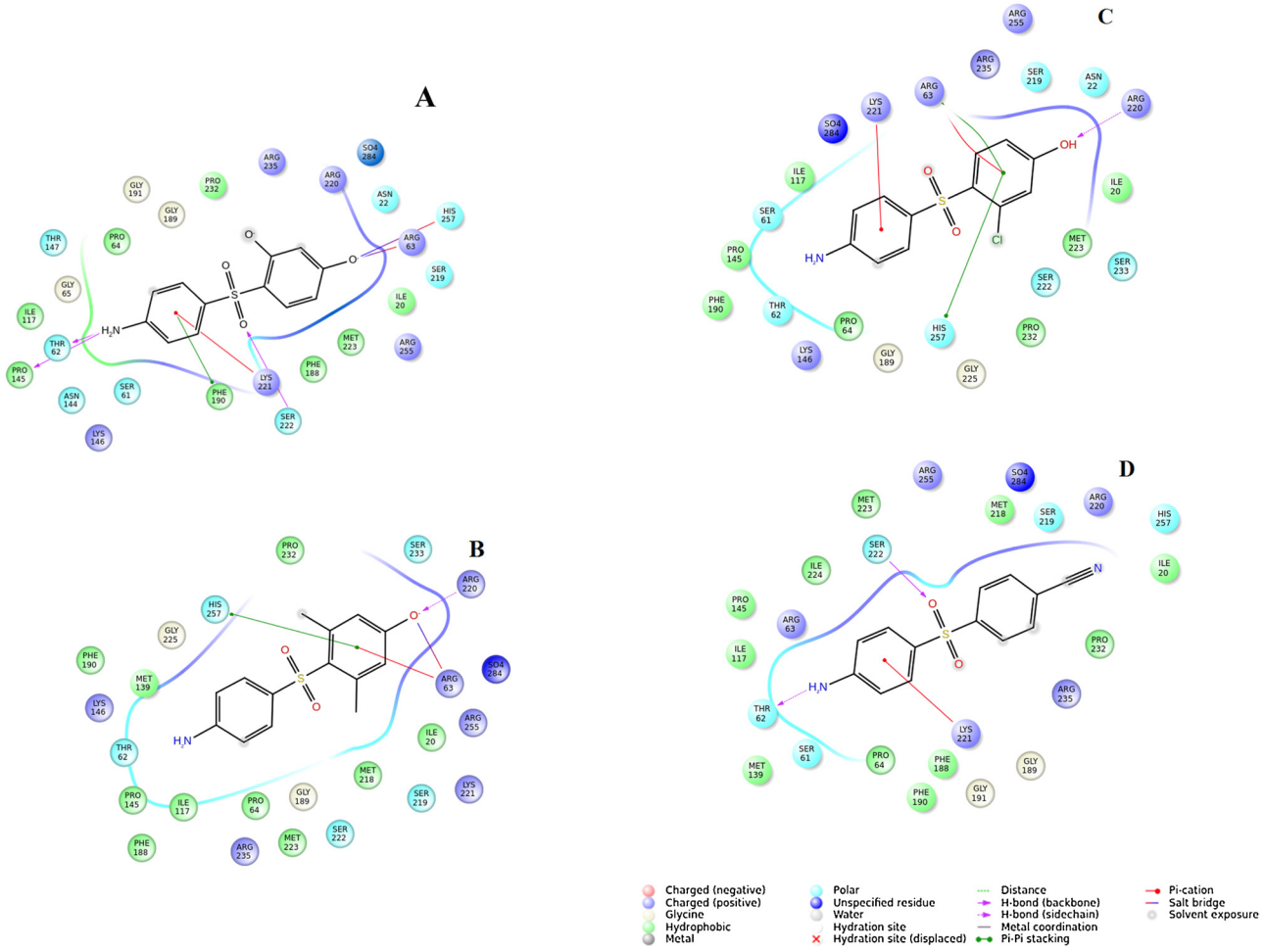


Fig. 4. 2D-ligand interaction diagram of A) compound 1, B) compound 3, C) compound 4 and D) compound 50 in the catalytic pocket of 1 AJ0.

(Fig. 4D). In addition **compound 50** was found to make steric clashes with the side chain of Lys 221 (**Supplementary** Fig. 3**D**). Fewer non-bonded interactions along with unfavourable steric clashes destabilize the **compound 50** in flexible *p*ABA binding site and cause lower potency of this compound.

The reliability of docking results was assessed by re-docking the co-crystallized ligand (*p*ABA) of protein DHPS (PDB ID: 1AJ0) into the binding site of the protein and calculating RMSD values between re-docked pose and the X-ray crystal structure. It is considered that RMSD values up to 2Å are reliable (Kramer et al., 1999). The docking pose and the interactions obtained after re-docking matches with the known crystal structure conformation with RMSD value of 1.21Å (**Supplementary** Fig. 4).

3.3. Frontier molecular orbital analysis

The HOMO, LUMO energies of highly active **compound 1**, **compound 4**, moderately active **compound 15** and less active **compound 21**, **compound 50** were calculated and documented Table 3. The aforementioned compounds were chosen to observe how functional group substitutions in R9 ring of diaryl sulfone compounds contribute for better stability of these derivatives in *p*ABA binding pocket of EcDHPS. **Compound 1** was chosen to observe the effect of two phenoxide groups (O^-) (anionic in nature) present at 2' and 4' position of R9 ring. **Compound 4** was chosen to see the effect of OH group (hydrogen bond donor group) located at 4' position of R9 ring. Moderately active **compound 15** was found to be important to investigate the effect of CH_3 substitution at 2', 4' and 6' position of the same ring. We also studied the effect of electron withdrawing group in **compound 21**, **compound 50** having chloro (Cl) group at 2', 4' position and cyano group (CN) at 4' position of R9 ring respectively. The variation in conformational energy due to substitutions in R9 was assessed from the differences of HOMO, LUMO orbital energies of chosen sulfone compounds and listed in Table 3.

From Table 3 it is evident that HOMO, LUMO energy differences, E_{gap} (eV) of highly potent and less potent compounds are in same range. This shows that substitution of various functional groups at R9 ring does not play much role in improving conformational stability and recognition ability. However the location of HOMO, LUMO orbitals and the total energy will be different when these ligands are bound to the protein. Therefore, it will be interesting to study the position of HOMO, LUMO and their energies in interacting amino acids and ligand complex. We have presented the HOMO, LUMO energy details of the complexes in Table 3 which will be discussed later. The distribution of HOMO orbitals on the sulfone molecules are illustrated in Fig. 5. Red and blue colours represent the positive and negative lobes of HOMO orbitals located at the sulfone compounds.

It is clear from Fig. 5**-A,B** that HOMO orbitals cover the entire R8 and R9 position for highly potent **compound 1** and **compound 4**. Moderately active **compound 15** found to have some portion of HOMO orbitals distributed over R9 ring (Fig. 5C). The accumulation of HOMO orbitals on R9 ring indicates the influence of electron donating group in increasing the electron density over this aromatic ring. In comparison to highly active and moderately active compounds, HOMO orbitals were absent in R9 position for less active **compound 21**, **compound 50** (Fig. 5**D** and **E**). HOMO orbitals were mainly located at R8 and NH_2

Table 3
HOMO- LUMO Energy details of Ligands at their bound state and unbound state.

HOMO- LUMO Energy Gap Details								
Sl. No	Compound $E_{compound}$ (eV)	$^{*}E_{HOMO}$ (eV)	$^{*}E_{LUMO}$ (eV)	$^{*}E_{gap}$ (eV)	Complex $E_{complex}$ (eV)	$^{*}E_{HOMO}$ (eV)	$^{*}E_{LUMO}$ (eV)	$^{*}E_{gap}$ (eV)
1	−33133.41	−5.977	−1.240	4.737	−146150.07	−0.774	−0.456	0.318
4	−43591.84	−5.877	−1.388	4.489	−135677.67	−0.508	−0.167	0.341
15	−32248.50	−5.856	−1.318	4.538	−90284.429	−0.444	−0.095	0.349
21	−54051.36	−6.095	−1.892	4.203	−126509.33	−0.212	0.083	0.295
50	−31550.25	−6.236	−2.220	4.016	−86158.636	−1.499	−1.142	0.357

$^{*}E_{HOMO}$ and E_{LUMO} defines the energy of HOMO, LUMO orbitals and Egap = E_{LUMO}- E_{HOMO}.

portion of **compound 21** and **50**. This is due to the presence of electron withdrawing group at R9 ring which decreases the electron density over this ring. Fig. 6 illustrates the distribution of LUMO orbitals over the interacting amino acid residues located at *p*ABA binding pocket of EcDHPS. The pink and green colours represent the positive and negative lobes of LUMO orbitals. According to the frontier molecular orbital theory, we know that HOMO of the ligands interacts with LUMO located at the catalytic pocket of protein. Therefore, it would be interesting to see the position of the LUMO orbitals present in the catalytic pocket. We have shown the orientation of the interacting amino acids over the sulfone compounds in **Supplementary** Fig. 3. Fig. 6 shows that LUMO orbitals are mainly located on the carbocation of pterin moiety, Arg220, Lys 221 and His257.

The presence of LUMO orbitals at the carbocation of pterin moiety (Fig. 6) and HOMO at NH_2 portion of SO (Fig. 5) compounds indicates the formation of a pterin-sulfa complex by the nucleophilic attack. The HOMO positions at the R9 ring and the LUMO located at Arg220 and His257 indicates that the substitutions at R9 ring interacts with those amino acid residues and stabilizes the ligands at the *p*ABA pocket. The closer proximity of HOMO of the ligand and the LUMO of the interacting amino acid in **compound 1**, **compound 4** and **compound 10** strongly suggests the reason of higher binding affinity of the protein-ligand complex due to the presence of electron donating, anionic groups in R9 ring. Absence of HOMO orbitals at the R9 ring of less active **compound 21** and **compound 50** gives less stability to these ligands at the pABA binding pocket. This can be correlated with our docking studies. Since, it is computationally expensive to do the DFT calculation for the whole protein molecule, so we chose the interacting amino acids and the ligand to see the HOMO, LUMO orbitals. We calculated the total energy of each atom clusters consist of each aforementioned sulfone compounds and interacting amino acid residues obtained in docking studies using self-consistent field approach. The HOMO, LUMO energy values are listed in Table 4 and the HOMO, LUMO energy gap are also calculated. The lower energy gap for each atom cluster indicates the easy movement of electrons between HOMO, LUMO orbitals. Fig. 7 shows the distribution of HOMO, LUMO orbitals at each complex. The total energy of the cluster containing the atoms of interacting amino acid residues with highly potent compound and the least potent

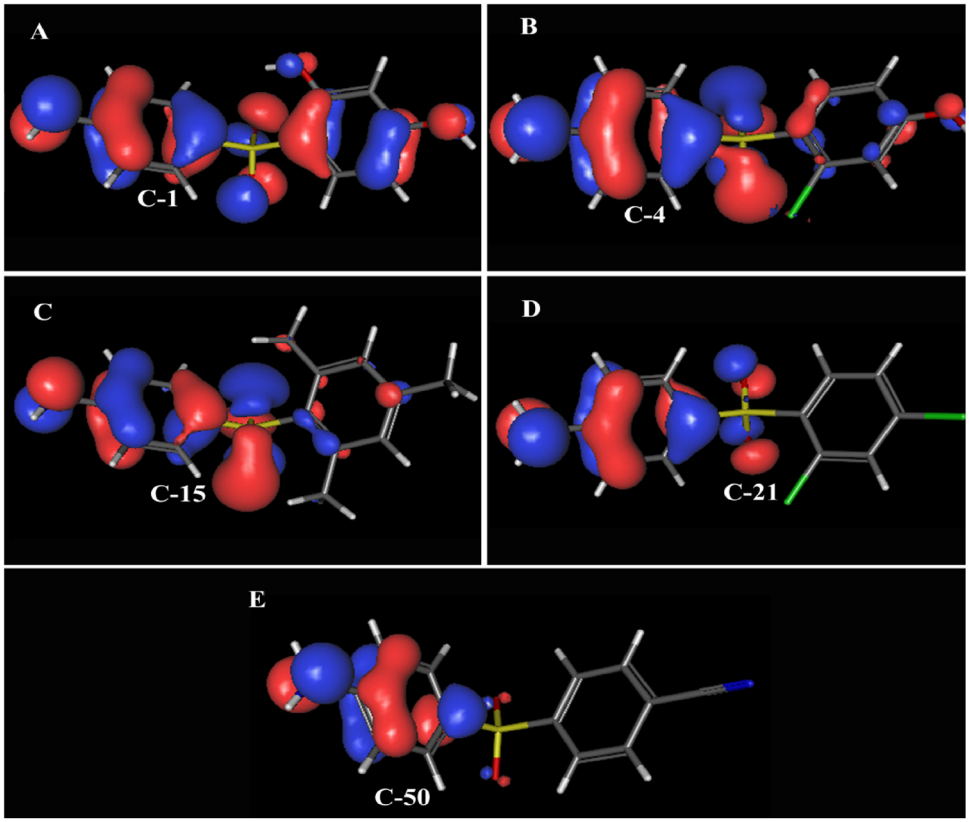


Fig. 5. The distribution of HOMO orbitals on selected sulfone drug molecules. (A) **Compound 1**, (B) **compound 4**, (C) **compound 15**, (D) **compound 21** and (E) **compound 50** respectively.

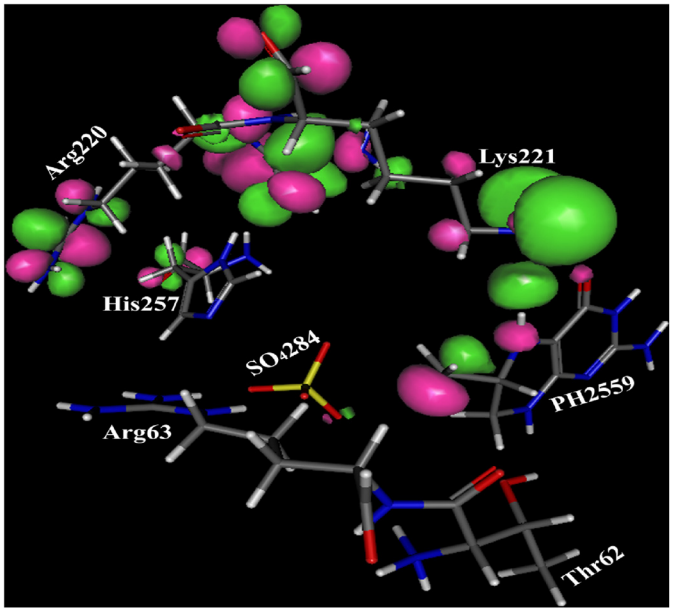


Fig. 6. The location of LUMO orbitals at the pABA binding site without diaryl sulfone compounds. PH2559 and SO_4284 represent the Pterin and Sulfate ion respectively. The location of LUMO orbital at the carbocation of PH2559 indicates the site prone to nucleophilic attack.

compound was found to be -33133.41 and -31550.25 eV respectively which indicates the greater stability of **compound 1** than **compound 50** at *p*ABA binding pocket. Fig. 7**A** illustrates the location of HOMO orbitals in **compound 1** and the LUMO region on pABA binding site. The presence of LUMO orbitals at Arg63, Lys221 and His257 indicated their interaction with the HOMO orbitals located at R9 ring and phenoxide groups of **compound 1**. These interactions found to stabilize the **compound 1** between the two flexible loops of *p*ABA binding pocket. Particularly, the presence of LUMO orbitals on Arg63 indicated its role in stabilizing sulfone compound after its binding at the catalytic pocket because LUMO orbitals did not appear in the amino acid cluster of the catalytic pocket of the protein alone (Fig. 6). Fig. 7**B** also shows that location of HOMO orbitals on R8 and R9 ring of **compound 4**. The location of LUMO orbitals in this complex was found to be similar as that of the amino acid complex containing **compound 1**. HOMO located on both R8 and R9 ring indicates the π-interactions of compound 4 with LUMO located at Lys221, His 257 and Arg 63 of *p*ABA pocket. Fig. 7**C** illustrates the position of HOMO, LUMO orbitals on the complex consist of **compound 15** and interacting amino acid residues. The location of LUMO orbitals over Arg63 and Lys 221 indicates π-interaction of these amino acids with the HOMO positioned in R8 and R9 ring of **compound 15**. This interaction stabilizes **compound 15** at pABA binding pocket and play an important role in improving the potency of this compound.

The positions of HOMO, LUMO orbitals on the complex containing **compound 21** and **compound 50** with interacting amino acid are depicted in Fig. 7**D** and **E** respectively. In both the complexes the HOMO orbitals were absent at R9 ring in comparison to highly active and moderately active compounds. In these cases, HOMO orbitals are located in R8 and NH_2 position sulfone compounds which interact with LUMO lies on Lys221 and Phe190. The absence of HOMO orbitals at R9 ring can be a strong reason for destabilizing the compounds between two flexible loops of *p*ABA binding pocket.

3.4. FEP calculations

FEP captures free energy difference between two ligands at the catalytic pocket of protein when ligand A changes to ligand B due to functional group mutation. The equation (Kollman et al., 1993; Zwanzig, 1954) for calculating relative free energy is given as:

$$G = \Delta G_A - \Delta G_B = -k_B T \ln [\exp(-(U_B - U_A)/k_B T)]\ A \quad (1)$$

Where $\boldsymbol{U_A}$ and $\boldsymbol{U_B}$ correspond to the potential energy of A and B states, $\boldsymbol{k_B}$ is the Boltzmann constant,

$\boldsymbol{G}$ is relative free energy difference when ligand A transforms to B and T is absolute temperature.

This process is valid only when the two ligands share the common core (Kollman et al., 1993), such that they overlap significantly in phase-space. The thermodynamic cycle for determining relative free energy calculation is depicted in Fig. 8. The relative binding free energy between the lowest potent (**Compound 50**, $pEII_{50}$ = 4.17, **compound 41**, $pEII_{50}$ = 4.89) and highly potent compounds in the data list (**compound 1**, $pEII_{50}$ = 8.302 and **compound 4**, $pEII_{50}$ = 6.59) were calculated using the FEP/REST method (Wang et al., 2015 2012). The experimental relative binding free energy of Ligand A and B was calculated using the following equation (Uciechowska et al., 2012)

$$\Delta G_A = -RT \ln(IC_{50}A) \quad (2)$$

$$\Delta G_B = -RT \ln(IC_{50}B) \quad (3)$$

$$\Delta\Delta G_{exp} = \Delta G_A - \Delta G_B \quad (4)$$

Where IC_{50}^{A} and IC_{50}^{B} are the biological activities (Inhibitory Concentration fifty) of ligand A and B respectively.

We took **compound 50** as our ligand A and **compound 1** as ligand B. **Compound 50** found to be involved in steric clashes with Ile20 and His257 (**First ligand,** Fig. 9**A-i**), but when the cyano group was mutated into phenoxide ($-O^-$) group at 2' and 4' position of R9 ring, gain in favourable interaction was achieved (**Second-ligand,** Fig. 9**A-ii**) by forming salt bridges with Arg 63 and His 257 residues at the pABA binding site of DHPS. The relative free energy gain was calculated based on Eq. (2). Mutating the cyano group (**-C≡N**) to phenoxide (**-O⁻**) group, the experimental binding free energy ($\boldsymbol{\Delta\Delta G_{EXP}}$) was found to be increased by -5.85 Kcal/mol due to gain in favourable salt bridge interactions. From FEP simulation, the computed relative free energy ($\boldsymbol{\Delta\Delta G_{CAL}}$) difference between **compound 50** and **compound 1** was found to be -3.78 kcal/ mol. Similarly, the increase in experimental binding free energy ($\boldsymbol{\Delta\Delta G_{EXP}}$) of -3.42 Kcal/mol was observed when cyano group (**Compound 50,** Fig. 9**B-i**) at 4' position of R9 ring, was mutated into a hydroxyl group (**Compound 4**, Fig. 9**B-ii**) at the same position of sulfone compounds. The cyano group found to be involved in steric clashes with Ile20 and His257. When the cyano group was mutated to hydroxyl (−OH) group, the resulted free energy gain due to the formation of a hydrogen bond with Arg220 was found to be -2.58 kcal/ mol.

To see the unfavourable steric effect we calculated the relative free energy difference when **compound 41** is mutated to **compound 1,** illustrated in Fig. 9**C-i&ii**. The methoxy group at 4' and 6' position of **compound 41** was found to involve steric clashes with the side chain of Arg63, Arg220, and His257. When the methoxy groups at 2' and 4' position are mutated into phenoxide ion ($-O^-$) in **compound 1**, the gain in favourable hydrogen bonding and salt bridge interaction with Arg63 and arg220 results in an increase in experimental binding free energy by -4.83 kcal/mol. This effect is captured by our FEP calculation by computing binding free energy difference ($\boldsymbol{\Delta\Delta G_{CAL}}$) between these two compounds of -1.94 kcal/ mol.

The timeline representation of free energy difference was found to reach a plateau within the simulation time (**Supplementary** Fig. 5) which indicates the convergence of free energy calculations. The relative free energy profile during the course of the simulation is shown in **Supplementary** Fig. 5 and the exchange density of FEP replicas over window is presented in **supplementary** Fig. 6.

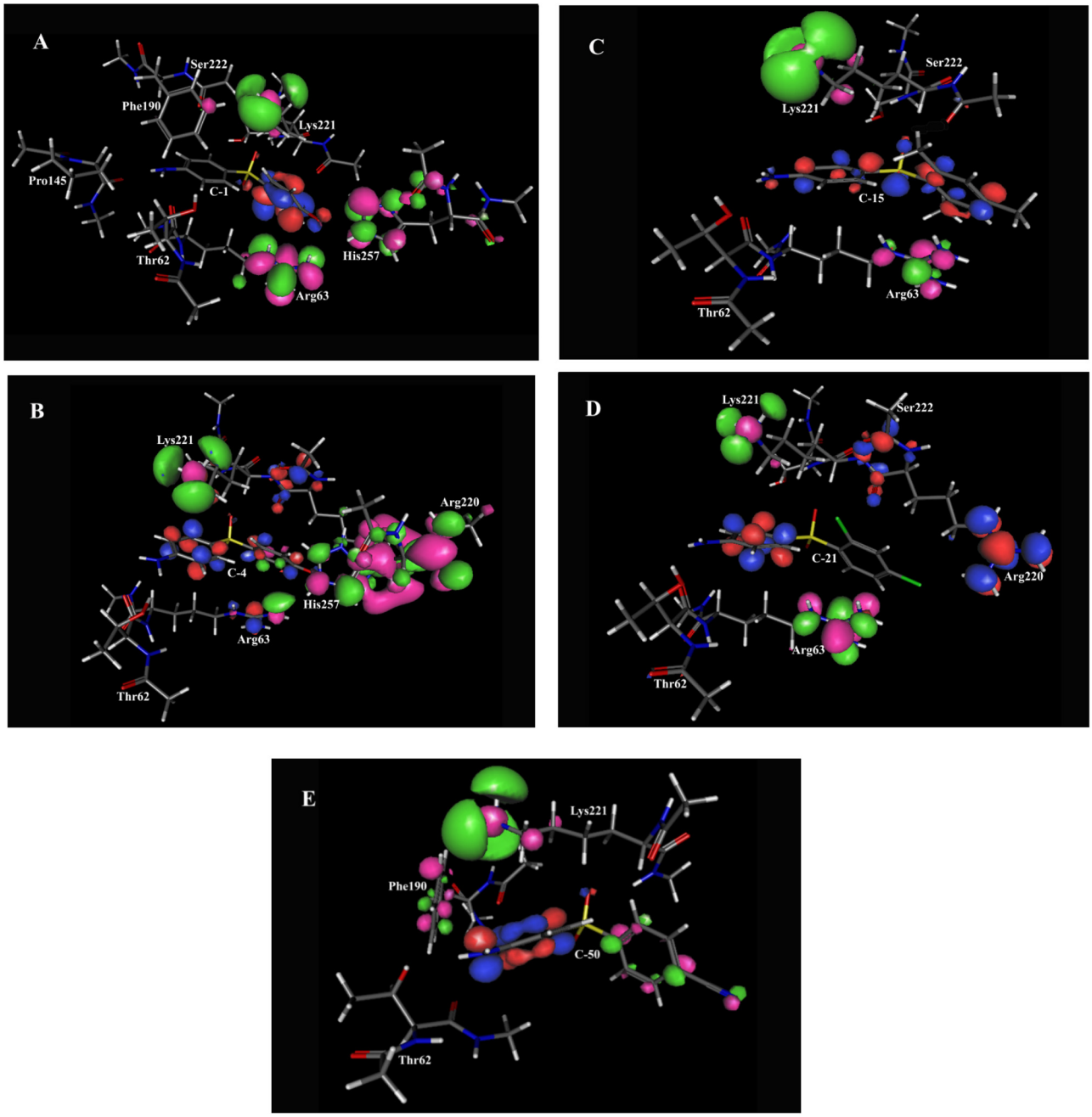


Fig. 7. The distribution of HOMO, LUMO orbitals on the atom cluster containing (A) **compound 1**, (B) **compound 4**, (C) **compound 15**, (D) **compound 21** and (E) **compound 50** and their respective interacting amino acid residues.

3.5. ADME/toxicity prediction of top scoring compounds

It is believed that only improving potency of a particular drug towards its target is not the primary objective in the drug development process, unless taking PK profile and toxicity in to consideration (Ekins et al., 2002). The various important pharmacokinetic properties of **compound 1** to **compound 10** and the permissible ranges are documented in **Supplementary** Table 3. The top scored compounds were thoroughly evaluated with the basic parameters of "Lipinski's rule of 5" and other pharmacokinetic parameters. Generally, violation of Lipinski's rule more the two is considered to be forbidden for orally active compounds. All the pharmacokinetic properties of the top scored compounds documented in **Supplementary** Table 3 were found to be under permissible limit.

The values of polar surface area and rotatable bond count of top scored compounds known to have great impact on their oral bioavailability. The number of rotatable bonds and the polar surface area in top ranked sulfone compounds are found to be in the range of 4–5 and 77-102 $Å^2$ respectively. These values found to be under recommended ranges thus expected to have good bioavailability. It can be observed from **Supplementary** Table 3 that the molecular weight of the top scored ligands are in the range of 263.31 to 283.72 which is acceptable for drug-like compounds(Lipinski et al., 1997). Moreover, the number of hydrogen bond donor groups and acceptor group found to be below the threshold limit (H-bond-donor should be less than or equal to 5, similarly the threshold value for number of H-bond acceptor is 10). This indicates good adsorption of diaryl sulfone in system circulation. Additionally, the present human oral absorption value found to be within 66.72–84.062%, indicating moderate to high adsorption of the top scoring compound. It is believed that a chemical compound will be drug

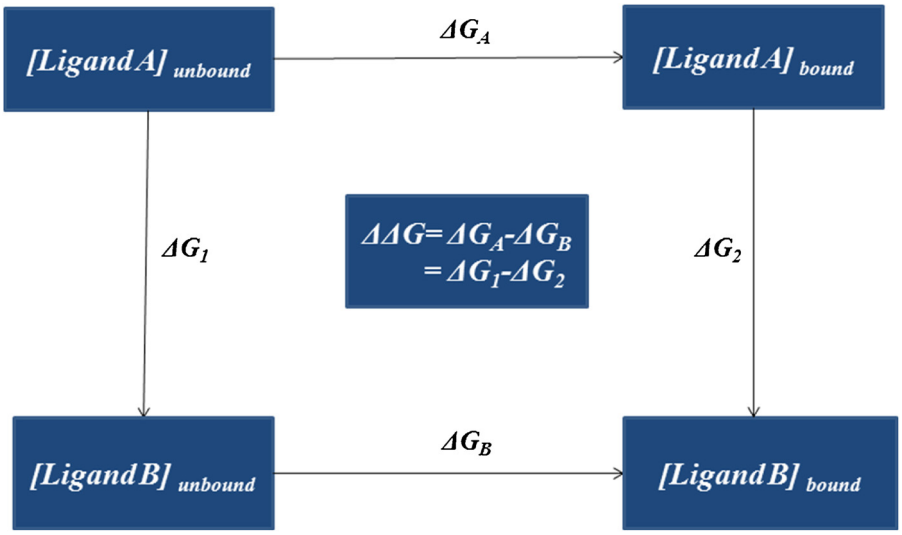


Fig. 8. The Thermodynamic cycle for calculating the relative free energy of binding between two states. ΔG_A and ΔG_B are the free energy difference for going unbound state to bound state. ΔG_1 and ΔG_2 are the free energy difference for transforming ligand A to ligand.

like if their water/octanol partition co-efficient is less than 5. It is obvious from Table 4 that the water/octanol partition co-efficient values are much below the threshold level. The values of solvent accessible surface area (SASA) and polar solvent area of all top scoring sulfone derivatives are found to be in accepted range (Duffy and Jorgensen, 2000). It is found from **Supplementary** Table 3 that all the top scored inhibitors did not violate the "Lipinski's rule of 5" (Lipinski et al., 1997) and "Jorgensen's rule of 3". Further, the human serum albumin binding affinity values and blood brain barrier coefficient values are found to be under permissible range. The predicted ADME properties of ten highly active drug compounds indicate acceptable pharmacokinetic and less toxicity profiles for Phase 1 clinical trials.

3.6. MD simulation

The dynamic behaviour of ligand at the binding pocket of enzyme is important to evaluate the stability of that particular ligand inside the active site. A 10 ns molecular dynamics simulation of 1 AJ0/**compound 4** provides further insights into molecular interaction of **compound 4** in motion. The Root Mean Square Deviation (RMSD) of the enzyme backbone with respect to its initial position increased up to 1.64 Å for first 4 ns and stabilized around 1.5 Å for rest of the trajectory. Moreover, the ligand movement cope well with the RMSD of binding site (Fig. 10A). The average RMSD of enzyme (Fig. 10A) backbone and heavy atoms are found to be 1.44 Å and 1.86 Å respectively. This reflects minute structural change of 1 AJ0 during simulation from the crystal structure of its own. The residue wise Root Mean Square Fluctuation (RMSF) of 1 AJ0 was illustrated in Fig. 10**B**. The detailed inspection of RMSF helped in characterizing the regional changes in the protein chain throughout the course of simulation. The maximum value of Cα, backbone, heavy atom (3.01, 3.07 and 3.33 Å respectively) RMSF was found for residue Glu277 which resides in a flexible region of Ser274-Glu282(C-terminal end). Additionally, Gly31 found to have high Cα RMSF of 2.50 Å at the N-terminal flexible region from Val23 to Ser36. Both of the highly fluctuating loops are found to reside away from binding pocket. Interestingly, the RMSF value of protein backbone residue at the catalytic site was found to be in the range of 0.51 Å–1.17 Å indicating the stability of catalytic region under motion.

The protein-ligand contacts during the course of simulation are illustrated in Fig. 11 and **Supplementary** Fig. 7. **Compound 4** was found to exhibit ionic interaction in the stable region (Fig. 11) of Thr62-Pro64 and Arg255-His257. Hydrogen bonding found to be formed majorly with Pro145, Ser222 and hydrophobic interaction was dominated by Phe190 throughout the simulation time. The region between Met223-Arg235 found to exhibit water mediated hydrogen bond with **compound 4**. The region from Gly65-Gly143 and Gly191-Ser219 no interaction was found due to higher fluctuation (Fig. 10**B** and **Supplementary** Fig. 7). During the course of simulation interaction observed with Arg63, Pro145, Phe190, Ser222, Met223, Arg255 and His 257 located to the active site are expected to be responsible for the inhibitory activity of **compound 4**. A low mean RMSD of 0.375 Å of **compound 4** indicates less conformational modification during simulation. The low solvent accessible surface area (SASA) (7.04-73.90 $Å^2$), high polar solvent area (PSA) (184-187 $Å^2$) and molecular surface area (MolSA) (233-238 $Å^2$) of **compound 4** further supports its stabilization during 10 ns molecular dynamics simulation (**Supplementary** Fig. 8).

4. Conclusion

In this present article, 3D-QSAR modelling, molecular docking, DFT calculations, FEP calculations and MD simulation were performed to explore the effect of different substitution on sulfone derivatives to stabilize them in *p*ABA binding loop of EcDHPS catalytic pocket. A highly predictive 3D-QSAR model was developed which helped in understanding how the structure of sulfone compounds affects the biological activity. Contour plot mapping was performed to observe the spatial arrangement of favourable and unfavourable functional groups and their contribution to inhibit the EcDHPS enzyme. Presence of electron donating group and hydrogen bond donor at 2' and 4' position of the R9 ring were found to increase the potency of sulfone compounds. Molecular docking was done to find the possible binding pose of sulfone compounds at the *p*ABA binding pocket. The docking study suggested that Thr62, Arg63, Pro145, Phe190, Lys 221, Ser222 and His257 amino acids are the key binding residues in the active site of EcDHPS. The highest potent **compound 1** contains two dissociated phenoxide ion which forms the salt bridge with Arg63 and His257. Possibly the mutations in Arg63 and His257 regions can disrupt the sulfone (SO) binding at DHPS catalytic pocket due to their active role in stabilizing the R9 ring. DFT studies showed the stabilization pattern and reaction mechanism of sulfone compound at electronic level. The HOMO, LUMO interaction pattern between sulfone compounds and *p*ABA binding site are also supported by docking studies. The presence of electron withdrawing group at 4' position directly affects the electron density over the R9 ring and an SO_2 moiety of the ligand which resulted **compound 50** to be least potent. Presence of hydrophobic groups found to have both favourable and unfavourable contribution

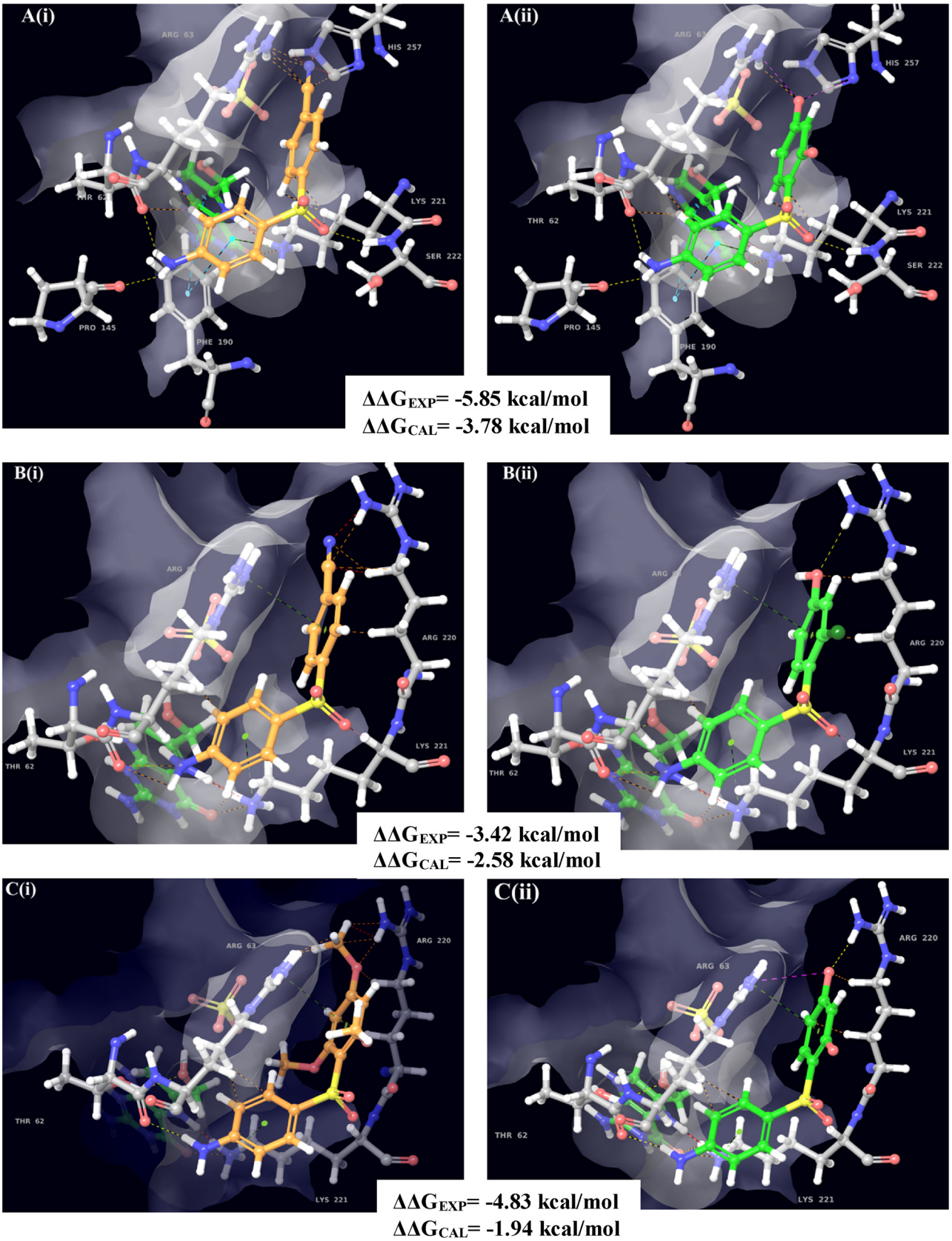


Fig. 9. Representative figures of different types of interactions captured by FEP. (A i&ii) Shows gain in favourable salt-bridge interaction due to the presence of phenoxide group (mutating **compound 50** to **compound 1**). (B i&ii) Shows gain in favourable salt-bridge interaction due to the presence of hydroxyl group (mutating **compound 50** to **compound 4**). (C i&ii) Shows gain in favourable salt-bridge and hydrogen bonding interaction due to the presence of phenoxide ion (mutating **compound 41** to **compound 1**).The hydrogen bond, Salt-bridge, π-cation, π-stacking and unfavourable interaction was illustrated by yellow, pink, green, sky and brown dotted line respectively.

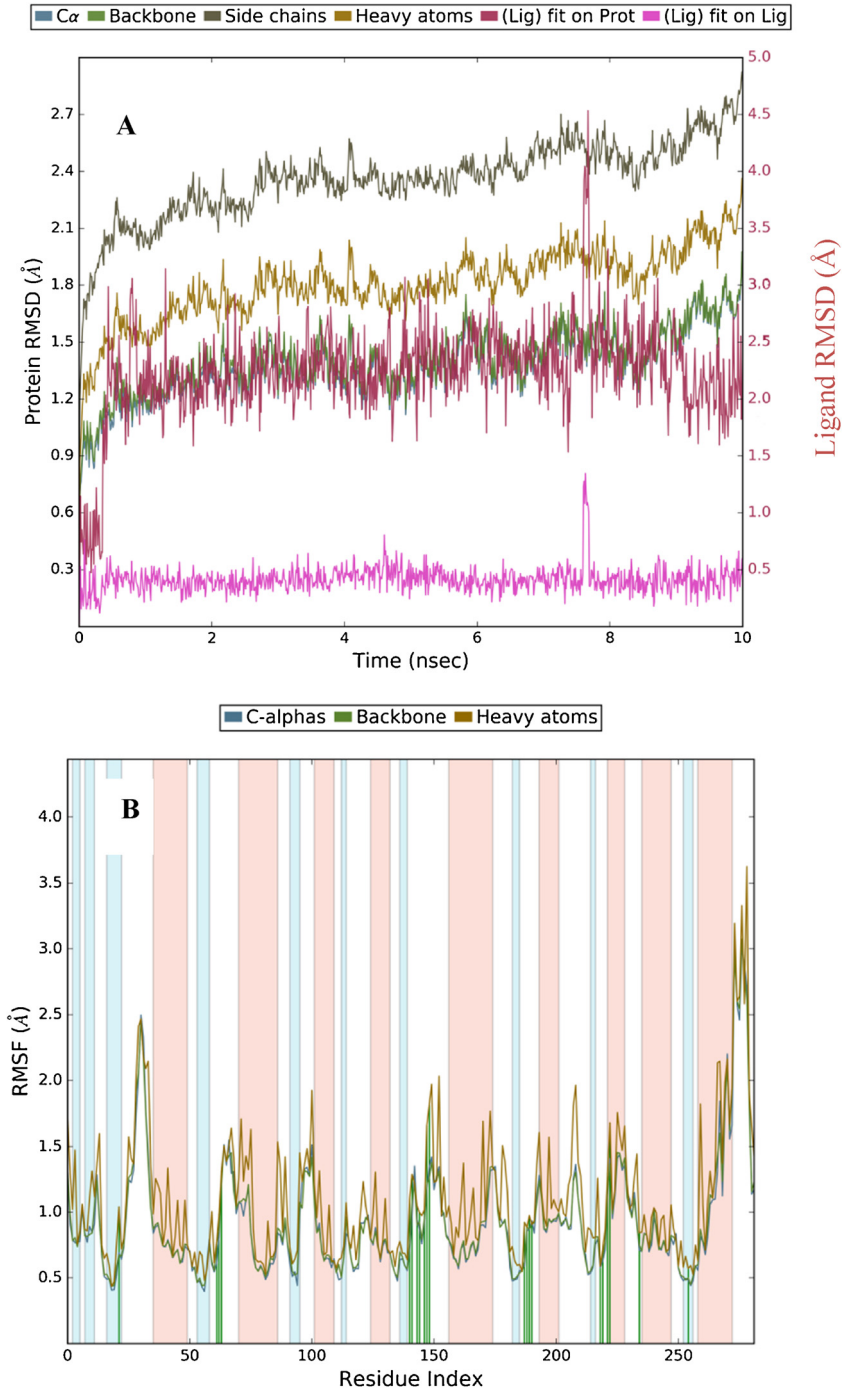


Fig. 10. (A)The RMSD (Å) of the simulated positions of 1 AJ0 backbone atoms from their initial coordinate in 10 ns MD trajectory. (B) The residue wise RMSF profile of 1 AJ0 illustrates the regional changes along the protein chain throughout the trajectories.

depending on the size of the group and site of substitution. Further, our conclusion was validated by FEP calculations by capturing relative binding free energies between less potent and highly potent compounds. The predicted ADME of top scored drug candidates are found to be in acceptable ranges. Lastly a 10 ns MD simulation of highly active **compound 4** indicates the fixation of this inhibitor in pABA catalytic pocket. This combinatorial computational study contributes a set of useful structural guidelines focusing on *p*ABA binding pocket, which will greatly help in designing sulfone-based and novel pterin-sulfa compounds for the EcDHPS-selective inhibitors to prevent infection caused by pathogenic and antibiotic-resistant *E.coli* strains.

Declaration of interest

The authors declare no conflict of interest.

Acknowledgment

Funding from DST, SERB (ECR/2016/000707) is highly acknowledged. Scholarship of BKD is acknowledged from DST. We would also like to thank Schrödinger Inc., Department of Chemistry and Department of Physics, NITK Surathkal for their constant support.

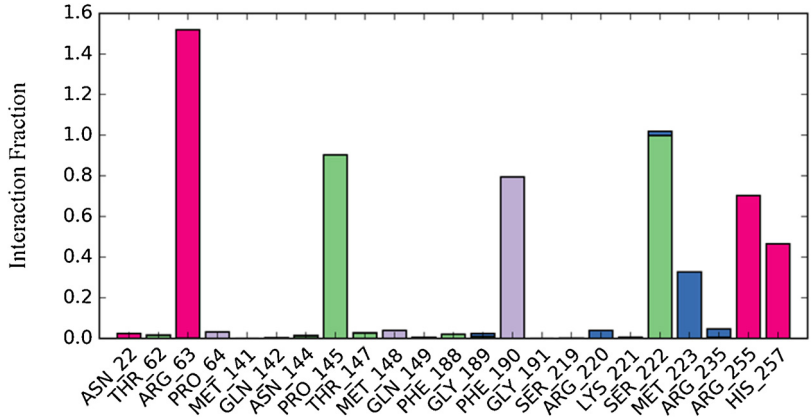


Fig. 11. Histogram represents interaction of compound 4 with different amino acids of 1 AJ0 during 10 ns MD trajectory. The pink colour represents the ionic interaction, green colour represents hydrogen bond, violet colour represents hydrophobic interaction and blue colour stands for water mediated hydrogen bonding.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.compbiolchem.2018.11.005.

References

Achari, A., Somers, D.O., Champness, J.N., Bryant, P.K., Rosemond, J., Stammers, D.K., 1997. Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. Nat. Struct. Biol. 4, 490–497.

Alun, Bermingham, Derrick Jeremy, P., 2002. The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. BioEssays 24, 637–648. https://doi.org/10.1002/bies.10114.

Bash, P.A., Singh, U.C., Brown, F.K., Langridge, R., Kollman, P.A., 1987. Calculation of the relative change in binding free energy of a protein-inhibitor complex. Science 235, 574–576.

Bemis, G.W., Murcko, M.A., 1996. The properties of known drugs. 1. Molecular frameworks. J. Med. Chem. 39, 2887–2893. https://doi.org/10.1021/jm9602928.

Bennett, C.H., 1976. Efficient estimation of free energy differences from Monte Carlo data. J. Comput. Phys. 22, 245–268. https://doi.org/10.1016/0021-9991(76)90078-4.

Bourne, C.R., 2014. Utility of the biosynthetic folate pathway for targets in antimicrobial discovery. Antibiotics 3, 1–28. https://doi.org/10.3390/antibiotics3010001.

Bowers, K.J., Chow, E., Xu, H., Dror, R.O., Eastwood, M.P., Gregersen, B.A., Klepeis, J.L., Kolossvary, I., Moraes, M.A., Sacerdoti, F.D., Salmon, J.K., Shan, Y., Shaw, D.E., 2006. Scalable algorithms for molecular dynamics simulations on commodity clusters. Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. https://doi.org/10.1145/1188455.1188544.

Campbell, M.L., Marchaim, D., Pogue, J.M., Sunkara, B., Bheemreddy, S., Bathina, P., Pulluru, H., Chugh, N., Wilson, M.N., Moshos, J., Ku, K., Hayakawa, K., Martin, E.T., Lephart, P.R., Rybak, M.J., Kaye, K.S., 2012. Treatment of methicillin-resistant Staphylococcus aureus infections with a minimal inhibitory concentration of 2 μg/mL to vancomycin: old (trimethoprim/sulfamethoxazole) versus new (daptomycin or linezolid) agents. Ann. Pharmacother. 46, 1587–1597. https://doi.org/10.1345/aph.1R211.

Chakraborty, S., Gruber, T., Barry, C.E., Boshoff, H.I., Rhee, K.Y., 2013. Para-aminosalicylic acid acts as an alternative substrate of folate metabolism in mycobacterium tuberculosis. Science 339, 88–91. https://doi.org/10.1126/science.1228980.

Chen, D., Ranganathan, A., IJzerman, A.P., Siegal, G., Carlsson, J., 2013. Complementarity between in silico and biophysical screening approaches in fragment-based lead discovery against the A2A adenosine receptor. J. Chem. Inf. Model. 53, 2701–2714. https://doi.org/10.1021/ci4003156.

Correa-Basurto, J., Ramos-Morales, F.R., Matus, M.H., Rosales-Hernández, M.C., Mancilla-Percino, T., Trujillo-Ferrara, J., Ilizaliturri-Flores, I., 2012. Docking and DFT studies to explore the topoisomerase II ATP pocket employing 3-substituted 2,6-piperazindiones for drug design. Mol. Simul. 38, 1072–1084. https://doi.org/10.1080/08927022.2012.690877.

De Benedetti, P.G., Iarossi, D., Menziani, C., Caiolfa, V., Frassineti, C., Cennamo, C., 1987. Quantitative structure-activity analysis in dihydropteroate synthase inhibition of sulfones. Comparison with sulfanilamides. J. Med. Chem. 30, 459–464. https://doi.org/10.1021/jm00386a004.

De Benedetti, P.G., Iarossi, D., Folli, U., Frassineti, C., Menziani, M.C., Cennamo, C., 1989. Quantitative structure-activity relationships in dihydropteroate synthase inhibition by multisubstituted sulfones. Design and synthesis of some new derivatives with improved potency. J. Med. Chem. 32, 2396–2399. https://doi.org/10.1021/jm00130a028.

Dixon, S.L., Smondyrev, A.M., Knoll, E.H., Rao, S.N., Shaw, D.E., Friesner, R.A., 2006. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. J. Comput. Aided Mol. Des. 20, 647–671. https://doi.org/10.1007/s10822-006-9087-6.

Domagk, G., 1935. Ein Beitrag zur Chemotherapie der bakteriellen Infektionen. DMW - Dtsch. Med. Wochenschr. 61, 250–253. https://doi.org/10.1055/s-0028-1129486.

Duan, L.L., Tong, Y., Mei, Y., Zhang, Q.G., Zhang, J.Z.H., 2007. Quantum study of HIV-1 protease-bridge water interaction. J. Chem. Phys. 127, 145101. https://doi.org/10.1063/1.2770720.

Duffy, E.M., Jorgensen, W.L., 2000. Prediction of properties from simulations: Free energies of solvation in hexadecane, octanol, and Water. J. Am. Chem. Soc. 122, 2878–2888. https://doi.org/10.1021/ja993663t.

Ekins, S., Boulanger, B., Swaan, P.W., Hupcey, M.A.Z., 2002. Towards a new age of virtual ADME/TOX and multidimensional drug discovery. Mol. Divers. 5, 255–275.

Ferguson, David M., 1995. Parameterization and evaluation of a flexible water model. J. Comput. Chem. 16, 501–511. https://doi.org/10.1002/jcc.540160413.

Forgacs, P., Wengenack, N.L., Hall, L., Zimmerman, S.K., Silverman, M.L., Roberts, G.D., 2009. Tuberculosis and trimethoprim-sulfamethoxazole. Antimicrob. Agents Chemother. 53, 4789–4793. https://doi.org/10.1128/AAC.01658-08.

Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Rob, M.A., Cheeseman, J.R., et al., 2009. Gaussian 09. Gaussian, Inc., Wallingford.

Fukunishi, H., Watanabe, O., Takada, S., 2002. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. J. Chem. Phys. 116, 9058–9067. https://doi.org/10.1063/1.1472510.

Gasteiger, J., Marsili, M., 1980. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. Tetrahedron 36, 3219–3228. https://doi.org/10.1016/0040-4020(80)80168-2.

Golbraikh, A., Tropsha, A., 2000. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. Mol. Divers. 5, 231–243. https://doi.org/10.1023/A:1021372108686.

Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y.-D., Lee, K.-H., Tropsha, A., 2003. Rational selection of training and test sets for the development of validated QSAR models. J. Comput. Aided Mol. Des. 17, 241–253. https://doi.org/10.1023/A:1025386326946.

Goldfeld, D.A., Murphy, R., Kim, B., Wang, L., Beuming, T., Abel, R., Friesner, R.A., 2015. Docking and Free energy perturbation studies of ligand binding in the kappa opioid receptor. J. Phys. Chem. B 119, 824–835. https://doi.org/10.1021/jp5053612.

Griffin, M.J., Brown, G.M., 1964. The biosynthesis of folic acid. iii. enzymatic formation of dihydrofolic acid from dihydropteroic acid and of tetrahydropteroylpolyglutamic acid compounds from tetrahydrofolic acid. J. Biol. Chem. 239, 310–316.

Hampele, I.C., D'Arcy, A., Dale, G.E., Kostrewa, D., Nielsen, J., Oefner, C., Page, M.G.P., Schönfeld, H.-J., Stüber, D., Then, R.L., 1997. Structure and function of the dihydropteroate synthase from staphylococcus aureus 11 Edited by R. Huber. J. Mol. Biol 268, 21–30. https://doi.org/10.1006/jmbi.1997.0944.

Hevener, K.E., Yun, M.-K., Qi, J., Kerr, I.D., Babaoglu, K., Hurdle, J.G., Balakrishna, K., White, S.W., Lee, R.E., 2010. Structural studies of pterin-based inhibitors of dihydropteroate synthase. J. Med. Chem. 53, 166–177. https://doi.org/10.1021/jm900861d.

Jorgensen, W.L., Duffy, E.M., 2000. Prediction of drug solubility from Monte Carlo simulations. Bioorg. Med. Chem. Lett. 10, 1155–1158. https://doi.org/10.1016/S0960-894X(00)00172-4.

Jorgensen, W.L., Jenson, C., 1998. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: seeking temperatures of maximum density. J. Comput. Chem. 19, 1179–1186. https://doi.org/10.1002/(SICI)1096-987X(19980730)19:10<1179::AID-JCC6>3.0.CO;2-J.

Jorgensen, W.L., Ravimohan, C., 1998. Monte Carlo simulation of differences in free energies of hydration. J. Chem. Phys. 83, 3050. https://doi.org/10.1063/1.449208.

Jorgensen, William L., Maxwell, David S., Tirado-Rives, J., 1996. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids [WWW Document]. https://doi.org/10.1021/ja9621760.

Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., Jorgensen, W.L., 2001. Evaluation and reparametrization of the OPLS-AA force Field for proteins via comparison with accurate quantum chemical calculations on peptides. J. Phys. Chem. B 105, 6474–6487.

https://doi.org/10.1021/jp003919d.
Keränen, H., Gutiérrez-de-Terán, H., Åqvist, J., 2014. Structural and energetic effects of A2A adenosine receptor mutations on agonist and antagonist binding. PLoS One 9, e108492. https://doi.org/10.1371/journal.pone.0108492.
https://pubs.acs.org/doi/abs/10.1021/cr00023a004.
Lawrence, M.C., Iliades, P., Fernley, R.T., Berglez, J., Pilling, P.A., Macreadie, I.G., 2005. The three-dimensional structure of the bifunctional 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase/dihydropteroate synthase of Saccharomyces cerevisiae. J. Mol. Biol. 348, 655–670. https://doi.org/10.1016/j.jmb.2005.03.021.
Lee, C., Yang, W., Parr, R.G., 1988. Development of the colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B 37, 785–789. https://doi.org/10.1103/PhysRevB.37.785.
Lenselink, E.B., Louvel, J., Forti, A.F., van Veldhoven, J.P.D., de Vries, H., Mulder-Krieger, T., McRobb, F.M., Negri, A., Goose, J., Abel, R., van Vlijmen, H.W.T., Wang, L., Harder, E., Sherman, W., IJzerman, A.P., Beuming, T., 2016. Predicting binding affinities for GPCR ligands using free-energy perturbation. ACS Omega 1, 293–304. https://doi.org/10.1021/acsomega.6b00086.
Levin, I., Giladi, M., Altman-Price, N., Ortenberg, R., Mevarech, M., 2004. An alternative pathway for reduced folate biosynthesis in bacteria and halophilic archaea. Mol. Microbiol. 54, 1307–1318. https://doi.org/10.1111/j.1365-2958.2004.04339.x.
Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug. Deliv. Rev., In Vitro Models For. Selection Development Candidates 23, 3–25. https://doi.org/10.1016/S0169-409X(96)00423-1.
Liu, P., Kim, B., Friesner, R.A., Berne, B.J., 2005. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. Proc. Natl. Acad. Sci. 102, 13749–13754. https://doi.org/10.1073/pnas.0506346102.
Lopez de Compadre, R.L., Pearlstein, R.A., Hopfinger, A.J., Seydel, J.K., 1987. A quantitative structure-activity relationship analysis of some 4-aminodiphenyl sulfone antibacterial agents using linear free energy and molecular modeling methods. J. Med. Chem. 30, 900–906.
Malkhasian, A.Y.S., Howlin, B.J., 2016. Docking and DFT studies on ligand binding to Quercetin 2,3-dioxygenase. J. Biomol. Struct. Dyn. 34, 2453–2461. https://doi.org/10.1080/07391102.2015.1123190.
Manyando, C., Njunju, E.M., D'Alessandro, U., Van geertruyden, J.-P., 2013. Safety and efficacy of Co-trimoxazole for treatment and prevention of plasmodium falciparum malaria: a systematic review. PLoS ONE 8. https://doi.org/10.1371/journal.pone.0056916.
Martyna, G.J., Klein, M.L., Tuckerman, M., 1992. Nosé–Hoover chains: the canonical ensemble via continuous dynamics. J. Chem. Phys. 97, 2635–2643. https://doi.org/10.1063/1.463940.
Martyna, G.J., Tobias, D.J., Klein, M.L., 1994. Constant pressure molecular dynamics algorithms. J. Chem. Phys. 101, 4177–4189. https://doi.org/10.1063/1.467468.
Matherly, L.H., Goldman, D.I., 2003. Membrane transport of folates. Vitam. Horm. 66, 403–456.
Morgan, R.E., Batot, G.O., Dement, J.M., Rao, V.A., Eadsforth, T.C., Hunter, W.N., 2011. Crystal structures of Burkholderia cenocepacia dihydropteroate synthase in the apo-form and complexed with the product 7,8-dihydropteroate. BMC Struct. Biol. 11, 21. https://doi.org/10.1186/1472-6807-11-21.
Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791. https://doi.org/10.1002/jcc.21256.
Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., Olson, A.J., n.d. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. J. Comput. Chem. 19, 24.
Nzila, A., 2006. Inhibitors of de novo folate enzymes in plasmodium falciparum. Drug Discov. Today 11, 939–944. https://doi.org/10.1016/j.drudis.2006.08.003.
Paliwal, H., Shirts, M.R., 2011. A benchmark test set for alchemical free energy transformations and its use to quantify error in common free energy methods. J. Chem. Theory Comput. 7, 4115–4134. https://doi.org/10.1021/ct2003995.
Pohorille, A., Jarzynski, C., Chipot, C., 2010. Good practices in Free-energy calculations. J. Phys. Chem. B 114, 10235–10253. https://doi.org/10.1021/jp102971x.
Polański, J., Gieleciak, R., Bąk, A., 2002. The comparative molecular surface analysis (COMSA) − a nongrid 3D QSAR method by a coupled neural network and PLS system: predicting pKa values of benzoic and alkanoic acids. J. Chem. Inf. Comput. Sci. 42, 184–191. https://doi.org/10.1021/ci010031t.
Qi, J., Virga, K.G., Das, S., Zhao, Y., Yun, M.-K., White, S.W., Lee, R.E., 2011. Synthesis of bi-substrate state mimics of dihydropteroate synthase as potential inhibitors and molecular probes. Bioorg. Med. Chem., Imaging Probes 19, 1298–1305. https://doi.org/10.1016/j.bmc.2010.12.003.
Richey, D.P., Brown, G.M., 1969. The biosynthesis of folic acid. IX. Purification and properties of the enzymes required for the formation of dihydropteroic acid. J. Biol. Chem. 244, 1582–1592.
Roland, S., Ferone, R., Harvey, R.J., Styles, V.L., Morrison, R.W., 1979. The characteristics and significance of sulfonamides as substrates for Escherichia coli dihydropteroate synthase. J. Biol. Chem. 254, 10337–10345.
Rossi, M., Amaretti, A., Raimondi, S., 2011. Folate production by probiotic bacteria. Nutrients 3, 118–134. https://doi.org/10.3390/nu3010118.
Russo, T.A., Johnson, J.R., 2003. Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. Microbes Infect. 5, 449–456.
Sastry, G.M., Adzhigirey, M., Day, T., Annabhimoju, R., Sherman, W., 2013. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J. Comput. Aided Mol. Des. 27, 221–234. https://doi.org/10.1007/s10822-013-9644-8.
Shah, U.A., Deokar, H.S., Kadam, S.S., Kulkarni, V.M., 2010. Pharmacophore generation and atom-based 3D-QSAR of novel 2-(4-methylsulfonylphenyl)pyrimidines as COX-2 inhibitors. Mol. Divers. 14, 559–568. https://doi.org/10.1007/s11030-009-9183-3.
Shivakumar, D., Deng, Y., Roux, B., 2009. Computations of absolute solvation free energies of small molecules using explicit and implicit solvent model. J. Chem. Theory Comput. 5, 919–930. https://doi.org/10.1021/ct800445x.
Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., Sherman, W., 2010. Prediction of absolute solvation Free energies using molecular dynamics Free energy perturbation and the OPLS force Field. J. Chem. Theory Comput. 6, 1509–1519. https://doi.org/10.1021/ct900587b.
Srivastava, P., Tiwari, A., 2017. Critical role of computer simulations in drug discovery and development. Curr. Top. Med. Chem. 17, 2422–2432. https://doi.org/10.2174/1568026617666170403113541.
Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K - The Journal of Physical Chemistry A (ACS Publications) [WWW Document], n.d. URL https://pubs.acs.org/doi/10.1021/jp003020w (Accessed 9.24.18).
Takao, Otsuka, Noriaki, Okimoto, Makoto, Taiji, 2015. Assessment and acceleration of binding energy calculations for protein–ligand complexes by the fragment molecular orbital method. J. Comput. Chem. 36, 2209–2218. https://doi.org/10.1002/jcc.24055.
Tao, P., Fisher, J.F., Shi, Q., Vreven, T., Mobashery, S., Schlegel, H.B., 2009. Matrix metalloproteinase 2 (MMP2) inhibition: QM/MM studies of the inhibition mechanism of SB-3CT and its analog. Biochemistry 48, 9839. https://doi.org/10.1021/bi901118r.
Then, R., Angehrn, P., 1973. Sulphonamide-induced 'thymineless death' in Escherichia coli. Microbiology 76, 255–263. https://doi.org/10.1099/00221287-76-2-255.
Tomasi, J., Mennucci, B., Cammi, R., 2005. Quantum mechanical continuum solvation models. Chem. Rev. 105, 2999–3094. https://doi.org/10.1021/cr9904009.
Tripuraneni, N.S., Azam, M.A., 2016. Pharmacophore modeling, 3D-QSAR and docking study of 2-phenylpyrimidine analogues as selective PDE4B inhibitors. J. Theor. Biol. 394, 117–126. https://doi.org/10.1016/j.jtbi.2016.01.007.
https://pubs.rsc.org/en/content/articlelanding/2012/md/c1md00214g#!divAbstract.
Verma, J., Khedkar, V.M., Coutinho, E.C., 2010. 3D-QSAR in drug design - a review. Curr. Top. Med. Chem. 10, 95–115. https://doi.org/10.2174/156802610790232260.
Walzer, P.D., Kim, C.K., Foy, J.M., Linke, M.J., Cushion, M.T., 1988. Inhibitors of folic acid synthesis in the treatment of experimental pneumocystis carinii pneumonia. Antimicrob. Agents Chemother. 32, 96–103.
Wang, L., Friesner, R.A., Berne, B.J., 2011. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). J. Phys. Chem. B 115, 9431–9438. https://doi.org/10.1021/jp204407d.
Wang, L., Berne, B.J., Friesner, R.A., 2012. On achieving high accuracy and reliability in the calculation of relative protein–ligand binding affinities. Proc. Natl. Acad. Sci. 109, 1937–1942. https://doi.org/10.1073/pnas.1114017109.
Wang, L., Wu, Y., Deng, Y., Kim, B., Pierce, L., Krilov, G., Lupyan, D., Robinson, S., Dahlgren, M.K., Greenwood, J., Romero, D.L., Masse, C., Knight, J.L., Steinbrecher, T., Beuming, T., Damm, W., Harder, E., Sherman, W., Brewer, M., Wester, R., Murcko, M., Frye, L., Farid, R., Lin, T., Mobley, D.L., Jorgensen, W.L., Berne, B.J., Friesner, R.A., Abel, R., 2015. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. J. Am. Chem. Soc. 137, 2695–2703. https://doi.org/10.1021/ja512751q.
Whetstine, J.R., Flatley, R.M., Matherly, L.H., 2002. The human reduced folate carrier gene is ubiquitously and differentially expressed in normal human tissues: identification of seven non-coding exons and characterization of a novel promoter. Biochem. J. 367, 629–640. https://doi.org/10.1042/BJ20020512.
Woods, D.D., 1940. The relation of p-aminobenzoic acid to the mechanism of the action of Sulphanilamide. Br. J. Exp. Pathol. 21, 74–90.
Yun, M.-K., Wu, Y., Li, Z., Zhao, Y., Waddell, M.B., Ferreira, A.M., Lee, R.E., Bashford, D., White, S.W., 2012. Catalysis and sulfa drug resistance in dihydropteroate synthase. Science 335, 1110–1114. https://doi.org/10.1126/science.1214641.
Zhao, Y., Shadrick, W.R., Wallace, M.J., Wu, Y., Griffith, E.C., Qi, J., Yun, M.-K., White, S.W., Lee, R.E., 2016. Pterin-sulfa conjugates as dihydropteroate synthase inhibitors and antibacterial agents. Bioorg. Med. Chem. Lett. 26, 3950–3954. https://doi.org/10.1016/j.bmcl.2016.07.006.