BIOINFORMATION
Discovery at the interface of physical and biological sciences

Open access www.bioinformation.net Hypothesis

Volume 12(3)

A QSAR model of benzoxazole derivatives as potential inhibitors for inosine 5`-monophosphate dehydrogenase from *Cryptosporidium parvum*

Pratibha Teotia, Surya Prakash Dwivedi & Neeraja Dwivedi*

School of Biotechnology, IFTM University, Moradabad, Utter Predesh, India; Neeraja Dwivedi – Email: neer.miet@gmail.com, *Corresponding author

Received November 8, 2015; Revised November 21, 2015; Accepted November 24, 2015; Published June 15, 2016

Abstract:
Cryptosporidium parvum is the common enteric protozoan pathogen causing cryptosporidiosis in human. Available drugs to treat cryptosporidiosis are ineffective and there is yet no vaccine against *C. parvum*. Therefore, it is of interest to design an improved yet effective drug against *C. parvum*. Here, we docked benzoxazole derivatives (collected from literature) with inosine 5`-monophosphate dehydrogenase (IMPDH) from *Cryptosporidium parvum* using the program AutoDock 4.2. The docked protein - inhibitor complex structure was optimized using molecular dynamics simulation for 5 ps with the CHARMM-22 force field using NAMD (NAnoscale Molecular Dynamics program) incorporated in visual molecular dynamics (VMD 1.9.2) and then evaluating the stability of complex structure by calculating RMSD values. NAMD is a parallel, object-oriented molecular dynamics code designed for high-performance simulation of large biomolecular systems. A quantitative structure activity relationship (QSAR) model was built using energy-based descriptors as independent variable and pIC_{50} value as dependent variable of fifteen known benzoxazole derivatives with *C. parvum* IMPDH protein, yielding correlation coefficient r^2 of 0.7948. The predictive performance of QSAR model was assessed using different cross-validation procedures. Our results suggest that a ligand-receptor binding interaction for inosine 5`-monophosphate dehydrogenase using a QSAR model is promising approach to design more potent inosine 5`-monophosphate dehydrogenase inhibitors prior to their synthesis.

Keywords: *Cryptosporidium parvum*; docking; inosine 5`-monophosphate dehydrogenase; AutoDock 4.2; benzoxazole derivatives.

Background:
Cryptosporidiosis is the most common food and waterborne diseases with worldwide spread, acting as a common cause of diarrhoea in animals and man **[1]**. Among the five common *Cryptosporidium* species in humans, *Cryptosporidium parvum* (*C. parvum*) and *Cryptosporidium hominis* (*C. hominis*) are responsible for more than 90% of human cases of cryptosporidiosis **[2]**. *Cryptosporidium* is one of the most important parasitic diarrheal disease among young children in developing nations, and is problematic as an opportunistic co-infection with HIV due to increased morbidity and mortality **[3, 4]**. Currently available drugs are not effective for treating cryptosporidiosis and vaccine therapy is lacking, so new drugs are needed. The sequencing of the genomes of *Cryptosporidium parvum* revealed a highly streamlined anabolic metabolism with potential choke points that might be exploited in drug design **[5]**. One such vulnerability lies in the pathway that supplies purine nucleotides for the synthesis of DNA and RNA. Like all protozoan parasites, *Cryptosporidium* is incapable of de novo purine synthesis and relies on salvage of purines from the host **[5]**. Adenosine is converted into guanine nucleotides in a streamlined pathway that relies on inosine 5`-monophosphate dehydrogenase (IMPDH) catalyzing the conversion of inosine 5`-monophosphate (IMP) to xanthosine 5`-monophosphate (XMP) **[6]**.

©2016 BIOMEDICAL INFORMATICS

In this study, we docked experimentally verified 15 benzoxazole-based inhibitors having inhibitory value IC_{50} in nM with *C. parvum* IMPDH using AutoDock 4.2, which resulted in energy-based descriptors. Molecular dynamics (MD) simulation studies of inhibitor – protein complex were performed and after that, we have built quantitative structure activity relationship (QSAR) model using Multiple Linear Regression.

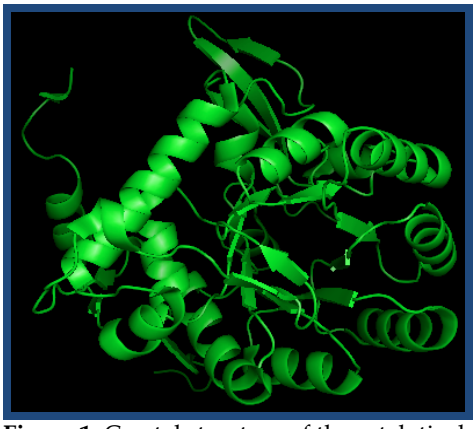


Figure 1: Crystal structure of the catalytic domain of the inosine monophosphate dehydrogenase from cryptosporidium parvum.

Methodology:

Protein target structure

The 3D coordinates of the crystal structure of the catalytic domain of the Inosine 5`-monophosphate dehydrogenase from *cryptosporidium parvum* (PDB Id: 4IXH) was retrieved from Protein Databank (http://www.rcsb.org/) and is shown in **Figure 1**. This is used as a target model for flexible docking. The structure was optimized using the chimera tool **[7]**.

Inhibitors dataset

Fifteen benzoxazole derivatives with known pIC50 were obtained from Gorla *et al*. (2013) **[8]**. The 3D structures of known 15 inhibitors were downloaded in .sdf format from pubchem compound database. They were later converted in .pdb format with the help of open babel **[9]** tool. All the ligands were subjected to energy minimization using the HyperChem software **[10]**.

Molecular docking

Docking of fifteen benzoxazole derivatives screened from literature against *C. parvum* IMPDH structure were done using molecular docking program AutoDock **[11].** Gasteiger charges are added to the ligand and maximum 6 numbers of active torsions are given to the lead compounds using AutoDock tool **[12]**. Kollman charges and the solvation term were added to the protein structure. The Lamarckian genetic algorithm implemented in Autodock was used for docking.

Molecular dynamics simulations

Molecular dynamics simulations were done using the NAMD (NAnoscale Molecular Dynamics program; v2.7) graphical interface module **[13]** incorporated visual molecular dynamics (VMD 1.9.2) **[14].** The protein-ligand complex was immersed in the center of a 50 Å box of water molecules where all water molecule atoms (H-O-H) were closer than 1.5 Å and a CHARMM (Chemistry at HARvard Macromolecular Mechanics) 22 parameter file for proteins and lipids; phi and psi cross-term map correction were used in the force field for complexes. For the minimization and equilibration of complex in the water box, we assumed force-field parameters excluding scaling of 1.0 Å and a cutoff of Coulomb forces with a switching function starting at 12 Å, reaching zero at a distance of 10 Å, ending at 14 Å with a margin of 3.0 Å, and all atoms, including those of hydrogen, were illustrated explicitly. A protein structure file (psf) stores structural information of the protein, such as various types of bonding interactions. The psf was created from the initial pdb and topology files using psfgen package of VMD. After running psfgen,two new files were generated protein pdb and protein psf and by accessing PSF and PDB files; NAMD generated the trajectory DCD file. After the simulations, the results were analyzed in VMD by calculating the Root mean square deviation (RMSD) of the complex using rmsd tcl source file from the Tk console and finally rmsd.dat was saved and accessed in Microsoft office excel 2007.

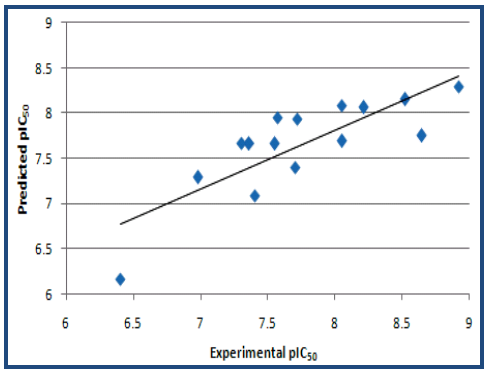


Figure 2: Relationship between experimental (x-axis) and predicted (y-axis) pIC50 values with an r^2 value 0.7948 is shown in a QSAR model developed using multiple linear regression analysis.

2D QSAR

A QSAR based model was generated having correlation coefficient r^2 value 0.7948 was developed using multiple linear regression analysis. An equation was developed for the inhibitory activities represented as pIC_{50} values using the six types of energy values as variable descriptors such as Binding Energy (BE), Intermolecular Energy (IME), Internal Energy (IE), Torsional Energy (TorE), vdW + Hbond + desolv Energy (VdwE) and electrostatic energy (EE). A correlation coefficient (r^2) of 0.7948 was obtained for 15 benzoxazole derivatives as shown below in equation 1.

Predicted pIC_{50} = 1.0291- 26.7693 (BE) + 25.9634(IME) + 0.7866(IE) + 25.9788 (TorE) - 0.0536 (VdwE) - 1.7919(EE) (1)

Several cross-validation procedures were adopted to assess the predictive performance of the QSAR model. In leave-one-out strategy (LOOCV), one molecule was removed from the dataset as a test compound and the remaining 14 molecules were used to build the model. This process was repeated 15 times with each inhibitor as a test molecule.

©2016 BIOMEDICAL INFORMATICS

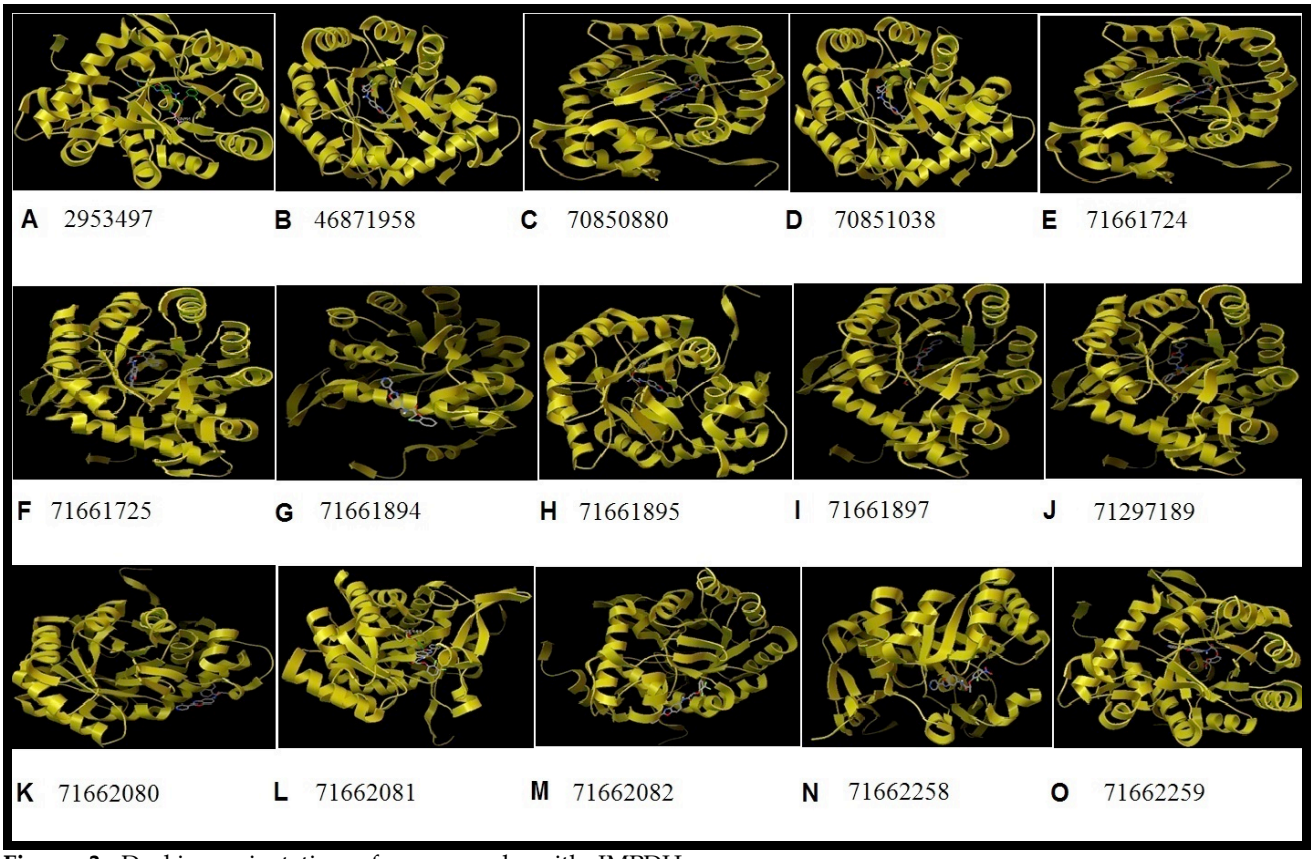


Figure 3: Docking orientation of compounds with IMPDH protein.

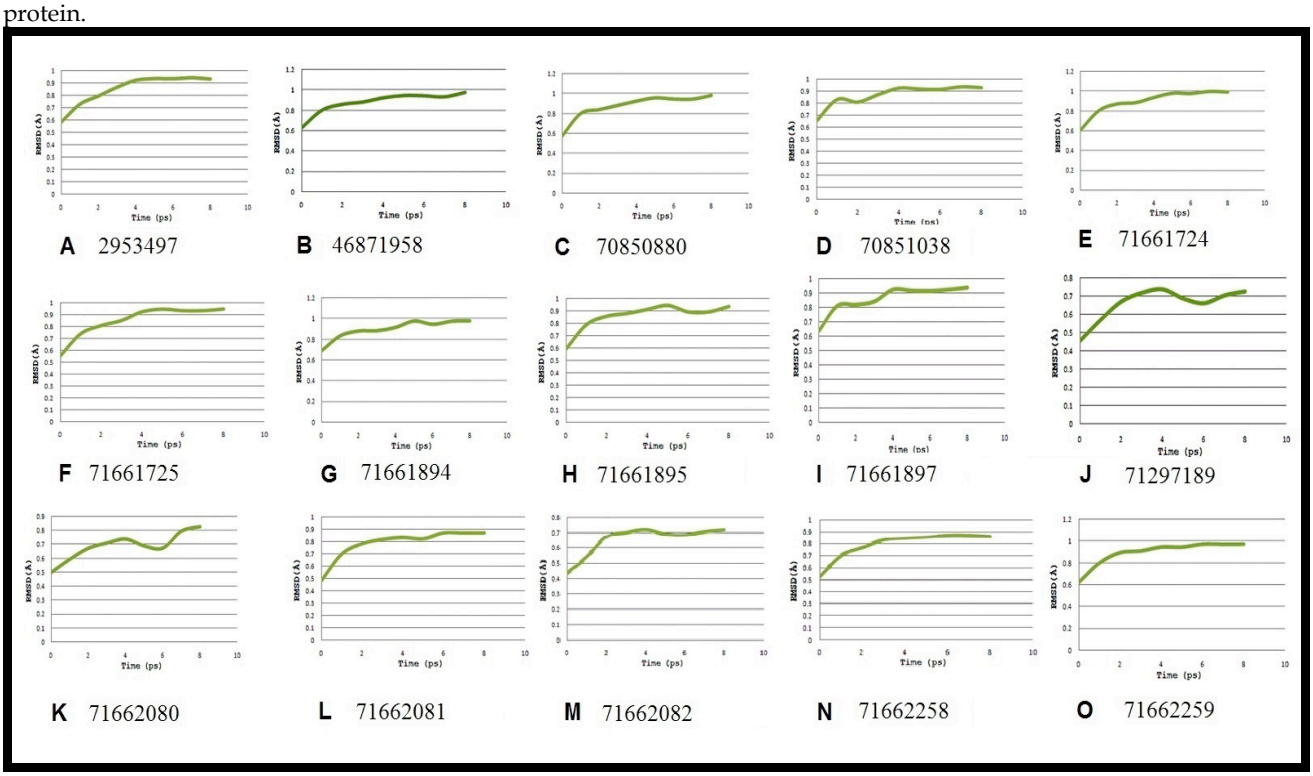
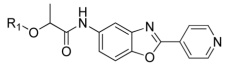


Figure 4: Graph displaying root mean square deviation (RMSD) of compounds – protein complex versus time (5 ps) at 310 K, resulted in highest peak at 0.98 Å.

Table 1: Benzoxazole derivatives of *Cryptosporidium parvum* inosine 5`-monophosphate dehydrogenase on the basis of different R_1 group.

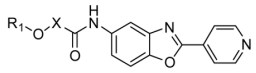
Benzoxazole derivatives



Sl. No.	PubChem CID	R1	Experimental pIC50
1	2953497	2,4-di-ClPh	7.36
2	71661724	2-ClPh	7.72
3	71661725	4-ClPh	6.98
4	70850880	Ph	7.40
5	71661894	4-OMePh	7.55
6	71661895	3-ClPh	7.70
7	46871958	2,3-di-ClPh	8.52
8	70851038	1-naphthyl	8.05
9	71661897	1-(4-Cl-naphthyl)	7.57

Table 2: Benzoxazole derivatives of *Cryptosporidium parvum* inosine 5`-monophosphate dehydrogenase on the basis of different R_1 and X group.

Benzoxazole derivatives



Sl. No.	PubChem CID	X	R1	Experimental pIC50
1	71662080	(R)-CHMe	1-naphthyl	8.92
2	71297189	(S)-CHMe	1-naphthyl	8.21
3	71662081	(S)-CHMe	2,3-di-ClPh	7.3
4	71662082	(S)-CHMe	2-Cl,3-CF3Ph	8.05
5	71662258	(S)-CHMe	2-Cl,3-NO2Ph	8.64
6	71662259	(S)-CHMe	2,3-di-OMePh	6.4

Table 3: Docking results of benzoxazole derivatives with IMPDH structure with activity (pIC_{50} = - $logIC_{50}$).

No	PubChem CID	Experimental pIC_{50}	Predicted pIC_{50}	BE	IME	IE	TorE	VdwE	EE
1	2953497	7.36	7.66	-9.03	-10.22	-1.38	1.19	-10.29	0.07
2	71661724	7.72	7.94	-8.59	-9.79	-0.6	1.19	-9.68	-0.11
3	71661725	6.98	7.29	-8.61	-9.8	-1.58	1.19	-9.8	-0.01
4	70850880	7.40	7.09	-8.41	-9.6	-1.54	1.19	-9.62	0.02
5	71661894	7.55	7.66	-8.51	-10.0	-1.15	1.49	-9.92	-0.07
6	71661895	7.70	7.40	-8.98	-10.18	-1.39	1.19	-10.06	-0.11
7	46871958	8.52	8.15	-8.97	-10.16	-1.1	1.19	-10.06	-0.11
8	70851038	8.05	8.08	-9.53	-10.72	-1.55	1.19	-10.72	0.0
9	71661897	7.57	7.95	-9.69	-11.18	-2.1	1.49	-11.1	-0.08
10	71662080	8.92	8.29	-9.69	-10.88	-1.53	1.19	-10.85	-0.03
11	71297189	8.21	8.06	-9.72	-10.92	-1.43	1.19	-10.92	0.01
12	71662081	7.3	7.67	-9.32	-10.51	-1.73	1.19	-10.56	0.05
13	71662082	8.05	7.69	-8.82	-10.32	-0.76	1.49	-10.24	-0.08
14	71662258	8.64	7.75	-9.13	-10.62	-1.66	1.49	-10.04	-0.58
15	71662259	6.4	6.16	-8.17	-9.96	-2.47	1.79	-10.0	0.04

BE = Binding Energy; IME: Intermolecular Energy; IE = Internal Energy; TorE= Torsional; Energy; VdwE = vdW + Hbond + desolv Energy; EE= Electrostatic energy.

©2016 BIOMEDICAL INFORMATICS

Results & Discussion:

Based on R_1 and X groups at different positions, benzoxazole derivatives of *C. parvum* IMPDH were retrieved from literature **[8]** and are shown in **Table 1 & 2**. In docking studies of benzoxazole derivatives with IMPDH protein, best autodock score was used as criteria to interpret the best conformation among the 30 conformations, generated by AutoDock 4.2 program. The docking results of the benzoxazole derivatives with IMPDH protein were shown in **Table 3.** Further, the docked complexes were analyzed through Python Molecular Viewer software **[15]** for their interaction studies and were shown in **Figure 3**. Thus from the Complex scoring and binding ability it's deciphered that these compounds are promising inhibitors for IMPDH protein. MD simulation is a well-known theoretical technique and is mainly used for evaluating the stability of any predicted 3D model. Therefore, the constructed 3D model of protein-ligand complexes were processed for MD simulation for a 5 ps timescale with Langevin dynamics to control the kinetic energy, temperature, and/or pressure of the system. The RMSD values of complexes contain alpha carbon atoms, and all atoms were calculated by taking structure with reference conformation points. The RMSD values of complex versus time were shown in **Figure 4.** Relationship between experimental and predicted pIC_{50} values of Benzoxazole derivatives is shown in **Figure 2.**

Conclusion:

A QSAR model using pIC50 values for fifteen known benzoxazole derivatives binding with *C. parvum* IMPDH protein as dependent variable and molecular docking based predicted pIC50 with a correlation coefficient r2 is 0.7948 was reported.

Acknowledgement:

This article is a part of Ph.D. studies carried out by the author registered at IFTM University, Moradabad, India. The authors are grateful to Prof. R. M. Dubey, Vice Chancellor, IFTM University, Moradabad, India for providing an institutional research promotion grant and their generous help and encouragement to proceed with the studies in the School of Biotechnology IFTM University, Moradabad, India.

References:

[1] Rossle NF *et al. Asian Pac J Trop Biomed*. 2013 **3:** 916 [PMCID: PMC3793167]
[2] Xiao L &, Ryan UM, *Curr Opin Infect Dis*. 2004 **17**: 483 [PMID: 15353969]
[3] Parsons MB *et al. PLoS Negl Trop Dis*. 2015 **9**: e0003529 [PMID: 25700265]
[4] Kotloff KL *et al. Lancet* 2013 **382**: 209 [PMID: 23680352]
[5] Sharling L *et al. PLoS Negl Trop Dis*. 2010 **4**: e794 [PMID: 20706578]
[6] Hedstrom L, *Chem Rev*. 2009 **109**: 2903 [PMID: 19480389]
[7] Pettersen EF *et al. J Comput Chem*. 2004 **25:** 1605 [PMID: 15264254]
[8] Gorla SK *et al. J Med Chem*. 2013 **56:** 4028 [PMID: 23668331]
[9] O'Boyle NM *et al. J Cheminform*. 2011 **3:** 33 [PMID: 21982300]
[10] HyperChem (TM) Release 7.5, Hypercube, Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA.
[11] Morris *et al. J Computational Chemistry*. 1998 **19**: 1639
[12] http://autodock.scripps.edu/resources/adt
[13] Phillips JC *et al. J Comput Chem*. 2005 **26**: 1781 [PMID: 16222654]
[14] Humphrey W *et al. J Mol Graph*. 1996 **14**: 33 [PMID: 8744570]
[15] Sanner MF, *J Mol Graph Model*. 1999 **17**: 57 [PMID: 10660911]

Edited by P Kangueane
Citation: **Teotia *et al.*** Bioinformation 12(3): 119-123 (2016)
License statement: This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License.