

Virtual screening of combinatorial library of novel benzene-sulfonamides on mycobacterial carbonic anhydrase II

Original research article

Dikant, F., Gáplovský, M., Garaj, V.✉

Comenius University in Bratislava, Faculty of Pharmacy,
Department of Pharmaceutical Chemistry,
Bratislava, Slovak Republic

Received 22 June, 2016, accepted 28 September, 2016

Abstract Combinatorial library of novel benzenesulfonamides was docked (Schrodinger Glide) into mycobacterial carbonic anhydrase (mtCA II) and human (hCA II) isoforms with an aim to find drug candidates with selective activity on mtCA II. The predicted selectivity was calculated based on optimized MM-GBSA free energies for ligand enzyme interactions. Selectivity, LogP (o/w) and interaction energy were used to calculate the selection index which determined the subset of best scoring molecules selected for further evaluation. Structure-activity relationship was found for fragment subsets, showing us the possible way regarding how to influence lipophilicity without affecting ligand-enzyme binding properties.

Keywords Tuberculosis – carbonic anhydrase inhibitors – virtual screening – ADME prediction

INTRODUCTION

Carbonic anhydrases (CAs) are metalloenzymes present in almost all living organisms. There are five genetically distinct CA families. Mammalian CAs (hCAs) belong to the α -family and are present in fifteen isoforms, which differ in molecular features, oligomeric arrangement, cellular localization, distribution in organs and tissues, expression levels, and kinetic properties (Supuran, 2008). So far, 3 isoforms of carbonic anhydrase, coded by genes Rv1284 (mtCA I), Rv3588c (mtCAII) and Rv3273 (mtCA III), have been identified in *Mycobacterium tuberculosis* (<http://tuberculist.epfl.ch/>). Of these 3 isozymes, mtCA I is strongly upregulated in starvation conditions and mtCA II was reported as essential for bacteria survival (Sassetti & Rubin, 2003).

Mycobacterium tuberculosis and related Mycobacteria infections (e.g., *Mycobacterium avium*) affect a large number of the world population, with an estimated 9.2 million new cases each year, many of which lead to deaths (WHO, 2013). Multi-drug resistant and extensively multi-drug resistant tuberculosis (TB) worsens the situation even more, as these strains are now present in more than 50 countries, posing serious concern to the global healthcare system, as this disease is largely unresponsive to the presently available drugs.

Indeed, the combination therapy used to treat TB is based on agents developed in the 60–80s, with no new drugs launched for the last 30 years (Covarubias, 2005). Mycobacterial CA has been identified as a target for anti-mycobacterial agents (Sassetti & Rubin, 2003) (Maresca, 2009). Sulfonamides can act as effective inhibitors of both α and β families of CA (Supuran, 2010). This creates a challenge to find agents that will work as selective carbonic anhydrase inhibitors (CAIs) of mycobacterial CA (mtCA) not acting on human CA isoforms. Some derivatives of phenylcarbamic acid showed good antituberculous activity (Keckesova, 2009) (Waisser, 2006). We have prepared a virtual library of similar phenylcarbamic acid derivatives containing zinc binding group (all of which are sulfonamides) and performed virtual screening on mtCA II against hCA II isoform, which has the highest spread amongst humans.

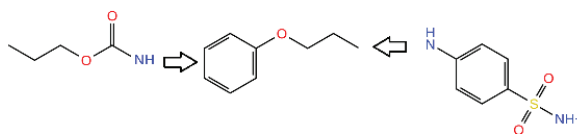
MATERIAL AND METHODS

Preparation of combinatorial library

Three groups of fragments were combined in CombiGlide from Schrodinger Maestro program pack to create structures

* E-mail: garaj@fpharm.uniba.sk

© European Pharmaceutical Journal



Scheme 1. Visualization of molecule combination (a)-carbamates or aliphatic alcohols orienting into *o*-, *m*- and *p*-; (b)-central fragments; (c)-sulfonamide fragments

according to the pattern shown in Scheme 1. Three groups of fragments were used. Derivatives of benzene with short alkyl, alkoxy and phenylcarbamate sidechains were used as the central fragment (**b**). Methyl, ethyl, propyl and butyl carbamates from group (**a**) were bound to *ortho*-, *meta*- and *para*- positions on benzene ring of group (**b**) fragments **A**, **D**, **E**, **F**, **G**, resulting in the creation of phenyl carbamates. Aliphatic primary alcohols up to butanol were bound to *ortho*-, *meta*- and *para*- positions on benzene ring of group (**b**) compounds **B** and **C**, resulting in the creation of phenyl esters. The last carbon on distant side of alkyl/alkoxy of phenylcarbamate sidechain was used as the second binding position on the central fragment (**b**). This terminal carbon was bound to the free nitrogen from amino group on sulfonamide fragment (**c**). Five different benzenesulfonamides were used (Tab. 1). This approach yielded 732 molecules including different enantiomers and ionization states stable at $\text{pH}=7\pm 2$ predicted by Epik (Shelley, 2007). The geometry of all the resulting molecules used for docking was optimized by means of OPLS 2005 force field.

Preparation of target proteins

Crystallographic protein structures were obtained from PDB (Protein Data Bank – www.rcsb.org). Crystal structure of tetrameric form M. tuberculosis beta carbonic anhydrase, Rv3588c with resolution of 2.20 Å (PDB ID 2A5V) (Covarrubias, 2006) was used. To predict the selectivity to mycobacterial β -carbonic anhydrase II (mtCA II) in relation to human cytosolic CA (hCA II), we chose ubiquitous α -CA II (PDB ID 3K34) with resolution 0.90 Å (Behnke, 2010). We processed the protein structure to remove any solvent molecules located in the PDB co-ordinates. We chose subunit A with Zn^{2+} in the active site to create a docking grid box, we defined the docking site size and location around the central zinc ion, and applied constrain on the creation of bond between nitrogen from sulfonamide functional group and central Zn^{2+} .

Glide Docking

Docking of all 732 molecules was performed in program Schrodinger Glide with XP precision (Friesner, 2006). The extra-precision (XP) mode of Glide combines a powerful sampling protocol with the use of a custom scoring function

Table 1. List of fragments used in the creation of combinatorial library

Fragment group	Fragment name	Fragment code
Carbamate and alcohol fragments (a)	Methyl carbamate	km
	Ethyl carbamate	ke
	Propyl carbamate	kp
	Butyl carbamate	kb
	Methanol	em
	Ethanol	ee
	Propan-1-ol	ep
	Butan-1-ol	eb
Central fragments (b)	1-phenoxypropan-2-ol	A
	2-hydroxypropyl phenylcarbamate	B
	2-hydroxypropyl benzoate	C
	1-phenylethanol	D
	Ethoxybenzene	E
	Propoxybenzene	F
	Butoxybenzene	G
Sulfonamide fragments (c)	4-aminobenzenesulfonamide	1
	4-(aminomethyl) benzenesulfonamide	2
	4-(2-aminoethyl) benzenesulfonamide	3
	4-piperazin-1-ylbenzenesulfonamide	4
	4-piperidin-4-ylbenzenesulfonamide	5

designed to identify the ligand poses that would be expected to have unfavourable energies, based on well-known principles of physical chemistry. Among other things, to model the solvation, explicit water molecules are docked into a list of protein-ligand complexes that otherwise receive good GlideScores, and descriptors based on the interaction of these water molecules with various charged and polar groups of the ligand and protein are used as a measure of whether the complex is physically realistic. All the generated molecules were docked to mycobacterial CA II (mtCA II) resulting in 186 molecules successfully fitted into mtCA II. All were subsequently docked into human CA II (hCA II). The positions of all 186 molecules docked in both enzymes were optimized using Schrodinger Prime MM-GBSA, providing us with predicted free interaction energy between the molecule and enzyme. MM-GBSA uses a novel force field method VSGB 2.0 (Li, 2011), which features an optimized implicit solvent model, as well as physics-based corrections for hydrogen bonding, π - π interactions, self-contact interactions and hydrophobic interactions. Flexibility of aminoacid residues within 5 Å from ligand was allowed during the optimization procedure. The lower MM-GBSA interaction energy stood for stronger interaction between molecule and enzyme.

In the following table (Tab. 2), we multiplied the value by (-1) to get to positive numbers, and to make it more convenient to evaluate and include the values into calculations. The physical and chemical properties, along with ADMET properties, were predicted in QikProp.

RESULTS AND DISCUSSION

Selectivity

Selectivity helps to mitigate unwanted adverse effects of therapy on human isoforms of carbonic anhydrase. From the 732 new molecules created in CombiGlide, 186 were successfully fitted with program Glide into mycobacterial CAII. Their optimized MM-GBSA free interaction energy was in interval ranging from -73.7 to 15.3, where the lower free interaction energy represented potentially stronger bond between molecule and protein. The same set of molecules was docked into human CAII, and optimized with Prime MM-

GBSA. Their interaction energies ranged from -71.8 to -10.3. The selectivity was calculated using the following formula:

$$selectivity (Sel) = \frac{MMGBSA \Delta G_{mtCA II}}{MMGBSA \Delta G_{hCA II}}$$

The calculated selectivity ranged from 3.1 to -0.3.

The selectivity of 101 molecules was higher than 1.0.

Bioavailability

First evaluated criterion of interest is the permeability of molecule through cell membrane. Bioavailability of a drug can be predicted using Lipinski rules (Lipinski, 2001): molar weight (<500Da), $\log P_{o/w}$ (≤ 5), number of acceptors (≤ 10) and donors (≤ 5) of hydrogen bonds, as well as Veber rules (Veber, 2002) including the number of freely rotatable bonds (≤ 10) and polar surface of the molecule ($\leq 140 \text{ \AA}^2$). All molecules fulfil at least three of the four Lipinski rules and both Veber rules. Bioavailability predicted in the program QikProp ranged from

Table 2. List of molecules with highest selection score selected for further evaluation

Molecule codename	Absolute value of ΔG [mtCA II] MM-GBSA	Selectivity [mtCA II/ hCA II]	QPlogPo/w	Selection index [ΔG [mtCA II] * selectivity * $\log P/100$]
m-E; [kp]; [5]	73.7	1.5	2.6	2.8
o-A; [kp]; [3]	70.5	3.1	1.3	2.8
m-G; [kp]; [4]	67.6	1.5	2.7	2.7
o-G; [kb]; [5]	61.4	1.2	3.6	2.6
o-C; [ep]; [5]	57.2	2.0	2.2	2.5
o-F; [kb]; [5]	59.1	1.1	3.3	2.1
p-E; [kb]; [5]	66.7	1.3	2.4	2.0
o-B; [eb]; [5]	70.4	1.2	2.3	2.0
m-E; [kb]; [5]	67.4	1.1	2.7	2.0
o-G; [kp]; [5]	59.2	1.1	3.1	1.9
m-B; [eb]; [5]	70.2	1.2	2.3	1.9
o-E; [kb]; [5]	56.1	1.2	2.8	1.9
p-F; [kp]; [5]	56.1	1.1	3.0	1.8
m-E; [ke]; [5]	70.5	1.2	2.1	1.8
m-F; [kb]; [4]	66.3	0.9	2.8	1.7
m-G; [kb]; [3]	43.9	1.4	2.8	1.7
m-C; [eb]; [5]	54.0	1.3	2.4	1.7
m-B; [ee]; [5]	58.9	1.8	1.6	1.7
o-F; [ke]; [5]	54.7	1.2	2.6	1.7
m-E; [ke]; [5]	58.7	1.3	2.2	1.7
o-G; [kb]; [2]	59.5	1.2	2.4	1.7
p-E; [kb]; [2]	58.8	1.6	1.7	1.7
m-C; [ep]; [5]	52.5	1.5	2.1	1.7
o-F; [kp]; [1]	54.3	1.4	2.1	1.7
m-C; [eb]; [5]	58.8	1.1	2.5	1.6
m-D; [kb]; [5]	71.9	1.2	1.9	1.6
o-E; [kb]; [4]	58.5	1.2	2.3	1.6

40-78%, which means that none of the candidates should be disqualified based on this parameter.

Penetration into mycobacteria

Good penetration into mycobacterial cell can significantly lower the dose, and therefore, help reduce the probability of adverse effects. A distinguishing characteristic of all *Mycobacterium* species is that their cell wall is thicker than that of many other bacteria. It consists of a hydrophobic mycolate layer and a peptidoglycan layer held together by a polysaccharide, arabinogalactan. This creates a barrier that is hard to penetrate for many active agents, and makes the mycobacteria invulnerable to most antibiotics. A molecule attacking the intracellular proteins that aspires to be active in low concentration needs higher $\log P_{o/w}$ partition coefficient to be able to effectively penetrate the mycobacterial cell wall. For our molecule set, $\log P_{o/w}$ ranged from 0.092 to 3.604. Since many of the molecules have lipophilicity too low to be able to penetrate mycobacterial cell, this parameter was included into the calculation of Selection index.

Selection index

Selection index was used for two purposes. The first purpose was to select molecules that would be synthesized and further evaluated *in vitro* on mtCA II and culture of *M. Tuberculosis*. Second purpose was to study structure-activity relationship to determine which fragments scored better across the molecule subset, and to use this information to select new fragments that could improve the overall performance. To calculate the Selection index, we included partition coefficient, absolute value of predicted free MM-BGSA interaction energy and selectivity into the calculation. For all 3 properties, we desired the highest possible values. The statistical weight was not corrected. Other properties like oral bioavailability, blood-brain barrier penetration and toxicity were in acceptable ranges for all molecules, so they were excluded from the calculation (in supplementary information).

$$\text{Selection index} = \frac{(\text{Sel})(\text{Logp}) - (\text{mtcaII } \Delta G)}{100}$$

In general, this approach filtered out molecules which were outstanding in only one of the properties and were lacking in others. Also, it included molecules that would be otherwise disqualified due to low $\log P_{o/w}$ or selectivity based on interaction energy. The range of index was 2.81 for the best scoring molecule and 0.00 for the worst performers.

The chosen subset for synthesis and evaluation of biological activity took 15% (27 molecules) of all 186 successfully docked molecules. The highest Selection score was 2.81 (for compound: m-E; [kp]; [5] in Fig. 1 and 2) and lowest was 1.58 in the subset. No molecules were discarded manually.

The second assessment included statistical evaluation of frequency of occurrence of fragments throughout the full set

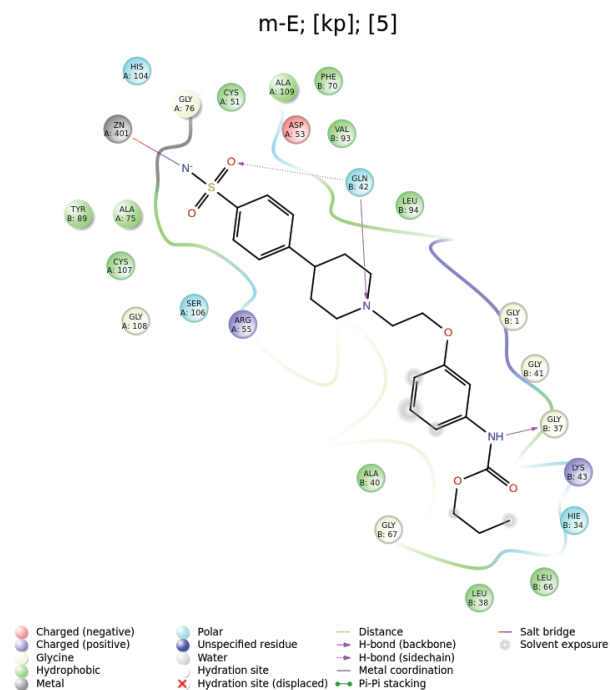


Figure 1. 2D interaction diagram of the molecule with the highest selection index interacting with mtCA II

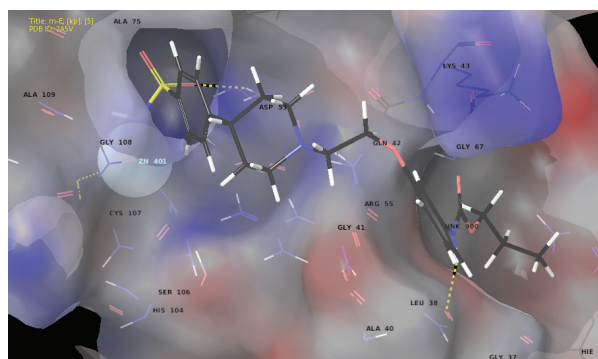


Figure 2. Molecule with the highest selection index docked in mycobacterial carbonic anhydrase II.

of 186 molecules. It did not give us exact information about what fragment combination can act as good pharmacophore, but it provided us with trends that included valuable information about which fragments were more likely to contribute to higher Selection score, and which parts of the molecule had no importance as a pharmacophore.

Structure-activity relationship

The suggested mechanism of action for phenylcarbamates, from which our set of sulfonamides originated, was inhibition of the proton pump of mycobacterial ATP synthase (Upadhayaya,

2009). While the hydroxyl group on the central linkage alkyl of those compounds seemed to be a pharmacophoric group for binding to the proton pump of mycobacterial ATP synthase, in the case of binding to mtCA II, it could be omitted without significant decrease of affinity. The removal of hydroxyl group would have probably led to lower affinity to FAS I, yet this could be overcome by benefits of increased lipophilicity and strongly diminished affinity to β -receptors, since these compounds clearly had pharmacophore of β -blockers.

All docked ligands had an expected coordination bond to the zinc dication in the active site through deprotonated nitrogen of sulfonamide group. For hydrogen bond, the interactions seemed to be very important Gln-42 (Fig. 1 and 2). While sidechain amide of Gln-42 was H-bond donor for oxygen of sulfonamide group, its backbone amide was H-bond donor for amine group connecting benzenesulfonamide fragment with central linkage. While alkylamines were mostly protonated at physiological pH, our results suggested that it was energetically favourable to keep neutral state, and be able to act as an acceptor of H-bond. Based on this, we can suppose that bioisosteric replacement of amine group by some other H-bond acceptors should preserve or meliorate the affinity.

The evaluation of fragment subset in relation to properties of the whole molecule gave us statistical insight into how fragments interact with protein, and which parts of the molecule tend to increase the activity or lipophilicity. From the group of sulfonamide fragments (**c**), fragment **5** (4-piperidin-4-ylbenzenesulfonamide) showed trend to affect the Selection index not only by higher $\text{LogP}_{\text{o/w}}$ but also by creating a stronger interaction with mtCA II resulting in higher absolute value of MM-GBSA free energy. None of the other four fragments **1-4** showed comparable behaviour (Fig. 3).

Alcohol and carbamate fragments showed that the length of their alkyl chain can affect $\text{LogP}_{\text{o/w}}$ without any major effect on selectivity or MM-GBSA free energy. Carbamate fragments **kb-km** showed a decrease in $\text{LogP}_{\text{o/w}}$ from average 2.12 (Var=0.54) for butyl chain to 0.90 (Var=0.41) for methyl derivative. Alcohol fragments **eb-em** showed very similar trend, where average $\text{LogP}_{\text{o/w}}$ for molecules containing butanol was 1.91 (Var=0.15) and 0.74 (Var=0.11) for methanol.

Central fragments showed no trends worth reporting, mainly due to a high number of derivatives in combination with different binding locations for fragments from (**a**) group in relation to the size of the subsets.

CONCLUSION

In this research, we created a combinatorial library of novel group of compounds aimed at inhibition of mycobacterial carbonic anhydrase II that was reported as a potential drug target, since it is essential for the survival of mycobacteria. We

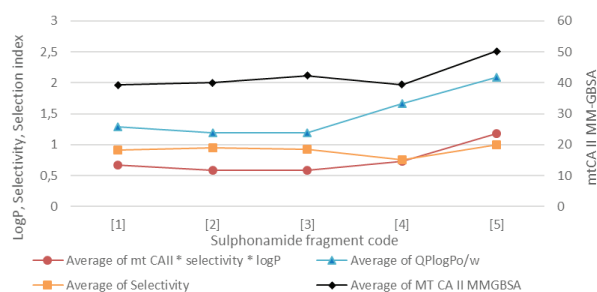


Figure 3. 4-piperidin-4-ylbenzenesulfonamide fragment [5] trends to elevate logP and strength of ligand-enzyme interaction

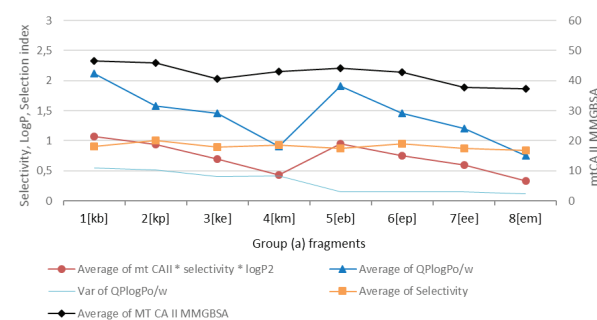


Figure 4. Average LogP is rising with longer alkyl chain for whole molecule, while selectivity and mtCA MM-GBSA haven't shown a significant relationship

used Selection index consisting of optimized MMBGSA free interaction energy (calculated with Schrodinger Prime) of molecules docked in mtCA II, Selectivity index and partition coefficient octanol/water ($\text{LogP}_{\text{o/w}}$). The best 27 molecules were selected as candidates for synthesis and further biological evaluation (Selection index 2.81-1.58). The biggest contribution to higher selection score showed 4-piperidin-4-ylbenzenesulfonamide fragment, where the impact of both MM-GBSA interaction energy and $\text{LogP}_{\text{o/w}}$ can be seen. The length of alkyl chains of carbamate and alcohol fragments had a major impact on $\text{LogP}_{\text{o/w}}$ and a minimal effect on selectivity and MM-GBSA energy. This part of the molecule can be effectively used to manipulate lipophilicity of the whole molecule in the desired way.

ACKNOWLEDGEMENTS

This contribution/publication is the result of the project implementation: Comenius University in Bratislava Science Park supported by the Research and Innovation Operational Programme funded by the ERDF Grant number : ITMS 26240220086 and grant VEGA 1/0743/13.

References

- [1] Behnke CA, Le Trong I, Godden JW, et al. Atomic resolution studies of carbonic anhydrase II. *Acta Crystallogr.* 2010;66:616-627.
- [2] Ceruso M, Vullo D, Scozzafava A, Supuran CT. Inhibition of human carbonic anhydrase isoforms I–XIV with sulfonamides incorporating fluorine and 1,3,5-triazine moieties. *Bioorg Med Chem.* 2013;21:6929-6936.
- [3] Covarrubias AS, Bergfors T, Jones TA, Högbom M. Structural mechanics of the pH-dependent activity of beta-carbonic anhydrase from *Mycobacterium tuberculosis*. *J Biol Chem.* 2006;281:4993-4999.
- [4] Covarrubias AS, Larsson AM, Högbom M, et al. Structure and Function of Carbonic Anhydrases from *Mycobacterium tuberculosis*. *J Biol Chem.* 2005;280:18782-18789.
- [5] Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J Med Chem.* 2006;49:6177-6196.
- [6] Keckesova S, Sedlarova E, Cizmarik J, et al. Antimycobacterial activity of novel derivatives of arylcarbonyloxyaminopropanols. *Ceska Slov Farm.* 2009;58:203-207.
- [7] Li J, Abel R, Zhu K, et al. The VSGB 2.0 Model: A Next Generation Energy Model for High Resolution Protein Structure Modelling. *Proteins* 2011;79:2794-2812.
- [8] Lipinski C, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001;46:3-26.
- [9] Maresca A. Carbonic anhydrase inhibitors. Inhibition of the Rv1284 and Rv3273 beta-carbonic anhydrases from *Mycobacterium tuberculosis* with. *Bioorg Med Chem Lett.* 2009;19:4929-4932.
- [10] Sassetti CM, Rubin EJ. Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A.* 2003;100:12989-12994.
- [11] Shelley JC, Cholleti A, Frye L, Greenwood JR, Timlin MR, Uchimaya M. Epik: a software program for pK_a prediction and protonation state generation for drug-like molecules. *J Comp-Aided Mo. Design.* 2007;21:681-691.
- [12] Showalter HD, Denny WA. A roadmap for drug discovery and its translation to small molecule agents in clinical development for tuberculosis treatment. *Tuberculosis (Edinb).* 2008;88(1 suppl):s3-s17.
- [13] Supuran CT. Carbonic anhydrases - An overview. *Curr Pharm Des.* 2008;14:603-614.
- [14] Supuran CT. Carbonic Anhydrase Inhibition/Activation: Trip of a Scientist Around the World in the Search of Novel Chemotypes and Drug Targets. *Curr Pharm Des.* 2010;16:3233-3245.
- [15] Upadhyaya RS, Kulkarni GM, Vasireddy NR, Vandavasi JK, Dixit SS, Sharma V, Chattopadhyaya J. Design, synthesis and biological evaluation of novel triazole, urea and thiourea derivatives of quinoline against *Mycobacterium tuberculosis*. *Bioorg Med Chem* 2009;17:4681-4692.
- [16] Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem.* 2002;45:2615-2623.
- [17] Waisser K, Dolezal R, Palat K, Cizmarik J, Kaustova J. QSAR study of antimycobacterial activity of quaternary ammonium salts of piperidinylethyl esters of alkoxy-substituted phenylcarbamic acids. *Folia Microbiologica.* 2006;51:21-24.
- [18] WHO. WHO Global tuberculosis report. WHO; 2013.