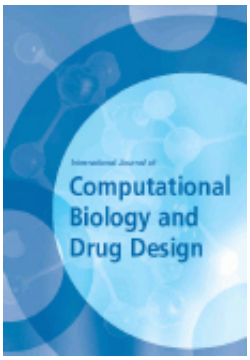


International Journal of Computational Biology and Drug Design

ISSN online: 1756-0764 - ISSN print: 1756-0756
https://www.inderscience.com/ijcbdd

Generation of 2D-QSAR and pharmacophore models for fishing better anti-leishmanial therapeutics

Clayton Fernando Rencilin, Joseph Christina Rosy, Krishnan Sundar

DOI: 10.1504/IJCBDD.2023.10055475

Article History:

Received:	28 December 2021
Last revised:	02 April 2022
Accepted:	04 April 2022
Published online:	17 April 2023

Generation of 2D-QSAR and pharmacophore models for fishing better anti-leishmanial therapeutics

Clayton Fernando Rencilin,
Joseph Christina Rosy
and Krishnan Sundar*

Department of Biotechnology,
School of Bio and Chemical Engineering,
Kalasalingam Academy of Research and Education,
Krishnankoil – 626126, Tamilnadu, India
Email: jehovahnissisjc@gmail.com
Email: christinarosy.j@gmail.com
Email: sundarkr@klu.ac.in
*Corresponding author

Abstract: Leishmaniasis, a life-threatening tropical disease that is endemic in nearly 100 countries, contributes to millions of deaths each year. However, very few antileishmanial compounds are available in the market and that too possess many drawbacks. Hence, the therapeutic arsenal requires potential and novel anti-leishmanial compounds to treat Leishmaniasis. In the present study, quantitative structure activity relationship (QSAR) model and Pharmacophore model were developed with a set of antileishmanial compounds collected from literature and commercial antileishmanial drugs. A ligand-based pharmacophore model was developed using active compound as template and it was used for searching the purchasable compound dataset of ZINC database for matching compounds. Thirteen novel, readily purchasable compounds were obtained from this approach, which shows good predicted activity, ADME and druglikeness. These compounds can be regarded as potential candidates to be developed as novel antileishmanial drugs with improved activity and reduced side effects.

Keywords: antileishmanial compounds; descriptor; pharmacophore; ZINCPharmar; pharmacophore search; QSAR; quantitative structure activity relationship.

Reference to this paper should be made as follows: Rencilin, C.F., Rosy, J.C., and Sundar, K. (2023) 'Generation of 2D-QSAR and pharmacophore models for fishing better anti-leishmanial therapeutics', *Int. J. Computational Biology and Drug Design*, Vol. 15, No. 4, pp.316–335.

Biographical notes: Clayton Fernando Rencilin has obtained his Masters' degree in Biotechnology from Kalasalingam Academy of Research and Education, Tamilnadu, India. He is currently pursuing his PhD at Indian Institute of Science, Bangalore, India in the field of Molecular Biophysics. He has coauthored four research articles.

Joseph Christina Rosy is currently pursuing her PhD at Kalasalingam Academy of Research and Education, Tamilnadu, India. She has completed her masters in Biotechnology and carrying out research on 'Targeting iron sequestration in

bacterial pathogens'. She has authored/ coauthored four publications and has presented posters in various national and international conferences.

Krishnan Sundar is currently a Professor of Biotechnology at Kalasalingam Academy of Research and Education, Tamilnadu, India. His research interest is on Infection and Immunity. He has coauthored more than 90 research articles in indexed journals. Research in his laboratory is supported by grants from Science and Engineering Research Board and Department of Biotechnology, Government of India.

1 Introduction

Leishmaniasis is one of the neglected tropical diseases affecting about 350 billion people all over the world (Kedzierski, 2010). This vector-borne disease is transmitted to humans by sandfly and caused by one of the several different species of the genus Leishmania, an obligate protozoan parasite. It holds the second place in mortality and fourth place in morbidity among all the tropical diseases according to the disease burden estimate (Bern et al., 2008). At present, 53 *Leishmania spp.* are known and 21 among them are reported to be pathogenic to humans (Akhoundi et al., 2016). Leishmanial infections occur in six different clinical forms: post-kala-azar dermal leishmaniasis (PKDL), mucocutaneous (MCL), mucosal (ML), diffuse cutaneous (DCL), cutaneous (CL) and visceral (VL) leishmaniasis. Among them, the most common form of infection reported is CL followed by VL. Worldwide annual incidence of visceral leishmaniasis (VL) and Cutaneous Leishmaniasis (CL) is estimated as ~0.3 million and ~0.95 million cases respectively (Alvar et al., 2012). It is now found that 70–100 countries are endemic to CL and VL (Desjeux et al., 2004; van Griensven and Diro, 2012; Burza et al., 2018; Kedzierski, 2010). *Leishmania major* is responsible for most of the cases of CL in the Mediterranean littoral, the Middle East, the Indian subcontinent, and central Asia and is endemic in many of the rural areas of various countries (Richard et al., 2005). Leishmania species causing CL were also reported to develop resistance to antileishmanial drugs (Molina et al., 2003; Croft et al., 2006).

During its life cycle, Leishmania occurs in tow forms namely:

i promastigote- flagellated, motile, non-dividing (metacyclic) organisms that live within sand-flies

ii amastigote- non motile form that lives in the host (Kaye and Scott, 2011).

There are number of antileishmanial drugs (ALD) available in the market to treat leishmanial infections. Antileishmanial drug resistance and toxicity were reported in many cases of leishmanial infections (Sangshetti et al., 2015). Also, these drugs were reported to have certain drawbacks that limit the usage of these drugs in therapy. The list of drugs available for treating Leishmaniasis and their drawbacks are listed in Table 1.

Clearly, the hunt for a better drug to treat Leishmaniasis is not yet over and the need for more effective and safer medications against Leishmania species is getting increased. Khraiwesh et al. have reported a set of 14 compounds that possess antileishmanial activity against *Leishmania major* (Khraiwesh et al., 2016). In the current work, these 14 compounds along with three commercial drugs and their IC50 values are used to derive

QSAR and Pharmacophore models. The pharmacophore model was then used for database search to obtain novel compounds that could potentially be developed as antileishmanial compounds against *Leishmania major*.

Table 1 List of commercially available anti-leishmanial drugs and their limitations

Antileishmanial drugs	*Limitations*	*References*
Pentamidine	Nephrotoxicity, hypotension, hypoglycaemia or local reactions	Goa and Campoli-Richards (1987)
Pentavalent antimonials	Nephrotoxicity	Veiga (1990)
Miltefosine	Gastrointestinal side effects	Jha et al. (1999)
Sitamaquine	Nephrotoxicity, Kidney failure causes methemeglobinemia	Yeates (2002)
Amphotericin B	Nephrotoxicity	Laniado-Laborín and Cabrales-Vargas (2009)

2 Materials and methods

2.1 Compounds (Dataset)

The antileishmanial compounds and their half maximal inhibitory concentration (IC50) were collected from literature (Khraiwesh et al., 2016). In addition, 4 commonly used Leishmanial drugs were also chosen. Hence, the dataset (Table 2) is composed of 14 antileishmanial compounds and 4 commercial antileishmanial drugs (Amphotericin B, Miltefosine, Paromomycin, Pentamidine). Out of these 18 compounds, 15 were regarded as training set while the remaining 3 were used as test set. The IC50 values of the compounds were obtained in nM unit and then were converted into negative logarithmic scale (pIC50 = –log IC50).

Table 2 Structure and IC50 values of compounds in dataset for ligand -based drug designing for Leishmaniasis

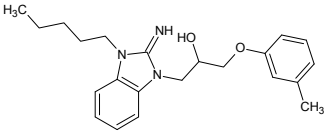
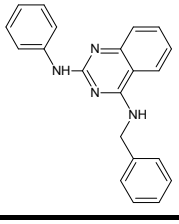
Compound ID	*2D Structures*	*IUPAC name*	*IC50 (nM)*
	Training dataset		
MMV000444		1-(2-imino-3-pentylbenzimidazol-1-yl)-3-(3-methylphenoxy)propan-2-ol	267.2
MMV006169		N-phenyl-N′-(phenylmethyl)quinazoline-2,4-diamine	466.6

Table 2 Structure and IC50 values of compounds in dataset for ligand -based drug designing for Leishmaniasis (continued)

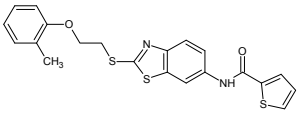
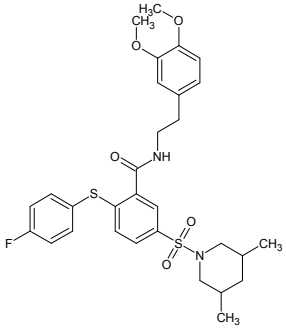
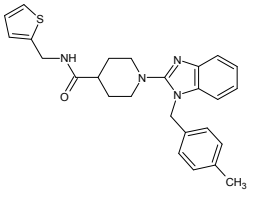
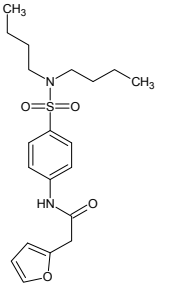
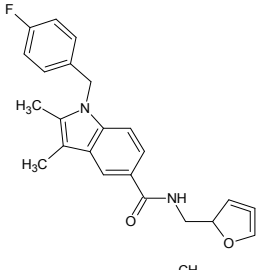
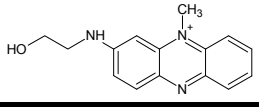
Compound ID	*2D Structures*	*IUPAC name*	*IC50 (nM)*
	Training dataset		
MMV007396		2-{2-[(4-nitrophenyl) methylidene]hydrazin-1-yl}-1H-1,3-benzodiazole	119.8
MMV007557		N-[2-(3,4-dimethoxyphenyl)ethyl]-5-(3,5-dimethylpiperidin-1-yl)sulfonyl-2-(4-fluorophenyl)sulfanylbenzamide	179.9
MMV007564		1-[1-[(4-methylphenyl)methyl]benzimidazol-2-yl]-N-(thiophen-2-ylmethyl)piperidine-4-carboxamide	60.8
MMV007881		N-[4-(dibutylsulfamoyl)phenyl]furan-2-carboxamide	417.7
MMV008149		1-[(4-fluorophenyl)methyl]-N-(furan-2-ylmethyl)-2,3-dimethylindole-5-carboxamide	228.9
MMV396693		2-[(10-methylphenazin-10-ium-2-yl)amino]ethanol	53.4

Table 2 Structure and IC50 values of compounds in dataset for ligand -based drug designing for Leishmaniasis (continued)

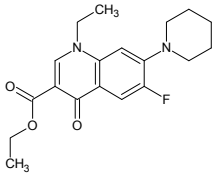
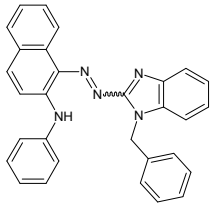
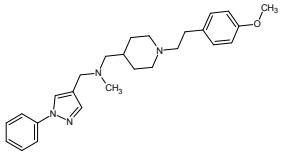
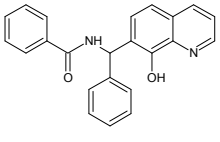
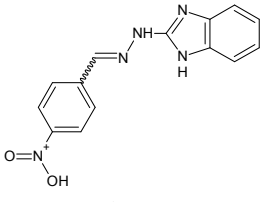
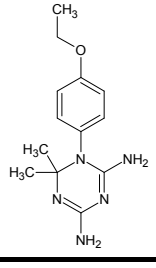
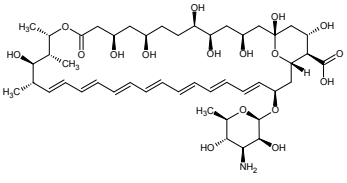
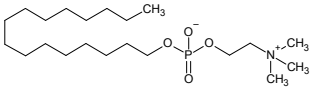
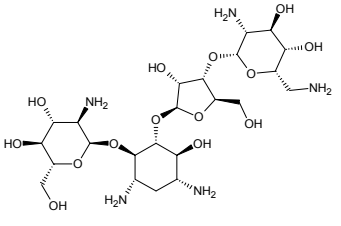
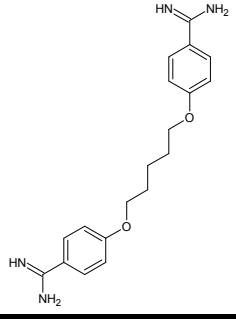
Compound ID	*2D Structures*	*IUPAC name*	*IC50 (nM)*
	Training dataset		
MMV665827		ethyl 1-ethyl-6-fluoro-4-oxo-7-piperidin-1-ylquinoline-3-carboxylate	477.7
MMV666023		2-{2-[(4-nitrophenyl)methylidene]hydrazin-1-yl}-1H-1,3-benzodiazole	91.7
MMV666069		({1-[2-(4-methoxyphenyl)ethyl]piperidin-4-ylmethyl)(methyl)[(1-phenyl-1H-pyrazol-4-yl)methyl]amine	366.5
MMV666080		N-[(8-hydroxyquinolin-7-yl)-phenylmethyl]benzamide	14.9
MMV666607		2-{2-[(4-nitrophenyl)methylidene]hydrazin-1-yl}-1H-1,3-benzodiazole	133.6
MMV667486		1-(4-ethoxyphenyl)-6,6-dimethyl-1,3,5-triazine-2,4-diamine	192.8

Table 2 Structure and IC50 values of compounds in dataset for ligand -based drug designing for Leishmaniasis (continued)

Compound ID	*2D Structures*	*IUPAC name*	*IC50 (nM)*
	Training dataset		
Amphotericin B		(1*R*,3*S*,5*R*,6*R*,9*R*,11*R*,15*S*,16*R*,17*R*,18*S*,19*E*,21*E*,23*E*,25*E*,27*E*,29*E*,31*E*,33*R*,35*S*,36*R*,37*S*)-33-[(2*R*,3*S*,4*S*,5*S*,6*R*)-4-amino-3,5-dihydroxy-6-methyloxan-2-yl]oxy-1,3,5,6,9,11,17,37-octahydroxy-15,16,18-trimethyl-13-oxo-14,39-dioxabicyclo[33.3.1]nonatriaconta-19,21,23,25,27,29,31-heptaene-36-carboxylic acid	27
	Test set		
Miltefosine		hexadecyl 2-(trimethylazaniumyl)ethyl phosphate	55
Paromomycin		(2*S*,3*S*,4*R*,5*R*,6*R*)-5-amino-2-(aminomethyl)-6-[(2*R*,3*S*,4*R*,5*S*)-5-[(1*R*,2*R*,3*S*,5*R*,6*S*)-3,5-diamino-2-[(2*S*,3*R*,4*R*,5*S*,6*R*)-3-amino-4,5-dihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-6-hydroxycyclohexyl]oxy-4-hydroxy-2-(hydroxymethyl)oxolan-3-yl]oxyoxane-3,4-diol	78
Pentamidine		4-[5-(4-carbamimidoylphenoxy)pentoxy]benzenecarboximidamide	342

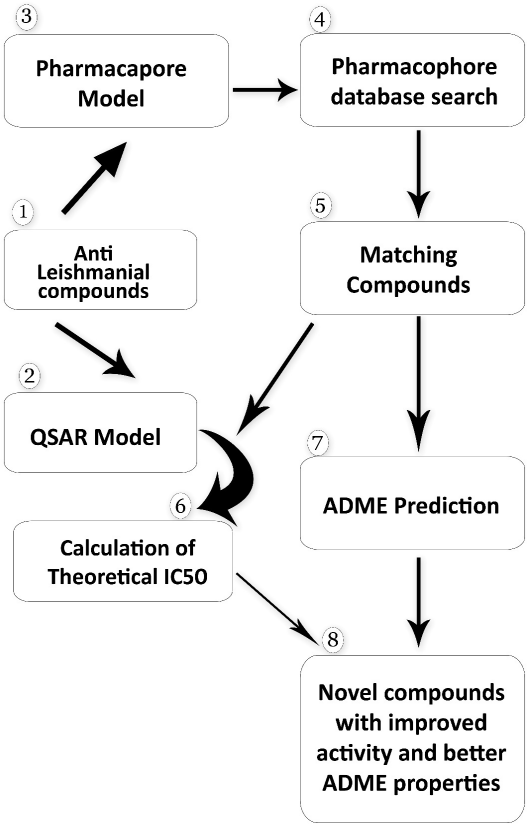
2.2 Computational data

The compounds were sketched using a freeware, ACD/ChemSketch (version C60E41). They were then converted into 3D structures using Discovery Studio 2019 (Discovery Studio Visualizer v19.1.0.18287). The structures of commercial drugs were obtained from PubChem database (https://pubchem.ncbi.nlm.nih.gov/). All these programs were administered on a machine with Core i3 2.30GHz processor running on Windows 10 operating system.

2.3 Generation of descriptors

SwissADME (http://www.swissadme.ch/index.php), a free web tool to compute physiochemical descriptors like ADME, druglikeness nature, pharmacokinetics properties, etc. provided by Molecular Modelling Group of Swiss Institute of Bioinformatics was used (Daina et al., 2017). From the computed properties, a set of 11 descriptors were picked for each compound for developing QASR models, which are listed in Table 3. The overall workflow is provided in Figure 1.

Figure 1 Overall workflow



2.4 QSAR – Activity predictions

The compounds were subjected to QSAR model generation by EasyQSAR software. The molecular descriptors and respective experimental IC50 values were used for modelling using multiple linear regression (MLR). For each model, descriptors were chosen randomly in the ratio of 5 : 1 (compounds: descriptors) (Rosy et al., 2016). Models were generated utilising all possible combinations of descriptors. The generated models were validated with the test set compounds and the best model was selected (Chowdhury et al., 2012; Das et al., 2017).

Table 3 Molecular descriptors computed by SwissADME for generation of QSAR models

Compounds	*Molecular weight (g/mol)*	*No. of heavy atoms*	*No. of arom. heavy atoms*	*No. of rotatable bonds*	*No. of H-bond acceptors*	*No. of H-bond donors*	*Molar refractivity*	*TPSA (Å²)*	*Log $P_{o/w}$ (iLOGP)*	*Log S (ESOL)*	*Log K_p (skin permeation) (cm/s)*
Training set											
MMV000444	367	27	9	9	3	2	109.96	63.17	3.62	–4.59	–5.55
MMV006169	326.39	25	5	5	2	2	102.87	49.84	3.35	–5.48	–4.58
MMV007396	426.57	28	8	8	3	1	119.22	133	4.05	–6.33	–4.57
MMV007557	586.74	40	11	11	7	1	158.87	118.62	4.41	–6.95	–5.53
MMV007564	444.59	32	7	7	2	1	134.88	78.4	3.77	–5.65	–5.58
MMV007881	392.51	27	12	12	5	1	107.15	88	3.35	–3.95	–6.25
MMV008149	376.42	28	6	6	3	1	107	43.26	3.62	–4.37	–6.12
MMV396693	254.31	19	3	3	2	2	78.11	49.03	1.44	–3.06	–6.39
MMV665827	346.4	25	5	5	4	0	99.77	51.54	3.16	–3.77	–6.36
MMV666023	453.54	35	6	6	3	1	143.15	54.57	4.27	–7.91	–3.43
MMV666069	418.57	31	9	9	4	0	130.36	33.53	4.72	–5.04	–5.7
MMV666080	354.4	27	5	5	3	2	105.74	62.22	2.84	–5.37	–5.02
MMV666607	282.28	21	4	4	4	3	78.65	93.38	–4.73	–3.62	–6.03
MMV667486	261.32	19	3	3	3	2	86.66	89.23	2.42	–1.98	–7.35
Amphotericin B	924.08	65	3	3	18	12	239.06	319.61	2.47	–5.37	–11.94
Test set											
Miltefosine	407.57	27	0	20	4	0	115.9	68.4	0.26	–5.32	–3.97
Paromomycin	615.63	42	0	9	19	13	133.56	347.32	1.08	2.44	–16.25
Pentamidine	340.42	25	12	10	4	4	100.7	118.2	2.24	–3.36	–6.56

2.5 Phamacophore model generation

All the compounds were converted and combined in a single mol2 file. Pharmacophore features were fabricated using PHARMAGIST (Schneidman-Duhovny et al., 2008). Various models were generated with all possible combinations of compounds, out of which best scoring pharmacophore model was chosen for further studies. The pharmacophore features of the best scoring pharmacophore were visualised using PyMOL molecular visualisation tool and analysed further by Discovery Studio 2019 visualiser.

2.6 Pharmacopohore database search

The pharmacophore generated was used to search in ZINC database for compounds that match the generated pharmacacophore model. The tool, ZINCPharmer was used for this purpose (Lipinski et al., 1997). The top matching compounds were chosen and their ADME and Druglikeness properties were predicted. The theoretical antileishmanial activities (IC50) were also predicted using the generated QSAR model.

Table 4 Druglikeness filters used in the study

Lipinski (Ghose et al., 1999)	*Ghose (Veber et al., 2002)*	*Veber (Egan et al., 2000)*	*Egan (Muegge et al. 2001)*	*Muegge (Koes and Camacho, 2012)*
MW ≤ 500	160 ≤ MW ≤ 480	Rotatable bonds ≤ 10	WLOGP ≤ 5.88	200 ≤ MW ≤ 600
MLOGP ≤ 4.15	–0.4 ≤ WLOGP ≤ 5.6	TPSA ≤ 140	TPSA ≤ 131.6	–2 ≤ XLOGP ≤ 5
N or O ≤ 10	40 ≤ MR ≤ 130			TPSA ≤ 150
NH or OH ≤ 5	20 ≤ atoms ≤ 70			Num. rings ≤ 7
				Num. carbon > 4
				Num. heteroatoms > 1
				Num. rotatable bonds ≤ 15
				H-bond acc. ≤ 10
				H-bond don. ≤ 5

2.7 ADME and druglikeness prediction

ADME properties and druglikeness were predicted for the top compounds obtained from pharmacophore database search. The ADME properties such as number of H-bond donors and acceptors, LogP, number of rotatable bonds, BBB permeation, GI absorption and Bioavailability were predicted using SwissADME. 'Drug-likeness' is the ability of a molecule to become an oral drug with respect to bioavailability. Drug likeness was predicted by SwissADME using rules as described by 6 different authors. These filters often originate from analyses by major pharmaceutical companies aiming to improve the quality of their proprietary chemical collections. The Lipinski filter is the pioneer rule, known as, rule-of-five which is used by Pfizer. Other filters, Ghose by Amgen, Veber by GSK, Egan by Pharmacia and Muegge by Bayer (Yeates,

2002). The rules describing all the above five methods are presented in Table 4. The compounds satisfying these rules are considered likely to be effectively developed into an oral drug.

3 Results

3.1 Generation of QSAR models

Totally, 9 QSAR models were generated from the 15 compounds that were used as training set and the models were validated using compounds included in the test set. R^2 values were calculated for each model and the best model was found to have a R^2 value of 0.4681. All the QSAR models generated and their corresponding R^2 values are presented in Table 5. The equation for the best scoring model is given below:

pIC50 = –9.321256040249E + 000 + –1.166262572987E + 000 * (LogS)
+ –1.062885390742 E + 000 * (LogKp)
+ –1.191587747759E-002 * (molwt)

The best model was derived using three descriptors namely LogS, LigKp and molecular weight (molwt) and all the three descriptors were found to be negatively influencing the activity as indicated by their coefficients in the equation.

Table 5 Generated QSAR models by random combination of molecular descriptors

Eqn no.	*Generated QSAR equations*	R^2
1	logIC50 = –2.150910758940E + 000 + 3.256306013600E-004*(molwt) + 1.113746379373E-002*(heavy) + –7.190811949380E-002*(aromatic)	0.3458
2	logIC50 = –2.077885476677E + 000 + –1.922606884875E-002*(rotatable) + –4.850549380525E-002*(Hacceptors) + 1.313890645734E-001*(Hdonors)	0.3296
3	logIC50 = –2.506269846845E + 000 + 3.322100644032E-003*(Molar) + 1.152261202649E-003*(TPSA) + –4.823078512426E-002*(logP)	0.216
4	logIC50 = –9.321256040249E + 000 + –1.166262572987E + 000*(LogS) + –1.062885390742E + 000*(LogKp) + –1.191587747759E-002*(molwt)	0.4681
5	logIC50 = –2.673077244687E + 000 + –1.314641633202E-001*(Hacceptors) + 1.819237777897E-001*(Hdonors) + 6.101124088349E-003*(Molar)	0.3953
6	logIC50 = –2.598550819236E + 000 + 2.415485477259E-003*(TPSA) + –4.885113221145E-002*(logP) + –7.786676103476E-002*(LogS)	0.2524
7	logIC50 = –2.778627140556E + 000 + –2.680891280987E-002*(LogKp) + –4.692767458284E-003*(molwt) + 8.080113123977E-002*(heavy)	0.175
8	logIC50 = –2.323341280251E + 000 + 9.692128808326E-002*(Hdonors) + 5.784062327190E-004*(Molar) + –1.008510610021E-003*(TPSA)	0.2584
9	logIC50 = –3.252607459717E + 000 + –6.230330795319E-002*(logP) + –1.462392190962E-001*(LogS) + –9.481312770372E-002*(LogKp)	0.2563

Validation of the models was done by predicting the activities of test set compounds and are presented in Table 6. It was found that the predicted values were approximately equal to the actual values. Model 4 predicted the activity of the test set compound Pentamidine as –2.49 which is almost equal to the actual activity value, –2.53. This indicates the model 4 is best out of the derived models and can further be used for analysis.

3.2 Generation and analysis of pharmacophore models

Pharmacophore models were generated with various combinations of compounds. The best scoring model was selected and visualised in PyMol. The features in the best scoring pharmacophores are given in Table 7. The pharmacophore that was generated with 8 aligned molecules and a score of 28.46 was selected as the best model and used for further studies. This indicates that the model 4 is best out of the derived models and can further be used for analysis (Figure 2). The positions of the features and distance between the features of the best model are shown in Figure 3(a)–(c).

Figure 2 Prediction by the best model (model 4) (see online version for colours)

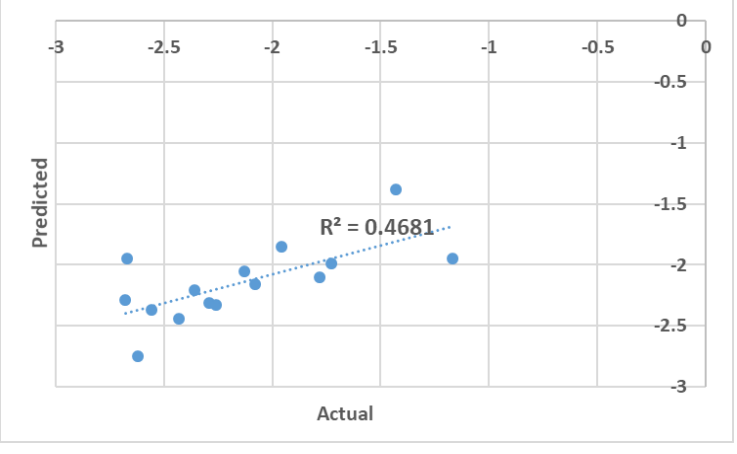
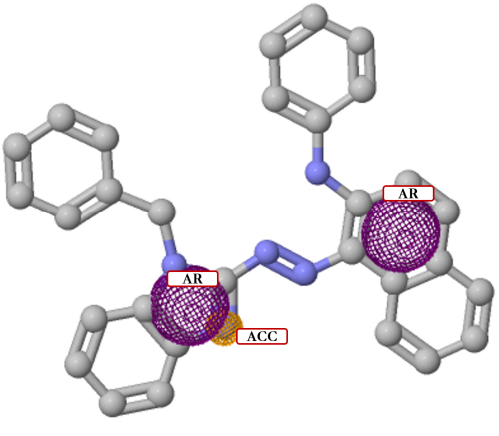
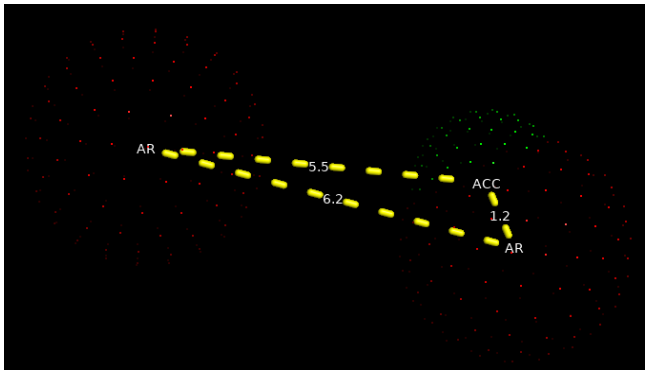


Figure 3(a) Features detected in the pharmacophore (see online version for colours)



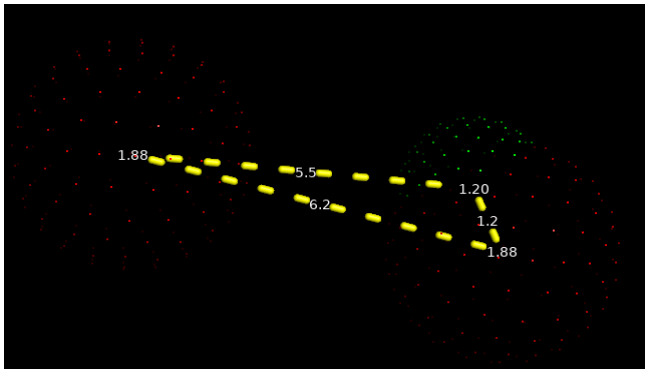
AR-Aromatic; ACC-Acceptor.

Figure 3(b) Spatial distance between the detected features (see online version for colours)



The spatial distance between the features in the pharmacophore is shown in the picture. The values are givein in Angstrom units. AR-Aromatic; ACC-Acceptor

Figure 3(c) Radius of detected features (see online version for colours)



The radius of each feature is shown in the picture. The values are givein in Angstrom units.

Table 6 Validation of QSAR models

Eqn. no.	R^2	*Miltefosine*		*Paromomycin*		*Pentamidine*	
		Actual	*Predicted*	*Actual*	*Predicted*	*Actual*	*Predicted*
1	0.3458	–1.7404	–1.72	–1.8921	–7.48	–2.534	–2.62
2	0.3296		–2.66		–1.46		–1.94
3	0.216		–2.05		–1.71		–2.14
4	0.4681		–3.75		–2.23		–2.49
5	0.3953		–2.49		–1.99		–1.86
6	0.2524		–2.03		–2		–2.16
7	0.175		–2.4		–1.84		–2.18
8	0.2584		–2.33		–1.34		–2
9	0.2563		–2.11		–2.14		–2.28

Table 7 Features present in the high scored pharmacophores

Features	*Model 1*	*Model 3*	*Model 3*
Number of aligned molecules	13	12	8
Score	24.875	27.042	28.46
Features	3	3	3
Spatial features	3	3	3
Aromatic	2	2	3
Hydrophobic	0	0	0
Donors	0	0	0
Acceptors	1	1	0
Negatives	0	0	0
Positives	0	0	0

*Values in the shaded column (Model 3) were selected for further studies

3.3 Searching for novel compounds with better activity

The best scoring pharmacophore was selected for fishing the ZINC database for matches. The search was performed using the tool ZINCPharmer and the tool retrieved a total of 50, 112, 847 hits matching the pharmacophore. Top compounds that match with the pharmacophore were selected with the RMSD cut off of ≤0.03. Thirteen compounds were found to be satisfying the criterion which are listed in Table 8. Theoretical activities of these compounds were also predicted using the QSAR model generated (Table 9). The ADME properties such as the number of H-bond donors and acceptors, LogP, number of rotatable bonds, BBB permeation, GI absorption and Bioavailability were predicted for these compounds. The druglikeness properties of these compounds were predicted based on the rules described by 6 different rules which are also presented in Table 9.

Table 8 Structures of top compounds matching with the pharmacophore filtered from ZINC database search results with the RMSD cut off of ≤0.003

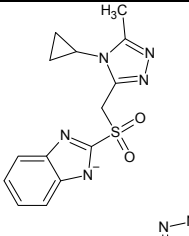
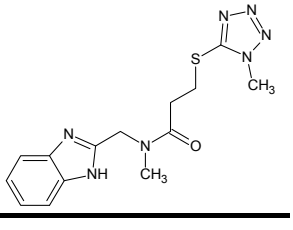
S. no.	*Name*	*RMSD*	*Structure*
1	ZINC91574499	0.002	
2	ZINC58290123	0.002	

Table 8 Structures of top compounds matching with the pharmacophore filtered from ZINC database search results with the RMSD cut off of ≤0.003 (continued)

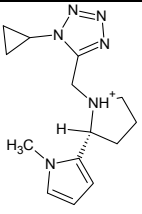
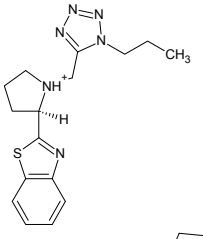
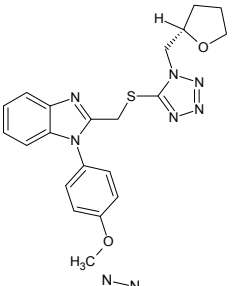
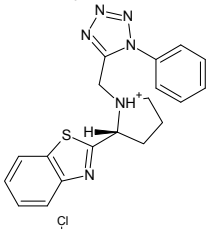
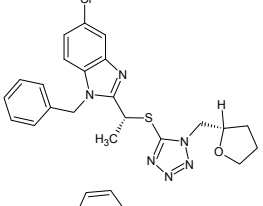
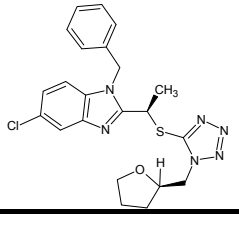
S. no.	*Name*	*RMSD*	*Structure*
3	ZINC71822573	0.002	
4	ZINC66562798	0.002	
5	ZINC28189912	0.003	
6	ZINC09638087	0.003	
7	ZINC29394394	0.003	
8	ZINC29394402	0.003	

Table 8 Structures of top compounds matching with the pharmacophore filtered from ZINC database search results with the RMSD cut off of ≤0.003 (continued)

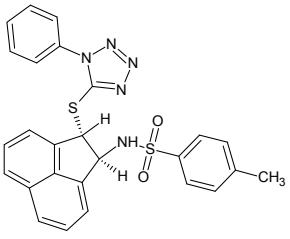
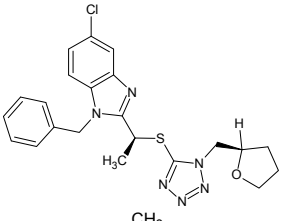
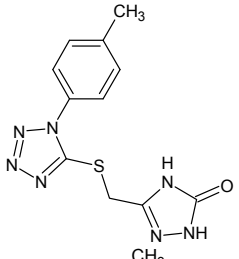
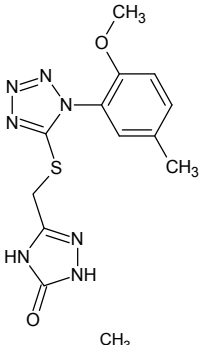
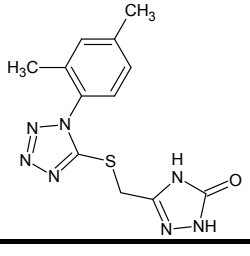
S. no.	*Name*	*RMSD*	*Structure*
9	ZINC02836915	0.003	
10	ZINC29394405	0.003	
11	ZINC92869862	0.003	
12	ZINC92869832	0.003	
13	ZINC91557386	0.003	

Table 9 Predicted IC50 values, ADME and druglikeness properties of novel compounds selected from pharmacophore-based database search

Compound ID	*Predicted IC50 (nM)*	*H- bond acceptors*	*H-bond donors*	*Rotatable bonds*	*$LogP_{o/w}$*	*GI Absorption*	*BBB permeation*	*Bioavailability score*	*Lipinski*	*Ghose*	*Veber*	*Egan*	*Muegge*
ZINC91574499	102.33	6	0	4	1.42	High	No	0.56	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC58290123	177.83	5	1	7	1.22	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC71822573	158.49	3	1	4	0.26	High	No	0.55	Yes; 0 violation	No; 1 violation: WLOGP←0.4	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC66562798	177.83	4	1	5	1.62	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC28189912	123.03	6	0	7	3.18	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC09638087	74.13	4	1	4	2.18	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC29394394	114.82	5	0	7	3.96	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation

Table 9 Predicted IC50 values, ADME and druglikeness properties of novel compounds selected from pharmacophore-based database search (continued)

Compound ID	*Predicted IC50 (nM)*	*H- bond acceptors*	*H-bond donors*	*Rotatable bonds*	*$LogP_{o/w}$*	*GI Absorption*	*BBB permeation*	*Bioavailability score*	*Lipinski*	*Ghose*	*Veber*	*Egan*	*Muegge*
ZINC29394402	114.82	5	0	7	4.03	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC02836915	66.07	6	1	6	4.38	Low	No	0.55	Yes; 1 violation: MLOGP > 4.15	No; 2 violations: MW > 480, MR > 130	Yes; 0 violation	Yes; 0 violation	No; 1 violation: XLOGP3>5
ZINC29394405	114.82	5	0	7	4.09	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC92869862	87.10	5	2	4	1.36	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation
ZINC92869832	107.15	6	2	5	1.35	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	No; 1 violation: TPSA > 131.6	Yes; 0 violation
ZINC91557386	91.20	5	2	4	1.71	High	No	0.55	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation	Yes; 0 violation

4 Discussion

Leishmaniasis is a group of diseases caused by intracellular protozoa of the genus *Leishmania*. It has become a major focus of concern and a serious problem of the developing countries affecting the poorer sections of the society (WHO: Leishmaniasis Fact sheet). As of now, no vaccines are available for this disease and drugs are the only way of treating Leishmaniasis (Croft and Coombs, 2003; Hussain et al., 2014). Though many anti-leishmanial drugs are available in the market, they fail to serve the purpose because of the vast number of side effects that they produce. The search for new anti-leishmanial drugs with a better biological activity and reduced side effects has started a long before; however, the search is not yet over as no novel compound with reduced side effect was found (Bhargava and Singh, 2012).

The current work was aimed to computationally identify a novel compound that could be tested further to be developed into a potential anti-leishmanial drug with no side effects. This was achieved by a strategical pipeline starting with the collection of reported anti-leishmanial compounds and drugs. In this ligand-based drug design approach, QSAR models were generated with 14 already reported anti-leishmanial compounds and 3 commercially available drugs. The model was validated using test sets and the best QSAR model was used for further studies. Pharmacophore model was also created which was used for a database similarity search. ZINC database was searched for similar compounds that are matching the pharmacophore. The 13 best matching compounds were filtered out, for which theoretical IC50, ADME properties and Druglikeness properties were predicted.

The compound 'ZINC02836915' was found to have the lowest theoretical IC50 value (66.07 nM); But this compound failed to satisfy 3 out of 6 rules for druglikeness. Whereas the compound ZINC09638087 was found to have a less IC50 value and poses a good ADME and Druglikeness properties. The bioavailability score of this compound was also found to be optimum. Though this compound cannot be used to treat VL, as it does not possess the capability of crossing the blood brain barrier (BBB), it could efficiently be used for treating CL.

Acknowledgements

KS would like to acknowledge the financial support by Science and Engineering Research Board of India (EMR/2016/003035). CFR acknowledges Kalasalingam Academy of Research and Education for a Post-Graduate Scholarship.

References

Akhoundi, M., Kuhls, K., Cannet, A., Votýpka, J., Marty, P., Delaunay, P. and Sereno, D. (2016) 'A historical overview of the classification, evolution, and dispersion of leishmania parasites and sandflies', *PLoS Negl. Trop. Dis.*, Vol. 10, pe0004349.

Alvar, J., Vélez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jean, J., Margriet den, B., the WHO Leishmaniasis Control Team (2012) 'Leishmaniasis worldwide and global estimates of its incidence', *PloS One*, Vol. 7, p.e35671.

Bern, C., Maguire, J.H. and Alvar, J. (2008) 'Complexities of assessing the disease burden attributable to leishmaniasis', *PLoS Negl. Trop. Dis.*, Vol. 2, p.e313.

Bhargava, P. and Singh, R. (2012) 'Developments in diagnosis and antileishmanial drugs', *Interdiscip. Perspect. Infect. Dis.*, Vol. 2012, p.626838.

Burza, S., Croft, S.L. and Boelaert, M. (2018) 'Leishmaniasis', *The Lancet*, Vol. 392, pp.951–970.

Chowdhury, A., Sen, S., Dey, P., Chetia, P., Talukdar, A.D., Bhattacharjee, A. and Choudhury, M.D. (2012) 'Computational validation of 3-ammonio-3-(4-oxido-1H-imidazol-1-ium-5-yl) propane-1, 1-bis (olate) as a potent anti-tubercular drug against mt-metAP', *Bioinformation*, Vol. 8, pp.875–878.

Croft, S.L. and Coombs, G.H. (2003) 'Leishmaniasis–current chemotherapy and recent advances in the search for novel drugs', *Trends in Parasitol.*, Vol. 19, pp.502–508.

Croft, S.L., Sundar, S. and Fairlamb, A.H. (2006) 'Drug resistance in leishmaniasis', *Clin. Microbiol. Rev.*, Vol. 19, pp.111–126.

Daina, A., Michielin, O. and Zoete, V. (2017) 'SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules', *Sci. Rep.*, Vol. 7, p.42717.

Das, S., Laskar, M.A., Sarker, S.D., Choudhury, M.D., Choudhury, P.R., Mitra, A., Jamil, S., Lathiff, S.M.A. and Abdullah, S.A., Basar, N., Nahar, L., Talukdar, A.D. (2017) 'Prediction of anti-alzheimer's Activity of Flavonoids Targeting Acetylcholinesterase *in silico*', *Phytochem. Anal*, Vol. 28, pp.324–331.

Desjeux, P. (2004) 'Leishmaniasis', *Nat. Rev. Microbiol.*, Vol. 2, pp.692–693.

Egan, W.J., Merz, K.M. and Baldwin, J.J. (2000) 'Prediction of drug absorption using multivariate statistics', *J. Med. Chem*, Vol. 43, pp.3867–3877.

Ghose, A.K., Viswanadhan, V.N. and Wendoloski, J.J. (1999) 'A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases', *J. Comb. Chem*, Vol. 1, pp.55–68.

Goa, K.L. and Campoli-Richards, D.M. (1987) 'Pentamidine isethionate. A review of its antiprotozoal activity, pharmacokinetic properties and therapeutic use in pneumocystis carinii pneumonia', *Drugs*, Vol. 33, pp.242–258.

Hussain, H., Al-Harrasi, A., Al-Rawahi, A., Green, I.R. and Gibbons, S. (2014) 'Fruitful decade for antileishmanial compounds from 2002 to late (2011) Chem. Rev., Vol. 114, pp.10369–10428.

Jha, T.K., Sundar, S., Thakur, C.P., Bachmann, P., Karbwang, J., Fischer, C., Voss, A. and Berman, J. (1999) 'Miltefosine, an oral agent, for the treatment of Indian visceral leishmaniasis', *New Engl. J. Med.*, Vol. 341, pp.1795–1800.

Kaye, P. and Scott, P. (2011) 'Leishmaniasis: Complexity at the host–pathogen interface', *Nat. Rev. Microbiol.*, Vol. 9, pp.604–615.

Kedzierski, L. (2010) 'Leishmaniasis vaccine: where are we today?', *J. Glob. Infect. Dis*, Vol. 2, pp.177–185.

Khraiwesh, M., Leed, S., Roncal, N., Johnson, J., Sciotti, R., Smith, P., Read, L., Paris, R., Hudson, T., Hickman, M. and Grogl, M. (2016) 'Antileishmanial activity of compounds derived from the medicines for malaria venture open access box against intracellular leishmania major amastigotes', *Am J. Trop Med Hyg*, Vol. 94, pp.340–347.

Koes, D.R. and Camacho, C.J. (2012) 'ZINCPharmer: Pharmacophore search of the ZINC database', *Nucleic Acids Res.*, Vol. 40, pp.W409–W414.

Laniado-Laborín, R. and Cabrales-Vargas, M.N. (2009) 'Amphotericin B: side effects and toxicity', *Rev Iberoam Micol*., Vol. 26, pp.223–227.

Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J. (2001) 'Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings', *Adv. Drug Deliv. Rev.*, Vol. 23, pp.3–26.

Molina, R., Gradoni, L. and Alvar, J. (2003) 'HIV and the transmission of leishmania', *Ann. Trop. Med. Parasitol.*, Vol. 97, pp.29–45.

Muegge, I., Heald, S.L. and Brittelli, D. (2001) 'Simple selection criteria for drug-like chemical matter', *J. Med. Chem.*, Vol. 44, pp.1841–1846.

Richard, L.G., David, H.W. and Peter, F. (2005) *Tropical Infectious Diseases*., 2nd ed., Elsevier, Churchill Livingstone.

Rosy, J., Sundar, K., Balamurali, S., Mary, J. and Shenbagarathai, R. (2016) 'Generation of 2D-qSAR model for angiogenin inhibitors: a ligand-based approach for cancer drug design', *Trends in Bioinformatics*, Vol. 9, pp.1–13.

Sangshetti, J.N., Khan, F.A.K., Kulkarni, A.A., Arote, R. and Patil, R.H. (2015) 'Antileishmanial drug discovery: comprehensive review of the last 10 years', *RSC Adv.*, Vol. 5, pp.32376–32415.

Schneidman-Duhovny, D., Dror, O., Inbar, Y., Nussinov, R. and Wolfson, H.J. (2008) 'PharmaGist: A webserver for ligand-based pharmacophore detection', *Nucleic Acids Res*., Vol. 36, pp.W223–W228.

van Griensven, J. and Diro, E. (2012) 'Visceral leishmaniasis', *Infect. Dis. Clin. North Am.*, Vol. 26, pp.309–322.

Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W. and Kopple, K.D. (2002) 'Molecular properties that influence the oral bioavailability of drug candidates', *J. Med. Chem.*, Vol. 45, pp.2615–2623.

Veiga, J.P.R. (1990) 'Pentavalent antimonial nephrotoxicity in the rat', *Rev. Inst. Med. Trop. S. Paulo*, Vol. 32, pp.304–309.

Yeates, C. (2002) 'Sitamaquine (GlaxoSmithKline/Walter reed army institute)', *Curr. Opin. Investig. Drugs*, Vol. 3, pp.1446–1452.

Website

World Health Organization: Leishmaniasis [Fact sheet]- 2020. Accessed from: https://www.who.int/news-room/fact-sheets/detail/leishmaniasis on 2 December, 2020.