
**SciForum**
**MOL2NET**

# Virtual screening of a cyclics imides to evaluate potential new multi-target agents against species of Leishmania

José Alixandre Luis[1,3,*], Normando Costa[2], Cristiane Luis[1], Luciana Scotti[3] and Marcus Scotti[3]

[1] *Federal University of Campina Grande, 58.175-000 Cuité, PB, Brazil; E-mails:* jalixluis@hotmail.com; criscosmosilva@hotmail.com

[2] *Chemistry Department, Federal University of Paraíba, 58.051-900 João Pessoa, PB, Brazil; E-mail:* normandoalex@uol.com.br

[3] *Post-Graduate Program in Natural and Synthetic Bioactive Products, Federal University of Paraíba, 58.051-900 João Pessoa, PB, Brazil; E-mails:* luciana.scotti@gmail.com; mtscotti@gmail.com

* Author to whom correspondence should be addressed; E-Mail: jalixluis@hotmail.com
Tel.: +55-83-99654-4529

*Received: / Accepted: / Published:*

**Abstract:** Leishmaniasis is a neglected disease that does not have adequate treatment. To try to solve this problem, we have tested a database with 33 cyclic imides and evaluated their potential anti-Leishmanial activity (*L. donovani*) through ligand-based and structure based virtual screening. A diverse set selected from CHEMBL databanks of 818 structures (*L. donovani*) with tested antileishmanial activity against promastigotes forms, were classified according pIC50 values in order to generate and validate Random Forest model that show higher statistical indices values. The structure of four different *L. donovani* enzymes were downloaded from PDB databank and imides structures were submitted to molecular docking. *In silico* study allowed us to suggest that the cyclic imide $5_{27}$ can be tested as a potential multitarget molecule for leishmanial treatment, presenting activity against four strategic enzymes to treatment with probability of activity of 60%.

**Keywords:** Cyclic imides, Virtual Screening, Molecular Docking, *Leishmania donovani*, antileishmanial activity.

**Mol2Net YouTube channel**: *http://bit.do/mol2net-tube*
**YouTube link:** *please, paste here the link to your personal YouTube video, if any.*

## 1. Introduction

Despite the great development of modern medicinal chemistry there are some microbial diseases that remain without adequate chemotherapeutic agents, either due to problems of toxicity or resistance, among them is Leishmaniasis. Leishmaniasis is a complex of infectious diseases caused by parasites of the family Trypanossomatidae and genus *Leishmania*[1, 2]. It affects around 12 million people around the world, there are reported cases in 98 countries spread across 5 continents, mainly in poor countries, making the disease be classified as a neglected disease by the World Health Organization[2-4]. In this context, strategies to obtain new, more active and less toxic drugs should be stimulated. Sources of natural products combined with synthetic and chemoinformatic methodologies are strategies used to obtain molecules that are most likely to be effective against a specific disease. Computer-Aided Drug Design (CADD) has become an indispensable tool to the pharmaceutical industry and academia in the last years and has been employed during various stages of the drug designs process. Initially, this method focuses on reducing the overall number of possible ligands; in the later stages, during lead-optimization, the emphasis shifts to reducing experimental costs and the duration of time required to make a discovery. Applied ligand-based virtual screening using Volsurf and Molegro descriptors and a random forest algorithm (a method of machine learning) were included in the structure-based virtual screening[5-7].

A group of compounds with various biological activities are cyclic imides, which present a large class of compounds obtained by organic synthesis including several subclasses, among them maleimides, succinimides, glutarimides, phthalimides and naphthalimides, as well as their respective derivatives[8]. Because they are electronically neutral and of a hydrophobic nature, they easily cross cell membranes, leading to the important pharmacological effects of these imides, such as anti-inflammatory, antitumor, antimicrobial activities among others, which may be related to the size and characteristics of the groups present in the imidic ring, which may alter the steric characteristics of the molecules and altering their activity[9-13].

Taking into account the great medicinal importance of cyclic imides and their derivatives, as well as the excellent results found and published to date, the planning of 33 cyclic imides (Figure 1) was carried out and a virtual chemical screening was carried out to select molecules with higher probability to show the desired effect against selected *Leishmania* targets.

## 2. Results and Discussion

The Volsurf (v 1.0.7) program generated 128 descriptors that, together with the dependent variables (binary classification) that described whether the compounds were active (A) or inactive (I), were used as input data in the Knime program (v. 3.4.0) to generate the RF model. For all compounds that comprised the training data sets, the generation of all 128 descriptors by

Volsurf+ was rapid, taking approximately 25 minutes using a computer with an i7 processor, running at 2.6 GHz, and equipped with 8 GB of RAM.

Table 1 summarizes the statistical indices of the RF model for the training, cross-validation, and test sets for compounds tested against forms Promastigotes of *L. donovani.* For the training set, the learning machine program gave the same hit rates for the inactive compounds and active compounds, which were 100%. However, for the cross-validation and test sets, the RF model was better, in the study, at predicting the inactive compounds; the specificity (true positive rate) was lower for the cross-validation and test sets (71.06% and 79.31%, respectively) than the sensitivity (true negative rate), which was measured to be 91.19% and 91.43%, respectively (Table 1). The ROC plot that was generated for the test set, which plotted the true positive (active) rate against the false positive rates had an area under the curve (AUC) value of approximately 0.91, which is significantly higher than 0.5. The Matthews Correlation Coefficient (MCC) values for training, cross-validation, and test sets were 1.000, 0.645 and 0.716, respectively. Because an MCC value of 1 represents a perfect prediction, 0 represents random prediction, and -1 represents total disagreement between prediction and observation, the RF model shows significant MCC values.

We evaluated the potential of 1-7 imides as antileishmanial leads using the RF model and docking on selected Leishmania enzyme targets. The results indicate that these compounds could show activity. The compounds showed good performance against TOPI and OASS.

Five structures were indicated as potentially active (compounds 4 and 5 were indicated as potentially inactive) using the RF mode. Docking results gave similar values for all compounds. Therefore, we evaluated a databank of 33 cyclic imides to obtain a qualitative structure-activity relationship using a combined approach of virtual screening, structure based and ligand-based, in order to select compounds with potential higher antileishmanial activity.

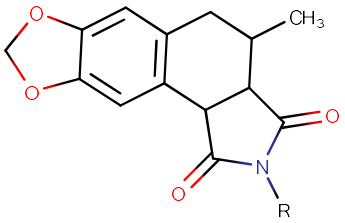
A computational chemistry multitarget model to predict the results of experimental tests for Leishmania with significant success has been reported in the literature[21], so we used our *in silico* results of the cyclic imides to select structures that presented lower energy binding (7 compounds) from each enzyme. Looking for multitarget compounds, we selected imides with activity against three or more enzymes. The compounds 27 and 32 presents activity against all enzymes: TOPI, NMT, Cyp and OASS. The compounds 11, 16, 18, 26 and 33 show low Moldock score energies against three different enzymes. From these structures, we selected only the compounds that were classified as active in the RF model. Therefore, our methodology was to apply two screening approaches simultaneously: ligand-based screening, using the RF model generated using Volsurf descriptors from the dataset of 818 compounds and structure-based screening using four enzymes. The ligand-based approach of those seven compounds shows six of them as active (probability over 50%)

corroborating with docking studies. Only the compound 11 present percentages close to the active line, with values of 48%.

The analysis of activity of selected imides demonstrates an affinity of this group of compounds with the enzyme NMT, where all presented a good interaction. Six compounds presented a good interaction with OASS, five compounds with TOPI and five with Cyp.

**Table 1.** Summary of training, internal cross-validation, test results, and corresponding match results, which were obtained using the RF algorithm on the total set of 818 compounds tested against forms Promastigotes of *L. donovani* (655 were in the training set and 163 in the test set).

| | Training | | | Validation | | Test | | |
|---|---|---|---|---|---|---|---|---|
| | **Samples** | **Match** | **%Match** | **Match** | **%Match** | **Samples** | **Match** | **%Match** |
| Active | 293 | 293 | 100 | 235 | 71.06 | 58 | 46 | 79.31 |
| Inactive | 525 | 525 | 100 | 420 | 91.19 | 105 | 96 | 91.43 |
| Overall | 818 | 818 | 100 | 655 | 83.97 | 163 | 142 | 87.12 |




1-33

1: R= 4-F,3-$NO_2$Ph
2: R= 4-Br,3-$NO_2$Ph
3: R= 4-Cl,3-$NO_2$Ph
4: R= 4-COOHPh
5: R= $SO_2NH_2$Ph
6: R= $CH_2$Ph
7: R= Ph
8: R= 4-$CH_3$Ph
9: R= 4-$CH_2CH_3$Ph
10: R= 4-$CH(CH_3)_2$Ph
11: R= 4-PhPh
12: R= 3-$CH_3$Ph
13: R= 3-$CH_2CH_3$Ph
14: R= 3-ClPh
15: R= 3-BrPh
16: R= 3-$NO_2$Ph
17: R= 3-COOHPh
18: R= 3-PhPh
19: R= 3-OHPh
20: R= 4-OHPh
21: R= 4-$NO_2$Ph
22: R=$CH_2$-4-$CH_3$Ph
23: R= $CH_2$-4-OHPh
24: R= $CH_2$-4-ClPh
25: R= $CH_2$-4-BrPh
26: R= $CH_2$-4-COOHPh
27: R= $CH_2$-4-$NO_2$Ph
28: R= $CH_2$-3-$CH_3$Ph
29: R= $CH_2$-3-OHPh
30: R= $CH_2$-3-ClPh
31: $CH_2$-3-BrPh
32: $CH_2$-3-$CH_2CH_3$Ph
33: $CH_2$-3-$CH(CH_3)_2$Ph

**Figure 1.** Cyclic imides tested

## 3. Materials and Methods

### Dataset

From the ChEMBL database, we selected a diverse set of 818 structures (https://www.ebi.ac.uk/chembl/), which had been screened (in vitro) to inhibit the promastigote *L. donovani* The compounds were classified using values of -log$IC_{50}$ (mol/L) = p$IC_{50}$, which led us to assign 293 actives (p$IC_{50}$ ≥ 5.0) and 525

inactives ($pIC_{50} \leq 5.0$). We used a border in the $pIC_{50}$ values looking for better prediction results. In this case, $IC_{50}$ represented the concentration required for 50% inhibition of promastigote *L. donovani*. The compounds with $pIC_{50}$ values between 4.7 and 5.0 were excluded to minimize the border effect and improve the discriminant power of the generated models. Our databank includes compounds 1-33. For all structures, SMILES codes were used as input data to Marvin 17.18.0.1784, 2017, ChemAxon (http://www.chemaxon.com). We used Standardizer software [JChem 17.18.0.1784, 2017; ChemAxon (http://www.chemaxon.com)] to canonize structures, add hydrogens, perform aromatic form conversions, clean the molecular graph in three dimensions, and save compounds in sdf format[14,15].

**Volsurf Descriptors**

Three-dimensional structures (3D) were used as input data in the Volsurf+ program v. 1.0.7 and were subjected to molecular interaction fields (MIFs) to generate descriptors using the following probes: N1 (amide nitrogen-hydrogen bond donor probe), O (carbonyl oxygen-hydrogen bond acceptor probe), $OH_2$ (water probe), and DRY (hydrophobic probe)[16]. Additional non-MIF-derived descriptors were generated to create a total of 128 descriptors. Volsurf descriptors have been previously used to predict antileishmanial activity of natural products on enzymes and predict activity of some molecules[17-18].

**Models**

Knime 3.4.0 software (KNIME 3.4.0 the Konstanz Information Miner Copyright, 2003-2017, (www.knime.org)[19] was used to perform all of the following analyses. The descriptors and class variables were imported from the Volsurf+ program, v. 1.0.7, and the data were divided using the "Partitioning" node with the "stratified sample" option to create a training set and a test set, encompassing 80% and 20% of the compounds, respectively. Although the compounds were selected randomly, the same proportion of active and inactive samples was maintained in both sets. For internal validation, we employed cross-validation using 10 randomly selected, stratified groups, and the distributions according to activity class variables were found to be maintained in all validation groups and in the training set. Descriptors were selected, and a model was generated using the training set and the Random Forest (RF) algorithm[20], using the WEKA nodes[21]. The parameters selected for RF included the following settings: number of trees to build = 1900, seed for random number generator = 1909501934341. The internal and external performances of the selected models were analyzed for sensitivity (true positive rate, i.e., active rate), specificity (true negative rate, i.e., inactive rate), and accuracy (overall predictability). In addition, the sensitivity and specificity of the Receiver Operating Characteristic (ROC) curve were found to describe the true performance with more clarity and accuracy. The plotted ROC curve shows the

true positive (active) rate either versus the false positive rates or versus sensitivity (1: specificity). In a two-class classification, when a variable that is being investigated cannot be distinguished between the two groups (i.e., when there is no difference between the two distributions), the area under the ROC curve equals 0.5, which is to say that the ROC curve will coincide with the diagonal. When there is a perfect separation of values between two groups (i.e., no overlapping of distributions), the area under the ROC curve equals 1, which is to say that the ROC curve will reach the upper left corner of the plot[22].

**Docking**

The structure of *L. donovani* enzymes Topoisomerase I (TOPI)[23], N-myristoyltransferase (NMT)[24], cyclophilin (Cyp)[24] and O-acetylserine sulfhydrylase (OASS)[25] downloaded from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). 1-33 structures were submitted to molecular docking using the Molegro Virtual Docker, v. 6.0.1 (MVD). All of the water molecules were deleted from the enzyme structure, and the enzyme and compound structures were prepared using the same default parameter settings in the same software package. The docking procedure was performed using a GRID of 15 Å in radius and 0.30 Å in resolution to cover the ligand-binding site of the enzyme's structures. The Moldock score algorithm was used as the score function[26]. For all enzymes the binding site was the same as the ligand present in the pdb file.

## 4. Conclusions

We have conducted a comparative ligand- and structure-based approach using Molegro Virtual Docking and machine learning RF to determine the antileishmanial potential of seven cyclic imides synthesized. In silico study allowed us to suggest that the cyclic imide 27 can be tested as a potential multitarget molecule for leishmanial treatment, presenting activity against four strategic enzymes to treatment with probability of activity of 60%.

**Acknowledgments**

We would like to thank the Post-Graduate Program in Natural and Synthetic Bioactive Products, Federal University of Paraíba.

**Author Contributions**

JAL, CL, NC built database; JAL performed all calculus; and JAL, MTS and LS wrote the paper. All authors read and approved the final manuscript.

**Conflicts of Interest**

The authors declare no conflict of interest.

**References and Notes**

1. J. Alvar, I.D. Vélez, C. Bern, M. Herrero, P. Desjeux, J. Cano, J. Jannin, M. den Boer, W.L.C. Team, Leishmaniasis worldwide and global estimates of its incidence, PloS one, 7 (2012) e35671.
2. K.A. da Franca Rodrigues, C.N. de Sousa Dias, P.L. do Nascimento Néris, J. da Câmara Rocha, M.T. Scotti, L. Scotti, S.R. Mascarenhas, R.C. Veras, I.A. de Medeiros, T.d.S.L. Keesen, 2-Amino-thiophene derivatives present antileishmanial activity mediated by apoptosis and immunomodulation in vitro, European journal of medicinal chemistry, 106 (2015) 1-14.

3. V.I. Bonano, J.K. Yokoyama-Yasunaka, D.C. Miguel, S.A. Jones, J.A. Dodge, S.R. Uliana, Discovery of Synthetic Leishmania Inhibitors by Screening of a 2-Arylbenzothiophene Library, Chemical biology & drug design, 83 (2014) 289-296.
4. C. Herrera Acevedo, L. Scotti, M. Feitosa Alves, M.D.F. Formiga Melo Diniz, M.T. Scotti, Computer-aided drug design using sesquiterpene lactones as sources of new structures with potential activity against infectious neglected diseases, Molecules, 22 (2017) 79.
5. M.A. Lill, M.L. Danielson, Computer-aided drug design platform using PyMOL, Journal of computer-aided molecular design, 25 (2011) 13-19.
6. V. Prates Lorenzo, A. Silvia Suassuna Carneiro Lúcio, L. Scotti, J. Fechine Tavares, M. Barbosa Filho, T. Keesen de Souza Lima, J. da Câmara Rocha, M. Tullius Scotti, Structure-and Ligand-Based Approaches to Evaluate Aporphynic Alkaloids from Annonaceae as Multi-Target Agent Against Leishmania donovani, Current pharmaceutical design, 22 (2016) 5196-5203.
7. V.P. Lorenzo, J.M. Barbosa Filho, L. Scotti, M.T. Scotti, Combined structure-and ligand-based virtual screening to evaluate caulerpin analogs with potential inhibitory activity against monoamine oxidase B, Revista Brasileira de Farmacognosia, 25 (2015) 690-697.
8. V. Cechinel Filho, Principais avanços e perspectivas na área de produtos naturais ativos: estudos desenvolvidos no NIQFAR/Univali, Quim. Nova, 23 (2000) 680-685.
9. V. Cechinel Filho, T. Pinheiro, R. Nunes, R. Yunes, A. Cruz, E. Moretto, Antibacterial activity of N-phenylmaleimides, N-phenylsuccinimides and related compounds. Structure-activity relationships, Farmaco (Societa chimica italiana: 1989), 49 (1994) 675-677.
10. J.A. Yunes, A.A. Cardoso, R.A. Yunes, R. Correa, F. de Campos-Buzzi, V. Cechinel Filho, Antiproliferative effects of a series of cyclic imides on primary endothelial cells and a leukemia cell line, Zeitschrift für Naturforschung C, 63 (2008) 675-680.
11. T. Zawadowski, J. Kossakowski, S. Rump, I. Jakowicz, A. Płaźnik, Synthesis and anxiolytic activity of N-substituted cyclic imides N-[4-[(4-aryl)-1-piperazinyl] alkyl]-5, 7-dioxabicyclo [2.2. 2] octane-2, 3-dicarboximide, Acta poloniae pharmaceutica, 52 (1995) 43-46.
12. A.-M. Alaa, Novel and versatile methodology for synthesis of cyclic imides and evaluation of their cytotoxic, DNA binding, apoptotic inducing activities and molecular modeling study, European journal of medicinal chemistry, 42 (2007) 614-626.
13. C.N. Berthold MR, Dill F, Data analysis, machine learning and applications. In: Preisach C, Burkhardt H, SchimidtThieme L, Decker R, , , Eds. Berlin: Springer 2007.
14. L. Breiman, Random forests, Machine learning, 45 (2001) 5-32.
15. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update, ACM SIGKDD explorations newsletter, 11 (2009) 10-18.
16. J.A. Hanley, B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology, 143 (1982) 29-36.
17. I. García, Y. Fall, G. Gómez, H. González-Díaz, First computational chemistry multi-target model for anti-Alzheimer, anti-parasitic, anti-fungi, and anti-bacterial activity of GSK-3 inhibitors in vitro, in vivo, and in different cellular lines, Molecular diversity, 15 (2011) 561-567.
18. G. Imre, G. Veress, A. Volford, Ö. Farkas, Molecules from the Minkowski space: an approach to building 3D molecular structures, Journal of Molecular Structure: THEOCHEM, 666 (2003) 51-59.
19. G. Cruciani, P. Crivori, P.-A. Carrupt, B. Testa, Molecular fields in quantitative structure–permeation relationships: the VolSurf approach, Journal of Molecular Structure: THEOCHEM, 503 (2000) 17-30.
20. L. Scotti, H. Ishiki, F. Mendonca, M. Da Silva, M. Scotti, In-silico analyses of natural products on leishmania enzyme targets, Mini reviews in medicinal chemistry, 15 (2015) 253-269.
21. L. Scotti, M. Tullius Scotti, Computer aided drug design studies in the discovery of secondary metabolites targeted against age-related neurodegenerative diseases, Current topics in medicinal chemistry, 15 (2015) 2239-2252.
22. D.R. Davies, A. Mushtaq, H. Interthal, J.J. Champoux, W.G. Hol, The structure of the transition state of the heterodimeric topoisomerase I of Leishmania donovani as a vanadate complex with nicked DNA, Journal of molecular biology, 357 (2006) 1202-1210.

23. M.D. Rackham, Z. Yu, J.A. Brannigan, W.P. Heal, D. Paape, K.V. Barker, A.J. Wilkinson, D.F. Smith, R.J. Leatherbarrow, E.W. Tate, Discovery of high affinity inhibitors of Leishmania donovani N-myristoyltransferase, MedChemComm, 6 (2015) 1761-1766.
24. V. Venugopal, A.K. Datta, D. Bhattacharyya, D. Dasgupta, R. Banerjee, Structure of cyclophilin from Leishmania donovani bound to cyclosporin at 2.6 Å resolution: Correlation between structure and thermodynamic data, Acta Crystallographica Section D: Biological Crystallography, 65 (2009) 1187-1195.
25. I. Raj, S. Kumar, S. Gourinath, The narrow active-site cleft of O-acetylserine sulfhydrylase from Leishmania donovani allows complex formation with serine acetyltransferases with a range of C-terminal sequences, Acta Crystallographica Section D: Biological Crystallography, 68 (2012) 909-919.
26. R. Thomsen, M.H. Christensen, MolDock: a new technique for high-accuracy molecular docking, Journal of medicinal chemistry, 49 (2006) 3315-3321.

Notes:
MDPI do not released one specific template for Sciforum conference. Consequentyl, this is not official template released by MDPI. In principle, the papers may be presented without specific format. However, the chairperson and the secretariat of the conference decided to create this template to give homogeneity to Mol2Net works. As is, this template should be used only for Mol2Net conference. Please, do not use this template for MDPI journals. Please, delete these notes before saving your final version. Type of the Paper (Proceeding, Letter, Expert Opinion, Communication, etc.). MDPI generates doi upon author request.
**YouTube link:** this option is only for those authors with welcome videos and/or oral presentations, plenary conferences uploaded to Mol2Net YouTube site.
**Reference list:** *We recommend the use of reference management software to prepare the references list (e.g., Endnote, http://endnote.com/styles/MDPI.ens ).*