JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Application of Computer-Aided Drug Repurposing in the Search of New Cruzipain Inhibitors: Discovery of Amiodarone and Bromocriptine Inhibitory Effects

Carolina L. Bellera,[†] Darío E. Balcazar,[‡] Lucas Alberca,[†] Carlos A. Labriola,[§] Alan Talevi,*[,†] and Carolina Carrillo[‡]

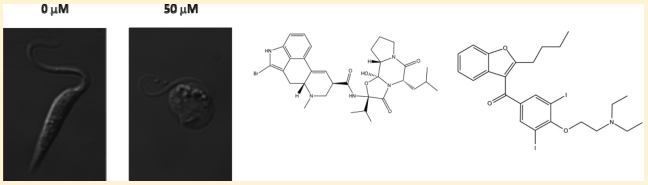
[†]Medicinal Chemistry, Department of Biological Sciences, Faculty of Exact Sciences, National University of La Plata, 47 y 115, La Plata (B1900AJI), Buenos Aires, Argentina

[‡]Instituto de Ciencia y Tecnología Dr. César Milstein (ICT Milstein), Argentinean National Council of Scientific and Technical Research (CONICET), Saladillo 2468, Ciudad Autónoma de Buenos Aires, Argentina

[§]Instituto de Investigaciones Bioquímicas de Buenos Aires, Argentinean National Council of Scientific and Technical Research (CONICET), Instituto es Av. Patricias Argentinas 435, Ciudad Autónoma de Buenos Aires, Argentina

Supporting Information

ABSTRACT: Cruzipain (Cz) is the major cystein protease of the protozoan *Trypanosoma cruzi*, etiological agent of Chagas disease. From a 163 compound data set, a 2D-classifier capable of identifying Cz inhibitors was obtained and applied in a virtual screening campaign on the DrugBank database, which compiles FDA-approved and investigational drugs. Fifty-four approved drugs were selected as candidates, four of which were acquired and tested on Cz and *T. cruzi* epimastigotes. Among them, the antiparkinsonian and antidiabetic drug bromocriptine and the antiarrhythmic amiodarone showed dose-dependent inhibition of Cz and antiproliferative activity on the parasite.



INTRODUCTION

Chagas disease (or American trypanosomiasis) is a tropical parasitic disease caused by the flagellate protozoan *Trypanosoma cruzi*. The *T. cruzi* life cycle includes both vertebrate (among them, man) and invertebrate (hematophagous triatomine bugs) hosts. Around 80−90% of infections in humans occur when the parasite feces come into contact with wounded skin or mucosae.[1] Other infection ways include blood transfusions and congenital transmission. Even though a series of control campaigns (with an emphasis on vector control) undertaken by World Health Organization (WHO), Pan American Health Organization (PAHO), and national authorities have dramatically reduced Chagas disease prevalence in the last 15 years, there are still almost 8 million infected people, 28 million people at risk, and more than 40,000 new cases annually.[2−4]

Current treatment against Chagas relies on only two agents developed during the 1960s and 1970s, namely, nifurtimox and benznidazole, which are not effective in the late chronic phase of the disease and present severe side effects and resistance issues.[5−7] This explains the urgent need of safer and more effective treatments for American trypanosomiasis. It is worth noting, however, that important advances have been made in the field of biochemistry and molecular biology of *T. cruzi* and novel antichagasic therapeutics.[4,8−11] Cystein protease inhibitors are among the most investigated candidates against *T. cruzi*.[11] Cruzipain (Cz), the major cystein protease of the parasite, has been particularly explored as new drug target. Cz is essential for replication of the intracellular form of *T. cruzi* and plays a role in host−parasite interactions.[12] Because it is autocatalytic, it is believed that Cz inhibition produces accumulation of the inactive precursor of the proteinase within the Golgi complex, which eventually leads to osmotic shock and cell death.[13]

Here, we present the development of a 2D classification model from a 163 compound data set that includes both Cz inhibitors and noninhibitors. The model has later been applied in a virtual screening (VS) campaign to explore the DrugBank small molecule database in order to identify novel Cz reversible inhibitors. Four of the selected candidates were acquired and tested both biochemically and biologically, validating the model.

It is worth stating that DrugBank compiles FDA-approved and experimental drugs (including biotech molecules and nutraceuticals).[14,15] It is therefore particularly useful to conduct VS campaigns aimed to drug repurposing (i.e., searching for second or further medical uses of known drugs). The application of VS to chemical libraries compiling known therapeutics can be considered as a form of knowledge-based rational drug repositioning,[16−19] (chemoinformatics- and bioinformatics-based, and others), which has been recently signaled as a relevant strategy to aid discovering novel treatment for rare and neglected conditions.[20−22]

Received: May 13, 2013

   dx.doi.org/10.1021/ci400284v | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

EXPERIMENTAL SECTION

Data Set Compilation and Splitting. A 163 compound balanced data set including 82 Cz reversible inhibitors and 81 noninhibitors was compiled from the literature.[23−35] In order to split the data set into representative training and test sets, the LibraryMCS v0.7 (ChemAxon) hierarchical clustering approach was applied in combination with the *k*-means clustering as implemented in Statistica 10 Cluster Analysis module (Statsoft Inc., 2011). LibraryMCS relies on the maximum common substructure (MCS, i.e., the largest subgraph shared by two chemical graphs) to cluster a set of chemical structures. The algorithm applies a similarity search to the pool of molecules, and the two structures with the highest similarity coefficient are considered more likely to share a large MCS. Once this likely MCS has been established, a substructure search is carried out in order to find the MCS of multiple structures efficiently, without exhaustive pairwise comparison. Certainly, it is possible that the two structures exhibiting the highest similarity coefficient do not share the largest MCS; thus, Library MCS leads to reproducible but approximate solutions.[36] As suggested by Everitt et al.,[37] hierarchical clustering has been applied here to define an initial partition of *n* objects into *g* groups, selecting a smallest common substructure of nine atoms, and the groups of compounds were later optimized by the *k*-means algorithm, minimizing the Euclidean distance to the group centers. A series of descriptors computed with Dragon 4.0 (Milano Chemometrics, 2003) representing different aspects of molecular structure (namely, molecular weight, log P, polar surface area, number of H bonds acceptors, information index of atomic content, and sum of atomic van der Waals volumes) were normalized and applied to calculate such distance. A randomly selected member of each cluster identified through hierarchical clustering has been selected as seed to be used in the *k*-means approach. Once the clusters were separately identified in the inhibitors and noninhibitors classes, 25% of each cluster was assigned to an independent test set for validation purposes, while the remaining 75% of the clusters were retained as training sets for modeling purposes. The structures of both training and test set compounds are provided in the Supporting Information.

Descriptor Calculation and Modeling. A total of 877 molecular 0D-2D molecular descriptors (constitutional and topological descriptors, functional group counts, and atom-centered fragments) were computed with Dragon 4.0 Academic version. The values of such descriptors are conformation-independent, thus being particularly suitable for their application in VS campaigns because no preprocessing of the database structures (e.g., conformational analysis or optimization) is required. From the 877 descriptors, 30 random subsets of no more than 254 descriptors were generated, and these subsets were used as descriptor pools for modeling purposes.

Linear discriminant analysis (LDA) was conducted in order to derive a classification model capable of distinguishing Cz inhibitors from noninhibitors. LDA is a qualitative supervised learning method aimed at finding a linear combination of independent variables to discriminate between two or more categories of objects. Each object class is associated to a given value (an integer value) of an arbitrary variable that serves as class label. In our case, only two object classes (ACTIVE − Cz inhibitors, and INACTIVE − noninhibitors) were considered; thus, the class label assumes two observed values (1 and −1, respectively). Because the output of the function being searched is not a continuous variable but only an object category, LDA and other classificatory techniques may be useful to handle noisy data, e.g., if a given experimental endpoint is associated to a large variability or if experimental data from a diversity of laboratories are compiled.[38]

The discriminant analysis module of Statistica 10 was used to derive the models. A tolerance value of 0.5 was selected in order to exclude highly correlated descriptors from the model. All the coefficients linked to the models descriptors were significant at a 0.05 level. A minimum ratio of 15 between the number of training set compounds and the number of independent variables was used in order to reduce the chances of overfitting. The Parsimony principle, Wilks' lambda, and performance of the model on the test set were used to select the best model. Standard validation approaches (leave-group out cross-validation, Fisher's randomization test, and external validation) were used to assess the model's robustness and predictive ability.[39] Stratified 24-fold cross-validation and 55 randomization tests were applied.

Virtual Screening. DrugBank 3.0 was used for VS. Only approved and experimental small molecules and nutraceuticals (6684 total compounds) were considered (biotech drugs were excluded a priori). PDD and separate receiver operating characteristic (ROC) curves were constructed for both the training and test sets, in order to select the discrimant function (score) threshold value to be used in the VS campaign.[40] To build the ROC curves, the MedCalc ROC curves analysis tool was used (Medcalc software, 2012).

On the basis of the ROC curves analysis, 0.29 was defined as the cutoff value to differentiate active from inactive compounds in the VS campaign. According to the ROC curves data, this corresponds to a sensitivity of 67% and a specificity of 95% in the training set and a sensitivity of 75% and a specificity of 100% in the test set. Considering the training set values, this corresponds to a positivity predictive value of 0.41 for a database whose yield of active compounds is 0.05 (see the Results section for further discussion).

Finally, the leverage approach was used to define whether a given prediction belonged or not to the model's applicability domain.[41] Briefly, the leverage for a compound *i* is defined as $h_i = x_i^T (X^T X)^{-1} x_i$, where x_i is the descriptor vector for compound *i* and *X* is the model matrix derived from the training set descriptor values. The warning leverage was fixed at $3k/n$, *k* being the number of model parameters and *n* the number of training set compounds.

Inhibitory Effect on Cz Activity Assay. Four candidates from the VS were acquired, and their ability to inhibit Cz was assessed. Amiodarone hydrochloride was a kind gift from Vannier Laboratories. Escitalopram oxalate was a kind gift from Bagó Laboratories. Bromocriptine and colchicine were acquired from Saporiti.

Parasite cultures were harvested, washed, and incubated with lyses buffer (Hepes 50 mM, pH 7.4, NaCl 200 mM, NP-40 1%). Cz activity was assayed with Bz-Pro-Phe-Arg-pNA (Bz-PFR-pNa, Sigma) as substrate, as described previously.[42] Briefly, a diluted aliquot from the cell extract was incubated in a buffer of 50 mM Tris-HCl, pH 7.6, 5 μM dithiothreitol (DTT), and 250 μM Bz-Pro-Phe-Arg-pNA (Sigma), in the presence or absence of the indicated compounds. The reaction was measured spectrophotometrically at room temperature at 410 nm for 5 min (Beckman CoulterTM DU530 Life Science UV−vis spectrophotometer.). The values obtained were converted into pmol of hydrolyzed substrate per min by

dx.doi.org/10.1021/ci400284v | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

using the extinction coefficients 8.800 M^{-1} cm^{-1} (p-nitroanilides). The inhibitory effect of the selected candidates was expressed as a percentage of residual activity of Cz respect of the assay without inhibitors.

To confirm the specificity of the observed effect of amiodarone and bromocriptine on Cz activity, the enzyme was partially purified by ammonium sulfate precipitation followed by affinity column chromatography on concavalin A-sepharosa (Sigma), as previously described.[43] The activity of the partially purified Cz was assayed using increasing amiodarone or bromocriptine concentrations.

Inhibitory Effects on Growth Curves of *T. cruzi* Epimastigotes. Epimastigotes of the *T. cruzi* strain Y were cultured at 28 °C in BHT medium with 20 mg/L Haemin, 20% heat-inactivated fetal calf serum, and antibiotics (100 μg/mL streptomycin and 100 U/mL penicillin),[44] adding the indicating amiodarone or bromocriptine concentration (0−200uM). Cultures were initiated at 10^6 cells/mL, and the proliferation was followed daily by cell counting in a hemocytometer chamber.

For microscopy, freshly grown trypanosome samples were washed twice in PBS. After letting the cells settle for 30 min at room temperature in poly-L-lysine-coated coverslips, parasites were fixed at room temperature for 20 min with 1% formaldehyde in PBS. Slides were mounted using Vectashield (Vector Laboratories). Cells were observed in an Olympus BX51 fluorescence microscope. Images were recorded with an Olympus XM10 camera. The percent of rounded parasites was determined using an Zeiss Axiovert 25 microscope.

RESULTS

Clustering procedure was applied on the data set using a combination of maximum common structure (MCS) hierarchical clustering and optimization by *k*-means clustering (see Experimental Section for details). Such a procedure revealed five groups of at least six compounds in the ACTIVE category and seven groups of at least seven compounds in the INACTIVE class. According to MCS clustering, there are four compounds in the ACTIVE class that can be considered outliers (they are clustered alone or in groups of only two compounds, meaning they have no MCS above the specified number of atoms), whereas the INACTIVE category presents 14 outliers. On the basis of the clustering procedure, 25% of each cluster was assigned to the test set for external validation purposes, while the remaining 75% of each cluster was assigned to the training set upon which the model was derived.

The following model was obtained through Linear Discriminant Analysis (LDA):

$$\text{Class} = -0.959 + 1.764 \times \text{GGI7} - 0.287 \times n\text{S} + 0.280 \times n\text{CN}$$

$$\text{Wilks}'\lambda = 0.56 \quad F(3, 117) = 31.24 \quad p < 0.0000$$

$$N = 121$$

where GGI7 represents Galvez topological charge index of order 7, *n*S represents the number of sulfur atoms, and *n*CN represents the number of aliphatic nitriles. The magnitudes of the beta coefficients of such descriptors are, in that order, 0.583, 0.183, 0.136, showing that GGI7 is the more relevant independent variable of the model. It should be highlighted that the model presents an excellent case per predictor ratio (above 40), which indicates a very low chance of overfitting, as confirmed later in the external validation results. Furthermore, the pairwise correlation between the descriptors included in the model is negligible, 0.147 being the correlation coefficient between the most correlated predictors. When using 0 as a score threshold to differentiate active from inactive compounds, the model presents 74% of good classifications among the training set inactive compounds, 80% of good classifications among the training set active compounds, and an overall of 77% good classifications. Regarding the test set, the model accurately classifies 81% of the active and 90.5% of the inactive compounds, with an overall good classification of 86%. These results seem to confirm that no overfitting has occurred because the performance on the test set is similar to (and in fact better than) the performance on the training set. The average performance of the randomized models was 63.6% (sd = 3.5) showing that the randomized models were significantly outperformed by the real model, as expected. The 24-fold cross-validation resulted in an average percentage of good classifications of 77.6%, which is very similar to the performance of the original model on the training set.

We resorted to pharmacological distribution diagrams (PDD) and ROC curves in order to optimize the chosen threshold score on a rational basis.[40,45] Figures 1 and 2 present, respectively, the PDD and training set ROC curve. The area under the curve (AUC) for the training and test sets ROC

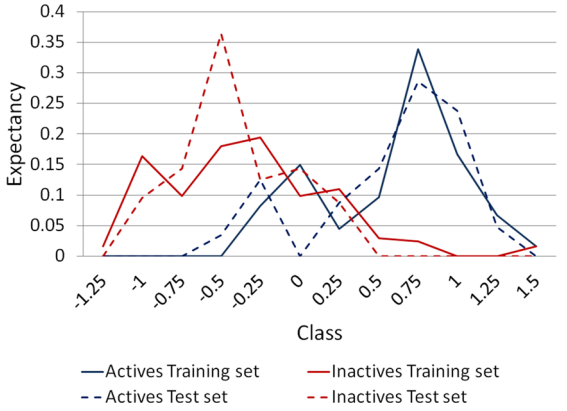


Figure 1. PDD showing the distribution of training and test set active and inactive compounds along the score values of the model. A fine superposition between the training and test set is observed.

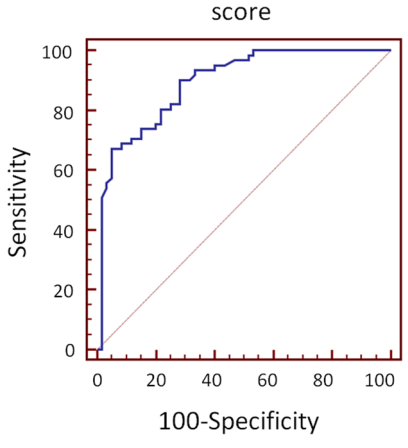


Figure 2. Training set ROC curve.

dx.doi.org/10.1021/ci400284v | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

curves were, respectively, 0.893 and 0.921 (1 represents perfect classification, while 0.5 represents random classification). Note that the distribution of active and inactive compounds in relation to the model score are almost identical (Figure 1). It is also interesting to underline that six of the misclassified compounds (four actives and two inactives) correspond to the ones identified as outliers during the hierarchical clustering procedure (that is, compounds which do not share a nine atom MCS with the rest). On the basis of the results, 0.29 was defined as the cutoff value to differentiate active from inactive compounds in the VS campaign. According to the ROC curves data, this corresponds to a sensitivity of 67% and a specificity of 95% in the training set and a sensitivity of 75% and a specificity of 100% in the test set. As stated by Triballeau in the original application of ROC curves to VS, the selection of a given balance between sensitivity and specificity is not a statistical matter but a context-dependent decision. In our case, due to a limited budget to acquire and test compounds, we have prioritized specificity (i.e., reducing the chance of false positives) over sensitivity. This means that potentially valuable active scaffolds will be lost during the screening in order to reduce the chance of sending an inactive compound (false positive) to experimental testing.

From 6684 small approved and investigational molecules of the DrugBank 3.0 database, 256 candidates belonging to the model's applicability domain presented a model score above the selected threshold; 54 of them correspond to approved drugs, which are the straightforward candidates for repositioning purposes. On the basis of their accessibility, four of them (Figure 3) were acquired and experimentally tested in

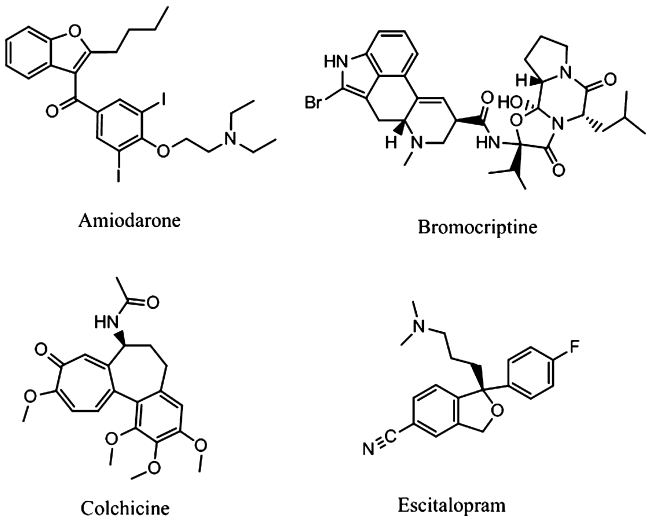


Figure 3. Molecular structures of the four candidates selected for enzymatic testing.

enzymatic assay on Cz crude extract. The acquired candidates were amiodarone (currently marketed as antiarrhythmic), bromocriptine (approved as antiparkinsonian and antidiabetic), colchicine (gout treatment), and escitalopram (antidepressant). Figure 4 shows the effect of 100 μM solutions of the four candidates on the Cz activity from *T. cruzi* crude extracts. Amiodarone and bromocriptine showed a significant inhibitory effect on Cz activity from cell extracts. Such inhibition proved to be dose-dependent on purified Cz (Figure 5), with a median inhibitory concentration (IC_{50}) approximately of 219.8 μM for amiodarone and 84.2 μM for bromocriptine

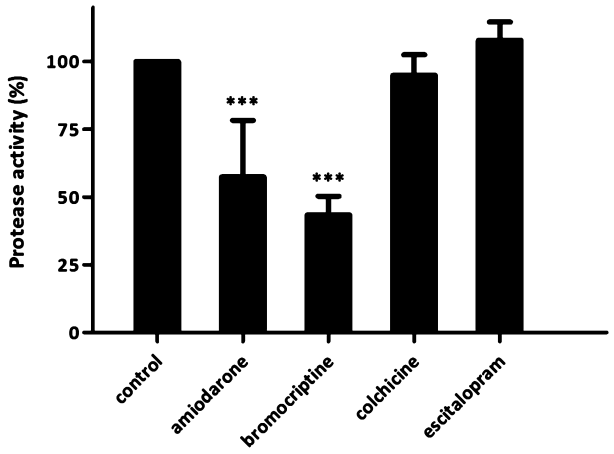


Figure 4. Inhibitory effect of the four selected candidates on Cz activity from *T. cruzi* crude extracts. The final concentration of each compound was 100 μM. Protease activity is expressed as percentage of the control condition (2% DMSO). Results represent the mean of three independent experiments. Asterisks indicate significant differences respect of the control (*** $p < 0.005$).

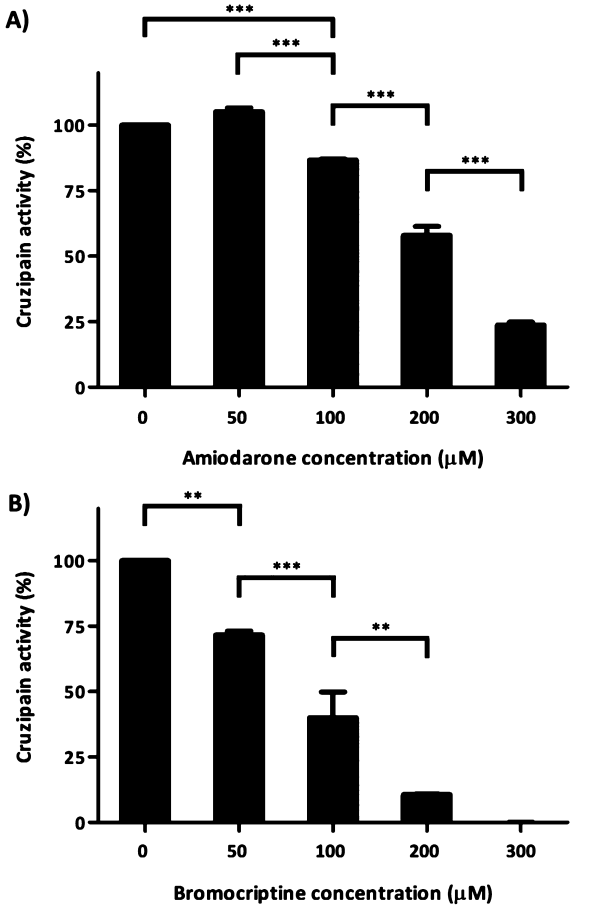


Figure 5. Dose-dependent inhibitory effect of amiodarone (A) and bromocriptine (B) on purified Cz activity. Both candidates were assayed in a concentration range of 0–300 μM. Remanent Cz activity was expressed as a percentage of the control (0 μM compound, 2% DMSO). Results represent the mean of two independent experiments. Asterisks indicate significant differences (** $p < 0.01$, *** $p < 0.005$).

Both candidates also showed a notorious effect on *T. cruzi* epimastigotes proliferation (Figure 6) presenting values of

dx.doi.org/10.1021/ci400284v | J. Chem. Inf. Model. XXXX, XXX, XXX–XXX

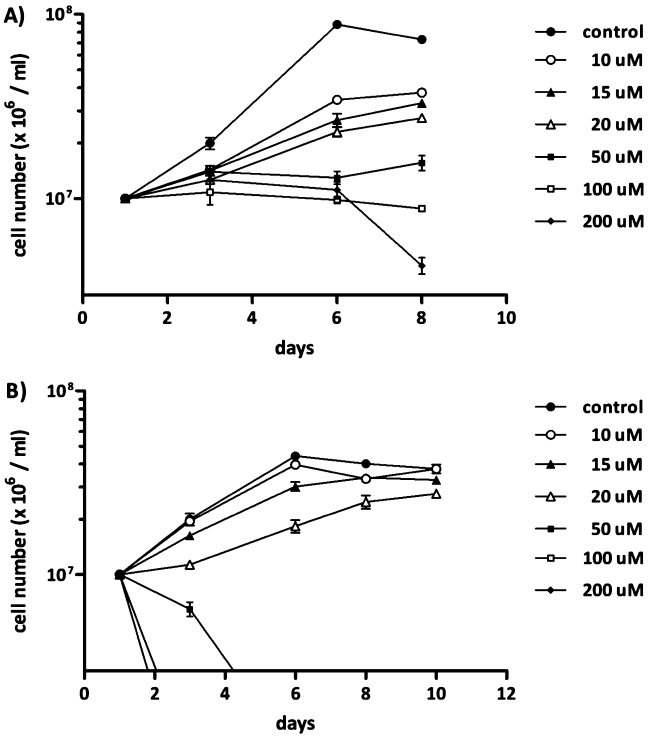


Figure 6. Effects on *T. cruzi* epimastigotes proliferation of (A) amiodarone and (B) bromocriptine. To determine the growth rate, 10^7 cells/mL were seeded in BHT medium and maintained at 28 °C for 10 or 8 days, respectively. The control condition was done with 2% DMSO. Parasites were counted using a hemocytometer chamber. Results represent the mean ± SD of a representative experiment.

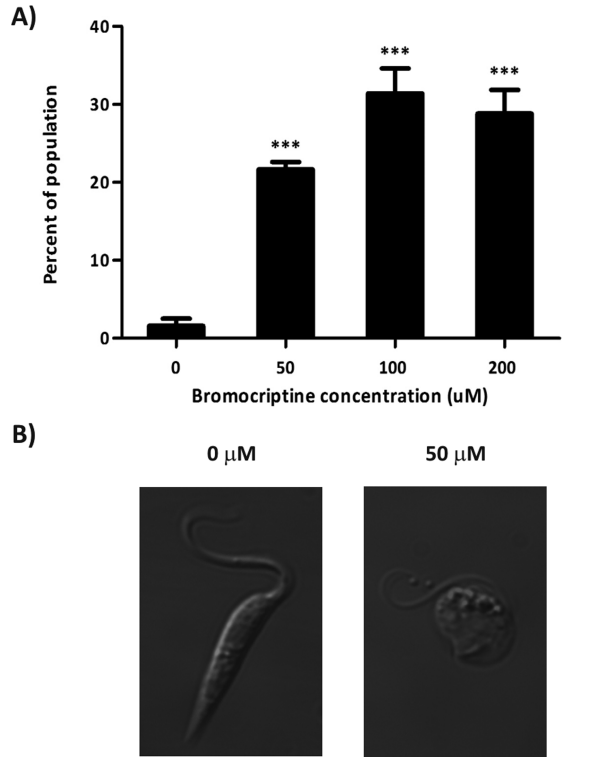


Figure 7. Effect of bromocriptine on *T. cruzi* morphology. (A) Quantitation of morphological changes in cells that have been cultured in the absence or presence of bromocriptine (50−200 μM) for four days. At least 100 cells were counted for each condition classifying in normal and rounded parasites. Results represent the means of the two independent experiments. (*** $p < 0.005$, respect to 0uM). (B) Differential interference contrast (DIC) photos of cells incubated for four days in the presence (normal morphology) or absence (rounded morphology) of 50 μM bromocriptine. Results represent images of an experiment representative.

median inhibitory dose (ID_{50}) around 22 μM for amiodarone and 15 μM for bromocriptine at the middle log phase of controls (4th day) . The fact that amiodarone presents higher effects on epimastigotes culture than on Cz activity could be explained by a pleiotropic effect of the drug, as it has been previously shown on parasites Ca^{2+} homeostasis and on ergosterol biosynthesis.[46] Even though, it should be highlighted that this is the first time that the antitrypanosomal effect of amiodarone, at least partially, is related to Cz activity.

Bromocriptine showed a significant effect on parasite morphology (Figure 7) as well as amiodarone (not shown).

DISCUSION AND CONCLUSIONS

A three-descriptor 2D classification model was derived from a 163 compound data set, which compiled Cz inhibitors and noninhibitors extracted from literature. The model presented an excellent case to descriptor ratio and similar performance on both the training and the test sets, which suggest good predictive ability and absence of overfitting. Because only conformation-independent descriptors were included in the model, it is particularly suitable for efficient exploration of drug libraries through VS campaigns without requiring any preprocessing of the library structures.

Having in mind the potential of knowledge-based drug repositioning to develop novel drugs for neglected and rare diseases, the model was applied in a VS campaign to select potential antichagasic drugs from the DrugBank database, which compiles approved and investigational active ingredients. PDD and ROC curve analysis were conducted in order to select a score cutoff value to differentiate active and inactive agents on a rational basis.

Four candidates were acquired and experimentally tested in enzymatic and inhibitory assays. Among them, amiodarone (approved as antiarrhythmic) and bromocriptine (traditionally used against Parkinson and more recently repurposed for the treatment of diabetes) showed a weak but dose-dependent inhibition on Cz activity with clear effects on *T. cruzi* proliferation and morphology. The results illustrate the possibilities of computer-aided drug repositioning in the search of novel medications for neglected diseases.

It should be mentioned that two of the four candidates showed no activity or almost no activity on the enzyme, thus suggesting that either the model should be optimized or a stricter cutoff value should be selected to reduce the false positive rate (e.g., score above 0.5). As discussed by Triballeau et al. in the original application of ROC curves to VS campaigns,[24] the probability of finding an active compound critically depends on the yield of active compounds of the screened database. The positivity predictive value, which represents the probability that a given candidate with a score above the selected threshold will be actually active, can be calculated through the following expression

$$PPV = Se \times Ya/[(Se \times Ya) + (1 - Sp) \times (1 - Ya)]$$

dx.doi.org/10.1021/ci400284v | J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

where *Se* represents the sensitivity of the model (or true positives rate) and *Sp* represents the specificity of the model (or true negatives rate). *Ya* symbolizes the yield of actives, that is, the number of active drugs in the screened database divided by the total number of compounds in the database. *Ya* is not known in real VS applications. The influence of *Ya* on the probability that a given selected compound is actually active means that even a highly specific and sensitive model may retrieve a high proportion of false positives from a screened library if the yield of actives in the library is low. As has been recently pointed out by Scior et al., the confirmed hit rate in VS real-world applications is often very low, between 0.01% and 0.14%.[25] It is noteworthy that if the selected cutoff value was 0.5, only bromocriptine among the four compounds would have overcome the screening and thus would have been experimentally tested.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Compounds that compose the training and test sets. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: atalevi@biol.unlp.edu.ar, alantalevi@gmail.com.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

C. L. Bellera and D. Balcazar are fellowship holders from the National Council of Scientific and Technical Research (CONICET). C. Carrillo, C. Labriola, and A. Talevi are members of CONICET. We are thankful to Bagó and Vannier for kindly providing samples of escitalopram and amiodarone, respectively. This work was supported by the Faculty of Exact Sciences of the National University of la Plata, the National University of La Plata (UNLP), the Commission of Scientific Research of Buenos Aires Province (CIC) (PIRPS Projects and Incentivos UNLP), UBACyT - GF 20020100200029, University of Buenos Aires, and PIP 2010 0685 from the National Council of Scientific and Technical Research (CONICET).

■ ABBREVIATIONS

LDA, linear discriminant analysis; MCS, maximum common substructure; ROC, receiver operating characteristic; VS, virtual screening

■ REFERENCES

(1) Rodrigues Coura, J.; de Castro, S. L. A critical review on Chagas disease chemotherapy. *Mem. Inst. Oswaldo Cruz* **2002**, *97*, 3−24.

(2) Moncayo, A.; Silveira, A. C. Current epidemiological trends for Chagas disease in Latin America and future challenges in epidemiology, surveillance and health policy. *Mem. Inst. Oswaldo Cruz* **2009**, *104* (Suppl. 1), 17−30.

(3) *Control of Chagas Disease*; Second report of the WHO Expert Committee; World Health Organization: Geneva , 2002.

(4) *Research Priorities for Chagas Disease, Human African trypanosomiasis and leishmaniasis*; TDR (Special Programme for Research and Training in Tropical Diseases); World Health Organization: Geneva, 2012.

(5) Urbina, J. A.; Docampo, R. Specific chemotherapy of Chagas disease: Controversies and advances. *Trends Parasitol.* **2003**, *19*, 495−501.

(6) Rodrigues Coura, J.; de Castro, S. L. A critical review on Chagas disease chemotherapy. *Mem. Inst. Oswaldo Cruz* **2002**, *97*, 3−24.

(7) Astelbauer, F.; Walochnik, J. Antiprotozoal compounds: State of the art and new developments. *Int. J. Antimicrob. Agents* **2011**, *38*, 118−124.

(8) Villalta, F.; Scharfstein, J.; Ashton, A. W.; Tyler, K. M.; Guan, F.; Mukherjee, S.; Lima, M. F.; Alvarez, S.; Weiss, L. M.; Huang, H.; Machado, F. S.; Tanowitz, H. B. Perspectives on the *Trypanosoma cruzi*−host cell receptor interactions. *Parasitol. Res.* **2009**, *104*, 1251−1260.

(9) Cuervo, P.; Domont, G. B.; De Jesus, J. B. Proteomics of trypanosomatids of human medical importance. *J. Proteomics* **2010**, *73*, 845−867.

(10) Sánchez-Sancho, F.; Campillo, N. E.; Páez, J. A. Chagas disease: Progress and new perspectives. *Curr. Med. Chem.* **2010**, *17*, 423−452.

(11) Duschak, V. G.; Couto, A. S. An insight on targets and patented drugs for chemotherapy of Chagas disease. *Recent Pat. Anti-Infect. Drug Discovery* **2007**, *2*, 19−51.

(12) Duschak, V. G.; Couto, A. S. Cruzipain, the major cysteine protease for vaccine development and drug target. A review. *Curr. Med. Chem.* **2009**, *16*, 3174−3202.

(13) Rodrigues, G. C.; Aguiar, A. P.; da Silva Gonçalves Vianez, J. L; Macrae, A.; de Melo, A. C. N.; Vermelho, A. B. Peptidase inhibitors as a possible therapeutic strategy for Chagas disease. *Curr. Enz. Inhib.* **2010**, *6*, 183−194.

(14) Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V.; Djoumbou, Y.; Eisner, R.; Guo, A. C.; Wishart, D. S. DrugBank 3.0: a comprehensive resource for 'Omics' research on drugs. *Nucelic Acids Res.* **2011**, *39*, D1035−D1041.

(15) Wishart, D. S.; Knox, C.; Guo, A. C.; Stothard, S. P.; Chang, Z.; Woolsey, J. DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acid Res.* **2006**, *34*, D668−D672.

(16) Glan, F.; Cao, B.; Wu, D.; Chen, Z.; Hou, T.; Mao, X. Exploring old drugs for the treatment of hematological malignancies. *Curr. Med. Chem.* **2011**, *18*, 1509−1514.

(17) Talevi, A.; Castro, E. A.; Bruno-Blanch, L. E. Virtual screening: An Emergent, Key Methodology for Drug Development in an Emergent Continent − A Bridge Towards Patentability. In *Advances Methods and Applications in Chmoinformatics: Research Progress and New Applications*, 1st ed.; Castro, E. A., Haghi, A. K., Eds.; IGI Global: Hershey, PA, 2011, 229−246.

(18) Deftereos, S. N.; Andronis, C.; Friedla, E. J.; Persidis, A.; Persidis, A. Drug repurposing and adverse event prediction using high-throughput literature analysis. *Wiley Interdiscip. Rev. Sys. Biol.* **2011**, *3*, 323−334.

(19) Lussier, Y. A.; Chen, J. L. The emergence of genome-based drug repositioning. *Sci. Transl. Med.* **2011**, *3*, 96ps35.

(20) Ekins, S.; Williams, A. J.; Krasowski, M. D.; Freundlich, J. S. In silico repositioning of approved drugs for rare and neglected diseases. *Drug Discovery Today* **2011**, *16*, 298−310.

(21) Sardana, A.; Zhu, C.; Zhang, M.; Gudivada, R. C.; Yang, L.; Jegga, A. G. Drug repositioning for orphan diseases. *Brief. Bioinform.* **2011**, *12*, 346−356.

(22) Pollastri, M. P.; Campbell, R. K. Target repurposing for neglected diseases. *Future Med. Chem.* **2011**, *3*, 1307−1315.

(23) Bender, A.; Tresadern, G.; Medina-Franco, J. L.; Martínez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D. K. Recognizing pitfalls in virtual screening: a critical review. *J. Chem. Inf. Model.* **2011**, *52*, 867−881.

(24) dos Santos Filho, J. M.; Lima Leite, A. C.; Galdino de Oliveira, B.; Magalhães Moreira, D. R.; Lima, M. S.; Pereira Soares, M. B.; Leite, L. F. C. C. Design, synthesis and cruzain docking of 3-(4-substituted-aryl)-1,2,4-oxadiazole-N-acylhydrazones as anti-*Trypanosoma cruzi* agents. *Bioorg. Med. Chem.* **2009**, *17*, 6682−6691.

(25) Zaldini Hernandes, M.; Montenegro Rabello, M.; Lima Leite, A. C.; Oliveira Cardoso, M. V.; Magalhães Moreira, D. R.; Brondani, D. J.; Simone, C. A.; Campos Reis, L.; Assis Souza, M.; Alves Pereira, V. R.; Salgado Ferreira, R.; McKerrow, J. H. Studies toward the structural optimization of novel thiazolylhydrazone-based potent antitrypanosomal agents. *Bioorg. Med. Chem.* **2010**, *18*, 7826−7835.
(26) Rodrigues, C. R.; Flaherty, T. M.; Springer, C.; McKerrow, J. H.; Cohen, F. E. CoMFA and HQSAR of acylhydrazide cruzain inhibitors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1537−1541.
(27) Huang, L.; Brinen, L. S.; Ellman, J. A. Crystal structures of reversible ketone-based inhibitors of the cysteine protease cruzain. *Bioorg. Med. Chem.* **2003**, *11*, 21−29.
(28) Choe, Y.; Brinen, L. S.; Price, M. S.; Engel, J. C.; Lange, M.; Grisostomi, C.; Weston, S. G.; Pallai, P. V.; Cheng, H.; Hardy, L. W.; Hartsough, D. S.; McKain, M.; Tilton, R. F.; Baldino, C. M.; Craik, C. S. Development of α-keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease. *Bioorg. Med. Chem.* **2005**, *13*, 2141−2156.
(29) Beaulieu, C.; Isabel, E.; Fortier, A.; Massé, F.; Mellon, C.; Méthot, N.; Ndao, M.; Nicoll-Griffith, D.; Lee, D.; Park, H.; Black, W. C. Identification of potent and reversible Cruzipain inhibitors for the treatment of Chagas disease. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 7444−7449.
(30) Du, X.; Hansell, E.; Engel, J. C.; Caffrey, C. R.; Cohen, F. E.; McKerrow, J. H. Aryl ureas represent a new class of anti-trypanosomal agents. *Chem. Biol.* **2000**, *7*, 733−742.
(31) Brak, K.; Doyle, P. S.; McKerrow, J. H.; Ellman, J. A. Identification of a new class of nonpeptidic inhibitors of cruzain. *J. Am. Chem. Soc.* **2008**, *130*, 6404−6410.
(32) Mott, B. T.; Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Kean-Ang, K. K.; Leister, W.; Shen, M.; Silveira, J. T.; Doyle, P. S.; Arkin, M. R.; McKerrow, J. H.; Inglese, J.; Austin, C. P.; Thomas, C. J.; Shoichet, B. K.; Maloney, D. J. Identification and optimization of inhibitors of trypanosomal cysteine protease: Cruzain, rhodesain, and TbCatB. *J. Med. Chem.* **2010**, *53*, 52−60.
(33) Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Eidam, O.; Mott, B. T.; Keiser, M. J.; McKerrow, J. H.; Maloney, D. J.; Irwin, J. J.; Shoichet, B. K. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. *J. Med. Chem.* **2010**, *53*, 4891−4905.
(34) Ang, K. K. H.; Ratnam, J.; Gut, J.; Legac, J.; Hansell, E.; Mackey, Z. B.; Skrzypczynska, K. M.; Debnath, A.; Engel, J. C.; Rosenthal, P. J.; McKerrow, J. H.; Arkin, M. R.; Renslo, A. R. Mining a cathepsin inhibitor library for new antiparasitic drug leads. *PLoS Neglected Trop. Dis.* **2011**, *5*, e1023.
(35) Scheidt, K. A.; Roush, W. R.; McKerrow, J. H.; Selzer, P. M.; Hansell, E.; Rosenthal, P. J. Structure-based design, synthesis and evaluation of conformationally constrained cysteine protease inhibitors. *Bioorg. Med. Chem.* **1998**, *6*, 2477−2494.
(36) Hariharan, R.; Janakiraman, A.; Nilakantan, R.; Singh, B.; Varghese, S.; Landrum, G.; Schuffenhauer, A. MultiMCS: A fast algorithm for the maximon common substructure problem on multiple molecules. *J. Chem. Inf. Model.* **2011**, *51*, 788−806.
(37) Everitt, B. S.; Landau, S.; Leese, M.; Stahl, D. Optimization clustering techniques. In *Cluster Analysis*, 5th ed.; Balding, D. J., Cressie, N. A. C., Fitzmaurice, G. M., Goldstein, H., Molenberghs, G., Scott, D. W., Smith, A. F. M., Tsay, R. S., Weisberg, S., Eds.; John Wiley & Sons: Chichester, 2011; pp 111−142.
(38) Talevi, A.; Bellera, C. L.; Di Ianni, M. E.; Duchowicz, P. R.; Bruno-Blanch, L. E.; Castro, E. A. An integrated drug development approach applying topological descriptors. *Curr. Comput.-Aided Drug Des.* **2012**, *8*, 172−181.
(39) Yasri, A.; Hartsough, D. Toward an optimal procedure for variable selection and QSAR model building. *J. Chem. Inf. Comp. Sci.* **2001**, *41*, 1218−1227.
(40) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J. P.; Bertrand, H. O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J. Med. Chem.* **2005**, *48*, 2534−2547.
(41) Tropsha, A.; Gramatica, P.; Gombar, V. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci.* **2003**, *22*, 69−77.
(42) Cazzulo, J. J.; Cazzulo Franke, M. C.; Martinez, J.; Franke de Cazzulo, B. M. Some kinetic properties of a cysteine proteinase (Cruzipain) from *Trypanosoma cruzi*. *Biochim. Biophys. Acta* **1990**, *1037*, 186−191.
(43) Labriola, C.; Sousa, M.; Cazzulo, J. J. Purification of the major cysteine proteinase (Cruzipain) from *Trypanosoma cruzi* by affinity chromatography. *J. Biol Res.* **1993**, *26* (1−2), 101−7.
(44) Cazzulo, J. J.; Franke de Cazzulo, B. M.; Engel, J. C.; Cannata, J. J. End products and enzyme levels of aerobic glucose fermentation in trypanosomatids. *Mol. Biochem. Parasitol.* **1985**, *16*, 329−343.
(45) Gálvez, J.; García-Domenech, R.; Alapont de Gregorio, C.; de Julián-Ortiz, J. V.; Popa, L. Pharmacological distribution diagrams: A tool for de novo drug design. *J. Mol. Graph.* **1996**, *14*, 272−276.
(46) Benaim, G.; Sanders, J. M.; García-Marchán, Y.; Colina, C.; Lira, R.; Caldera, A. R.; Payares, G.; Sanoja, C.; Burgos, J. M.; Leon-Rossell, A.; Concepción, J. L.; Schijman, A. G.; Levin, M.; Oldfield, E.; Urbina, J. A. Amiodarone has intrinsic anti-*Trypanosoma cruzi* activity and acts synergistically with posaconazole. *J. Med. Chem.* **2006**, *49*, 892−899.