SciForum
Mol2Net

Fragment-Based Approach for Affinity and Selectivity of *Pf*dUTPase inhibitors: Insights for Design of New Anti-Malarial Agents

Marília N. Nascimento, Marina R. Martins, Rodolpho C. Braga, Bruno J. Neves, Vinícius M. Alves and Carolina H. Andrade *

Labmol – Laboratory for Molecular Modeling and Drug Design, Faculty of Pharmacy, Federal University of Goias, Goiania, Goiás, 74605-170, Brazil.

*Author to whom correspondence should be addressed; E-Mail: carolina@ufg.br.
Tel: + 55 62 3209-6451; Fax: + 55 62 3209-6037.

Published: 4 December 2015

Abstract: Malaria is one of the leading causes of death by infectious disease worldwide. The widespread of resistance of *Plasmodium falciparum* to the current antimalarial drugs makes urgent the search and discovery of new targets and new drugs. A potential target for the development of new antimalarial drugs is deoxyuridine triphosphatase (dUTPase), and it has been validated for other organisms such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Mycobacterium smegmatis*. This enzyme plays an important role in maintaining the balance between 2'-deoxyuridine 5'-triphosphate (dUTP) and 2'-deoxythymidine 5'-triphosphate (dTTP), in order to avoid the erroneous incorporation uracil in the DNA tape. In this study, we developed robust conformation-independent fragment-based quantitative structure–activity (QSAR) and structure–selectivity relationship (QSSR) models for a series of β-branched acyclic nucleotides inhibitors of *Plasmodium* and human dUTPase, aiming to design new antimalarial agents. The Hologram QSAR and QSSR models generated showed good robustness and external predictability, and is capable of predict affinity and selectivity of untested compounds inside the applicability domain. Therefore, the generated models can be used in virtual screening campaigns in the search of new potent and selective *Pf*dUTPase inhibitors.

Keywords: malaria; *Plasmodium falciparum*; drug design; dUTPase; QSAR; QSSR

1. Introduction

About 3.2 billion of people in 97 countries and territories are under the risk to contract and develop malaria. It is estimated that 198 million cases of malaria occurred in 2013, and the disease caused 584,000 deaths. 90% of cases of malaria deaths occur in Africa, and of those, 78% in

children below 5 years-old[1]. Recently, notable reduction in cases of malaria have been achieved by vector control, proper case management, and combination of drugs, but the emergence of drug resistant *Plasmodium falciparum* creates a frequent need to continue the search for new antimalarial drugs [2].

The enzyme 2'-deoxyuridine 5'-triphosphate nucleotide hydrolase (dUTPase) is responsible for the hydrolytic cleavage of dUTP (deoxyuridine triphosphate) in dUMP (deoxyuridine monophosphate) and pyrophosphate [3]. Although the enzyme DNA glycosylase can excise the uracil base of DNA, many repairs can destabilize the DNA strand resulting in the breaking of the tape, that is fatal to the cell[4]. Because of that, dUTPase is a potential target for development of new drugs, and there are experimental findings that dUTPase is essential for some organisms such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Mycobacterium smegmatis* [5–7]. Moreover, the *P. falciparum* enzyme (*Pf*dUTPase) has relatively low sequence similarity with human ortholog (*Hs*dUTPase) (28,4% identity) [4], making it an attractive target for the development of selective inhibitors as potential new antimalarial drugs.

The aim of this work was to develop robust conformation-independent fragment-based quantitative structure–activity (QSAR) and structure–selectivity relationship (QSSR) models for a series of β-branched acyclic nucleotides inhibitors of *Plasmodium* and human dUTPase, in order to obtain structural information on the requirements for affinity and selectivity of the enzyme *Pf*dUTPase and to design new antimalarial agents.

2. Results and Discussion

The fragment-based QSAR and QSSR models were derived for a series 127 inhibitors of *Pf*dUTPase and 47 inhibitors with biological data for both enzymes *Pf*dUTPase and *Hs*dUTPase (as measured by K_i) were determined under the same experimental conditions [4,8–13]. The K_i values were converted to the corresponding pK_i ($-\log K_i$) and used as dependent variables in the QSAR investigations. The selectivity parameter (S) was defined as the ratio of the binding affinities of the *Plasmodium* and human enzyme (*Pf*dUTPase K_i / *Hs*dUTPase K_i), whose values were then converted to the corresponding log S. The generation of reliable QSAR models is dependent on the creation of appropriate modeling and evaluation sets. Therefore, we used a modified Kennard-Stone (KS-MD) algorithm to guide an appropriate compound selection in such a way that structurally diverse molecules, possessing activities of a wide range, were included in both sets.

Table 1 shows the results for the best HQSAR and HQSSR models developed for affinity and selectivity for *Pf*dUTPase, with different combinations of atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA) as fragment distinctions for obtaining the molecular fragments. According to Table 1, the best statistical results for HQSAR were obtained for the combination A/B fragment distinction (q^2 = 0.80; r^2 = 0.89). For HQSSR, the best model was obtained using C/H/Ch/DA as fragment distinction (q^2 = 0.79; r^2 = 0.98). Moreover, the best HQSAR and HQSSR models showed substantial predictive power with values of Q^2_{ext} equal to 0.79 and 0.83, respectively. These values indicate the reliability of the models in predicting the affinity and selectivity of untested compounds. We have also performed an analysis of the modified correlation coefficient (r^2_m). This method is considered a parameter for an

additional external validation, because it is based on the actual difference between the predicted and experimental values, and is not influenced by the division of modeling and evaluation sets [14–16]. All r_m^2 metrics for our models were within the recommended range values, confirming the robustness of the models.

Besides predicting the biological property of untested molecules, HQSAR models should also provide important hints as to what molecular fragments are directly related to the biological property. This can be achieved through a careful interpretation of the structural fragments incorporated to the hologram-based QSAR models. The contribution maps show color scales that indicate the magnitude of the contribution of each atom/fragment. The colors next to red end of the spectrum (red, orange and reddish orange) indicate unfavorable or negative contributions, while colors in the green region (yellow, green-blue and green) indicate favorable or positive contributions. Atoms with intermediate contributions are colored white.

The most important fragments for the most potent and less potent inhibitors of *Pf*dUTPase and (*Hs/Pf*)dUTPase can be viewed in their individual contributions maps (Figure 1). It can be noted that the trityl ring has a favorable for the biological properties (affinity and selectivity) (colored in green). This result suggests that the trytil group could be important for the biological activity. Previous molecular modeling and crystallography studies indicated that two of the three trityl rings have significant interaction with hydrophobic site of the dUTPase enzyme[13].

Table 1. Fragment-based QSAR and QSSR models for affinity and selectivity of *Pf*dUTPase.

	q^2	r^2	Q^2_{ext}	$\bar{r}_m^2$	Δr_m^2
Model	**Affinity *Pf*dUTPase**				
A	0.72	0.85	0.85	0.81	0.09
A/B	0.80	0.89	0.79	0.76	0.11
B/C	0.79	0.90	0.77	0.75	0.15
Model	**Selectivity**				
C/H/Ch/DA	0.79	0.98	0.83	0.50	0.25
H/Ch/DA	0.75	0.96	0.88	0.55	0.22
Ch/DA	0.75	0.96	0.88	0.55	0.22

q^2: Cross-validated correlation coefficient; r^2: non cross-validated correlation coefficient; Q^2_{ext}: determination coefficient of the evaluation set; $\bar{r}_m^2$: average regression coefficient; Δr_m^2: difference between regression coefficient; Fragment distinction: A, atoms; B, bonds; C, connections; H, hydrogen atoms; Ch, chirality, DA, donor and acceptor.

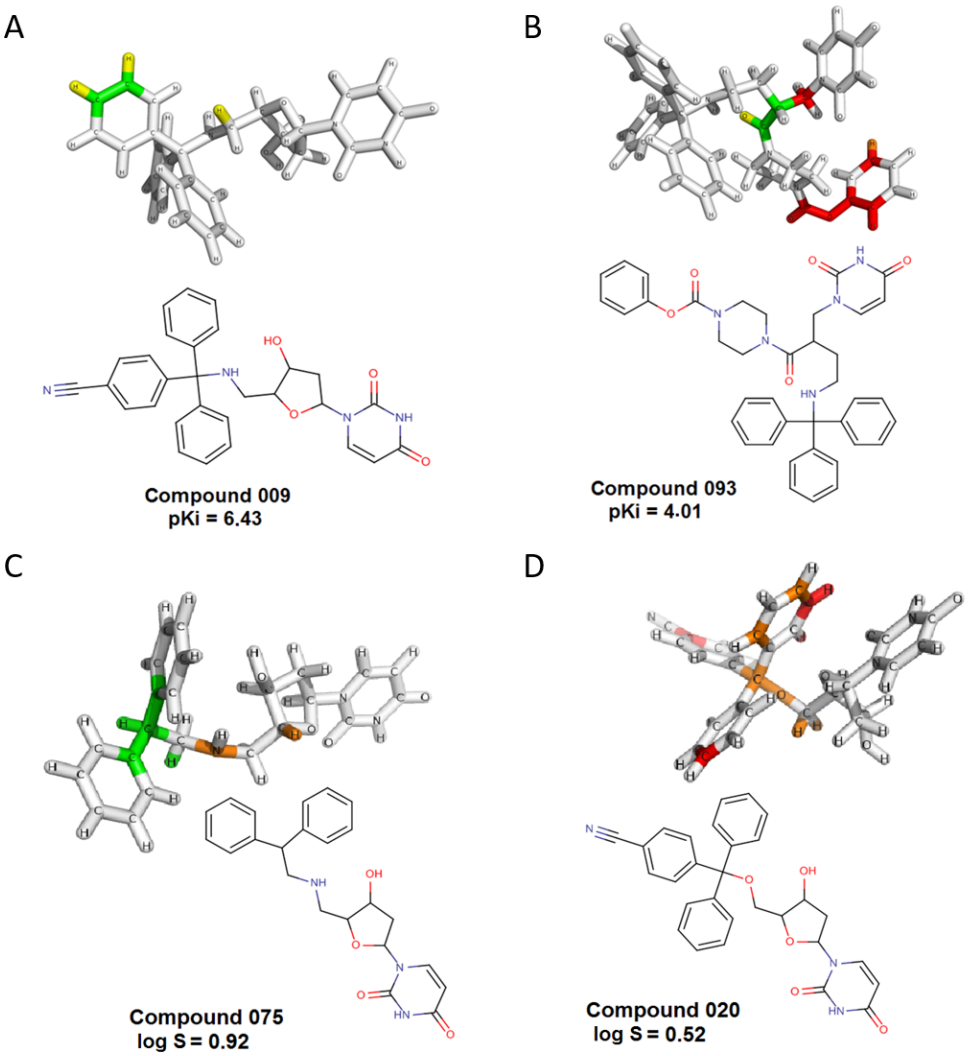

Figure 1. Structural features required for affinity, as highlighted by the fragment-based HQSAR models. Contribution maps for the most potent inhibitor of *Pf*dUTPase (A), and one of the least potent inhibitor of *Pf*dUTPase (B). Structural requirements for *Pf*dUTPase selectivity, as highlighted by the HQSSR models. Contribution maps for the most selective compound (C) and one less selective compound (D).

3. Materials and Methods

The dataset was compiled and integrated from a series of publications of Prof. Ian Gilbert (University of Dundee, UK) [4,8–13], and contained of 127 compounds along with *Ki* values against *Pf*dUTPase. From these, only 47 compounds had data reported against both enzymes from *Plasmodium* and human. This data was used to calculate the selectivity (S) (Eq. 1), which was later used to generate QSSR models. The *Ki* (inhibition constant) values were converted to the corresponding p*Ki* (-log*Ki*) and used as dependent variables in fragment-based QSAR and QSSR modeling.

Equation 1.

$$S = log \frac{HsdUTPase\ Ki}{PfdUTPase\ Ki}$$

The dataset was curated using the protocol described by Fourches and co-workers [17]. Briefly, counterions were removed, whereas specific chemotypes such as aromatic and nitro groups were normalized using the ChemAxon Standardizer (v.5.3, ChemAxon, Budapest, Hungary, http://www.chemaxon.com). The presence of duplicates, *i.e.,* identical compounds reported several times in the same dataset, is known to lead to over-optimistic estimations of the predictivity for developed QSAR models. Thus, after structural standardization, the duplicates were identified using ISIDA Duplicates [18] and HiT QSAR [19] software and carefully analyzed. If the experimental properties associated with two duplicated structures were identical, then one compound was deleted. However, if their experimental properties were significantly different, we deleted both records from the dataset. The dataset was then divided into modeling and evaluation sets, based on modified Kennard-Stone (KS-MD) algorithm. The applicability domain (AD) was assessed using the qsaR in R package.

Hologram QSAR (HQSAR) and QSSR (HQSSR) models were generated using SYBYL-X v.1.2 software (Tripos, Inc., St. Louis, MO, USA). QSAR and QSSR models were constructed using the HQSAR approach as previously described in several drug design studies [20–22]. Briefly, molecular holograms were generated for each molecule of the dataset, using different combinations of parameters concerning hologram generation. Holograms were generated using 6 distinct fragment sizes over the 12 default series of hologram lengths (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353, and 401 bins). Several combinations of fragment distinction parameters were considered, such as atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA). The patterns of fragment counts were then related to the dependent variables using the partial least squares (PLS) regression analyses to derive the HQSAR and HQSSR models. Cross-validation procedures (q^2 leave-one-out, q^2_{LOO}) were used to determine the number of components that yielded optimally predictive models and to assess the stability and statistical significance of the models.

4. Conclusions

The HQSAR and HQSSR models presented good internal consistency, with values of q^2>0.6 and r^2>0.7. The best models were used to predict the affinity of external sets and were rigorously evaluated using several metrics, demonstrating high prediction accuracy. Therefore, the best models should be useful in predicting the affinity and selectivity of untested compounds inside the applicability domain in virtual screening campaigns, in the search of new potent and selective *Pf*dUTPase inhibitors.

Acknowledgments

We are grateful to Brazilian funding agencies, Coordination for the Improvement of Higher Education Personnel (CAPES), the National Counsel of Technological and Scientific Development (CNPq), and the State of Goiás Research Foundation (FAPEG) for their financial support and fellowships. We are also grateful to ChemAxon (Budapest, Hungary) for providing us with the academic license for their software.

Author Contributions

Conceived and designed the experiments: MNN, BJN, VMA, RCB, CHA. Performed the experiments: MNN, MRM. Analyzed the data: MNN, MRM, BJN, VMA, RCB, CHA. Contributed analysis tools: MNN, BJN, VMA, RCB, CHA. Wrote the paper: MNN, BJN, VMA, RCB, CHA.

Conflicts of Interest

The authors declare no conflict of interest.

References and Notes

1. WHO. World malaria Report 2014. Available online: http://www.who.int/malaria/publications/world_malaria_report_2014/report/en/ (accessed Oct 19, 2015).

2. Njogu, P. M.; Guantai, E. M.; Pavadai, E.; Chibale, K. Computer-Aided Drug Discovery Approaches against the Tropical Infectious Diseases Malaria, Tuberculosis, Trypanosomiasis, and Leishmaniasis. *ACS Infectious Diseases* **2015**, 1–24.

3. Nyman, P. O. Introduction. *Current Protein and Peptide Science* **2001**, *2*, 277–285.

4. Whittingham, J. L.; Leal, I.; Nguyen, C.; Kasinathan, G.; Bell, E.; Jones, A. F.; Berry, C.; Benito, A.; Turkenburg, J. P.; Dodson, E. J.; Ruiz Perez, L. M.; Wilkinson, A. J.; Johansson, N. G.; Brun, R.; Gilbert, I. H.; Gonzalez Pacanowska, D.; Wilson, K. S. dUTPase as a platform for antimalarial drug design: structural basis for the selectivity of a class of nucleoside inhibitors. *Structure* **2005**, *13*, 329–38.

5. El-hajj, H. H.; Zhang, H. U. I.; Weiss, B. Lethality of a dut (Deoxyuridine Triphosphatase) Mutation in Escherichia coli. *Journal of Bacteriology* **1988**, *170*, 1069–1075.

6. Gadsden, M. H.; Mclntosh, E. M.; Game, J. C.; Wilson, P. J.; Haynes, R. H. dUTP pyrophosphatase is an essential enzyme in Saccharomyces cerevisiae. *The EMBO Journal* **1993**, *12*, 4425–4431.

7. Pecsi, I.; Hirmondo, R.; Brown, A. C.; Lopata, A.; Parish, T.; Vertessy, B. G.; Tóth, J. The dUTPase enzyme is essential in Mycobacterium smegmatis. *PloS one* **2012**, *7*, e37461.

8. Nguyen, C.; Kasinathan, G.; Leal-Cortijo, I.; Musso-Buendia, A.; Kaiser, M.; Brun, R.; Ruiz-Pérez, L. M.; Johansson, N. G.; González-Pacanowska, D.; Gilbert, I. H. Deoxyuridine triphosphate nucleotidohydrolase as a potential antiparasitic drug target. *Journal of Medicinal Chemistry* **2005**, *48*, 5942–54.

9. Nguyen, C.; Ruda, G. F.; Schipani, A.; Kasinathan, G.; Leal, I.; Musso-Buendia, A.; Kaiser, M.; Brun, R.; Ruiz-Pérez, L. M.; Sahlberg, B.-L.; Johansson, N. G.; Gonzalez-Pacanowska, D.; Gilbert, I. H. Acyclic nucleoside analogues as inhibitors of Plasmodium falciparum dUTPase. *Journal of Medicinal Chemistry* **2006**, *49*, 4183–95.

10. McCarthy, O.; Musso-Buendia, A.; Kaiser, M.; Brun, R.; Ruiz-Perez, L. M.; Johansson, N. G.; Pacanowska, D. G.; Gilbert, I. H. Design, synthesis and evaluation of novel uracil acetamide derivatives as potential inhibitors of Plasmodium falciparum dUTP nucleotidohydrolase. *European Journal of Medicinal Chemistry* **2009**, *44*, 678–88.

11. Baragaña, B.; McCarthy, O.; Sánchez, P.; Bosch-Navarrete, C.; Kaiser, M.; Brun, R.; Whittingham,

J. L.; Roberts, S. M.; Zhou, X.-X.; Wilson, K. S.; Johansson, N. G.; González-Pacanowska, D.; Gilbert, I. H. β-Branched acyclic nucleoside analogues as inhibitors of Plasmodium falciparum dUTPase. *Bioorganic & Medicinal Chemistry* **2011**, *19*, 2378–91.

12. Ruda, G. F.; Nguyen, C.; Ziemkowski, P.; Felczak, K.; Kasinathan, G.; Musso-Buendia, A.; Sund, C.; Zhou, X. X.; Kaiser, M.; Ruiz-Pérez, L. M.; Brun, R.; Kulikowski, T.; Johansson, N. G.; González-Pacanowska, D.; Gilbert, I. H. Modified 5'-trityl nucleosides as inhibitors of Plasmodium falciparum dUTPase. *ChemMedChem* **2011**, *6*, 309–20.

13. Hampton, S. E.; Baragaña, B.; Schipani, A.; Bosch-Navarrete, C.; Musso-Buendía, J. A.; Recio, E.; Kaiser, M.; Whittingham, J. L.; Roberts, S. M.; Shevtsov, M.; Brannigan, J. a.; Kahnberg, P.; Brun, R.; Wilson, K. S.; González-Pacanowska, D.; Johansson, N. G.; Gilbert, I. H. Design, synthesis, and evaluation of 5'-diphenyl nucleoside analogues as inhibitors of the Plasmodium falciparum dUTPase. *ChemMedChem* **2011**, *6*, 1816–31.

14. Mitra, I.; Roy, P. P.; Kar, S.; Ojha, P. K.; Roy, K. On further aplication of rm2 as a metric for validation of QSAR Models. *Journal of Chemometrics* **2010**, *24*, 22–33.

15. Roy, K.; Chakraborty, P.; Mitra, I.; Ojha, P. K.; Kar, S.; Das, R. N. Some case studies on application of "r(m)2" metrics for judging quality of quantitative structure-activity relationship predictions: emphasis on scaling of response data. *Journal of Computational Chemistry* **2013**, *34*, 1071–82.

16. Roy, K.; Mitra, I. On the use of the metric rm^2 as an effective tool for validation of QSAR models in computational drug design and predictive toxicology. *Mini Reviews in Medicinal Chemistry* **2012**, *12*, 491–504.

17. Fourches, D.; Muratov, E.; Tropsha, A. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. *Journal of Chemical Information and Modeling* **2010**, *50*, 1189–204.

18. Varnek, A.; Fourches, D.; Horvath, D.; Klimchuk, O.; Gaudin, C.; Vayer, P.; Solov'ev, V.; Hoonakker, F.; Tetko, I.; Marcou, G. ISIDA - Platform for Virtual Screening Based on Fragment and Pharmacophoric Descriptors. *Current Computer Aided-Drug Design* **2008**, *4*, 191–198.

19. Kuz'min, V. E.; Artemenko, a G.; Muratov, E. N. Hierarchical QSAR technology based on the Simplex representation of molecular structure. *Journal of Computer-Aided Molecular Design* **2014**, *22*, 403–21.

20. Andrade, C. H.; Salum, L. D. B.; Castilho, M. S.; Pasqualoto, K. F. M.; Ferreira, E. I.; Andricopulo, A. D. Fragment-based and classical quantitative structure-activity relationships for a series of hydrazides as antituberculosis agents. *Molecular diversity* **2008**, *12*, 47–59.

21. Salum, L. B.; Andricopulo, A. D.; Honório, K. M. A fragment-based approach for ligand binding affinity and selectivity for the liver X receptor beta. *Journal of molecular graphics & modelling* **2012**, *32*, 19–31.

22. Salum, L. B.; Dias, L. C.; Andricopulo, A. D. Fragment-Based QSAR and molecular modeling studies on a series of discodermolide analogs as microtubule-stabilizing anticancer agents. *QSAR & combinatorial science* **2009**, *28*, 325–337.