

Cheminformatics Modeling of Closantel Analogues for Treating River Blindness

Melaine A. Kuenemann, Phyo Phy Kyaw Zin, Sravya Kuchibhotla, and Denis Fourches*

Department of Chemistry, Bioinformatics Research Center, North Carolina State University,
Raleigh, NC, USA.

* To whom correspondence should be sent. Email: *dfourch@ncsu.edu*

Abstract

Onchocerciasis (also known as river blindness) is a neglected tropical disease caused by the *Onchocerca volvulus* parasitic nematode. Currently, the only approved drug for treating this disease is ivermectin, which is a broad-spectrum antiparasitic agent. However, signs of resistance towards ivermectin have started to emerge. New therapeutic agents are thus urgently needed. The OvCHT1 chitinase enzyme from *O. volvulus* has been established as a relevant biological target for combatting river blindness. The veterinary anthelmintic drug closantel has been found to be a potent, micro-molar OvCHT1 inhibitor. Herein, we investigated the chemical space of closantel and all its synthesized analogues, focusing on the analysis of their potential binding modes towards OvCHT1. First, we conducted an unsupervised hierarchical clustering to group highly similar analogues and explore structure-activity relationships. Second, we conducted a structure-based molecular docking to predict and study the binding modes of all 57 closantel analogues in the active site of OvCHT1. Third, we screened more than 4 million lead-like compounds from the ZINC library to identify other structurally similar ligands that could potentially bind to OvCHT1. The cheminformatics analysis of the closantel analogues illustrated how minor structural changes in closantel analogues can impact their OvCHT1 activity.

Keywords: drug design, structure-activity relationships, molecular modeling, neglected disease, docking.

1. Introduction

Onchocerciasis, often referred as River Blindness (RB), is an eye and skin disease affecting more than 37 million people worldwide (mainly Africa and South America).^[1] The parasitic nematode *Onchocerca volvulus* worm has been found to cause the disease to humans via bites of a blackfly.^[1,2] Those worms evolve inside the human body and migrate to the skin and eye regions, causing severe itching and various lesions. Chronic exposure typically result in permanent blindness and/or harmful skin diseases. Recent international efforts are now aiming at the elimination of onchocerciasis by 2030 in Africa.^[3-5]

Ivermectin^[2,6-8] is the only approved drug to treat RB and has been used for the past twenty years for millions of infected individuals^[1], including US soldiers and NGO workers. However, the *Onchocerca volvulus* worms have begun to develop resistance against ivermectin. The latter compound is indeed stimulating genetic changes in the parasite, resulting in irresponsiveness of the drug for a growing compendium of RB patients^[9,10]. In a study by Mottier and Prichard^[11], more variations within a tubulin gene was observed in RB patients exposed to repeated doses of ivermectin as compared to the parasites obtained from the same patients before exposure to ivermectin. Moreover, it has been shown that more exposure to ivermectin correlated to higher alteration of the parasite's genetic make-up, which led to further resistance against the drug^[12]. Thus, it is now imperative to search for new drugs capable of combatting RB, before the parasite develops a complete and irreversible resistance. In 2018, the US Food and Drug Administration approved another macrocyclic lactone, moxidectin, in order to combat RB. Initially developed for animal use to fight parasites, this compound is structurally similar to ivermectin but have shown very promising results with RB patients compared to ivermectin.^[13]

Vital biosynthesis and degradation processes of *Onchocerca Volvulus* worm are regulated by the chitin synthase and chitinase enzymes. Although the exact role of chitinase in the development of the worm is unclear, this enzyme is sought to play a crucial biological function. Additionally, due to the absence of chitinase in humans, this enzyme potentially represents a viable therapeutic opportunity for treating RB via the blockade of this unique new pathway. The chitinase from *O. volvulus* (OvCHT1) has been characterized and shown to belong to the glycoside hydrolase family 18, which is involved in a substrate-assisted hydrolysis mechanism^[14–16]. OvCHT1 was found to be expressed in the infective L3 larvae with a potential involvement in host transmission, molting, and the developmental processes. The Janda's research group at Scripps identified the small molecule closantel (**Figure 1A**) as a potent, micro-molar and specific inhibitor of OvCHT1^[17]. Until then, closantel had been known as a veterinary anthelmintic drug for the treatment of sheep and cattle infected with liver fluke^[18,19]. When tested *in vitro* against OvCHT1, closantel afforded a very promising IC₅₀ value of 1.6 μ M (pIC₅₀ = 5.80) and high specificity for filarial chitinases. Additionally, closantel was found to completely inhibit the molting of the L3 larvae^[17]. As a result, multiple experimental screening efforts were conducted by Janda and co-workers to synthesize and test more than 50 closantel analogues. The latter compounds were characterized by experimental pIC₅₀ ranging from 3.7 to 6.4 (see refs^[17,20–22]).

In this cheminformatics study, we aimed at (1) characterizing the chemical space of all tested closantel analogues and cluster them based on their structural similarity to identify structure-activity relationships, (2) using 3D molecular docking to compute the binding mode of each analogue in the active site of OvCHT1 and study the key molecular interactions formed by closantel analogues with OvCHT1 residues, and (3) virtually screening the ZINC database to identify analogues predicted to have enhanced OvCHT1 inhibition potencies.

2. Materials and Methods

2.1. Compilation and pre-processing of the dataset

In total, 57 closantel analogues were compiled from the recent publications^[17,20–22] of the Janda's research group at the Scripps Institute. The structures of these compounds were prepared according to the standard curation protocols developed earlier^[23–25]. All experimental IC₅₀ values were converted into pIC₅₀ (*e.g.*, pIC₅₀ = 9 for a nanomolar inhibitor). Closantel analogues were pre-processed using the LigPrep program^[26] from the Schrodinger suite. This process generated multiple entries for each compound, depending on the possible protonation states at a specific pH range (7 ± 2), tautomers, and ring conformations (1 ring conformer per compound). Then, all the compounds were minimized using the OPLS3 force field to obtain reasonable 3D conformations for each analogue.

2.2. Structural characterization using molecular descriptors

After structural standardization^[24,25], we computed 117 two-dimensional RDKit descriptors using Knime^[27] to characterize the constitutional and structural properties of the closantel analogues. A distribution analysis using R (v3.3.1) was performed for the hydrophobicity (SlogP), molecular weight (MW), number of hydrogen bond acceptors (NumHBA), and number of hydrogen bond donors (NumHBD).

2.3. Cluster analysis and identification of activity cliffs

A circular dendrogram was created using the *ggtree* package^[28] in R (v3.3.1). The clustering was conducted with a curated set of 66 RDKit descriptors after all correlated, low

variance, constant, and near-constant descriptors were removed. The dendrogram was created according to the Euclidean distance between molecules and the Ward linkage^[29] between clusters. Each dendrogram node was then colored according to the experimental pIC₅₀ value of the corresponding compound.

2.4. Homology modeling and protein preparation

As there was no X-ray structure of OvCHT1 in the Protein Data Bank^[30] at the time of this study, we built a homology model of this target using Prime^[31] from the Schrodinger Suite (v2016-3). Our previous case studies using 3D docking with homology models of proteins showed that reasonable docking poses can be obtained.^[32,33] OvCHT1 sequence was collected from the Universal Protein Resource^[34] (Uniprot code Q25615). The homology model was created based upon a single protein template (PDB code 1WAW - CHIT1 chitinase 1) that shared 37% sequence identity with OvCHT1 (**Figure 4A**). The homology model was created using Prime's energy-based procedure and all loops were minimized using the VSGB solvation model and the OPLS3 force field. Then, the protein structure was further pre-processed using the Protein Preparation Wizard^[35] from Schrodinger Maestro suite (v2016-3). Molecules of water beyond 5.00Å from any hetero groups were removed. Explicit hydrogen atom addition and H-bond assignments were conducted at pH=7 using PROPKA. Finally, the homology model was minimized using restrained minimization with the OPLS3 force field. The corresponding curated PDB file is available in the Supplementary Material.

2.5. Structure-based molecular docking using Glide

A grid box for molecular docking was created using the Schrodinger Suite with the center of grid coordinates equal to $x=1.45$, $y=3.45$ and $z=19.33$. The size of the grid was set up to accept and dock ligands with a length lower than 30\AA . Glide docking was performed using the SP scoring function (standard precision mode) to generate multiple poses per compound. Then, all the SP poses were re-docked using the Glide XP scoring function (extra precision mode). At last, all the results were analyzed according to the XP scoring data.

2.6. Virtual screening of the ZINC library

Over 4,333,000 lead-like compounds from the ZINC database were screened to identify structurally-similar chemical structures that could potentially inhibit OvCHT1. Our protocol was divided in four different steps: First, we conducted a similarity search based upon the structural features of the Top-5 most potent closantel analogues that also obtained the best XP docking scores. To do so, we calculated the MACCS fingerprint^[36] of all the ZINC lead-like compounds using the CDK fingerprint node in the KNIME platform^[37]. Then, only the compounds presenting a Tanimoto structural similarity higher than 80% with the Top-5 selected closantel analogues were kept. These selected compounds were then pre-processed using LigPrep^[26] from the Schrodinger suite (same protocol than for the closantel analogues). The prepared compounds were then docked in the OvCHT1 binding site using both SP and XP scoring functions^[38,39]. Only the compounds that presented a XP docking score below -6 kcal/mol and an eModel score below -40 kcal/mol were kept for further investigation.

3. Results & Discussion

3.1. Characterization of the chemical space of closantel analogues

After chemical curation, we obtained a set of 57 unique closantel analogues. This set presented a relatively broad range of experimental activity from $IC_{50} = 0.34 \mu M$ up to $200 \mu M$ (**Figure 1B**) with the activity of closantel itself being $1.6 \mu M$ ($pIC_{50} = 5.80$) (**Figure 1A**). The most active compound (*24918716_3m*) is characterized by a pIC_{50} equal to 6.4 ($IC_{50} = 0.34 \mu M$) (**Figure 2A**). The five most active compounds are represented in **Figure 2**. Three compounds out of the five most active inhibitors (**Figure 2A-C**) contain the 2-hydroxy-3,5-diiodo-N-phenylbenzamide substructure. Interestingly, the two other top active compounds (**Figure 2D-E**) present a different scaffold (tri-substituted triazole) while still affording a similar activity at $pIC_{50} \sim 6.4$.

To further characterize the chemical space of the 57 closantel analogues, 117 two-dimensional RDKit structural descriptors were calculated. In particular, we analyzed the distributions of four molecular properties: the hydrophobicity (SlogP), the molecular weight (MW), the number of Hydrogen-Bond Acceptors (NumHBA) and Donors (NumHBD). Closantel analogues presented an average hydrophobicity ($mean_{SlogP}$) equal to 5.04 ± 1.44 , an average molecular weight ($mean_{MW}$) equal to $439 \text{ g/mol} \pm 140.93$, an average number of H-bond donors ($mean_{NumHBD}$) equal to 1.54 ± 0.73 and an average number of H-bond acceptors ($mean_{NumHBA}$) equal to 2.57 ± 1.31 . Interestingly, according to the distribution of these physico-chemical properties, we found a positive correlation between the potency of the closantel analogues with their molecular weight and the hydrophobicity. This result is illustrated in **Figure 1C-D**, especially with the distributions of MW and SlogP colored by pIC_{50} activity class (low: $pIC_{50} < 5$, intermediate: $5 < pIC_{50} < 6$ and high: $pIC_{50} > 6$). However, NumHBA and NumHBD did not appear

to discriminate the analogues with the highest pIC_{50} values (**Figure 1E-F**). Noticeably, many closantel analogues have properties fulfilling the Lipinski's rule of five^[40] with 38% of the dataset passing the rules with zero violations, 66% with one violation and 100% with two violations (**Figure 1C-F**). This result is not surprising as those analogs are very similar to the orally bioavailable drug closantel.

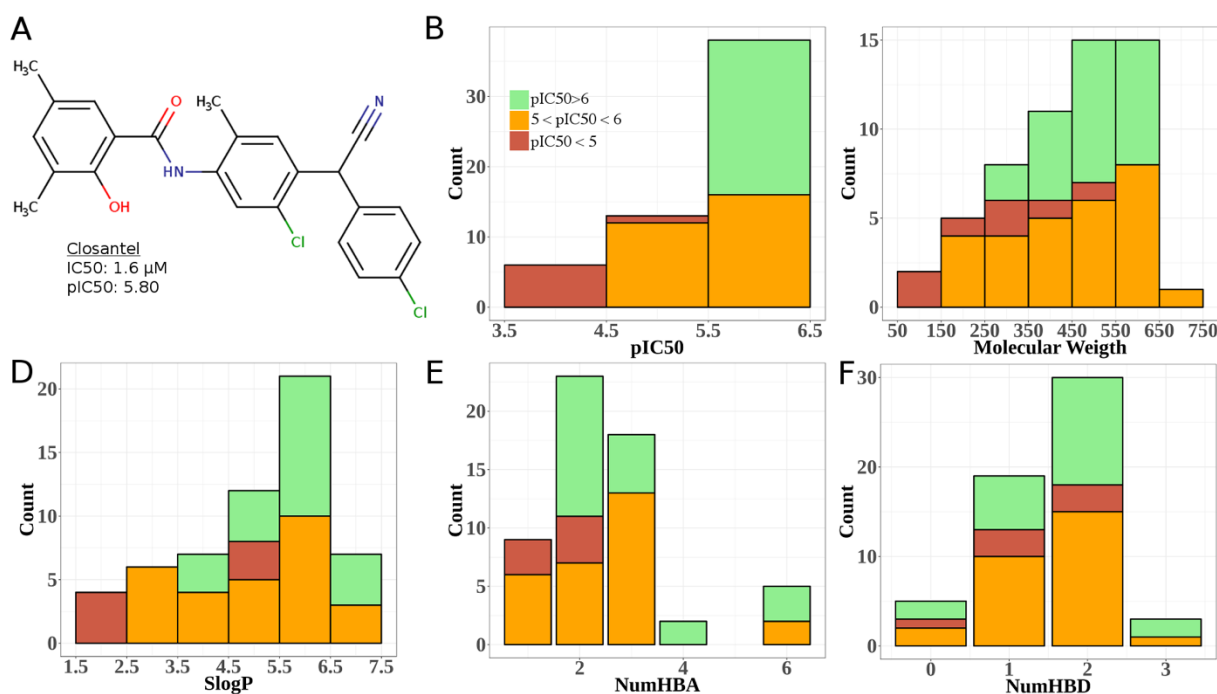


Figure 1: Structure and activity for closantel (A), Distribution of closantel analogues according to their pIC_{50} (B), Molecular Weight (C), SlogP (D), Number of Hydrogen Bond Acceptor (E), and Number of Hydrogen Bond Acceptor (F). Bins for each histogram are colored according to their pIC_{50} , red if pIC_{50} below 5, orange if pIC_{50} between 5 and 6, and green if pIC_{50} higher than 6 (most active compounds).

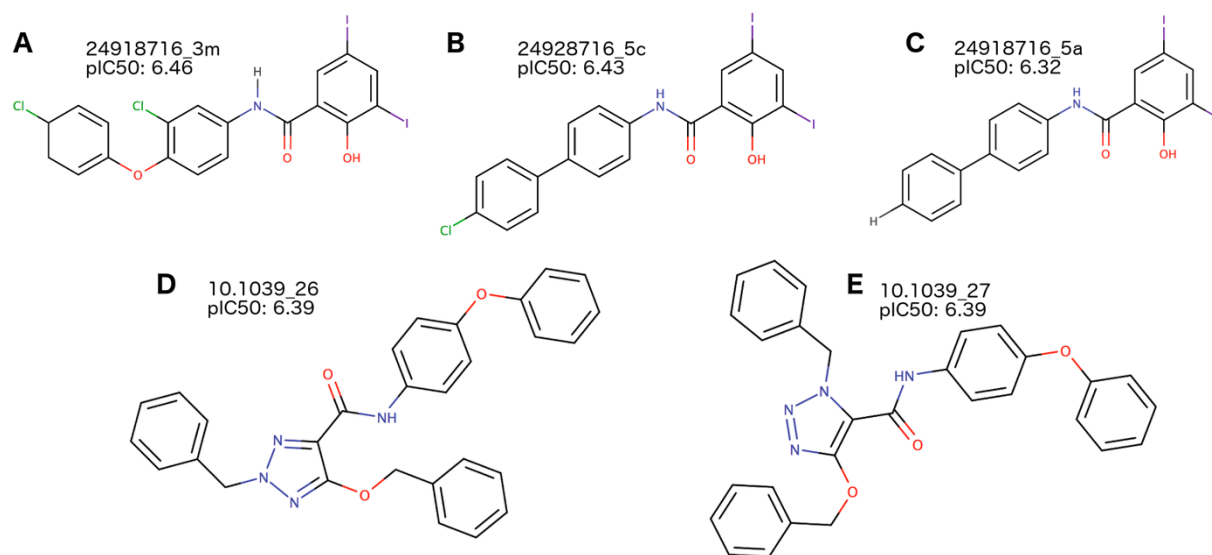


Figure 2: Structure and inhibition potency of the five most active closantel analogues.

3.2. Hierarchical clustering of OvCHT1 inhibitors

Then, we conducted an unsupervised hierarchical clustering to group structurally similar analogues using Euclidean distance, Ward Linkage^[29], and a downsized set of 66 RDKit descriptors (all correlated, low variance, constant, and near-constant descriptors were removed). The resulting circular dendrogram is provided in **Figure 3**. Each dendrogram node was colored according to the experimental pIC₅₀ value afforded by its corresponding analogue. This procedure allowed us to identify all the groups of analogues having similar potencies towards OvCHT1. For instance, **cluster1** (**Figure 3**) involved four compounds (20142509_8, 20142509_5, 20142509_4, and 20142509_3), all characterized by low potencies (pIC₅₀ = 3.70 - 4.17). The dendrogram also revealed **cluster2** containing three compounds (10.1039_27, 10.1039_25 and 10.1039_26) with excellent pIC₅₀ values (pIC₅₀ = 6.28 - 6.39). These three compounds have the same scaffolding with different substituents at their triazole ring. The two compounds 10.1039_27 and 10.1039_26 are constitutional isomers as the benzyl substituent has been swapped between positions 1 and 2 of the triazole. However, the third compound (10.1039_25) presents a methyl-cyclohexyl

substituent at position 2 of the triazole ring. The **cluster3** presents an *activity cliff*^[41,42] (i.e., pair of highly similar compounds with significantly different activity values). Here, the two compounds (20142509_6 and 21534605_3b) differ by the presence/absence of a chlorine and the ortho/para position of a methyl group but present a one-fold activity difference (pIC₅₀ = 4.48 versus 5.76 respectively). Interestingly, these two compounds came from two different studies^[17,20]. This cluster analysis shows that useful structure-activity relationships can be extracted from such fully integrated set of compounds and help us understand what structural changes (even minor) can influence the activity toward OvCHT1.

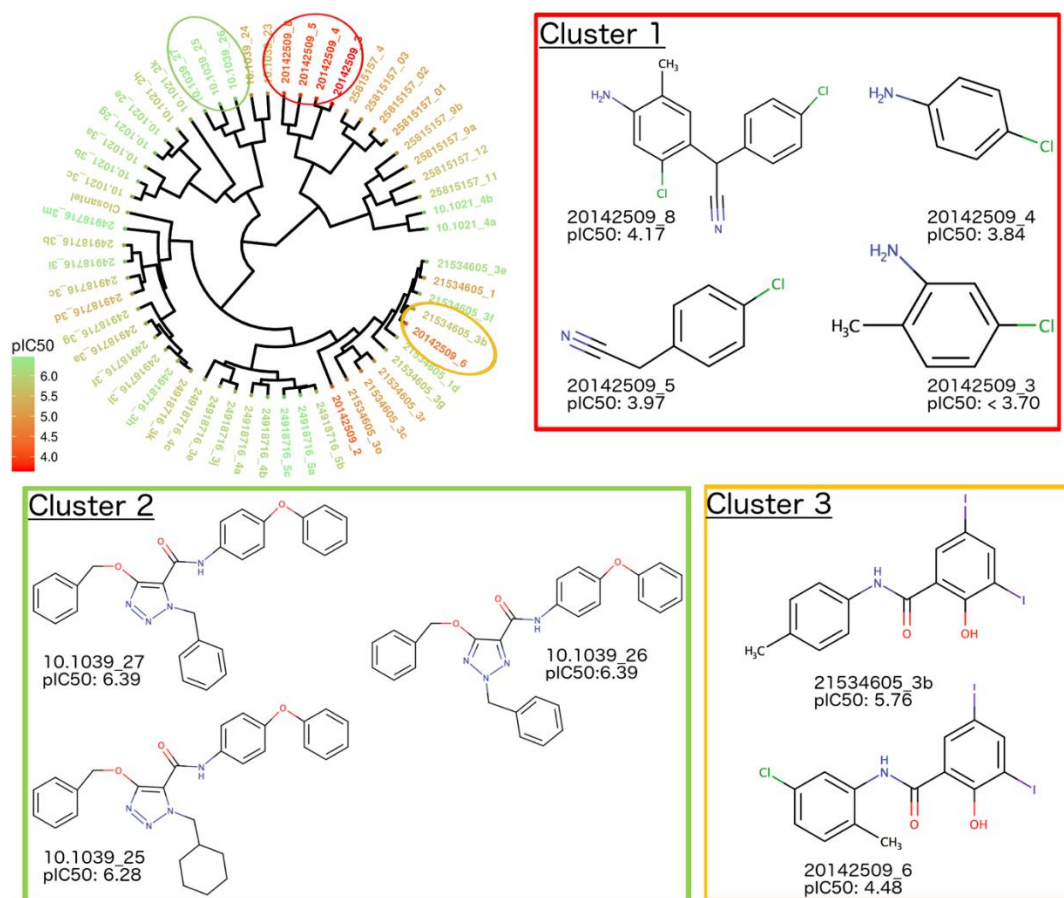


Figure 3: Circular dendrogram obtained from the hierarchical clustering of the set of 57 closantal analogs using 66 RDKit descriptors space, Ward linkage, and Euclidian distance. Compound nodes & names are colored according to their pIC₅₀. Three examples of cluster based on the hierarchical clustering.

3.3. Molecular docking of closantel analogues

We conducted the structure-based docking of closantel analogues using Glide XP to compute and study their predicted binding modes in the active site of the OvCHT1 enzyme. This was an essential step in order to better understand the key OvCHT1-inhibitor interactions and attempt to validate molecular docking as a predictive technique for reliably discriminating the most potent OvCHT1 inhibitors.

All compounds were docked in the same pocket (**Figure 4A**). Resulting docking scores were ranging between -6.50 and -1.72 kcal/mol, whereas the eModel scores were from -76.11 and -15.37 kcal/mol (in both cases, the lower, the better). As represented in **Figure 4B**, the molecular docking procedure was able to discriminate the most active ligands from less active and inactive compounds using the combination of the two scores. One should note that, despite the separation between experimentally-confirmed active versus less active compounds, the docking scores afforded by those analogues were not excellent (higher than -7 kcal/mol, the empirical threshold for selecting micromolar inhibitors in a virtual screening campaign). The range of docking scores for the Top-5 most active molecules (pIC_{50} = 6.32 - 6.46, **Figure 2**) is from -5.07 to -3.04 kcal/mol and -74.74 to -40.50 kcal/mol for eModel score. Moreover, these 5 compounds presented different binding modes but they shared important interactions with OvCHT1, especially the H-bond with Arg270. Compounds *10.1039_26* and *10.1039_27* also shared a π - π stacking with Trp31.

In particular, we decided to evaluate the binding modes of two active compounds, *10.1021_3a* and *10.1021_2g* (pIC_{50} equal to 6.08 and 6.30, respectively) that afforded the best docking scores (-6.50 and -6.16 kcal/mol, respectively). The compound *10.1021_3a* (**Figure 4C**) presents a cation- π interaction between its benzene ring and Arg270. The predicted binding mode also depicted four hydrogen bonds, one donor between the side chain of Tyr148 and the oxygen

from the amide group, two hydrogen bond donors between the oxygen on the cycloheptane and Ala187 and Gly188 backbones. And a hydrogen bond acceptor between the ligand's amide group and the side chain of Asp214. The second selected closantel analogue (10.1021_2g) also presented a π - π interaction with Trp361 (like 10.1039_26 and 10.1039_27). Another hydrogen bond was predicted to be formed with Trp31 (**Figure 4D**). A cation- π interaction with the benzene group in 10.1021_2g and the sidechain of Arg35 was also found. Finally, the predicted pose from 10.1021_2g presented a π - π interaction with Phe365.

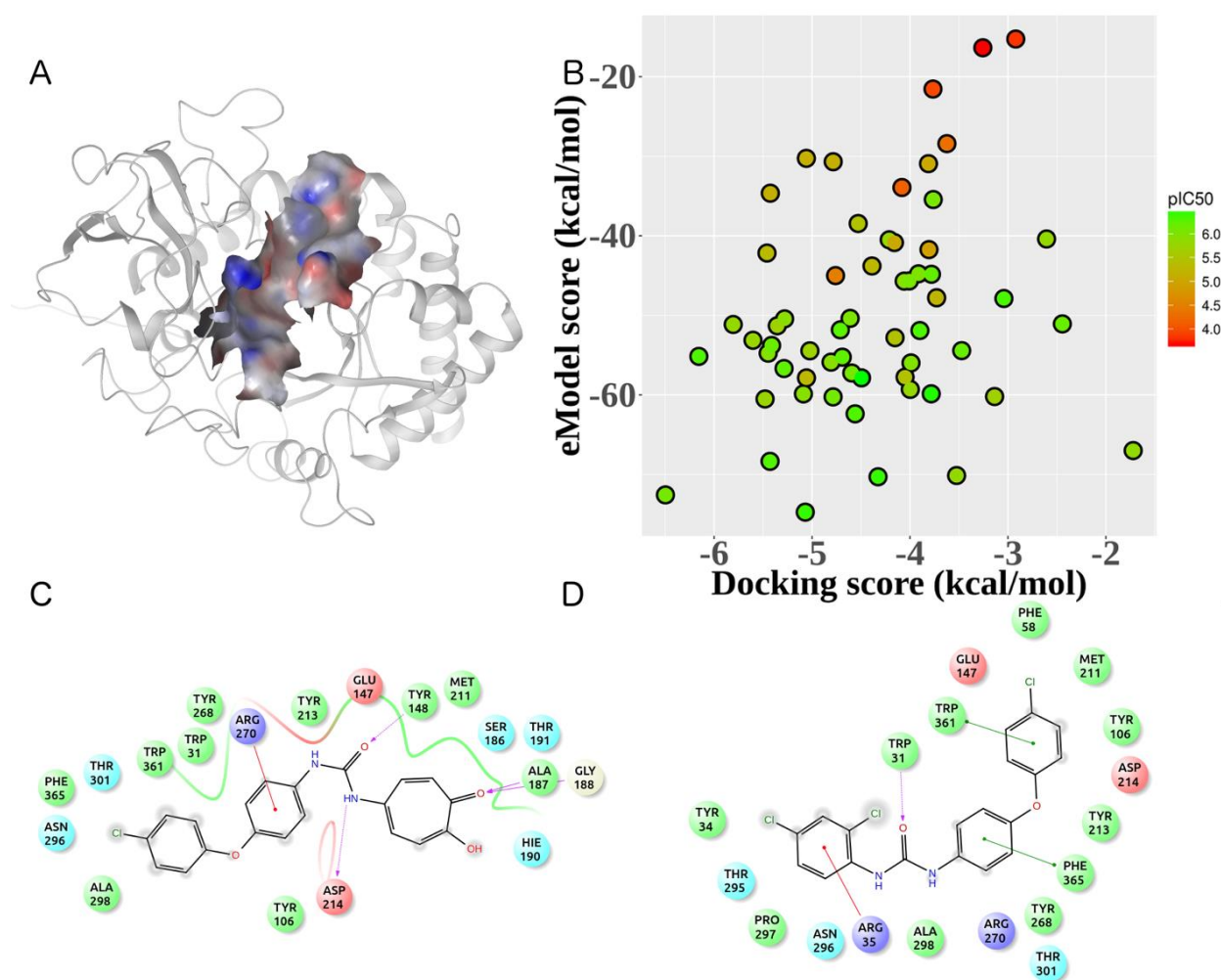


Figure 4: (A) The predicted 3D structure of OvCHT1 according to our homology model, (B) scatter plot of XP docking score vs eModel score for the closantel analogs, the 2D ligand interaction of compounds (C) 10.1021_3a (pIC50 = 6.07) and (D) 10.1021_2g (pIC50 = 6.30)

Interestingly, these two compounds presented two different binding modes, likely due to the large binding pocket of OVCHT1. The two ligands do share a common area of the pocket but also interact with two different side parts of the pocket as seen in **Figure 5**. This result could enable the rational design of new analogues better occupying the different parts of the OVCHT1 binding pocket.

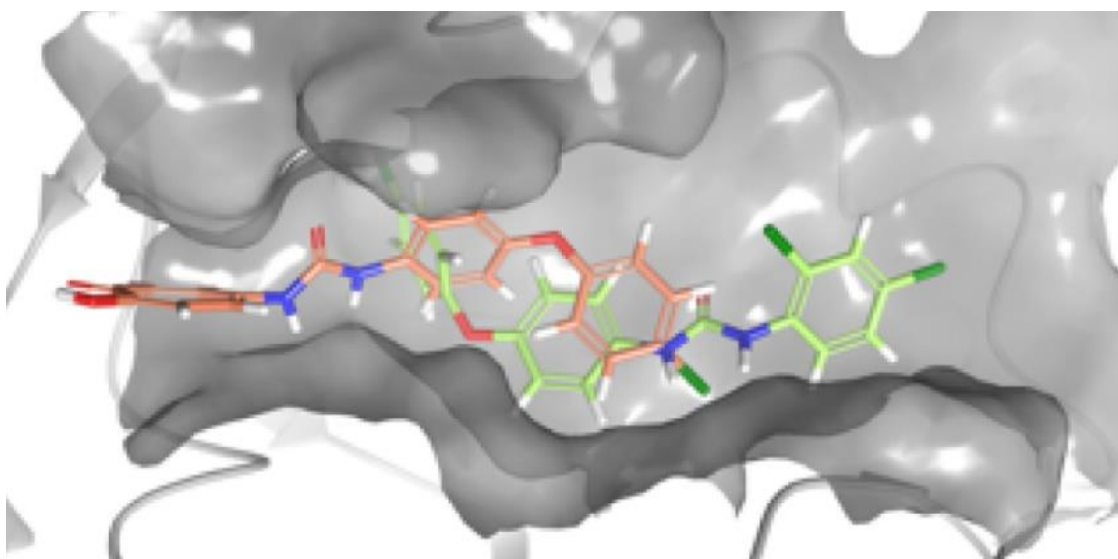


Figure 5: Superimposition of the docking poses for 10.1021_3a ($pIC_{50} = 6.07$, *red*) and 10.1021_2g ($pIC_{50} = 6.30$, *green*).

3.4. ZINC virtual screening

At last, we performed a virtual screening (**Figure 6**) of the lead-like subset of the ZINC library (>4 million unique compounds). The first step of our protocol was to select only the analogs of the five most active closantel analogues (compounds with a Tanimoto distance higher than 0.8 using MACCS fingerprints). After the first selection, 3,283 similar compounds were prepared with Ligprep and were docked inside the binding site of OvCHT1. The virtual screening was first performed using the Glide SP scoring function and then re-docked using the XP extra-precision

scoring function from Glide. Then only compounds presenting a XP docking below -6 kcal/mol and an eModel score below -40 kcal/mol were kept, representing 121 selected compounds.

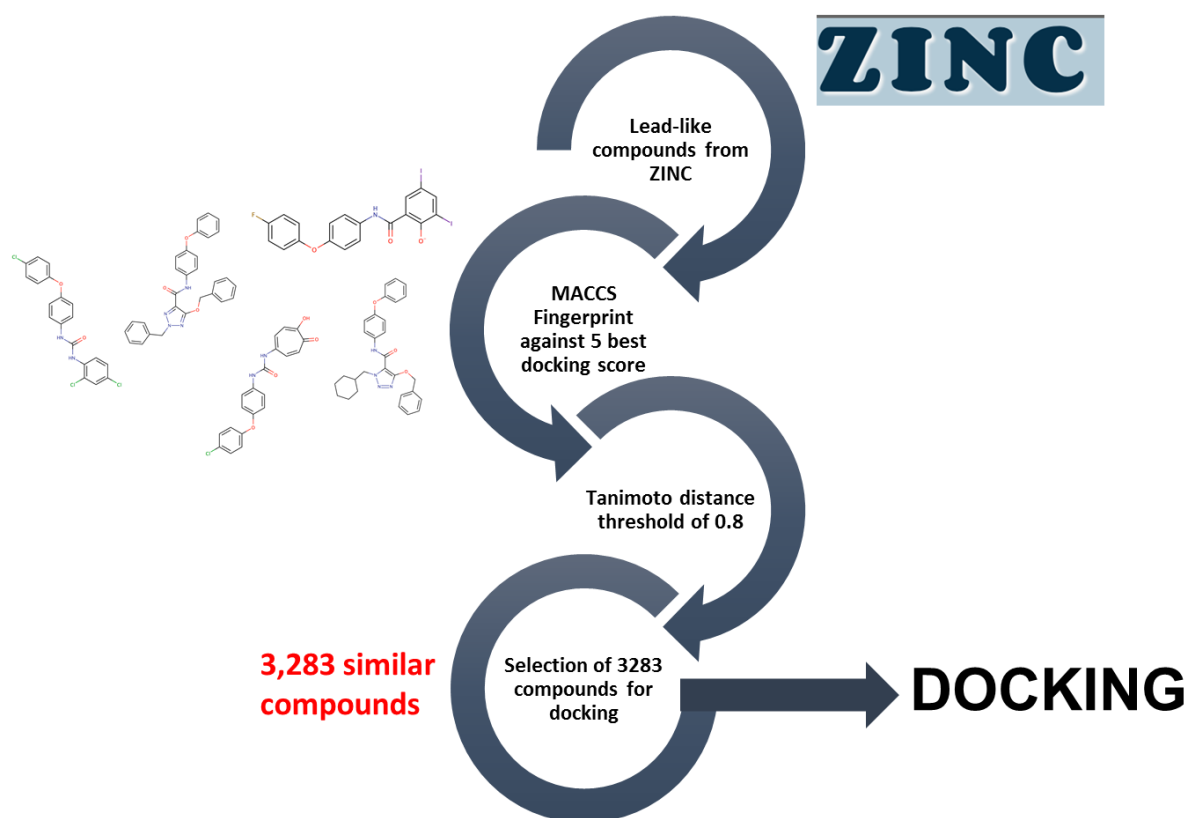


Figure 6: General scheme representing our protocol used for ZINC virtual screening.

Interestingly, two ZINC compounds (ZINC43271075 and ZINC62272215) afforded a better docking score than any of the closantel analogs. The predicted binding poses of these two compounds are represented in **Figure 7**. The first compound (ZINC43271075) afforded a docking score equal to -7.39 kcal/mol and eModel score equal to -56.11 kcal/mol. The predicted binding mode of ZINC43271075 presented a π - π interaction with Trp361, a hydrogen bond between an oxygen and the Trp31 and one cation- π stacking with Arg35 (**Figure 7A/C**). These three interactions were already predicted in the binding mode of 10.1021_3a (**Figure 4C**). Moreover,

ZINC43271075 presented two other interactions with OvCHT1, two hydrogen bonds: one between the Glu147 and a nitrogen and one between the Tyr 106 and an oxygen. The second highlighted compound (ZINC62272215) was predicted with a less favorable docking score (-6.67 kcal/mol, eModel score equal to -49.02 kcal/mol) but still better than any of the scores obtained by the 57 closantel analogues. ZINC62272215 was predicted to interact with the same amino acid with a π - π interaction with Trp361 and a salt bridge with Glu147 (in **Figure 7B/D**). These two compounds present new interactions (especially with Glu147 and Tyr106) that could open a new way of inhibiting OvCHT1 and help us designing new inhibitors.

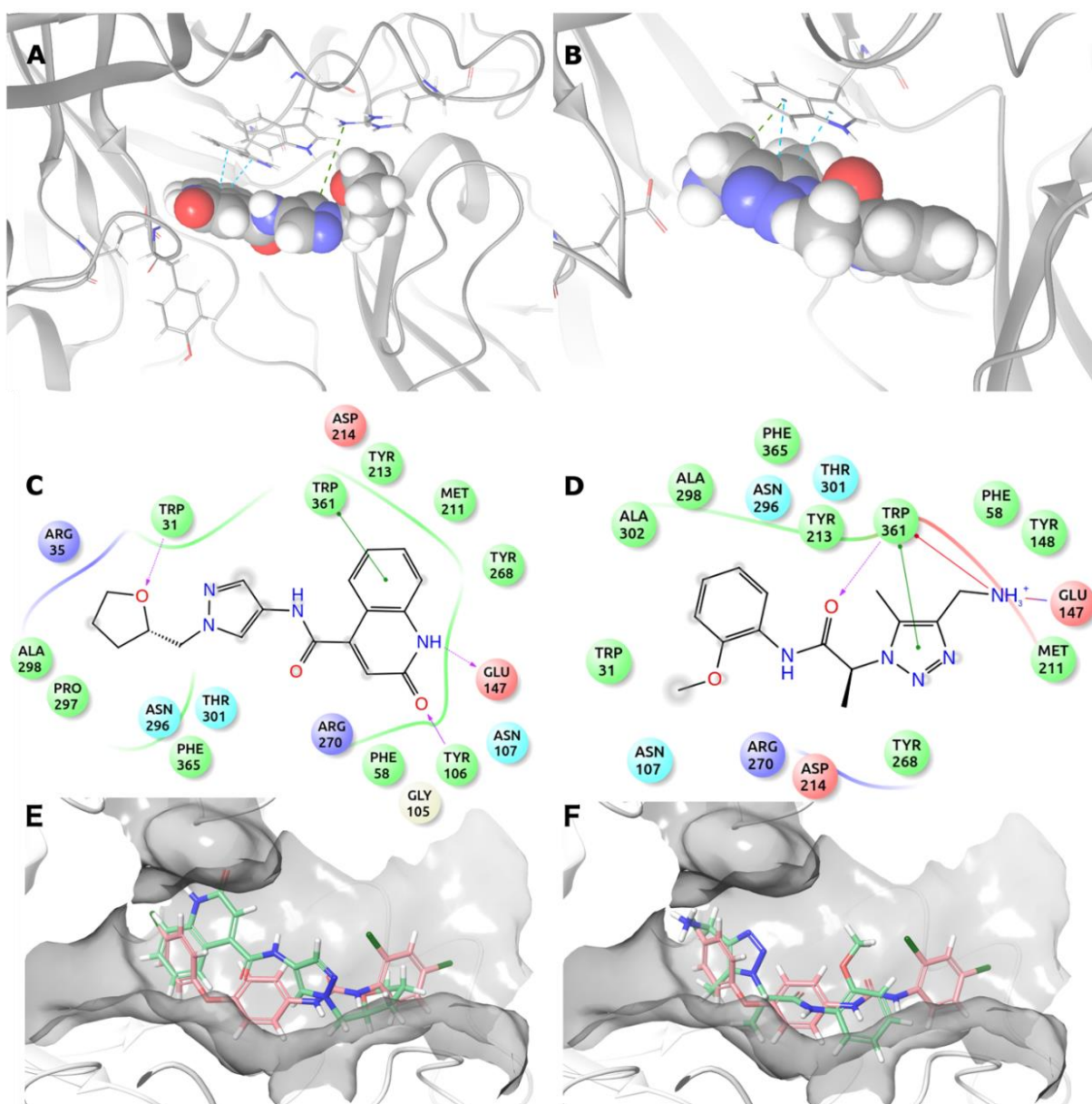


Figure 7: Predicted binding mode of (A) ZINC43271075 and (B) ZINC62272215 with the interacting amino acid highlighted. 2D interaction plot of the (C) ZINC43271075 and (D) ZINC62272215, (E) superposition of the ZINC43271075 in green and 10.1021_2g binding pose in pink, (F) superposition of the ZINC62272215 in green and 10.1021_2g binding pose in pink.

4. Conclusions

This analysis revealed that molecular docking could help differentiating the most active OvCHT1 inhibitors from the less active and inactive ligands. Subsequent experimental screening

must be accomplished to validate the hits we identified from the ZINC library, notably ZINC43271075 and ZINC62272215. The analysis of the binding site revealed several regions for which no closantel analogue has been predicted to interact with. Closer collaborations between experimentalist and computational chemists could thus help to design new bioactive OvCHT1 inhibitors of high interest for combatting river blindness.

Supporting Information

The curated homology model (PDB) as well as all chemical structures (modeling set and ZINC hits) are available for free in the Supporting Material.

Conflict of Interest

None declared.

Acknowledgements

MK and DF gratefully acknowledge the NC State Chancellor's Faculty Excellence Program and the NC State Comparative Medicine Institute.

References

- [1] M.-G. Basanez, S. D. S. Pion, T. S. Churcher, L. P. Breitling, M. P. Little, M. Boussinesq, *PLoS Med.* **2006**, 3, 1454–1460.
- [2] M. G. Basáñez, S. D. Pion, E. Boakes, J. A. Filipe, T. S. Churcher, M. Boussinesq, *Lancet Infect. Dis.* **2008**, 8, 310–322.
- [3] R. Colebunders, W. A. Stolk, J. N. Siewe Fodjo, C. D. Mackenzie, A. Hopkins, *Infect. Dis. poverty* **2019**, 8, 83.
- [4] R. Colebunders, M. G. Basáñez, K. Siling, R. J. Post, A. Rotsaert, B. Mmbando, P. Suykerbuyk, A. Hopkins, *Infect. Dis. Poverty* **2018**, DOI 10.1186/s40249-018-0406-7.
- [5] A. Hopkins, *Community eye Heal.* **2019**, 31, 94.
- [6] W. L. Shoop, *Parasitol. Today* **1993**, 9, 154–159.
- [7] R. Laing, V. Gillan, E. Devaney, *Trends Parasitol.* **2017**, 33, 463–472.
- [8] W. C. Campbell, *Angew. Chemie - Int. Ed.* **2016**, DOI 10.1002/anie.201601492.
- [9] C. Bourguinat, S. D. S. Pion, J. Kamgno, J. Gardon, B. O. L. Duke, M. Boussinesq, R. K. Prichard, *PLoS Negl. Trop. Dis.* **2007**, 1, e72.
- [10] S. Lustigman, J. P. McCarter, *PLoS Negl. Trop. Dis.* **2007**, 1, e76.
- [11] M. de Lourdes Mottier, R. K. Prichard, *Pharmacogenet. Genomics* **2008**, 18, 129–140.
- [12] M. Y. Osei-Atweneboana, J. K. Eng, D. A. Boakye, J. O. Gyapong, R. K. Prichard, *Lancet* **2007**, 369, 2021–2029.
- [13] M. Boussinesq, *Lancet* **2018**, DOI 10.1016/S0140-6736(18)30101-6.
- [14] Y. Wu, G. Egerton, A. P. Underwood, S. Sakuda, A. E. Bianco, *J. Biol. Chem.* **2001**, 276, 42557–42564.
- [15] D. M. F. Van Aalten, D. Komander, B. Synstad, S. Gåseidnes, M. G. Peter, V. G. H.

- Eijsink, *Proc. Natl. Acad. Sci.* **2001**, 98, 8979–8984.
- [16] Y. Wu, R. Adam, S. A. Williams, A. E. Bianco, *Mol. Biochem. Parasitol.* **1996**, 75, 207–219.
- [17] C. Gloeckner, A. L. Garner, F. Mersha, Y. Oksov, N. Tricoche, L. M. Eubanks, S. Lustigman, G. F. Kaufmann, K. D. Janda, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 3424–9.
- [18] R. J. Martin, *Vet. J.* **1997**, 154, 11–34.
- [19] H. Van Den Bossche, H. Verhoeven, O. Vanparijs, H. Lauwers, D. Thienpont, *Arch. Int. Physiol. Biochim.* **1979**, 87, 851–853.
- [20] A. L. Garner, C. Gloeckner, N. Tricoche, J. S. Zakhari, M. Samje, F. Cho-Ngwa, S. Lustigman, K. D. Janda, *J. Med. Chem.* **2011**, 54, 3963–3972.
- [21] M. Gooyit, T. L. Harris, N. Tricoche, S. Javor, S. Lustigman, K. D. Janda, *ACS Infect. Dis.* **2015**, 1, 198–202.
- [22] A. C. Pippione, F. Dosio, A. Ducime, A. Federico, K. Martina, S. Sainas, B. Frølund, M. Gooyit, K. D. Janda, D. Boschi, et al., *Med. Chem. Commun.* **2015**, 6, 1285–1292.
- [23] D. Fourches, E. Muratov, A. Tropsha, *Nat. Chem. Biol.* **2015**, 11, 535–535.
- [24] D. Fourches, E. Muratov, A. Tropsha, *J. Chem. Inf. Model.* **2011**, 50, 1189–1204.
- [25] D. Fourches, E. Muratov, A. Tropsha, *J. Chem. Inf. Model.* **2016**, 56, 1243–1252.
- [26] LigPrep Schrödinger LLC New York NY, **2016**.
- [27] M. R. Berthold, N. Cebron, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, in *KNIME Konstanz Inf. Min.* (Eds.: C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker), Springer Berlin Heidelberg, Berlin, Heidelberg, **2008**, pp. 319–326.

- [28] G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Y. Lam, *Methods Ecol. Evol.* **2017**, 8, 28–36.
- [29] J. H. Ward Jr, *J. Am. Stat. Assoc.* **1963**, 58, 236–244.
- [30] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, 28, 235–242.
- [31] Schrödinger Release 2015-4, **2015**.
- [32] M. A. Kuenemann, D. Fourches, *Mol. Inform.* **2018**, DOI 10.1002/minf.201700138.
- [33] M. A. Kuenemann, P. A. Spears, P. E. Orndorff, D. Fourches, *Mol. Inform.* **2018**, DOI 10.1002/minf.201800004.
- [34] U. Consortium, others, *Nucleic Acids Res.* **2017**, 45, D158--D169.
- [35] Schrödinger Release 2016-3, **2016**.
- [36] J. L. Durant, B. A. Leland, D. R. Henry, J. G. Nourse, *J. Chem. Inf. Comput. Sci.* **2002**, 42, 1273–1280.
- [37] M. Berthold, N. Cebron, F. Dill, *Data Anal. Mach. ...* **2008**, 11, 319–326.
- [38] Schrödinger Release 2015-4, **2015**.
- [39] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, et al., *J. Med. Chem.* **2004**, 47, 1739–1749.
- [40] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Deliv. Rev.* **2001**, 46, 3–26.
- [41] G. M. Maggiora, *J. Chem. Inf. Model.* **2006**, 46, 1535.
- [42] J. Pérez-Villanueva, O. Méndez-Lucio, O. Soria-Arteche, J. L. Medina-Franco, *Mol. Divers.* **2015**, 19, 1021–1035.