

Using graph-based structural signatures and machine learning algorithms for molecular docking assessment of histone deacetylases and small ligands

Alessandra G. Cioletti¹, Rafael P. Lemos¹, Lucas M. dos Santos¹, Diego Mariano¹,
Raquel C. de Melo-Minardi¹

¹Laboratório de Bioinformática e Sistemas (LBS)
Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

raquelcm@dcc.ufmg.br

Abstract. *Histone deacetylases (HDACs) are enzymes that play an essential role in regulating gene expression, with recent studies linking their inhibition to autism spectrum disorders (ASD). As a result, there is growing interest in understanding the effects of HDAC inhibition. In this paper, we used molecular docking to investigate the binding between HDACs and small ligands, focusing on two enzymes involved in embryonic development: Histone deacetylase 1 (H1) and Histone deacetylase 2 (H2). Using a graph-based structural signature algorithm, we extracted features from the resulting complexes and employed machine learning algorithms to distinguish natural ligands from decoys, achieving 72% of accuracy in the classification test.*

1. Introduction

Autism and autism spectrum disorders (ASD) are complex disorders in the neurological development of children, characterized by dysfunctions in social interactions, communication, and repetitive and stereotyped behaviors. It is a condition that mainly affects boys and has clinical and etiological heterogeneity [Waye and Cheng 2017].

In the past, the incidence of ASD was thought to be rare and genetic, but in recent decades, this figure has increased worldwide. In the United States, the figure increased from 1 in 150 children aged 8 in 2000 to 1 in 44 in 2018 [Maenner et al. 2020]. Moreover, recent research has shown that epigenetic modifications are important for creating the pathophysiology of ASD. Epigenetic modifications are heritable genetic changes that do not alter the DNA sequence and can occur throughout life [Saxena et al. 2020].

These changes in gene expression, obtained through DNA methylation and/or histone modifications, result from transcriptional regulatory influences of environmental factors, such as nutritional deficiencies, pollution, alcohol, cigarettes, certain foods containing phenols, immunological effects, and drugs [Saxena et al. 2020, Waye and Cheng 2017]. These changes may occur mainly during pregnancy due to the mother's exposure to risk factors. In the fetal phase, the brain is without a protective barrier and becomes more susceptible to the effects of external substances.

One of the best-known and most studied factors in the cause of ASD is valproic acid (VPA), a drug commonly used to treat epilepsy and bipolar disorders [Schiavi et al. 2019, Waye and Cheng 2017]. It targets histone deacetylase enzymes (HDACs). Histone deacetylases catalyze the removal of the acetyl group from histones.

They act in transcriptional activation. HDACs are involved in the development and progression of several diseases, which has led to research in recent years both in the clinical and physiological fields [Tartaglione et al. 2019] and in structural issues to understand which factors are important for inhibiting the catalytic activity of these enzymes. For these reasons, molecules that bind to HDAC have been studied.

In a recent study, [Sixto-López et al. 2020] analyzed the structural, molecular, and energetic characteristics through molecular docking and molecular dynamics simulation between valproic acid and different classes of histone deacetylases. They found that VPA binds to the hydrophobic active site channel, where it can act as a competitive inhibitor, preventing substrate binding with subsequent loss of enzyme activity. It reaches the catalytic site in HDAC 1-3 and 7. In HDAC 4 and 8, it binds to other sites. On the other hand, in HDAC6, VPA reaches only the catalytic tunnel. This explains why VPA preferentially inhibits class I and IIa over class IIb.

In this study, we used molecular docking approaches to verify the binding between HDAC and small ligands. We performed two case studies: H1 (Histone deacetylase 1 - HDAC1) and H2 (Histone deacetylase 2 - HDAC2). We chose class I histone deacetylases (HDAC1 and HDAC2) to study because they are involved in embryonic development. We then extracted features from the formed complexes using a graph-based structural signature algorithm. Structural signatures are techniques capable of converting three-dimensional structures of macromolecules into numerical vectors, obtaining differences and similarities between them. Structural signatures have been used to represent structures of protein-peptide complexes [Martins et al. 2023], detect mutations in enzymes used for biofuel production [Mariano et al. 2019], and analyze mutations in SARS-CoV-2 variants [Moreira et al. 2024]. Finally, we used several machine learning algorithms to build a model that differentiates natural and nonreal ligands (decoys).

2. Material and Methods

2.1. Data Collection

Protein structures were collected from the Protein Data Bank (PDB). Two structures were collected from the PDB: 4BKX [Millard et al. 2013] for the case study H1 and 4LXZ [Lauffer et al. 2013] for H2.

The ligands were obtained from the MUBD-HDAC database provided by [Xia et al. 2015]. This database contains data on ligands, decoys, and receptors of various HDACs. It has a confidence score ≥ 4 and $IC_{50} \leq 1 \mu M$. The decoys follow the DUD-E methodology with $T_c > 0.75$, with 39 for each ligand. For the case study H1, we collected 184 ligands and 7,023 decoy molecules. For the case study H2, we collected 62 ligands and 2,457 decoy molecules.

2.2. Molecular Docking

2.2.1. Docking Protocol

Defining the docking protocol to be used was challenging due to the zinc atom in the binding site. Histone deacetylases are metalloproteins whose enzymatic reaction is catalyzed by the interaction of the substrate with the zinc ion present at the catalytic site [Tao and Cheng 2020]. It is important to predict the interaction of the ligands with zinc.

To do so, it is necessary to use a protocol that extends the AutoDock force field to include force maps capable of describing the energetic and geometric components of these interactions [Santos-Martins et al. 2014]. Thus, we used the zinc specific protocol and performed a re-docking experiment to establish the docking protocol used in the automated process (Figure 1).

The receptor was initially pre-prepared using the PyMOL tool. We selected one of the chains for both structures and extracted the ligands. We then removed the water and other molecules and added hydrogens. The ligands for re-docking were obtained directly from the smiles format in the PDB and saved in an smi file with all the re-docking ligands. ACT was the ligand for histone deacetylases HDAC1 (H1), and SHH was the ligand for the histone deacetylase HDAC2 (H2).

In the re-docking experiment, the original ligand of the protein is extracted from the 3D structure file. Then, a similar molecule is designed using Avogadro [Hanwell et al. 2012]. This is done to ensure that the coordinates of the original ligand do not bias the result. Then, a docking protocol is applied to obtain the binding poses of the ligand relative to the receptor. Finally, the result is compared with the position of the original ligand using the RMSD (Root Mean Square Deviation) and visual analysis. The formula for RMSD for 3D molecular coordinates is given by:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N [(x_i - x'_i)^2 + (y_i - y'_i)^2 + (z_i - z'_i)^2]}$$

where N is the total number of atoms in the molecules, x_i, y_i, z_i are the $x, y,$ and z coordinates of atom i in the first set of coordinates, x'_i, y'_i, z'_i are the $x, y,$ and z coordinates of atom i in the second set of coordinates, and $\sum_{i=1}^N$ represents the summation of the squared differences for $i = 1$ to N .

The lower the RMSD value, the better the docking protocol can model the protein's real representation. We also consider a visual analysis of the structural overlap between the original ligand and the re-docked ligand. If the binding pose of the re-docked molecule is not similar to the original, the protocol is modified, and the process is repeated. Finally, the protocol that best represents the original pose is used in the next docking steps.

2.2.2. Virtual Screening

Figure 2 summarizes the virtual screening process and the following steps.

The ligands and decoys were extracted from the MUBD.HDAC database in the format sdf. The structure was converted to smiles format and saved in the smi format. For H1, we collected 184 ligands and 7,020 decoys. For H2, we collected 62 ligands and 2,457 decoys.

The 2D and 3D structures were obtained using MarvinView and Avogadro, respectively. In MarvinView, the isomer with the highest percentage at pH 7.4 was selected, and then the tautomer with the highest predominance at pH 7.4 was saved in the mol2

Docking protocol validation using re-docking

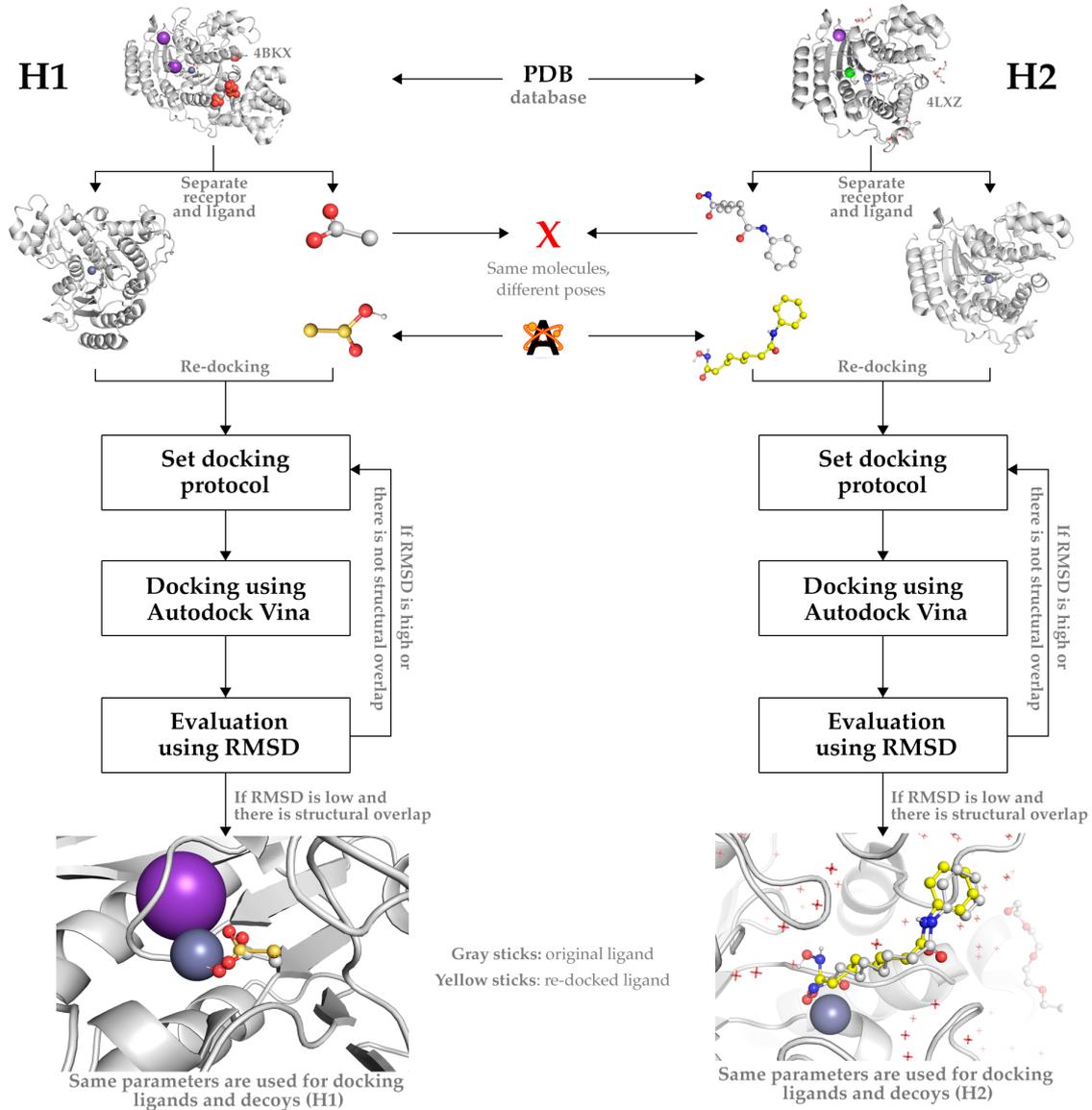


Figure 1. Protocol Validation using Re-docking. The workflow illustrates the two datasets selected (H1, left; and H2, right). After the collection of the PDB files, the ligands were separated from the receptors, and both were submitted to a re-docking protocol. This protocol uses Autodock Vina, with evaluation using the RMSD values. If these values were high or there was no structural overlap, we iterated through these steps again. If the experiment was successful, the parameters were used for the other experiments.

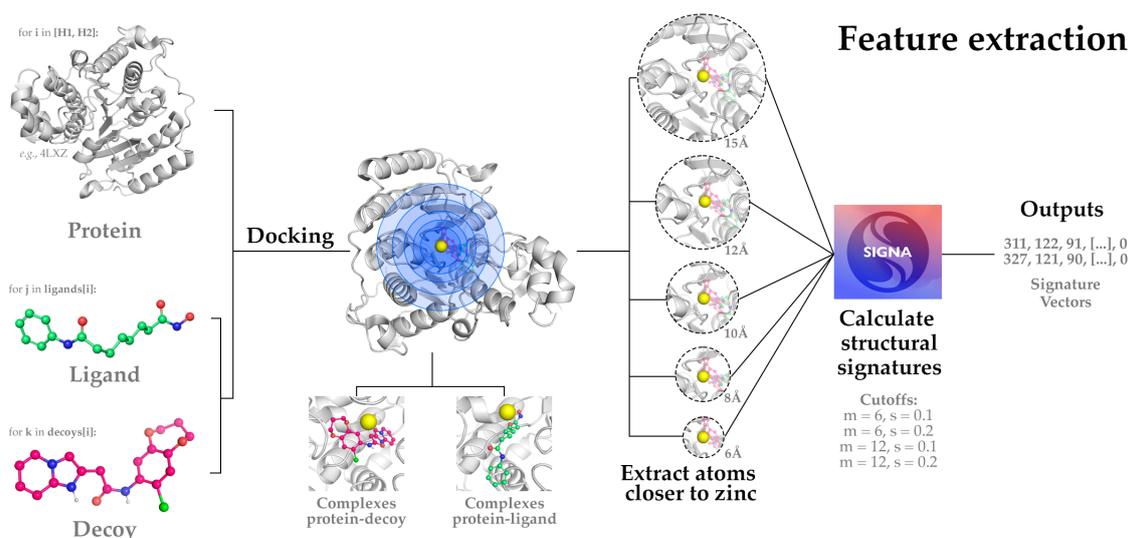


Figure 2. Methodology for molecular docking and feature extraction using graph-based signatures. The proteins and the ligands or decoys were docked, resulting in protein-decoy or protein-ligand complexes. Then, atoms close to zinc were extracted using different cutoff values (6, 8, 10, 12, and 15), and submitted to the SIGNA library to calculate the structural signatures. H1 (PDB ID: 4BKX); H2 (PDB ID: 4LXZ); H1-ligands (n=184); H1-decoys (n=7,020); H2-ligands (n=62); H2-decoys (n=2,457).

format. Then, the structure was opened in Avogadro. The force field was set to MMFF94. The optimization geometry was set to 500. We used the SteepestDescent algorithm and convergence $10 e^{-7}$. A conformer search was performed using the random rotor search and a maximum of 100 conformers. The structure was saved in a mol2 file.

Docking was performed using the specific tutorial for zn metalloproteins (https://autodock-vina.readthedocs.io/en/latest/docking_zinc.html). The preparation of the ligands and decoys was performed using the `prepare_ligand4.py` script. The receptor preparation was initially done by transforming the file into `.pdbqt`, followed by the addition of tetrahedral zinc pseudoatoms (TZ) around the zinc in the receptor, and saved as `receptor_tz.pdbqt`, using the `prepare_receptor4.py` and `zinc_pseudo.py` files.

In the third step, force maps for the receptor are created. For this step, the ligand, receptor, grid (60,60,60), and box (center of mass) data calculated previously are provided. The output is a `.gpf` file that records the potentials of unbounded pairs, among other information. Finally, we encode the force field information into the affinity maps. For these two steps, the `prepare_gpf4zn.py` and `autogrid4.py` scripts were used, respectively.

Finally, docking was performed, informing the affinity maps and `ad4` scoring as parameters. The results were saved in `.pdbqt` and `.log` files for each receptor-ligand/decoy set. The first pose of each docking and complexed with the receptor was selected and saved in a `pdb` file. After virtual screening, we selected the first binding pose defined by the Vina tool and joined it to the receptor protein chain. Finally, we calculate the structural signature vector of the complex formed to extract the features that would be used in the machine-learning steps.

2.3. Structural Signatures Calculation

For each complex obtained in the previous step, several structural signatures were calculated using SIGNA (<https://github.com/LBS-UFMG/signa>). To generate signatures, a study was initially carried out with the protein complex, where signatures were generated with a cutoff limit of 12 Å and a cutoff step of 0.1 Å. Then, it was submitted to machine learning models. Based on the result, it was decided to work with the cut-off of the receptor-pose complex. For this, the signatures were constructed only with residues close to the binding site. Atoms were extracted at a distance of 6, 8, 10, 12, and 15 Å from the zinc atom (Figure 2).

The structural signatures were extracted using the aCSM algorithm [Pires et al. 2013]. The following parameters were used: simple atomic signature, cutoff limit of 6 and 12 Å, and cutoff step of 0.1 and 0.2 Å. The structural signature vectors obtained were used to train machine learning models.

2.4. Machine Learning

Figure 3 summarizes the steps taken in this stage of the project. We used the signature vectors to build models to identify whether there are patterns that differentiate real ligands from decoys. To balance the databases, we selected the same number of decoy molecules as ligands. Hence, for H1, we selected 184 decoys, and for H2, we selected 62 decoys.

Classical machine learning models were built using the Python scikit-learn library (Figure 3). The following algorithms and parameters were used (Table 1): GaussianNB (Naive Bayes), RandomForestClassifier (Random Forest), ExtraTreesClassifier (Extra Trees), XGBClassifier (XGBoost), MLPClassifier (Multiperceptron) and KNeighborsClassifier (KNN). GaussianNB implements the Gaussian Naive Bayes algorithm, which assumes that the probability of the features is Gaussian. Random Forest, XGBoost, and Extra Trees are ensemble methods that use trees for classification. Both random forest and extra trees use random algorithms, differing in that Extra Trees presents an additional layer of randomness, while XGBoost relies on a parallelization layer. KNN relies on neighbors to classify, and MLP is an algorithm that uses the perceptron neuron. All models were trained using the default scikit-learn parameters. We divided the datasets into training (80%) and testing (20%) subsets. Training was performed using the cross-validation technique (k=5). Finally, we evaluated the results and discussed which models best represented the proposed problem.

3. Results and Discussion

3.1. Docking Results

To evaluate the results, we manually checked some outputs of the ligand and decoy docking. In all results, the adopted protocol performed the docking in the correct position [Santos-Martins et al. 2014]. Even the decoys, which we expected not to fit perfectly in the binding site, were allocated in a position close to zinc. This happens because the docking protocol forces the ligand to fit in the region specified by the box. Some work described in the literature suggests that the presence of a chelation of the zinc atom is important to be a ligand [Wang et al. 2005, Drakontaeidi and Pontiki 2023]. However, we expected that the binding energy predicted by the docking scoring algorithm would be lower for the decoys and higher for the real ligands.

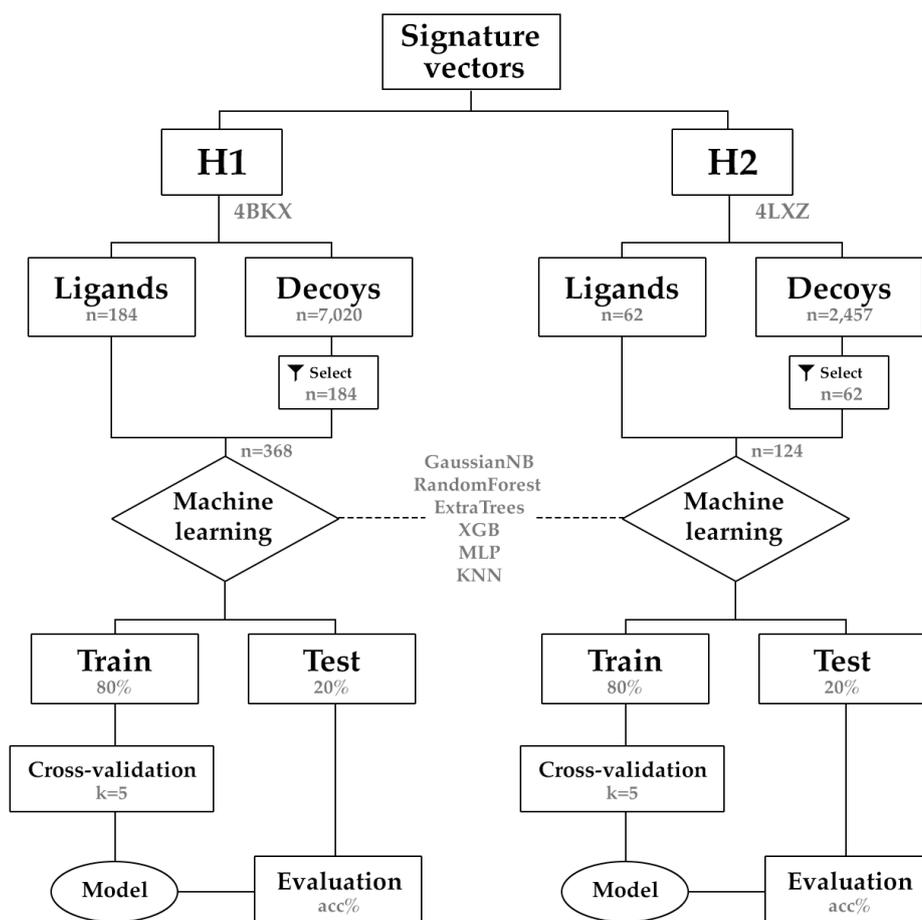


Figure 3. Models to detect other HDAC inhibitory molecules. After filtering of decoys to match the number of ligands for each dataset, the complexes signatures (n=368 and n=124 for H1 and H2, respectively) were evaluated using six different Machine Learning methods. For each method, datasets were divided between train and test.

Table 1. Main Models Parameters

Model	Main Parameters
GaussianNB	priors=None (adjusted according to the data), var. smoothing=1e-09
RF	num. estimators=100, criterion='gini', max.depth=None, nodes=None, bootstrap=True, min. samples split=2, min. samples leaf=1, max features='sqrt'
ExtraTrees	num. estimators=100, criterion='gini', max.depth=None, nodes=None, bootstrap=None, min. samples split=2, min. samples leaf=1, max features='sqrt'
XGB	objective='binary:logistic', loss='log_loss', num. estimators=100, criterion='friedman mse', max.depth=3, subsample=1.0, min. samples split=2, min. samples leaf=1, learning rate=0.1, max features=None
MLP	hidden layer sizes=100, activation='relu', solver='adam', alpha=0.0001, batch size='auto' (min=200), learning rate='constant', learning rate init=0.001
KNN	neighbors=5, weights='uniform', algorithm='auto', leaf size=30, p=2, metric='minkowski'

For example, for ligand 1 of H1 (Figure 4A), AutoDock Vina predicted a binding

affinity of -14.91 kcal/mol. Meanwhile, for decoy 1 (Figure 4B) from the same receptor (PDB ID: 4BKY), the predicted binding energy was -11.29 kcal/mol. Note that the lower the binding energy, the higher the affinity. Therefore, in this case, the ligand obtained a higher affinity, as expected. It is also observed that the ligand is probably capable of chelating zinc due to the presence of two oxygens at the end of the molecule, while the decoy offers only a possible interaction with zinc, but no possibility of chelation.

For H2 ligand 1 (Figure 4C), AutoDock Vina predicted a binding affinity of -8.604 kcal/mol. Meanwhile, for decoy 1 (Figure 4D) of the same receptor (PDB ID: 4LXZ), the predicted binding energy was -10.15 kcal/mol. In this case, the decoy obtained a higher affinity, which is different from what was expected. This may indicate that some decoys may have a higher affinity than ligands described in the literature.

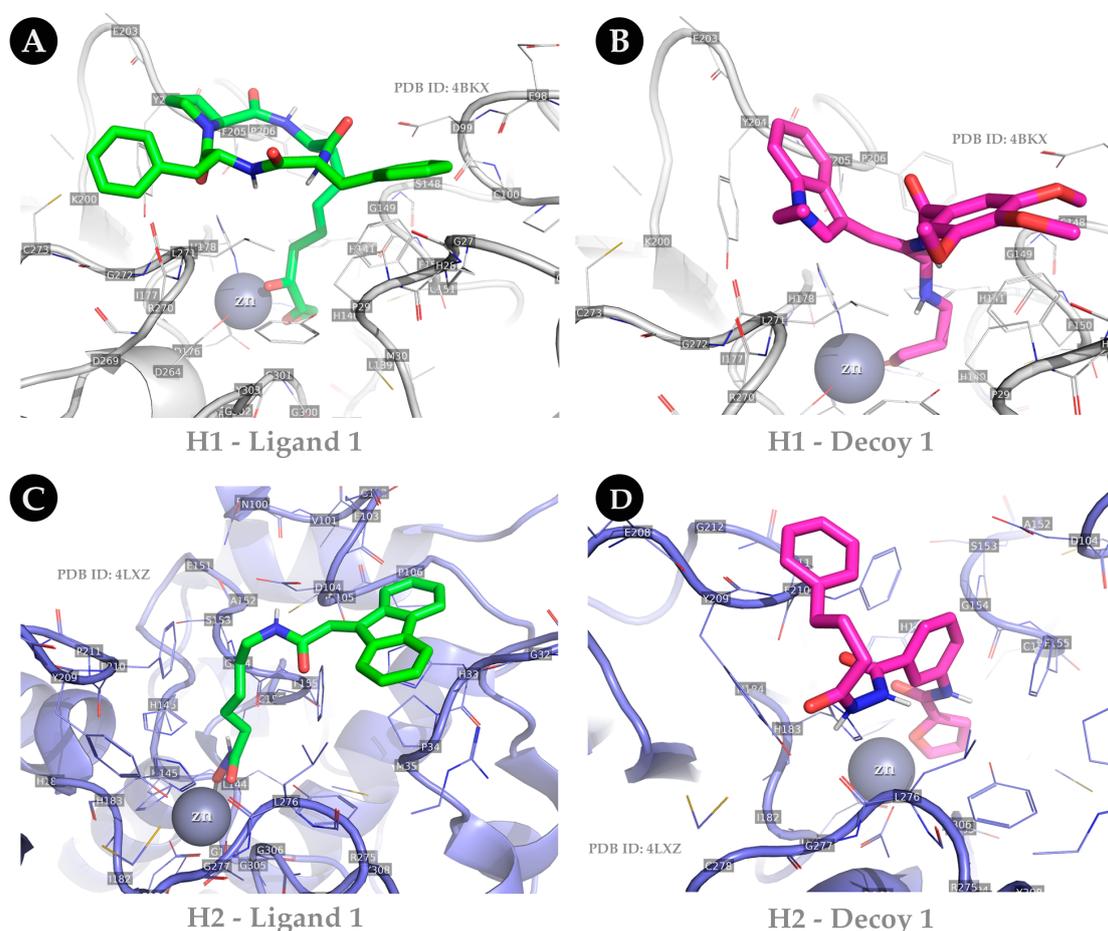


Figure 4. Example of ligands (green sticks) and decoys (magenta sticks) docked at the histone deacetylase site. (A) Ligand docked to H1; (B) Decoy docked to H1; (A) Ligand docked to H1; (B) Decoy docked to H1.

However, we cannot measure our experiment's unexpected result by evaluating a single result. To do this, we checked the binding energy of all ligands and all decoys. The docking results indicate that, on average, the ligands had a lower surface energy than the decoys in both case studies. A lower affinity energy represents a stronger binding. Thus, these results corroborate our expectations that real ligands would bind better to the receptor (Figure 5).

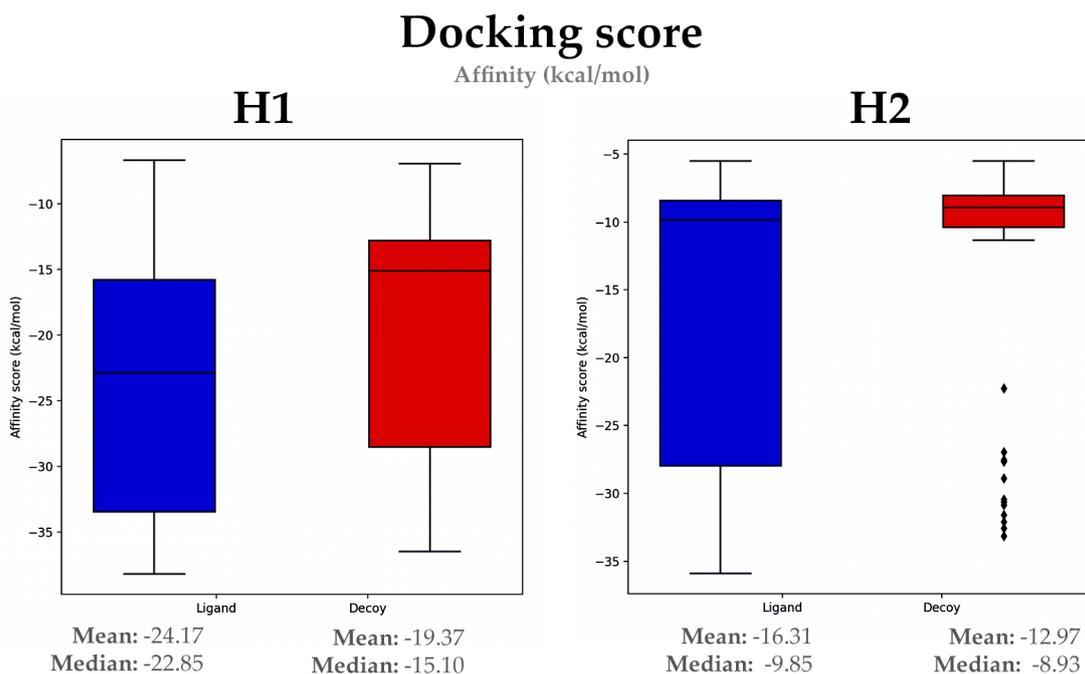


Figure 5. Boxplots for the affinity energy calculated by Autodock Vina software. Blue bars show the ligand affinity score values (in kcal/mol) and red bars show the decoy values. The lower the value, the stronger the binding. Left: H1; Right: H2. The mean and median values for each bar are shown at the bottom. The experiments were statistically significant (details included in the supplementary material).

3.2. Evaluating Constructed Models

We analyzed the ROC AUC score of the models and the score model and test accuracy to select which signatures and cavities (Figure 6) would have the best results. We can observe that distances from the zinc atom of 15 Å for H1 and 10 Å for H2 zinc presented the best results, indicating that cutoff points very close to the zinc atom, such as 6 angstroms, do not seem to be adequate to separate the ligand from the decoys. Regarding the signature, the results were better for the cutoff limit of 12 Å with a step of 0.2 Å for both. In the case of H1, there was a predominance of the cutoff of 12 Å while for H2, the cutoff limit of 12 Å with a step of 0.2 presented a good result, but very close to the cutoff limit of 6 Å with a step of 0.1. It is worth noting that in the case of H1, the result for the cutoff threshold 12Å with a step of 0.2 was very close to that of 0.1, indicating that, in this case, there was no significant difference.

Analyzing the model data for the best results, it was observed that the model trained with the Random Forest and ExtraTrees were the best models for H1 and H2. Both are models that are based on randomness, and both presented good results for H1 and H2, which indicates that these conditions seem to help in classification. For H1, the model built with Random Forest obtained an accuracy of 0.918 in cross-validation and 0.716 in the test, being very close to Extra Tree, but with higher accuracy (Table 2). Interestingly, MLP achieved higher accuracy in the test (0.770) than in the cross-validation (0.718).

For H2, the model built with ExtraTrees obtained an accuracy of 0.889 in cross-validation and 0.720 in the test. Random Forest obtained higher accuracy in cross-

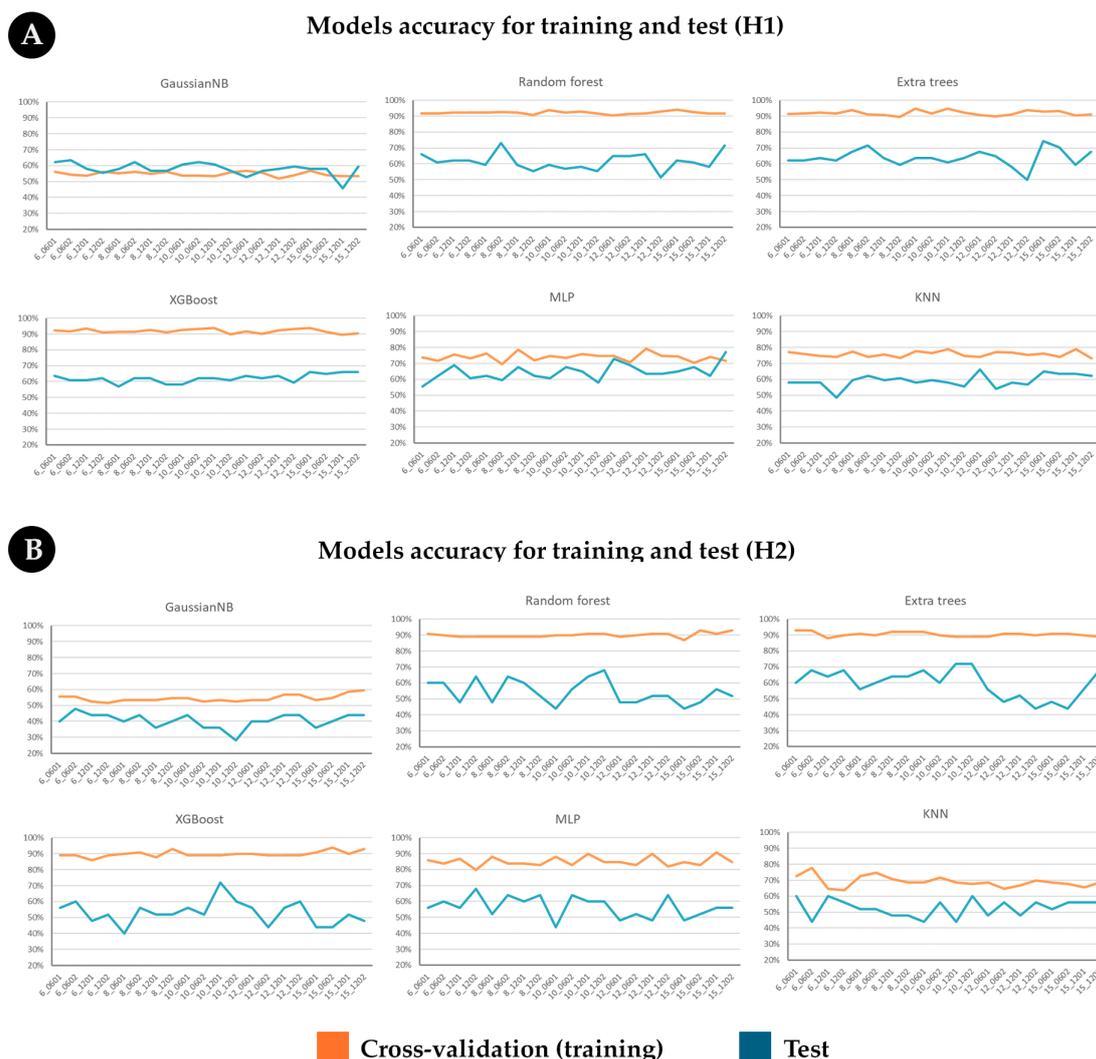


Figure 6. Accuracy for training and test datasets for H1 (A) and H2 (B). Orange lines indicate the accuracy obtained in the cross-validation in different parameters. Blue lines indicate the accuracy obtained in the tests. Y-axis: accuracy; X-axis: different signature parameters (size of the extracted region, cutoff-max, and cutoff-step).

Table 2. Machine Learning models for H1. Parameters - Cutoff ray: 15; cutoff max: 12; cutoff step: 0.2

Model	Accuracy (training)	Accuracy (test)	Precision	Recall	MCC	ROC AUC
GaussianNB	0.534	0.595	0.593	0.457	0.182	0.647
RF	0.918	0.716	0.719	0.657	0.430	0.786
ExtraTrees	0.912	0.676	0.657	0.657	0.350	0.767
XGB	0.905	0.662	0.639	0.657	0.324	0.719
MLP	0.718	0.770	0.781	0.714	0.539	0.771
KNN	0.731	0.622	0.733	0.314	0.263	0.730

validation than ExtraTrees; however, it performed worse in the test accuracy (Table 3).

It is important to note that our experiments have some limitations. First, we only use the first pose returned by the AutoDock Vina docking tool. This was done due to the

Table 3. Machine Learning models for H2. Parameters - Cutoff ray: 12; cutoff max: 12; cutoff step: 0.2

Model	Accuracy (training)	Accuracy (test)	Precision	Recall	MCC	ROC AUC
GaussianNB	0.525	0.280	0.467	0.412	-0.560	0.360
RF	0.909	0.68	1.00	0.529	0.514	0.849
ExtraTrees	0.889	0.720	0.917	0.647	0.487	0.750
XGB	0.899	0.60	0.818	0.529	0.263	0.757
MLP	0.848	0.60	1.00	0.412	0.428	0.787
KNN	0.677	0.60	1.00	0.412	0.428	0.713

difficulty of dealing with multiple docking poses in machine-learning modeling. Despite the challenges, using various poses could provide a more comprehensive evaluation, so we intend to perform other experiments using multiple docking poses in the future. Also, the choice of the best pose is based on the affinity energy value returned by the docking tool. The scoring function varies according to the docking tool used and may not correspond to the best docking position. An alternative would be reproducing the experiments using different docking tools.

Another problem is the difference in accuracy values between cross-validation and testing in several models. This may be an indicator of overfitting. To deal with this problem, testing other types of signatures that extract more features or even test different machine learning algorithms may be necessary. We intend to do this in future work.

Finally, the constructed models can be used in future experiments to detect new ligands for histone deacetylases. However, further studies and experimental validations are needed to draw better conclusions.

4. Conclusion

Using a signature based on the classical aCSM model, we obtained an accuracy of 72% in both tests. The aCSM signature uses only the coordinates of atoms to build the representative graph. This tends to be lower than more complex signature models, such as aCSM-all, which also consider atomic types. Our results indicate that geometric complementarity is sufficient to obtain a high accuracy. Thus, we could obtain a better result if we consider atomic types. We hope to redo these experiments in the future, considering the atomic types of the ligands and decoys. In this way, we believe that we would have a more representative signature of the protein binding region, which would allow the development of a simple model for identifying potential HDAC inhibitors and, consequently, substances capable of causing an autism spectrum disorder.

Acknowledgements The authors would like to thank the research funding agencies CAPES, FAPEMIG, and CNPq. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Supplementary Material Available at <https://github.com/LBS-UFMG/HDAC-docking>

References

Drakontaeidi, A. and Pontiki, E. (2023). A review on molecular docking on hdac isoforms: Novel tool for designing selective inhibitors. *Pharmaceuticals*, 16(12):1639. 10.3390/ph16121639.

- Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4:1–17.
- Lauffer, B. E., Mintzer, R., Fong, R., Mukund, S., Tam, C., Zilberleyb, I., Flicke, B., Ritscher, A., Fedorowicz, G., Vallerio, R., et al. (2013). Histone deacetylase (hdac) inhibitor kinetic rate constants correlate with cellular histone acetylation but not transcription and cell viability. *Journal of Biological Chemistry*, 288(37):26926–26943.
- Maenner, M. J. et al. (2020). Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network. *MMWR. Surveillance Summaries*, 69(4):1–12. 10.15585/mmwr.ss6904a1.
- Mariano, D., Santos, L., Machado, K., Werhli, A., de Lima, L., and de Melo-Minardi, R. (2019). A computational method to propose mutations in enzymes based on structural signature variation (ssv). *International journal of molecular sciences*, 20(2):333.
- Martins, P., Mariano, D., Carvalho, F. C., Bastos, L., Moraes, L., Paixão, V., and de Melo-Minardi, R. (2023). Propedia v2.3: A novel representation approach for the peptide-protein interaction database using graph-based structural signatures. *Frontiers in Bioinformatics*, 3:1103103.
- Millard, C. J., Watson, P. J., Celardo, I., Gordiyenko, Y., Cowley, S. M., Robinson, C. V., Fairall, L., and Schwabe, J. W. (2013). Class i hdacs share a common mechanism of regulation by inositol phosphates. *Molecular cell*, 51(1):57–67.
- Moreira, E. U., Mariano, D. C., and de Melo-Minardi, R. C. (2024). Computational analysis of mutations in sars-cov-2 variants spike protein and protein interactions. In *Features, Transmission, Detection, and Case Studies in COVID-19*, pages 123–139. Elsevier.
- Pires, D. E. V. et al. (2013). acsm: noise-free graph-based signatures to large-scale receptor-based ligand prediction. *Bioinformatics*, 29(7):855–861. 10.1093/bioinformatics/btt058.
- Santos-Martins, D. et al. (2014). Autodock4zn: An improved autodock force field for small-molecule docking to zinc metalloproteins. *Journal of Chemical Information and Modeling*, 54(8):2371–2379. 10.1021/ci500209e.
- Saxena, R. et al. (2020). *Role of environmental factors and epigenetics in autism spectrum disorders*, page 35–60. Elsevier. 10.1016/bs.pmbts.2020.05.002.
- Schiavi, S. et al. (2019). Reward-related behavioral, neurochemical and electrophysiological changes in a rat model of autism based on prenatal exposure to valproic acid. *Frontiers in Cellular Neuroscience*, 13. 10.3389/fncel.2019.00479.
- Sixto-López, Y. et al. (2020). Exploring the inhibitory activity of valproic acid against the hdac family using na mmgsba approach. *J Comput Aided Mol Des.*, 34(8):85–878. 10.1007/s10822-020-00304-2.
- Tao, H. N. and Cheng, H. Y. (2020). Docking belinostat into hdac 8 using autodock tool. *Can Tho University Journal of Science*, 12(2):1–8. 10.22144/ctu.jen.2020.009.
- Tartaglione, A. M. et al. (2019). Prenatal valproate in rodents as a tool to understand the neural underpinnings of social dysfunctions in autism spectrum disorder. *Neuropharmacology*, 159:107477. 10.1016/j.neuropharm.2018.12.024.
- Wang, D. et al. (2005). Toward selective histone deacetylase inhibitor design: homology modeling, docking studies, and molecular dynamics simulations of human class i histone deacetylases. *Journal Medicinal Chemistry*, 48(22):6936–6947. 10.1021/jm0505011.
- Waye, M. M. Y. and Cheng, H. Y. (2017). Genetics and epigenetics of autism: A review. *Psychiatry and Clinical Neurosciences*, 72(4):228–244. 10.1111/pcn.12606.
- Xia, J. et al. (2015). Comparative modeling and benchmarking data sets for human histone deacetylases and sirtuin families. *Journal of Chemical Information and Modeling*, 55(2):374–388. 10.1021/ci5005515.