JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Application Note

AlzyFinder: A Machine-Learning-Driven Platform for Ligand-Based Virtual Screening and Network Pharmacology

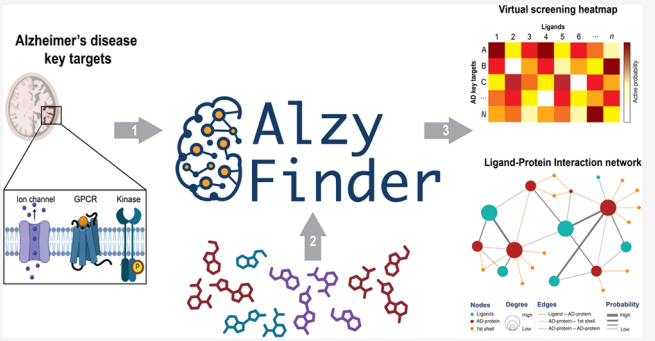
Jessica Valero-Rojas,[+] Camilo Ramírez-Sánchez,[+] Laura Pacheco-Paternina, Paulina Valenzuela-Hormazabal, Fernanda I. Saldivar-González, Paula Santana, Janneth González, Tatiana Gutiérrez-Bunster, Alejandro Valdés-Jiménez, and David Ramírez*

Cite This: https://doi.org/10.1021/acs.jcim.4c01481

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Alzheimer's disease (AD), a prevalent neurodegenerative disorder, presents significant challenges in drug development due to its multifactorial nature and complex pathophysiology. The AlzyFinder Platform, introduced in this study, addresses these challenges by providing a comprehensive, free web-based tool for parallel ligand-based virtual screening and network pharmacology, specifically targeting over 85 key proteins implicated in AD. This innovative approach is designed to enhance the identification and analysis of potential multitarget ligands, thereby accelerating the development of effective therapeutic strategies against AD. AlzyFinder Platform incorporates machine learning models to facilitate the ligand-based virtual screening process. These models, built with the XGBoost algorithm and optimized through Optuna, were meticulously trained and validated using robust methodologies to ensure high predictive accuracy. Validation included extensive testing with active, inactive, and decoy molecules, demonstrating the platform's efficacy in distinguishing active compounds. The models are evaluated based on balanced accuracy, precision, and F1 score metrics. A unique soft-voting ensemble approach is utilized to refine the classification process, integrating the strengths of individual models. This methodological framework enables a comprehensive analysis of interaction data, which is presented in multiple formats such as tables, heat maps, and interactive Ligand−Protein Interaction networks, thus enhancing the visualization and analysis of drug−protein interactions. AlzyFinder was applied to screen five molecules recently reported (and not used to train or validate the ML models) as active compounds against five key AD targets. The platform demonstrated its efficacy by accurately predicting all five molecules as true positives with a probability greater than 0.70. This result underscores the platform's capability in identifying potential therapeutic compounds with high precision. In conclusion, AlzyFinder's innovative approach extends beyond traditional virtual screening by incorporating network pharmacology analysis, thus providing insights into the systemic actions of drug candidates. This feature allows for the exploration of ligand−protein and protein−protein interactions and their extensions, offering a comprehensive view of potential therapeutic impacts. As the first open-access platform of its kind, AlzyFinder stands as a valuable resource for the AD research community, available at http://www.alzyfinder-platform.udec.cl with supporting data and scripts accessible via GitHub https://github.com/ramirezlab/AlzyFinder.



INTRODUCTION

Nowadays, the prevalence of neurodegenerative diseases (NDD) is increasing, in part due to modern living standards which extend life length and contribute to the world population's aging.[1] Among NDDs, Alzheimer's disease (AD) is the most notable form of dementia in the world, and the one that most affects the elderly,[2] making it a worldwide public health problem. Currently, there are few therapeutic alternatives for AD (donanemab, lecanemab, aducanumab, donepezil, rivastigmine, galantamine, and memantine).[3−5] However, multiple clinical trials (Phase I − IV) of different drugs and bioactive compounds are being carried out.[6] Among the drug-targets in clinical trials, a polypharmacological profile can be observed. For instance, vorinostat (histone deacetylase (HDAC) inhibitor), an FDA-approved drug indicated for the treatment of cutaneous manifestations in patients with T-cell lymphoma[7] has proven to reverse memory deficits in a mouse model with AD;[8] therefore, it is being repositioned as anti-AD

Special Issue: Machine Learning in Bio-cheminformatics

Received: August 17, 2024
Revised: October 16, 2024
Accepted: October 21, 2024

(ClinicalTrials.gov Identifier: NCT03056495). Dasatinib (tyrosine kinase inhibitor) is currently approved for the treatment of lymphoblastic or chronic myeloid leukemia;[9] however, its effects on neuroinflammation mediated by microglial and/or astrocytic responses[10] has led to testing its effect in modulating AD progression (ClinicalTrials.gov Identifier: NCT04063124). Polypharmacological profiles of drug/targets could be used for drug repurposing, identifying, and validating new targets, and finding new bioactive ligands, among other applications.

In this study, we introduce the AlzyFinder Platform, a unique web-based tool designed to perform virtual screening that uses an array of machine learning models built for over 85 key targets associated with Alzheimer's disease. The platform's user-friendly interface facilitates the execution of multiple virtual screening tasks, utilizing ligand-based drug design approaches. Results are showcased through drug−protein interaction networks (DPIs), which are instrumental in decoding the impact of potential therapeutics on biological systems by delineating the interactions between screened ligands and Alzheimer's disease-related targets. To refine the displayed results, users can apply probability thresholds. While the models generate probabilities ranging from 0 to 1, selecting a threshold does not alter the predictions, but rather adjusts the data to be displayed based on the selected criteria. For instance, if a user selects a cutoff of 0.7, the server will display only the results with probabilities equal to or greater than 0.7. Uniquely, the AlzyFinder Platform is the first of its kind to be offered openly to the community, encapsulating machine learning models that have been rigorously tested and validated through different methodologies for virtual screening. AlzyFinder Platform aims to bridge this gap in AD drug design by providing a comprehensive tool for multitarget ligand identification. This unique free web-based tool not only predicts interactions of candidate molecules with essential targets but also, from a systems pharmacology view, allows to identify potential multitarget ligands, and to unravel the systemic actions of these ligands mediating their therapeutic effects in Alzheimer's disease. AlzyFinder Platform is freely accessible at http://www.alzyfinder-platform.udec.cl, where users can explore its functionalities. It is licensed under a Creative Commons Attribution 4.0 International. Additionally, all the models, as well as the scripts and other required specifications to use this platform in standalone format are available in the AlzyFinder GitHub repository available at https://github.com/ramirezlab/AlzyFinder.

SELECTION OF ALZHEIMER'S KEY PROTEINS

To find key targets in Alzheimer's disease (AD) we used Open Target Platform v23.02 (release 22 February 2023).[11] We search all targets related to AD under the Experimental Factor Ontology (EFO): MONDO_0004975 and/or Medical Subject Headings (MeSH): D000544 terms. Then, the *Associated Targets* option in the Open Target Platform was used to map all AD associated targets. Next, with the AD targets identified, we proceeded to search for all compounds with reported biological activity in the ChEMBL database (v32).

We use the pChEMBL value to classify interacting compounds as active (1) or inactive (0) according to established activity thresholds. pChEMBL expresses the potency of molecules determined from one of several semicomparable values from the ChEMBL database and is defined as the negative $\log_{10}$ (in molar concentration) of the IC_{50}, XC_{50}, EC_{50}, AC_{50}, K_i, K_d or potency.[12] To classify compounds as active or inactive we use the Illuminating the Druggable Genome Knowledge Management Center (IDG-KMC),[13] where a compound is considered active against an ion channel, G protein-coupled receptor (GPCR), nuclear receptor or kinase, if the pChEMBL is greater than, 5.00, 7.00, 7.00., and 7.52 respectively. If the target of a given compound is not in the above-mentioned target classification, it enters to the Non-IDG Family Targets classification, and to be considered active it must have a pChEMBL greater than 6.0. To classify compounds as inactive, we selected activity thresholds with 1-fold respect to the adobe active classification (Table 1). The SMILES of the compounds were filtered to remove salts using the RDKit Salt Stripper (https://www.rdkit.org/) node in KNIME.[14] To handle duplicated data, we applied a filter to select the molecules with the highest pChEMBL value. We also focused on single proteins, excluding protein complexes and molecules with missing pChEMBL values from the data set. Table 1 shows initial thresholds selected to build the ML models.

Table 1. Compound Bioactivity Classification According to the IDG[a]

	IC_{50}, EC_{50}, XC_{50}, AC_{50}, K_i, K_d or potency (M)		pChEMBL	
protein families	active (≤)	inactive (>)	active (≤)	inactive (>)
kinases	3×10^{-8}	3×10^{-7}	7.52	6.52
GPCRs	1×10^{-7}	1×10^{-6}	7.00	6.00
ion channels	1×10^{-5}	1×10^{-4}	5.00	4.00
non-IDG family targets	1×10^{-6}	1×10^{-5}	6.00	5.00

[a]IDG: Illuminating the Druggable Genome.[13] M: Molar. GPCRs: G-protein-coupled receptor. IC_{50}: Half maximal inhibitory concentration. EC_{50}: Half maximal effective concentration. XC_{50}: Stands in for either IC_{50} or EC_{50}, whichever is pertinent to that target and assay. AC_{50}: Concentration at which half maximal activity is observed. K_i: Inhibition constant. K_d: Equilibrium dissociation constant. pCHEMBL: −Log(molar IC_{50}, XC_{50}, EC_{50}, AC_{50}, K_i, K_d, or potency).

We then selected the proteins that had at least 100 compounds reported in each category (active and inactive) with respect to the defined activity thresholds (Table 1) and subsequently calculated the active: inactive ratio. Finally, for the final selection of proteins and ML model building, we kept the compounds that maintained a maximum ratio of 3:1 between both labels (active 1 and inactive 0) to avoid imbalance between active and inactive data and to improve the quality of the ML models. For some targets, which initially had unbalanced data between active and inactive, we manually moved the activity thresholds to better balance the data, but always maintaining a 1-fold factor for both active and inactive categories, and a ratio lower than 3:1 between the categories For more details about how the data was balanced see the Supporting Information - section "Compound data preparation for machine learning (ML) models".

DEVELOPMENT AND VALIDATION OF MODELS BASED ON MACHINE LEARNING

In the present work, and following the methodology described in the selection of key proteins in Alzheimer's disease, 85 proteins were selected. Among these proteins, targets widely studied and validated as key in AD are included, such as acetyl cholinesterase (ACHE),[15] beta-secretase 1 (BACE1),[16]

https://doi.org/10.1021/acs.jcim.4c01481
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

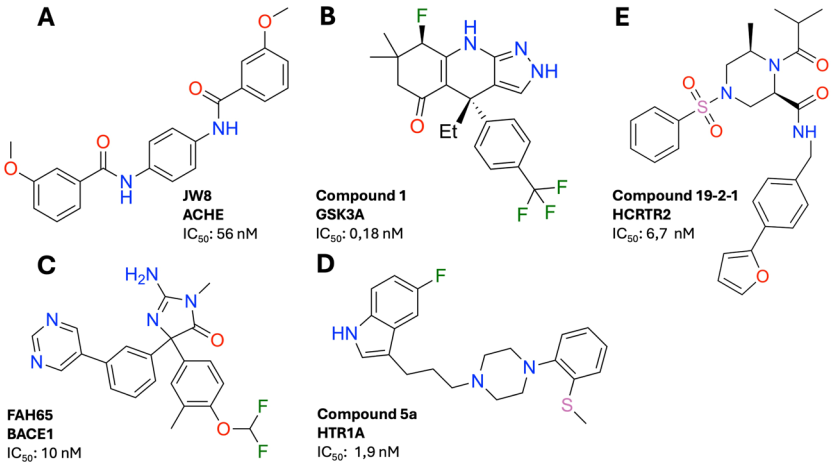


Figure 1. Selected molecules to evaluate the AlzyFinder Platform. (A) JW8,[21] (B) Compound **1**.[22] (C) FAH65.[23] (D) Compound **5a**,[24] and (E) Compound 19-2-1[25] active against ACHE (acetylcholinesterase), GSK3A (glycogen synthase kinase-3 alpha), BACE-1 (beta-secretase 1), HTR1A (5-hydroxytryptamine receptor 1A), and HCRTR2 (orexin receptor type 2), respectively.

amyloid-β precursor protein (APP),[15] glycogen synthase kinase-3 beta (GSK3β),[17] among others. Supplemental Table S1 shows the complete list of the selected AD key proteins, as well as the number of active, intermediate, and inactive compounds extracted from the ChEMBL database based on their pChEMBL value.

For each AD key protein, using the XGBoost classification and regression algorithm,[18] three different ML models were built and then evaluated based on balanced accuracy, precision and F1 score metrics, respectively. To validate the performance of the models, Supplementary Table S2 presents the metrics used for training and data set as well as the data of the confusion matrix. As an example, the validation of the ML model built for BACE1 protein (Uniprot ID: P56817) evaluated with the precision metric is shown in Supplemental Figures S3 and S4

Finally, since each protein was modeled using three different optimized classifications (one per selected metric balanced accuracy, precision and F1), a fourth integrative model was developed by using a soft-voting method[19] implemented by calculating the median of the classification probabilities provided by each of the three independent models' results. In this ensemble, each of the three models contributes a vote weighted by its confidence in the classification, calculated from the probability of belonging to the assigned class. For example, if two models predict a molecule as 'active' with high certainty probabilities, and the third classifies it as 'inactive' with a lower probability, the soft-voting XGBoost ensemble computes an arithmetic median of these weighted probabilities to assign the final classification to 'active'. This method capitalizes on the specific information on each model and mitigates the risk of isolated errors from a single model, thus enhancing the robustness and accuracy of the final classification. By employing this soft voting approach, the ensemble model ensures that the final verdict is based on an informed and balanced consensus, significantly improving the reliability of predictions in practical applications.

SIMILARITY CHECK BETWEEN INPUT LIGANDS AND ML MODEL TRAINING DATA SET

To evaluate the confidence of the results, each prediction generated by the server is categorized into four quartiles (Q1−Q4) based on the average Jaccard distance between the compound to be screened and the corresponding data set used to train the ML model, as previously described in.[20] This categorization is detailed in Supplementary Table S3 and is computed using a Morgan fingerprint with a radius of 2 (2048 bits). The Jaccard distance J(A/B) between two compounds A and B is defined as follows:

$$J(A, B) = 1 - \frac{c}{a + b - c}$$

where a and b represent the number of "on" bits for compounds A and B, respectively, and c denotes the number of "on" bits common to both compounds. Compounds with mean Jaccard distances below the minimum value in Q1 or exceeding the maximum value in Q4 are placed in an additional category labeled "OUT," indicating that these compounds fall present low confidence in predictions due to lower similarity with training data set. An example of this classification is showing on Supplemental Figure S2.

LIGAND-BASED VIRTUAL SCREENING USING ALZYFINDER: CASE STUDY

The targets in the web-based tool are categorized into four families: G-protein-coupled receptors, ion channels, kinases, and a miscellaneous group, according to the Illuminating the Druggable Genome (IDG) initiative. To evaluate the efficacy of AlzyFinder's classification capabilities, we selected five molecules with biological activity against key targets in Alzheimer's disease (Figure 1). These molecules were published in late 2023 or 2024 and were not included in the original set of molecules to build any ML models.

To perform virtual screening, AlzyFinder platform allows users to submit the molecules of interest in three different forms: (i) by uploading a file containing the SMILES of each molecule (up to 100 molecules), (ii) pasting the SMILES into the provided text box (up to 10 molecules), (iii) or sketching the molecular structure directly on the platform (Supplemental Figure S5). To virtually screen the five, know molecules (Figure 1) and predict the probability of being active against ACHE, GSK3A, BACE1, HTR1A, and HCRTR2, it was necessary to obtain the SMILES representation for each molecule and introduce them into the test box on the AlzyFinder platform (Supplemental Figure S5B).

https://doi.org/10.1021/acs.jcim.4c01481
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

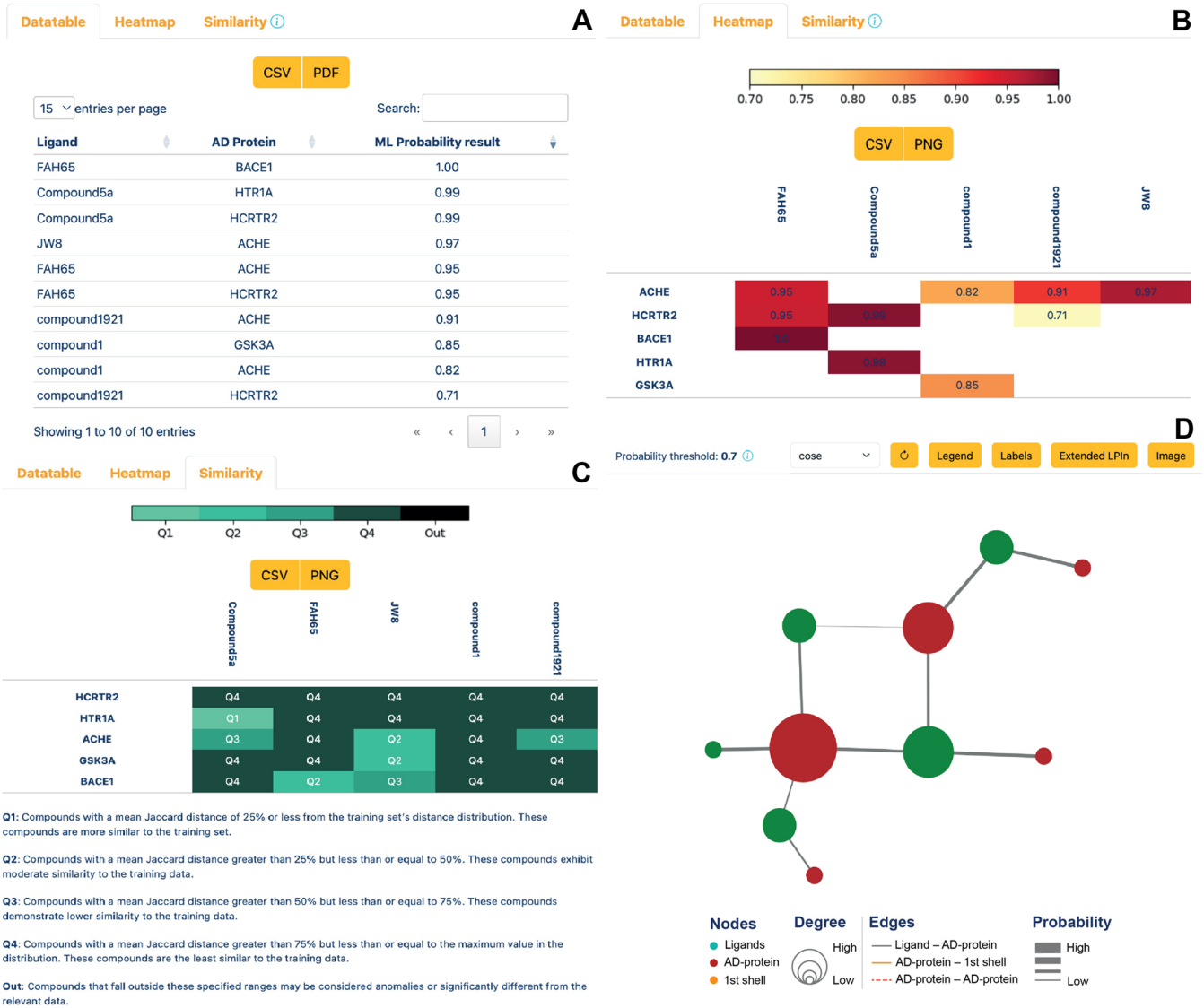


Figure 2. AlzyFinder virtual screening dynamic results. (A) Results are presented in a table that can be downloaded in CSV or PDF formats. (B) Graphical results as heat map displaying all proteins and ligands as well as the probability values. The heat map could be downloaded in CSV or PNG formats. (C) Similarity matrix for each compound predicted using the training sets. This information provides the user with insight into the similarity of the input compounds to the training sets. (D) Interactive ligand–protein interactions network (LPIn) displaying results using Cytoscape.js to render the graph in the web browser. Results are displayed with respect to the selected threshold, in this case 0.70. Users can use the orange buttons to display or undisplay the legend and labels and to save the network as an image in SVG format.

The platform offers a diverse array of targets associated with Alzheimer's disease, classified according to the IDG. After entering the molecules of interest, the user must select the targets to perform the virtual screening. The platform allows the user to select targets one by one, all targets that are part of the same group (e.g., all kinases), or all 85 targets if desired (Supplemental Figure S6). Furthermore, the platform enables users to adjust a threshold probability (ranging from 0.01 to 1.0). This probability threshold allows filtered results to be generated. The platform will only present compounds whose probability of activity against the selected targets is equal to or greater than the threshold set. According to the results, a compound with probability greater than 0.5 is considered as a "hit" or "active". However, using the probability threshold to show compounds with higher probability is highly recommended. This way, compounds with a better chance of being active in biological assays will be selected. Furthermore, users have the option of entering their email address to receive notifications upon the completion of virtual screening, and to access information stored on the server, which is retained for a period of 60 days.

After running the virtual screening, AlzyFinder presents the results in three different forms:

1. As a data table (Figure 2A), where users can get access to the ligand–AD protein prediction results (value between 0.01 and 1). The higher the value the better, since it means that the probability of the ligand being active against the selected targets is high. Since all ligands (in this case: JW8, Compound 1, FAH65, Compound **5a** and Compound 19-2-1, Figure 1) were screened against all selected targets (in this case: ACHE, GSK3A, BACE1, HTR1A, and HCRTR2, Supplemental Figure S6), the results show all predictions above the selected probability threshold (in this case 0.70,

https://doi.org/10.1021/acs.jcim.4c01481
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Supplemental Figure S6). Results can be downloaded as CSV and/or PDF formats. The data table also allows to find dynamically ligands and/or targets as well as probability result values.

2. As a heat map (Figure 2B), where users can access a complete and dynamic heat map to visualize the results. This visualization is very useful, since it allows to quickly spot and identify which ligands have a high probability of being multitarget, and which selected AD proteins have a high probability of being modulated by the same ligand. Results can be downloaded as CSV and/or PNG formats.
3. Similarity matrix (Figure 2C) allows users to visualize how virtually screened compounds compare to those in the training set based on their similarity. This is crucial for assessing whether the new compounds are closely related to known ones or if they differ significantly, which could either indicate potential benefits or highlight possible issues.
4. As a Ligand-Protein Interaction network (LPIn) (Figure 2D), where users can visualize the results as a graph, displaying the probability results as interactions (edges) between proteins and ligands. The size of each node (ligand and/or target) is proportional to the degree topology parameter (number of interactions), which makes it easier to quickly spot which nodes have the most relevance in the network. Users can use the upper buttons to display the network legend as well as labels of the nodes and edges. The network is generated with the Cytoscape.js graph theory library for visualization and analysis.[26] Therefore, users can also play around the LPIn network with the graphical interface to better understand the results and can download the LPIn visualization as a SVG image.

AlzyFinder is the first free and open access online server that allows virtual screening using machine learning models against key proteins in Alzheimer's disease, and presents the results dynamically, so that users can perform their analysis quickly and effectively. The results can be downloaded in multiple formats (images and data) so that users can continue their analysis independently. In this case study, we can observe that by performing the virtual screening of the molecules described in Figure 1, AlzyFinder was able to accurately predict (with a probability greater than 0.70) all five molecules as true positives (Table 2). It is important to note that machine learning models are not accurate, which is why we refer to predictions in terms of probability. As such, not all active molecules may be classified as active.

Table 2. AlzyFinder Results of Reported Compounds

AD protein	ligand	reported activity (nM)	ref	predicted activity[a]
ACHE	JW8	56	21	0.97
GSK3A	compound **1**	0.18	22	0.83
BACE1	FAH65	10	23	1.00
HTR1A	compound **5a**	1.9	24	0.99
HCRTR2	compound 19-2-1	6.7	25	0.73

[a]Machine Learning probability of being active (values between 0.01 and 1.0).

Finally, for users who wish to screen more than 100 molecules, the AlzyFinder GitHub repository (https://github.com/ramirezlab/AlzyFinder) provides all the machine learning models, their validations, and a script to perform the screening locally. The availability of these resources as open access is vital for the scientific community, as it facilitates the reproducibility of studies, fosters collaboration, and accelerates the discovery of new drugs. By offering free access to the necessary models and tools, AlzyFinder democratizes the virtual screening process and enhances researchers' ability to explore potential new therapies efficiently and effectively. This accessibility is particularly important in researching complex diseases like Alzheimer's, where collaboration and knowledge sharing are essential to advancing our understanding and treatment of the disease.

Our findings align with previous studies that emphasize the importance of multitarget approaches in AD treatment. Different authors have demonstrated that compounds targeting multiple pathways exhibited enhanced therapeutic effects, supporting the utility of AlzyFinder in identifying such candidates,[27−29] highlighting the important role of in silico methods in multitarget directed ligands design.[30]

NETWORK PHARMACOLOGY REVEALS SYSTEMIC ACTION MECHANISM OF DRUG CANDIDATES

To better understand the systemic impact that an active ligand can have Alzheimer's disease, we took the 85 proteins studied in this work and built a protein−protein interaction network (PPIn-AD). Protein−protein interactions in the PPIn-AD were mapped using the STRING database.[31] To ensure data robustness, interactions with scores of ≥0.7 were filtered based on their experimental and/or database STRING scores. Then, the PPIn-AD was expanded using NetworkX (https://networkx.org/) by finding the first neighbors (first shell) in the network constructed from the initial proteins. With this extended PPIn-AD users can expand the Ligand-Protein interaction network using the *Expended LPIn* button (Figure 2C). As a result, proteins that interact with selected AD targets (Supplemental Figure S5) will be added to the LPIn (orange nodes - Figure 3). This extension is highly relevant, as for the first time a web-based tool such as AlzyFinder employs machine learning to perform ligand-based virtual screening together with network pharmacology analysis to obtain information on the potential therapeutic impact of a hit molecule.

CONCLUSIONS

The AlzyFinder Platform represents a significant advancement in the field of Alzheimer's disease (AD) research, providing a comprehensive and user-friendly tool for ligand-based virtual screening and network pharmacology. Our findings demonstrate that AlzyFinder effectively predicts interactions between candidate molecules and over 85 key proteins implicated in AD, utilizing advanced machine learning models that have been rigorously validated. The platform's ability to identify potential multitarget ligands highlights its innovative approach to addressing the complex and multifactorial nature of AD, which is often inadequately tackled by traditional single-target therapies. The integration of robust validation techniques, including a soft-voting ensemble method, has shown to enhance the accuracy and reliability of predictions, thereby facilitating the identification of compounds that may exert

https://doi.org/10.1021/acs.jcim.4c01481
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

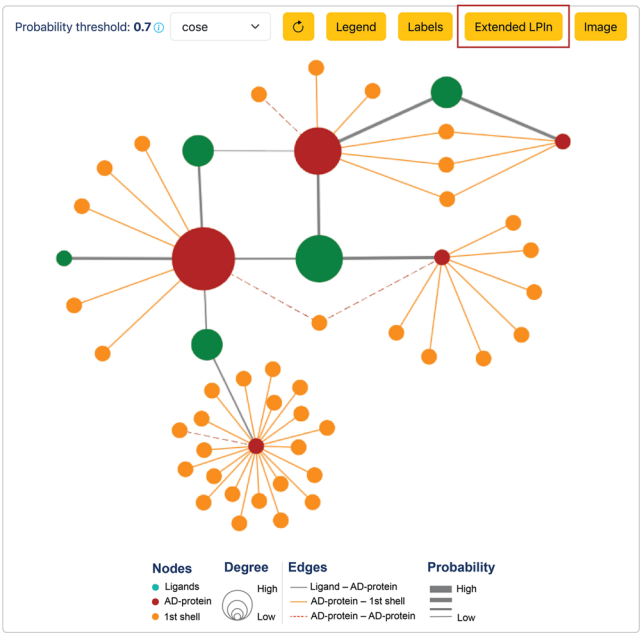


Figure 3. AlzyFinder virtual screening dynamic results with the ligand−protein interaction network (LPIn) expanded. Interactive LPIn displaying results using Cytoscape.js to render the graph in the web browser. Results are displayed with respect to the selected threshold, in this case 0.70.

systemic therapeutic effects. The outputs generated by AlzyFinder, including interaction data presented in various formats such as tables, heat maps, and interactive networks, provide researchers with valuable insights into the potential mechanisms of action of drug candidates.

Looking ahead, several avenues for future research can be identified. First, it is essential to conduct experimental validation of the predicted ligand-protein interactions to confirm the therapeutic potential of the identified compounds. This could involve in vitro and in vivo studies to assess the efficacy and safety of these multitarget ligands in relevant biological models of AD; also expanding the database of known active and inactive compounds within the AlzyFinder framework could further enhance the predictive power of the machine learning models. Incorporating data from recent studies and clinical trials will ensure that the platform remains up-to-date and relevant in the rapidly evolving field of AD research.

It is important to note that future iterations of AlzyFinder could benefit from the integration of more sophisticated network pharmacology analyses, including the exploration of protein−protein interactions and the impact of polypharmacology on therapeutic outcomes. This would provide a more holistic view of how potential drug candidates interact within biological systems, paving the way for the development of more effective and targeted therapeutic strategies. By enabling researchers worldwide to independently analyze and validate their findings, AlzyFinder not only democratizes the drug discovery process but also contributes to the collective effort to combat Alzheimer's disease. In summary, the AlzyFinder Platform stands as a valuable resource for the AD research community, setting a new standard for open-access tools in biomedical research. Its innovative integration of machine learning and network pharmacology opens new avenues for investigating the intricate mechanisms underlying AD, ultimately leading to the development of more effective therapeutic strategies.

■ ASSOCIATED CONTENT

Data Availability Statement

The AlzyFinder Platform can be accessed via the following link: http://www.alzyfinder-platform.udec.cl/. All the compounds to train and validate the machine learning models, their validations, and a script to perform the screening locally are accessible in the AlzyFinder GitHub repository (https://github.com/ramirezlab/AlzyFinder).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.4c01481.

Supplementary Table S1: List of the selected AD key proteins, as well as the number of active, intermediate, and inactive compounds extracted from the ChEMBL database based on their pChEMBL value. (XLSX)

Supplementary Figure S1: Architecture of the AlzyFinder platform. Supplementary Figure S2: Probability density plots with the distribution of pairwise Jaccard distances. Supplementary Figure S3 and S4: Construction and validation of machine learning models. Supplementary Figure S5: AlzyFinder interface to submit molecules to be virtually screening. Supplementary Figure S6: Selection of proteins associated with Alzheimer's disease (AD) for virtual screening. Supplementary Table S3: Distance-to-model of data sets to train models. (PDF)

Supplementary Table S2: Key AD proteins, evaluation metrics (AUC, Accuracy, Balanced Accuracy, F1, Precision, Recall, Specificity) for training and test data, as well as confusion matrix. (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

David Ramírez − *Departamento de Farmacología, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción 4070386, Chile;* orcid.org/0000-0003-0002-1189; Email: dramirezs@udec.cl

Authors

Jessica Valero-Rojas − *Departamento de Farmacología, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción 4070386, Chile; Facultad de Ciencias Químicas, Universidad de Concepción, Concepción 4070386, Chile;* orcid.org/0000-0003-3391-256X

Camilo Ramírez-Sánchez − *Facultad de Ingeniería, Diseño e Innovación, Institución Universitaria Politécnico Gran Colombiano, Bogotá 110231, Colombia*

Laura Pacheco-Paternina − *Departamento de Farmacología, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción 4070386, Chile*

Paulina Valenzuela-Hormazabal − *Departamento de Farmacología, Facultad de Ciencias Biológicas, Universidad de Concepción, Concepción 4070386, Chile;* orcid.org/0009-0008-6110-1573

Fernanda I. Saldivar-González − *Departamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México 04510, México;* orcid.org/0000-0002-0435-8662

https://doi.org/10.1021/acs.jcim.4c01481
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Paula Santana − *Facultad de Ingeniería, Instituto de Ciencias Aplicadas, Universidad Autónoma de Chile, Santiago 8910060, Chile*
Janneth González − *Departamento de Nutrición y Bioquímica, Facultad de Ciencias, Pontificia Universidad Javeriana, Bogotá 110231, Colombia*
Tatiana Gutiérrez-Bunster − *Departamento de Sistemas de Información, Facultad de Ciencias Empresariales, Universidad del Bío-Bío, Concepción 4051381, Chile*
Alejandro Valdés-Jiménez − *Departamento de Sistemas de Información, Facultad de Ciencias Empresariales, Universidad del Bío-Bío, Concepción 4051381, Chile*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.4c01481

Author Contributions

[+]J.V.-R. and C.R.-S. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Agencia Nacional de Investigación y Desarrollo (ANID) − Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) grant no. 1220656 to DR.

■ REFERENCES

(1) Durães, F.; Pinto, M.; Sousa, E. Old Drugs as New Treatments for Neurodegenerative Diseases. *Pharmaceuticals* **2018**, *11* (2), 44.
(2) Weller, J.; Budson, A. Current Understanding of Alzheimer's Disease Diagnosis and Treatment. *F1000Research* **2018**, *7*, 1161.
(3) Punyakoti, P.; Behl, T.; Sehgal, A.; Yadav, S.; Sachdeva, M.; Anwer, M. K.; Vargas-De-La-Cruz, C.; Venkatachalam, T.; Naqvi, M.; Verma, R.; et al. Postulating the Possible Cellular Signalling Mechanisms of Antibody Drug Conjugates in Alzheimer's Disease. *Cell. Signal.* **2023**, *102*, No. 110539.
(4) Van Dyck, C. H.; Swanson, C. J.; Aisen, P.; Bateman, R. J.; Chen, C.; Gee, M.; Kanekiyo, M.; Li, D.; Reyderman, L.; Cohen, S.; et al. Lecanemab in Early Alzheimer's Disease. *N. Engl. J. Med.* **2023**, *388* (1), 9−21.
(5) Mintun, M. A.; Lo, A. C.; Duggan Evans, C.; Wessels, A. M.; Ardayfio, P. A.; Andersen, S. W.; Shcherbinin, S.; Sparks, J.; Sims, J. R.; Brys, M.; et al. Donanemab in Early Alzheimer's Disease. *N. Engl. J. Med.* **2021**, *384* (18), 1691−1704.
(6) Elmaleh, D. R.; Farlow, M. R.; Conti, P. S.; Tompkins, R. G.; Kundakovic, L.; Tanzi, R. E. Developing Effective Alzheimer's Disease Therapies: Clinical Experience and Future Directions. *J. Alzheimer's Dis.* **2019**, *71* (3), 715−732.
(7) Mann, B. S.; Johnson, J. R.; Cohen, M. H.; Justice, R.; Pazdur, R. FDA Approval Summary: Vorinostat for Treatment of Advanced Primary Cutaneous T-Cell Lymphoma. *Oncologist* **2007**, *12* (10), 1247−1252.
(8) Kilgore, M.; Miller, C. A.; Fass, D. M.; Hennig, K. M.; Haggarty, S. J.; Sweatt, J. D.; Rumbaugh, G. Inhibitors of Class 1 Histone Deacetylases Reverse Contextual Memory Deficits in a Mouse Model of Alzheimer's Disease. *Neuropsychopharmacology* **2010**, *35* (4), 870−880.
(9) Steinberg, M. Dasatinib: A Tyrosine Kinase Inhibitor for the Treatment of Chronic Myelogenous Leukemia and Philadelphia Chromosome—Positive Acute Lymphoblastic Leukemia. *Clin. Ther.* **2007**, *29* (11), 2289−2308.
(10) Ryu, K.-Y.; Lee, H.; Woo, H.; Kang, R.-J.; Han, K.-M.; Park, H.; Lee, S. M.; Lee, J.-Y.; Jeong, Y. J.; Nam, H.-W.; et al. Dasatinib Regulates LPS-Induced Microglial and Astrocytic Neuroinflammatory Responses by Inhibiting AKT/STAT3 Signaling. *J. Neuroinflammation* **2019**, *16* (1), 190.
(11) Ochoa, D.; Hercules, A.; Carmona, M.; Suveges, D.; Gonzalez-Uriarte, A.; Malangone, C.; Miranda, A.; Fumis, L.; Carvalho-Silva, D.; Spitzer, M.; et al. Open Targets Platform: Supporting Systematic Drug-Target Identification and Prioritisation. *Nucleic Acids Res.* **2021**, *49* (D1), D1302−D1310.
(12) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; et al. The ChEMBL Database in 2017. *Nucleic Acids Res.* **2017**, *45* (D1), D945−D954.
(13) Bosc, N.; Atkinson, F.; Felix, E.; Gaulton, A.; Hersey, A.; Leach, A. R. Large Scale Comparison of QSAR and Conformal Prediction Methods and Their Applications in Drug Discovery. *J. Cheminform.* **2019**, *11*, 4.
(14) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Thiel, K.; Wiswedel, B. KNIME-the Konstanz Information Miner: Version 2.0 and Beyond. *AcM SIGKDD Explor. Newsl.* **2009**, *11* (1), 26−31.
(15) Arrué, L.; Cigna-Méndez, A.; Barbosa, T.; Borrego-Muñoz, P.; Struve-Villalobos, S.; Oviedo, V.; Martínez-García, C.; Sepúlveda-Lara, A.; Millán, N.; Márquez Montesinos, J. C. E.; et al. New Drug Design Avenues Targeting Alzheimer's Disease by Pharmacoinformatics-Aided Tools. *Pharmaceutics* **2022**, *14* (9), 1914.
(16) Nozal, V.; García-Rubia, A.; Cuevas, E.; Pérez, C.; Tosat-Bitrián, C.; Bartolomé, F.; Carro, E.; Ramírez, D.; Palomo, V.; Martínez, A. From Kinase Inhibitors to Multitarget Ligands as Powerful Drug Leads for Alzheime's Disease Using Protein-Templated Synthesis. *Arch. Anesthesiol. Crit. Care* **2021**, *60*, 19344.
(17) Hernandez, F.; Lucas, J. J.; Avila, J. GSK3 and Tau: Two Convergence Points in Alzheimer's Disease. *J. Alzheimer's Dis.* **2012**, *33* (s1), S141−S144.
(18) Tarwidi, D.; Pudjaprasetya, S. R.; Adytia, D.; Apri, M. An Optimized XGBoost-Based Machine Learning Method for Predicting Wave Run-up on a Sloping Beach. *MethodsX* **2023**, *10*, No. 102119.
(19) Wang, H.; Yang, Y.; Wang, H.; Chen, D. Soft-Voting Clustering Ensemble. *Multiple Classifier Systems: 11th International Workshop, MCS 2013, Nanjing, China, May 15−17, 2013. Proceedings 11* **2013**, *7872*, 307−318.
(20) Sánchez-Cruz, N.; Medina-Franco, J. L. Epigenetic Target Profiler: A Web Server to Predict Epigenetic Targets of Small Molecules. *J. Chem. Inf. Model.* **2021**, *61* (4), 1550−1554.
(21) Drozdowska, D.; Maliszewski, D.; Wróbel, A.; Ratkiewicz, A.; Sienkiewicz, M. New Benzamides as Multi-Targeted Compounds: A Study on Synthesis, AChE and BACE1 Inhibitory Activity and Molecular Docking. *Int. J. Mol. Sci.* **2023**, *24* (19), 14901.
(22) Amaral, B.; Capacci, A.; Anderson, T.; Tezer, C.; Bajrami, B.; Lulla, M.; Lucas, B.; Chodaparambil, J. V.; Marcotte, D.; Kumar, P. R.; et al. Elucidation of the GSK3$α$ Structure Informs the Design of Novel, Paralog-Selective Inhibitors. *ACS Chem. Neurosci.* **2023**, *14* (6), 1080−1094.
(23) Campagna, J.; Jagodzinska, B.; Wi, D.; Zhu, C.; Lee, J. T.; Cohn, W.; Jun, M.; Elias, C. J.; Descamps, O.; Zhang, Q.; et al. Discovery of an APP-Selective BACE1 Inhibitor for Alzheimer's Disease. *bioRxiv* **2023**, 552030.
(24) Wu, J.; Zhang, Z.; Zhang, Q.; Li, J. Design, Synthesis, and Biological Evaluation of Aralkyl Piperazine and Piperidine Derivatives Targeting SSRI/5-HT1A/5-HT7. *Bioorg. Med. Chem.* **2024**, *104*, No. 117698.
(25) Sabnis, R. W. Sulfonamide Compounds as Orexin Receptor Agonists for Treating Sleep Disorders, Namely, Narcolepsy and Hypersomnia. *ACS Med. Chem. Lett.* **2024**, *15* (1), 17−18.
(26) Franz, M.; Lopes, C. T.; Huck, G.; Dong, Y.; Sumer, O.; Bader, G. D. Cytoscape. Js: A Graph Theory Library for Visualisation and Analysis. *Bioinformatics* **2016**, *32* (2), 309−311.
(27) Kumar, N.; Kumar, V.; Anand, P.; Kumar, V.; Dwivedi, A. R.; Kumar, V. Advancements in the Development of Multi-Target Directed Ligands for the Treatment of Alzheimer's Disease. *Bioorg. Med. Chem.* **2022**, *61*, No. 116742.
(28) Kumar, B.; Thakur, A.; Dwivedi, A. R.; Kumar, R.; Kumar, V. Multi-Target-Directed Ligands as an Effective Strategy for the

Treatment of Alzheimer's Disease. *Curr. Med. Chem.* **2022**, *29* (10), 1757−1803.
(29) Chatrath, S. T.; Chapeaurouge, A.; Lin, Q.; Lim, T. K.; Dunstan, N.; Mirtschin, P.; Kumar, P. P.; Kini, R. M. Identification of Novel Proteins from the Venom of a Cryptic Snake Drysdalia Coronoides by a Combined Transcriptomics and Proteomics Approach. *J. Proteome Res.* **2011**, *10* (2), 739−750.
(30) Kumar, A.; Tiwari, A.; Sharma, A. Changing Paradigm from One Target One Ligand towards Multi-Target Directed Ligand Design for Key Drug Targets of Alzheimer Disease: An Important Role of in Silico Methods in Multi-Target Directed Ligands Design. *Curr. Neuropharmacol.* **2018**, *16* (6), 726−739.
(31) Szklarczyk, D.; Gable, A. L.; Nastou, K. C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N. T.; Legeay, M.; Fang, T.; Bork, P.; et al. The STRING Database in 2021: Customizable Protein-Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res.* **2021**, *49* (D1), D605−D612.

https://doi.org/10.1021/acs.jcim.4c01481
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX