



Article

Molecular Modeling and Analysis of Cannabinoid and Cannabinoid-like Molecules Combining K-Means Clustering with Pearson Correlation and PCA

Rafael Campos Vieira ¹ , Érica C. M. Nascimento ² and João B. L. Martins ^{1,2,*}

¹ Department of Pharmacy, Faculty of Health Sciences, University of Brasilia, Brasilia 70910-900, DF, Brazil; rafaelcamposunb@gmail.com

² Laboratory of Computational Chemistry, Institute of Chemistry, University of Brasilia, Brasilia 70910-900, DF, Brazil; ericacristinamoren@gmail.com

* Correspondence: lopes@unb.br

Abstract

More recently, cannabinoid molecules have been widely studied for their potential to treat various diseases. We used a multidisciplinary approach, combining molecular docking and machine learning tools, to identify cannabinoid-based molecules as potential acetylcholinesterase inhibitors. We brought together molecules from the classes of cannabinoids, stilbenoids, isoflavones, and other natural products, along with their electronic structure and absorption, distribution, metabolism, excretion and tolerable toxicity (ADMET) data. A novel machine learning framework (MolSimEx, Molecular Similarity Explorer) combining K-means clustering, Pearson correlation, and principal component analysis was developed to address the similarities of these groups. From the dataset, 30 molecules were selected based on docking scores below -11 kcal/mol. The K-means clustering yielded high classification accuracy on the dataset, correctly grouping the cannabinoid analogues. Additionally, these analogues clustered with classical acetylcholinesterase inhibitors such as huprine-X, huprine-W, and donepezil when considering ADMET and electronic descriptor data. Radulanin J showed the highest correlation (0.41) with donepezil's profile, suggesting the potential of cannabinoid-derived compounds as acetylcholinesterase inhibitors.

Keywords: cannabinoids; Alzheimer's diseases; Pearson correlation; K-means; electronic structure; machine learning



Academic Editor: Paolo Umari

Received: 26 September 2025

Revised: 31 October 2025

Accepted: 4 November 2025

Published: 27 November 2025

Citation: Vieira, R.C.; Nascimento, É.C.M.; Martins, J.B.L. Molecular Modeling and Analysis of Cannabinoid and Cannabinoid-like Molecules Combining K-Means Clustering with Pearson Correlation and PCA. *Int. J. Mol. Sci.* **2025**, *26*, 11520. <https://doi.org/10.3390/ijms262311520>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The medicinal use of *Cannabis sativa* L. dates back to ancient Chinese and Egyptian civilizations [1], with many compounds (more than 500) divided into flavonoids, terpenoids, and phytocannabinoids [2]. Δ -9-tetrahydrocannabinol (THC) and cannabidiol (CBD) are the most abundant natural molecules produced by cannabis [3]. Additionally, the plant produces more than 120 derivatives, generally called minor cannabinoids, in small quantities [4,5]. Recent research has shown potential antinociceptive [6], anti-inflammatory [7], and dermatological [8] activities of minor cannabinoids and other cannabinoid analogs, called cannabinoid-like molecules.

Cannabinoid-like molecules are described as compounds with cannabimimetic properties. These molecules are found in diverse plants and may share structural similarities with cannabinoids and minor cannabinoids. Examples include desmodianones and machaeriols [4].

An important field in medicinal cannabis research is the drug design for neurodegenerative diseases such as multiple sclerosis, Parkinson's, and Alzheimer's [9–11]. The multitarget effect of cannabinoid molecules is increasingly relevant for the design of novel medicines for neurodegenerative diseases. For example, in Alzheimer's disease, some of these compounds can competitively inhibit the human acetylcholinesterase enzyme (hAChE) and bind endocannabinoid receptors, producing a neuroprotective effect [12–15]. Consequently, theoretical studies of the electronic structure of these compounds could provide insights into their reactivity, ligand–receptor interactions, and therapeutic potential [16]. By using machine learning and data mining tools that combine molecular and electronic structure properties, it is possible to develop new medicines affordably.

Machine learning (ML) methods aim to minimize the cost function, meaning they reduce the difference between the variable predicted by the model and the observed variable [17]. Deep learning (DL) methods are machine learning algorithms that involve constructing so-called deep artificial neural networks, which have several applications, such as classifying molecules, generating new molecules, and predicting physicochemical properties [18,19]. For instance, in computational biology, the successful development of the AlphaFold algorithm has enabled accurate prediction of protein structures [20]. Additionally, in organic synthesis, ML algorithms have been developed to predict more favorable routes in retrosynthetic reactions [21].

Machine learning methods are widely used in drug design [22–24], encompassing the discovery of new potential drugs, the prediction of biological activity, and the determination of molecular descriptors with significant biological activity. These ML models may take inputs such as in-line molecular simplifications, e.g., simplified molecular-input line-entry system (SMILES) notation, pharmacokinetic property descriptors (e.g., absorption, distribution, metabolism, excretion, and toxicity—ADMET), and electronic structure descriptors (e.g., HOMO and LUMO energies, gap values, and atomic charges). It is important to normalize or scale the data to ensure the accuracy and integrity of ML models [25], especially when the input data have different measurement units. In most cases, the input data present considerably different units of measurement, as mentioned in the examples above.

We used classical AChEIs with well-established clinical profiles, including tacrine (THA), the first FDA-approved AChEI [26], despite its hepatotoxicity concerns; huperzine A (HUP), a natural alkaloid from *Huperzia serrata* with neuroprotective properties [27]; huprine X (HUX) and huprine W (HUW), synthetic hybrids combining structural features of tacrine and huperzine A [28,29]; galantamine (GNT), a natural phenanthrene alkaloid that also modulates nicotinic receptors [30]; physostigmine (PHY), a reversible carbamate hAChE inhibitor derived from *Physostigma venenosum* [31]; and ladostigil (LADO), a dual hAChE-monoamine oxidase inhibitor with neuroprotective properties [32]. Additionally, we included the natural hAChE substrate acetylcholine (ACh) and dopamine (DOP) as neurochemical references.

Donepezil (DNP) was chosen as a key reference compound because of its favorable pharmacokinetic profile and current status as a first-line treatment for Alzheimer's disease. We employed two machine learning algorithms to analyze molecular property relationships among cannabinoids, cannabinoid-like compounds, and these classical AChEIs: similarity matrix and K-means clustering. We implemented these algorithms specifically for the data structure to identify patterns within molecules of the same class and between different classes. Additionally, we used principal component analysis (PCA) to identify the molecular descriptors that most significantly contributed to the pharmacophoric profiles of the studied molecules. All algorithms were integrated into a Python package version 3.9 called MolSimEx (short for Molecular Similarity Explorer).

2. Results

The frontier molecular orbital values and electronic descriptors are provided in Tables S3 and S4 of the Supplementary Material. The molecular orbital energy values of cannabinoids and cannabinoid-like molecules did not significantly differ between the maximum and minimum values of each orbital of the studied molecules. The most significant difference was found in the LUMO orbital (1.70 eV) between D9CT (−0.17 eV) and DMDD (−1.87 eV), while the smallest difference was found in the LUMO+4 orbital (0.49 eV), between D9CT (0.47 eV) and RDLH (−0.02 eV), excluding the energy values of classical drugs. The slight difference between the highest and lowest values for the HOMO-4 orbital (1.45 eV) between DMDE (−6.68 eV) and CNBA (−8.13 eV) is similar to the difference observed for the LUMO orbital, reflecting the similar electronic distribution of the molecules. Owing to the electronic similarities among cannabinoids, distinguishing classes without machine learning tools is challenging. This topic is particularly important, as ES descriptors are relevant in modeling the pharmacophoric profile and defining drug classes [33].

The general descriptors and ADMET data (Tables S5 and S6) are highly dispersed, as shown in a similarity matrix depicting the ES and ADMET property vectors for each molecule, expressed as percentiles. The similarity matrix is shown in the Supplementary Material (Figures S1–S3). Figure 1 shows the correlation between the ADMET and ES data spaces across the set of molecules.

When the data are normalized (Figure 1), outliers can be interpreted as having significant positive or negative correlations (ρ) for molecules with the highest and lowest property vectors. Therefore, unique outliers above the upper limits, such as those observed in the DNP for ADMET data and in the THA for ES data, correspond to the autocorrelation values of these molecules in the similarity matrix. It is worth noting that HUW exhibits a substantial increase in correlation within the ES data space compared with the ADMET data space. Conversely, molecules such as GNT and MCRA show lower correlation in the ES data space than in the ADMET data space. In contrast, others, such as RDLJ, maintain a nearly constant correlation with both ADMET and ES data.

Figure S6 displays the results of applying the K-means algorithm to the ADMET and ES data separately, using a correlation matrix, \mathbf{M}_{corr} , as the data space, with $k = 5$. The feature space illustrated in the image is based on the first row (x -axis) and the last row (y -axis) of the dataset (Tables S3–S6 of the Supplementary Material).

A comparative analysis between ADMET and ES data, using $k = 5$, revealed significant differences (Figures S6 and S7 of the Supplementary Material) in the positioning of specific molecules within the data space, suggesting that similarities between molecules may change depending on the data type considered. For instance, OMCNB is positioned close to DNP in the ADMET data space, whereas it is close to THA in the ES data space. The ADMET data showed an even distribution of cluster members, whereas the ES data showed a concentration of molecules primarily in clusters 1 and 4 (Figure S6). Cluster 3 in the ES data has increased considerably, driven by the inclusion of cannabinoids such as THC, THCP, HD11, and D9CT, along with ACh. Moreover, the THA in the ES data was migrated to cluster 2. Although MCDC and CNBA remained grouped in cluster 1, other molecular groups varied between the datasets. For example, HUP clustered with cannabinoids such as THCV2, TCNT, and CMG, as well as the classical AChEI THA, which is in cluster 2 of the ES data. At the same time, classical inhibitors HUX and HUW co-clustered with cannabinoid-like molecules DMDD and DMDE, along with the prenyl bibenzyls RDLH, RDLI, and RDHJ. These associations suggest potential targets for drug repositioning studies.

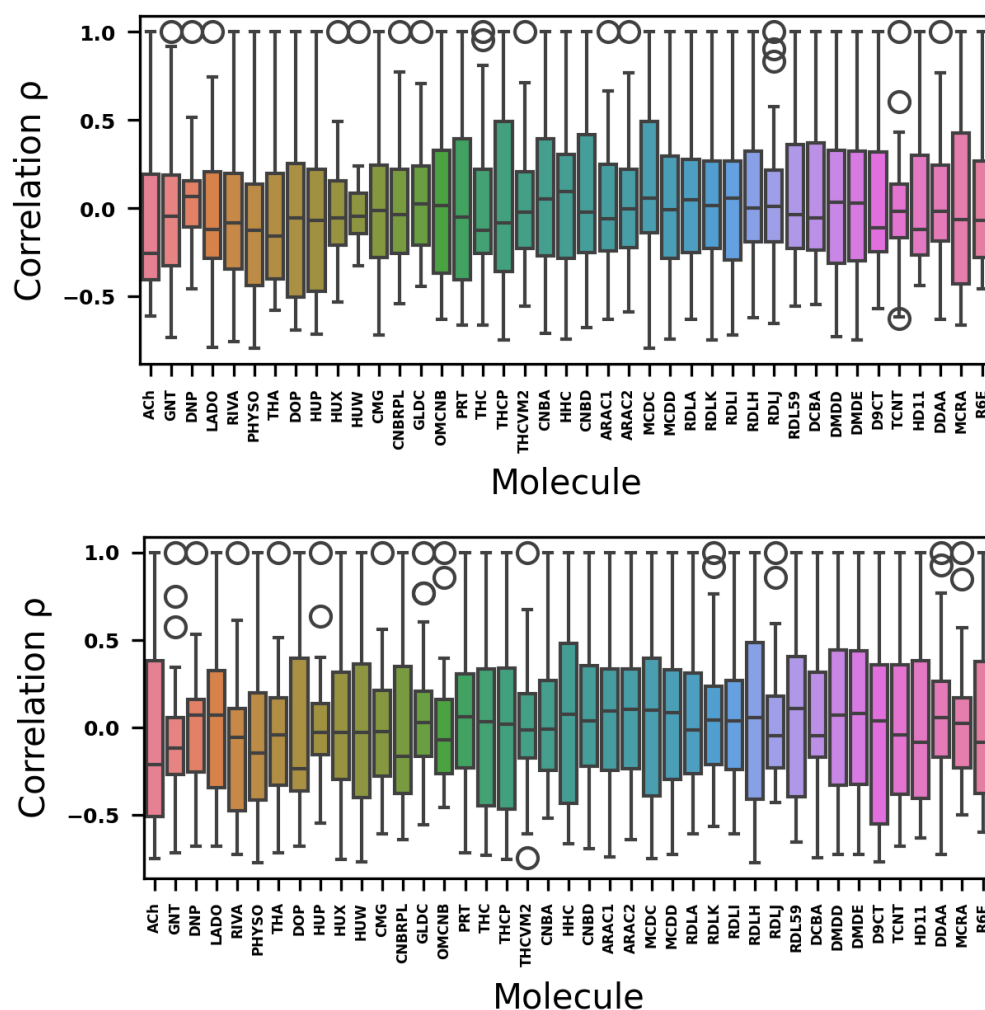


Figure 1. Boxplot with ADMET (**top**) and electronic structure data (**bottom**). Colors are used to facilitate the visual differentiation of molecules.

A set of molecules showed unexpected clustering patterns when the ADMET and ES data spaces were separated, with $k = 5$. Therefore, the next step was to use ADMET and ES data simultaneously for the K-means algorithm. For instance, separating RDLA from other prenyl bibenzyls was only possible when ADMET and ES data were combined, with $k \leq 5$, as shown in Table S7 and Figure 2. Figure S8 in the Supplementary Material displays the results of clustering using both ADMET and EE data.

As illustrated in Figure S9 of Supplementary Material, the Pearson similarity matrix has correctly grouped the classical acetylcholinesterase inhibitors (AChEIs) based on their ADMET and ES data. For example, there is a positive correlation of 0.51 between the alkaloid cholinomimetics physostigmine and galantamine. Furthermore, the synthetic compounds HUX and HUW, derived from a combination of tacrine and huperzine, also show a positive correlation of 0.57.

Table 1 shows the similarity index higher than or equal to 0.75 for the ADMET, ES, and ADMET + ES data. As expected, acetylcholine (ACh) shows no significant positive correlations with minor cannabinoids and cannabinoid analogs when analyzed with ADMET data (as illustrated in Figure S2 of Supplementary Material).

The largest vectors in loadings (Figure S6 of Supplementary Material) for the first five PCs were HOMO-4 and HOMO-3 (0.28) for PC1, LUMO+3 (0.36) for PC2, TPSA (0.47) for PC3, charge on heteroatoms and volume (0.36) for PC4, and drugscore (0.67) for PC5.

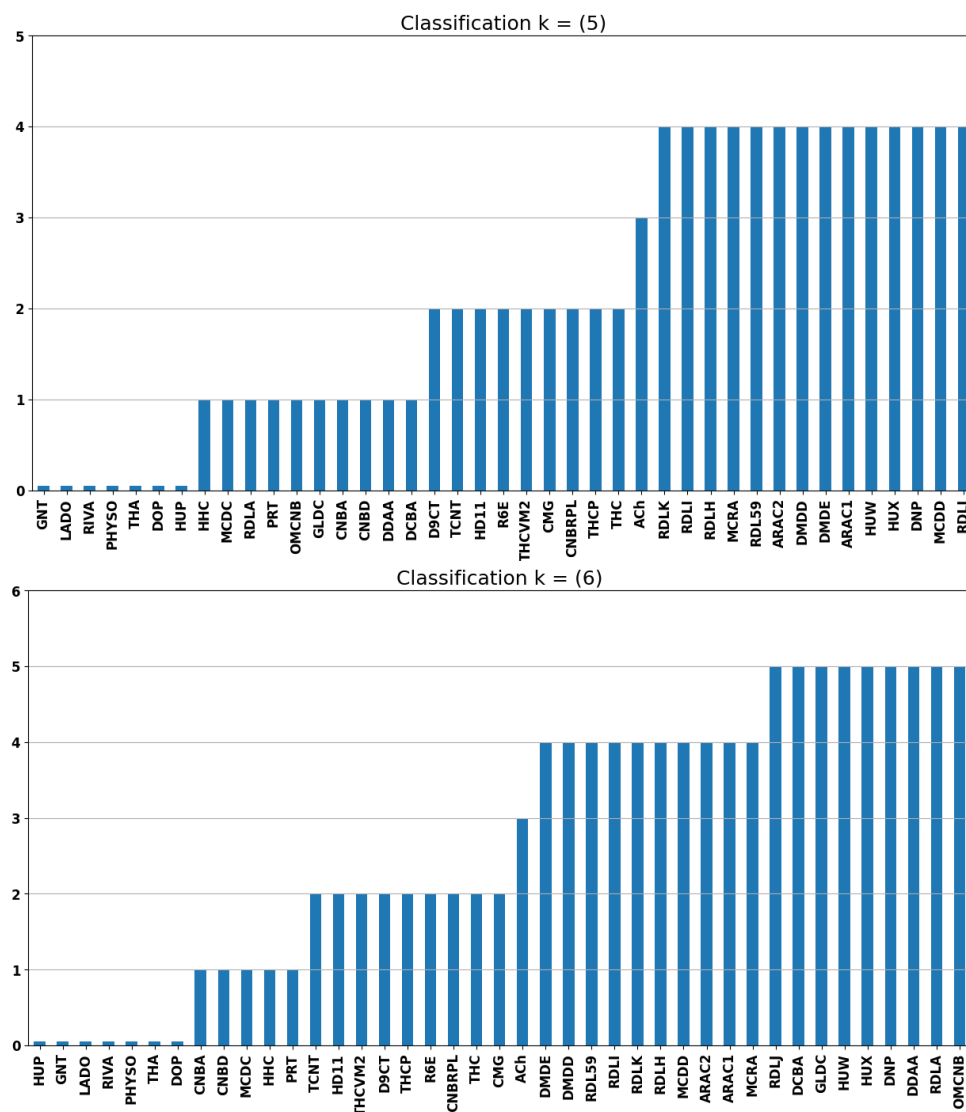


Figure 2. K-means ADMET and ES data with k = 5 (top) and k = 6 (bottom).

The distribution of molecules in the ADMET and ES data space can be visualized using the first three principal components (PCs) of the molecular space, as illustrated in Figure 3. This visualization uses bi-dimensional scatter plots generated from PCA, following dimensionality reduction through loading analysis. In Figure 3, each class is represented by a distinct color and symbol. Figure 3 illustrates the molecular distribution in the reduced-dimensional space of the remaining descriptors after PCA dimensionality reduction. The diagonal panels display the density distributions of each cluster projected onto individual principal components. Higher-variance principal components (particularly the first component) provide enhanced cluster separation, as evidenced by the approximately Gaussian distributions of individual clusters along these dimensions and the reduced overlap between different clusters.

The properties retained after PCA—such as molecular weight (MW(g/mol)), number of aromatic rings, blood–brain barrier (BBB) penetration (binary descriptor), and binding score (kcal/mol) (Table 2)—demonstrate a strong ability to cluster most of the molecules examined effectively.

Table 1. Similarity index of ADMET, ES and ADMET + ES data.

ADMET	ES	ADMET + ES
GNT-PHYSO, GNT-HUP, RIVA-PHYSO, PHYSO-HUP, CNBRPL-TCNT, CNBRPL-HD11, OMCNB-PRT, OMCNB-THCP, OMCNB-MCRA, PRT-MCRA, THC-THCP, THC-D9CT, THC-HD11, THCP-D9CT, CNBA-MCDC, RDLA-RDLI, RDLA-RDLJ, RDLK-RDLH, RDLK-RDL59, RDLI-RDLJ, RDLH-RDL59, RDLJ-DCBA, DCBA-DDAA, DMDD-DMDE, D9CT-HD11	LADO-HHC, RIVA-PHYSO, DOP-HD11, DOP-R6E, HUX-CNBA, HUW-HHC, HUW-RDL59, CNBRPL-TCNT, CNBRPL-HD11, GLDC-CNBD, GLDC-RDLA, GLDC-DCBA, OMCNB-RDLJ, PRT-RDLA, THC-THCP, THC-D9CT, THC-HD11, THC-R6E, THCP-D9CT, THCP-HD11, CNBD-DCBA, CNBD-DDAA, ARAC1-RDLI, ARAC2-MCDD, ARAC2-RDLK, MCDC-DMDD, MCDC-DMDE, RDLA-DCBA, RDLA-DDAA, RDLH-RDL59, RDLH-DMDE, DCBA-DDAA, DMDD-DMDE, D9CT-HD11, D9CT-R6E, HD11-R6E	GNT-PHYSO, RIVA-PHYSO, CNBRPL-TCNT, CNBRPL-HD11, PRT-MCRA, THC-THCP, THC-D9CT, THC-HD11, THCP-D9CT, ARAC2-RDLK, RDLA-DCBA, RDLA-DDAA, RDLH-RDL59, DCBA-DDAA, DMDD-DMDE, D9CT-HD11

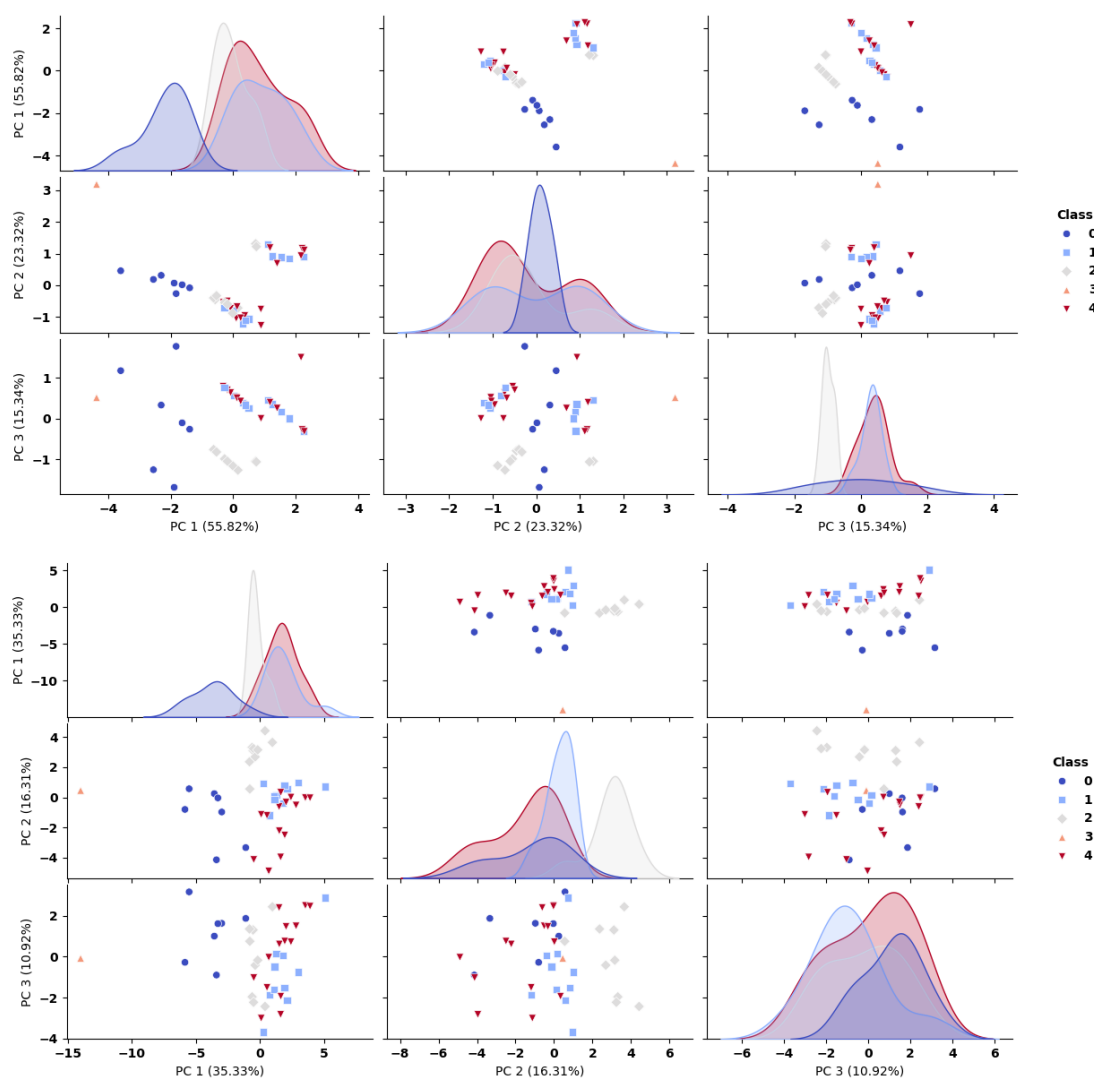
**Figure 3.** PCA scatter matrix for first 3 PCs with remaining descriptors (**top**) and with all descriptors (**bottom**) using classes (symbols and colors) predicted by k-means algorithm. The molecules in each class are presented in Table S7.

Table 2. PC loadings for remaining descriptors after loading analysis. Colors were used to help visualize the scale of the loadings.

	PC1	PC2	PC3
MW	0.60	0.06	−0.39
Num. Aromatic Rings	0.48	−0.21	0.85
BBB	0.33	0.88	0.05
Score (kcal/mol)	−0.55	0.42	0.35

3. Discussion

The Elbow and PCA methods were applied to evaluate the k value in the K-means algorithm. The elbow method was applied to the K-means algorithm, using separate ADMET and ES properties, while PCA was applied to all properties. In the PCA approach, the value of k was assumed to be equal to the number of principal components (PCs) that collectively contributed to almost all of the data variance ($k = 5$), as shown in Figure S4.

The elbow method is typically used to identify the ideal number of clusters in the K-means algorithm [34]. The value of k obtained by the elbow method was $k = 5$ for ADMET and ES data, with the maximum value of the WSS being greater (>2000) for the ADMET data than for the ES data, as shown in Figure S5. Therefore, using the variance from the dimensionality reduction of the principal components (PCs) to determine k (Figure S4) is a valid approach, which is supported by the established mathematical relationship between PCA and k-means clustering, where both methods optimize similar objective functions based on data variance minimization [35]. Specifically, PCA looks for directions (principal components) that maximize variance (minimize reconstruction error). Otherwise, k-means minimizes within-cluster variance. The eigenvalue spectrum from PCA shows a natural clustering structure, as principal components with higher eigenvalues correspond to directions of greater data separability, directly informing optimal cluster number selection.

Figure 2 suggests that the clusters up to group 4 share similarities, regardless of whether $k = 5$ or $k = 6$ for the ADMET + ES data in Table S7. In contrast to the formation of these groups, ACh is isolated from the other molecules, which is significant because it serves as a functional negative control in this study. It means that ACh does not inhibit hAChE since it is the natural substrate of the enzyme. This outcome demonstrates that using combined ADMET and ES data resulted in a clustering profile that aligns more closely with the expected outcomes. Moreover, the known molecules used in the treatment of Alzheimer's are grouped in cluster 0.

However, DNP, HUX, and HUW are grouped with RDLJ despite $k = 5$ or $k = 6$, indicating a promising possibility of RDLJ, a cyclopropanochroman derivative from *Radula javanica* with less psychoactive effects [36]. The RDLJ has the same methoxymethyl ether group also found in DNP. Another outcome is that ARAC1 and ARAC2 are grouped together with HUW, HUX, and DNP for $k = 5$. ARAC1 and ARAC2 are prenylated stilbenoids with the same skeleton as stilbene and other stilbenoids, which are linked to the treatment of Alzheimer's disease using natural products [37,38].

The highest correlation value observed for ACh, cannabinoid, and cannabinoid-like molecules was 0.33 with TCNT (Figure S4 in the Supplementary Material). However, when examining the similarity matrix using ES data (Figure S3 of Supplementary Material), ACh exhibited positive correlations with THC (0.45), THCP (0.45), THCVM2 (0.40), D9CT (0.55), HD11 (0.60), and R6E (0.57), as well as an increased correlation with TCNT (0.54). Interestingly, the correlation between hAChE inhibitors (AChEIs) decreases significantly when only the ADMET data is taken into account. These differences result in a different cluster classification in the K-means algorithm, as shown in Figure S8 for ADMET + ES properties.

The similarity index shown by the ADMET + ES (Table 1) groups most similar cannabinoid and cannabinoid-like molecules, while separating GNT, PHYSO, and RIVA. However, all cannabinoid and cannabinoid-like molecules have a positive correlation, albeit small, with the classical AChEIs in the ADMET + ES data. For this set, the largest values are found for THCMV2 (0.35 with GNT), RDLH (0.37 with HUP), RDLJ (0.41 with GNT), and TCNT (0.41 with DOP). The observed clustering of cannabinoid-like compounds with classical AChEIs can be rationalized through shared pharmacodynamic and structural features that may influence acetylcholinesterase inhibition.

Structurally, both classical AChEIs and several cannabinoid derivatives possess aromatic systems and hydrogen bond acceptor/donor groups, where the presence of such groups facilitates binding to the acetylcholinesterase active site [39,40]. Furthermore, cannabinoids have demonstrated cholinergic modulation, e.g., in the direct AChE inhibition, as reported for THC [41], and the indirect cholinergic enhancement via CB1 receptor-mediated modulation of acetylcholine release [42,43]. The prenylated compounds, especially those derived from *Radula* species, share structural motifs with established AChEIs, including phenolic hydroxyl groups and extended conjugated systems; such groups may contribute to enzyme–substrate interactions [44]. These similarities support the computational clustering results, suggesting that the identified correlations may reflect pharmacological relationships rather than statistical artifacts.

As illustrated in Figure 3, this is particularly suitable in the apparent grouping of ACh in class 3 (depicted by the triangle symbol with a salmon color palette). Moreover, a distinct separation is observed between class 0, which contains classical inhibitors like DNP, GNT, and HUP, and both class 2, which consists of cannabinoids such as CMG, CNBRPL, and D9CT (represented by the grey diamond symbol), and class 1, which includes cannabinoid-like molecules such as DCBA, DDAA, GLDC, and OMCNB (marked by the light blue square symbol). From Figure 3, we can also see that the molecules in cluster 4 (marked by a red inverted triangle), which consist of cannabinoid-like compounds such as ARAC1, ARAC2, and DMDD, as well as the molecules in cluster 2, are closely grouped with the classic AChEIs in cluster 0. These findings emphasize the need to treat these properties as essential descriptors of molecular characteristics, highlighting their pivotal role in molecular analysis.

Although the loading analysis retained only ADMET properties in our final model (Table 2), it is worth noting that all electronic structure descriptors exhibited high loadings (Figure S10, Supplementary Material). This feature indicates that electronic descriptors also contributed significantly to the principal components, even though they were not retained in the final dimensionally reduced model.

Furthermore, the boxplot analysis (Figure 1) showed that electronic structure data produced more positive and negative outliers than ADMET descriptors, as indicated by the Pearson correlation coefficient. These outliers represent molecules that exhibited high positive or negative correlations with the reference molecule in each boxplot analysis. This suggests that electronic structure descriptors can identify molecular similarity and dissimilarity within our dataset, thereby separating cannabinoids from classical AChE inhibitors.

The complementary nature of these descriptor types is a challenge. While ADMET properties provided the most consistent dimensionality reduction (as shown by the loadings analysis), electronic descriptors exhibit a broader distribution of correlations, suggesting that both are valuable for molecular characterization.

Future studies could use these two approaches—dimensionality reduction via loading analysis and correlation coefficient variation analysis—by developing more robust search algorithms, such as simplex optimization or Monte Carlo simulations, to better capture the combined effects of both ADMET and electronic structure descriptors in molecular clustering.

4. Materials and Methods

4.1. Docking and Electronic Structure Study

We employed molecular docking as a fast, low-cost, structure-based screening to estimate the relative binding affinities of cannabinoids and cannabinoid-like molecules towards hAChE. In the present work, docking was used strictly to generate comparable docking scores as ranking criteria and as features for subsequent machine-learning models, rather than to make definitive claims about absolute affinities or detailed binding mechanisms. By standardizing the receptor conformation and docking parameters, the resulting scores provide a consistent framework for prioritization that integrates seamlessly with our data-driven pipeline.

Our method focuses on initial screening based on the canonical binding pose for this well-studied enzyme [28,45–49]. Therefore, this approach is a well-suited simplification because more recent computational work using molecular dynamics has shown that donepezil's inhibitory mechanism involves multiple stable interaction modes within the active site of hAChE [45]. The molecular docking study was performed using the AutoDock4 Tools version 4.2.6 software package [50]. The bioreceptor selected was the human hAChE crystal structure, deposited in the PDB (Protein Data Bank) under the code 6O4W. The docking study protocol was applied in accordance with the literature [51].

AutoDock Vina searched for up to 1000 solutions at an exhaustiveness level of 32. The gridbox parameters were set to a width of 25.5 Å, a height of 25.5 Å, and a depth of 25 Å. The center point was located at $X = 87.6$, $Y = 84.7$, and $Z = -4.7$ Å. All calculations were performed using a rigid protein–ligand approach. This methodological sequence is supported by the known structural nature of classical AChE inhibitors, which are characterized by predominantly rigid molecular frameworks [52], thereby validating the use of a rigid model to approximate their primary bioactive state [53].

All water molecules of the 6O4W structure and other molecules, such as stabilizers and artifacts, were removed. The protonation states of histidine residues in the enzyme were manually configured in AutoDock Tools, based on PROPKA 3.5 software predictions [54,55]. To validate the docking protocol, a redocking study was performed using the crystallographic structure 6O4W (resolution 2.35 Å) complexed with its co-crystallized ligand E20 (DNP).

The native ligand was extracted from the binding site and subsequently redocked using the same parameters established for the virtual screening protocol. The redocking validation yielded an RMSD value of 0.26 Å between the docked pose and the crystallographic conformation, with a docking score of -11.60 kcal/mol. The low RMSD value, well below the 2 Å threshold considered acceptable for redocking protocols [56], and the docking score, proportional to experimental K_i values for donepezil reported in the literature [46], both validated the docking protocol. While redocking supports the chosen conformer, rigid docking overlooks receptor dynamics and conserved waters, so rankings should be viewed as first-pass enrichment, most reliable for relatively rigid ligands.

Following the validation protocol, we conducted virtual screening using AutoDock Vina [57] with 253 molecules, categorized into phytocannabinoids, flavonoids, and terpenoids, against the hAChE (6O4W) enzyme (Tables S1 and S2 of Supplementary Material). The custom dataset was assembled from curated public resources (PubChem, ChEMBL, ZINC, and KNApSACk) and based on literature on cannabis metabolites and plant-derived flavonoids/terpenoids [11,58,59]. Molecules with binding scores ≤ -11 kcal/mol (89 molecules) were selected for docking pose analysis via Discovery Studio software version 20.1. We used the binding score of -11 kcal/mol, as it is close to the redocking value [60]. All the docking poses were compared to those of donepezil.

Molecules exhibiting interactions similar to those of donepezil, especially with the residues Trp286, Trp86, Phe338, Phe295, Ser293, and Tyr341 of the peripheral anionic site (PAS) and with residue His447 of the catalytic active site (CAS), were chosen for further analysis. Only the beta portion of the hAChE enzyme was considered in the docking protocol. The van der Waals interactions were evaluated for distances between 3 and 5 Å, whereas hydrogen bond interactions were considered for distances between 2 and 3 Å. This choice is supported by prior QM/MM molecular dynamics simulations of the AChE system [49]. Following this analysis, the remaining 30 molecules were selected for electronic structure (ES) calculations (Figure 4).

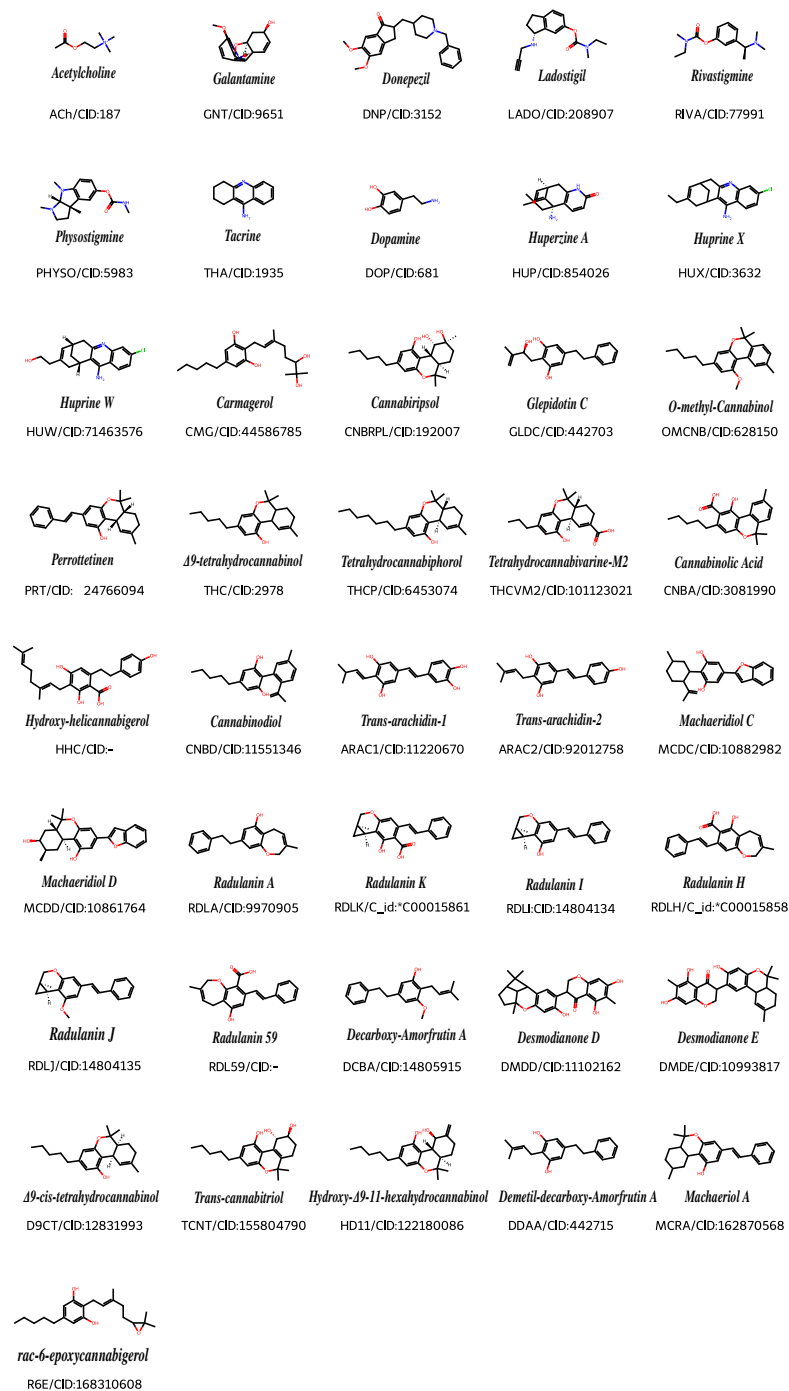


Figure 4. Molecules studied and the notation used. PubChem CID was also available. * C_id molecule identification found in KNApSACk [61] metabolite database.

The dataset was purposely kept smaller due to the computational cost of quantum-chemical calculations, which included electronic structure descriptors at the Density Functional Theory (DFT) level. Therefore, this level of theory requires a significant computational cost per molecule. This choice prioritizes quality and physically grounded interpretability over raw scale. In practice, it is expected that these descriptors—such as dipole moment, polarizability, frontier orbital energies (HOMO/LUMO), electronic gap, global hardness/softness, partial charges, and electrostatic potentials—capture electronic aspects directly related to molecular recognition interactions and reactivity, information that is often not fully reflected by purely empirical descriptors.

The analysis of orbitals from HOMO-4 to LUMO+4 is grounded in frontier molecular orbital (FMO) theory and conceptual DFT [62–64]: reactivity, polarization, and charge transfer in complexation are governed not solely by a single HOMO–LUMO pair, but by multiple neighboring occupied and low-lying virtual orbitals with comparable energies and appropriate symmetry/localization [65]. Conceptual DFT demonstrates that local reactivity—via Fukui functions and local softness—is distributed across near-frontier levels rather than confined to a single orbital [66,67]. In practice, several occupied orbitals near the HOMO and low-lying virtuals near the LUMO contribute significantly, particularly when the dominant HOMO/LUMO exhibit nodes or poor localization on the binding motif [68,69]. Consequently, a symmetric ± 4 window captures these contributions without excessive computational cost.

In line with these expectations, electronic descriptors have been shown to modulate the pharmacological activity of drugs, affecting the antiproliferative activity of multiple human cancer cell lines [70] and the anticoagulant effect of 17β -aminoestrogens [71].

The relevance of electronic descriptors is further highlighted in the field of AChEIs, where a QSAR analysis of tacrine-related inhibitors found that models incorporating quantum–mechanical and electronic features—such as orbital energies and charge distribution—achieved strong predictive performance for inhibitory activity [33,45,49,51,72], reinforcing the practical value of these properties for anticipating biological effects and guiding the rational design of these drugs [51,73,74].

We optimized the molecular geometries in vacuum and solvent (water) using the B3LYP [75–78] functional and the 6-311+G(d,p) basis set. The B3LYP/6-311+G(d,p) functional and basis set approach has recently been used with accurate results [51]. The implicit solvation method, based on the density model (SMD), was used (SMD) [79]. The vibrational modes, atomic charges, molar volume, and polarizability were also calculated. Vibrational modes were used to characterize the energy-minimum geometry structures. Atomic charges were calculated using the CHELPG method [80]. All electronic calculations were performed using Gaussian 16 software [81]. The descriptor distance between the most acidic hydrogens was analyzed using Chemcraft software Version 1.8, build 682 [82]. We combined similarity matrix, K-means, and PCA algorithms to analyze the results using machine learning and data mining packages such as Scikit-learn [83] and Pandas [84].

4.2. ADMET Data

ADMET data were obtained using the SwissADME web platform and the OSIRIS software version 2017 [85,86]. Weights were defined to obtain qualitative data on gastrointestinal (GI) absorption and blood–brain barrier (BBB) permeability.

4.3. MolSimEx Framework

Machine learning approaches have become increasingly popular for molecular clustering in drug discovery in recent decades [22,23]. Current computational methods for potential AChEI screening encompass various approaches, including XGBoost-based classifiers

such as vScreenML, which reduce false positives in structure-based virtual screening [87]; supervised models employing Random Forest and support vector machines for compound classification [88]; web-based platforms like Tropmol that combine ML models for prediction [89]; and pharmacophore-based 3D-QSAR modeling for structural guidance [90]. These methods have shown promising results in compound identification and classification. Otherwise, they typically focus on individual prediction tasks and lack integrated frameworks for comprehensive molecular descriptor analysis and interactive clustering exploration. To address these limitations, we developed MolSimEx, an integrated computational framework that combines analytical techniques in an interactive environment designed explicitly for molecular property analysis.

The MolSimEx pipeline introduces methodological innovations that distinguish it from conventional workflows. First, we have implemented a novel distance metric called ASSED (Amplified Squared Euclidean Distance). This metric enhances the separation of molecular clusters by amplifying penalties for considerable dissimilarities while preserving the relative ordering of molecular relationships. We apply this approach to molecular datasets that exhibit high variability in both similarity and dissimilarity patterns. Second, the framework integrates K-means initialization (K-means++, Bayesian, and random initialization), enabling comparisons of clustering approaches.

MolSimEx provides an interactive visualization interface that enables real-time parameter adjustment and dynamic exploration of clustering results, particularly for complex systems where standard similarity metrics may not adequately capture the cross-nature of molecular properties.

4.4. Similarity Matrix

Three correlation models were evaluated: Pearson, Kendall, and Spearman. Among the known similarity metrics, the choice of the Pearson correlation coefficient was motivated by its ability to capture linear relationships between molecular descriptors, which is particularly advantageous for subsequent K-means clustering analysis that benefits from broader correlation ranges. The Pearson correlation coefficient measures the correlation of two coefficients in the range of -1 to $+1$ [91,92]. The developed algorithm aims to generate a correlation matrix by calculating Pearson correlation coefficients between molecular descriptors. Given the set of \mathbf{S} molecules:

$$\mathbf{S} = (A_1, A_2, A_3, \dots, A_N), \quad A_i = \text{molecule } i, \quad (1)$$

where each molecule is defined as a set of its descriptors

$$A_i = (x_1, x_2, x_3, \dots, x_N), \quad x_i = \text{descriptor } i \quad (2)$$

The next step is to get a new set S' from the scaling of all descriptors:

$$\mathbf{S}' = (A'_1, A'_2, A'_3, \dots, A'_N), \quad A'_i = (z_1, z_2, \dots, z_N), \quad z_i = \frac{x_i - \mu}{\sigma}, \quad (3)$$

where μ is the mean value, and σ is the standard deviation. After scaling, the product-moment coefficients of molecules A and B are calculated.

$$\rho_{A'B'} = \frac{\text{cov}(A', B')}{\sigma_{A'}\sigma_{B'}} \approx \frac{\sum_{i=1}^n (z_i^{A'} - \bar{z}^{A'}) (z_i^{B'} - \bar{z}^{B'})}{\sqrt{\sum_{i=1}^n (z_i^{A'} - \bar{z}^{A'})^2 \sum_{i=1}^n (z_i^{B'} - \bar{z}^{B'})^2}}, \quad (4)$$

where $\bar{z}^{A'}$ and $\bar{z}^{B'}$ are the mean values of descriptors of the new scaled set of molecules A' and B' . After all calculations of product-moment correlation coefficients, the similarity matrix M_{corr} is generated (Algorithm 1).

$$\mathbf{M}_{corr} = \begin{bmatrix} \rho_{A'A'} & \rho_{A'B'} & \rho_{A'C'} & \cdots & \rho_{A'N'} \\ \rho_{B'A'} & \rho_{B'B'} & \rho_{B'C'} & \cdots & \rho_{B'N'} \\ \rho_{C'A'} & \rho_{C'B'} & \rho_{C'C'} & \cdots & \rho_{C'N'} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{N'A'} & \rho_{N'B'} & \rho_{N'C'} & \cdots & \rho_{N'N'} \end{bmatrix} \quad (5)$$

Algorithm 1 Similarity matrix algorithm

$\mathbf{D} \leftarrow$ input data, \mathbf{D} = structure data object with rows and columns
 $\mathbf{SD} \leftarrow S(\mathbf{D})$, S = Scalar operator
 $\mathbf{SDT} \leftarrow T(\mathbf{SD})$, T = Transpose operator
 $\mathbf{C} \leftarrow \mathbf{D}(\text{columns})$
 $\mathbf{ND} \leftarrow \{\mathbf{SDT}, \mathbf{C}\}$, \mathbf{ND} = New structure data
 $\mathbf{M}_{corr} \leftarrow MC(\mathbf{ND})$, MC = Correlation matrix operator

4.5. K-Means

K-means is an unsupervised machine learning algorithm in the cluster algorithms class. Stuart Lloyd of Bell Labs developed the initial version of the algorithm in 1957 [93]. The present work used the Lloyd implementation, the most common computational implementation of K-means. We employed a modified version of the traditional Euclidean norm (L2)

$$L2(\mathbf{a}, \mathbf{b}) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (6)$$

where \mathbf{a} and \mathbf{b} are n-dimensional feature vectors, as the distance metric for the K-means algorithm. The variation, named ASED, is defined as:

$$ASED(\mathbf{a}, \mathbf{b}) = \left(\sum_{i=1}^n (a_i - b_i)^2 \right)^2 \quad (7)$$

The purpose of ASED is to amplify the penalty for large distances while reducing the penalty for small distances. This approach is particularly relevant for molecular similarity assessment, as ADMET and ES properties of molecular datasets often exhibit high variability in both similarity and dissimilarity degrees. Consequently, this strategy improves navigation of the feature space induced by the Pearson-based similarity matrix, whose broader correlation spectrum expands the adequate search space for K-Means.

An additional advantage of ASED is that it constitutes a monotonic transformation of the L2 norm, which is beneficial because ROC (Receiver Operating Characteristic) analysis is invariant to monotonic transformations. This mathematical property ensures that the relative ordering of distances is preserved (see Figure 5). Notably, the ROC curves for ASED and L2 metrics are identical (fully overlapping).

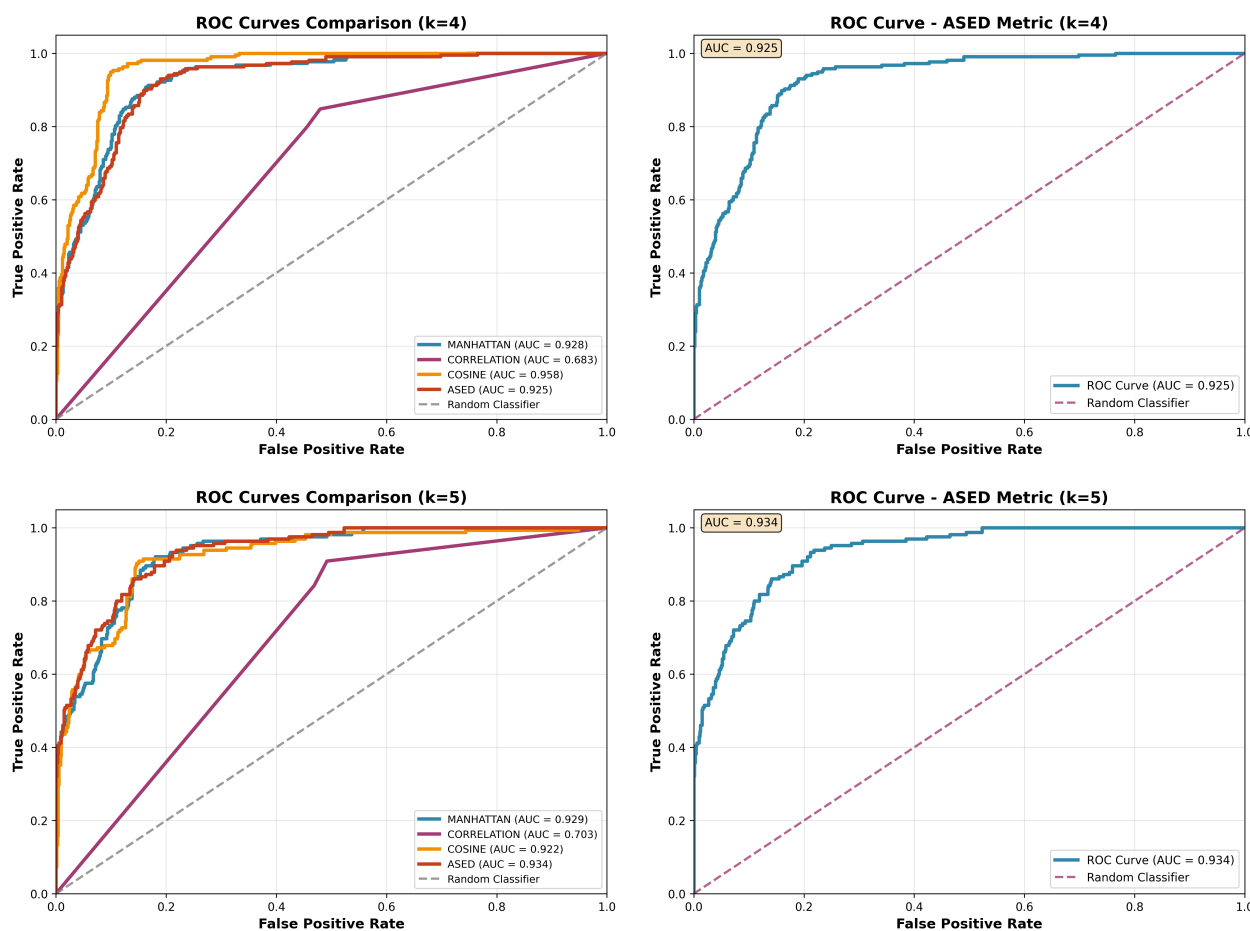


Figure 5. ROC curves for the ES/ADMET dataset with $k = 4$ (top row) and $k = 5$ (bottom row). The left column compares ASED (red) with Manhattan (blue), correlation (purple), cosine (yellow), and a random classifier (gray), while the right column shows ASED alone.

We employed ROC analysis to compare ASED against other distance metrics to verify whether alternative metrics could better model the studied data. This comparative analysis allows us to identify the most appropriate distance metric for capturing the underlying structure of molecular ADMET and ES properties, thereby optimizing the clustering performance.

In drug discovery studies, K-means algorithms are currently used to identify molecular clusters and descriptors [94,95]. Other applications of the K-means algorithm are also being explored, such as identifying the best substrates for chemical reactions in molecular synthesis [96]. These applications demonstrate the broad applicability of cluster machine learning algorithms in chemistry.

Initially, the total set of data A is divided into k -subsets or clusters:

$$A \rightarrow A^1 \cup A^2 \cup A^3 \dots A^k \quad (8)$$

Given a set of k predefined clusters, the goal of the algorithm is to obtain the centroids (c) of each cluster.

$$C = (c^1, c^2, \dots, c^k) \quad (9)$$

That minimizes the cost function, which for K-means is commonly named within-cluster sum of squares (WSS) or inertia, defined as the sum of the Euclidean distances between

each component of the property vector, \mathbf{x} , and its closest centroid. The leading information here is the mean distance among all points in a cluster.

$$WSS = f(\mathbf{A}, \mathbf{C}) = \sum_{a=1}^k \sum_{\mathbf{x} \in A^i} \Delta(\mathbf{x}, \mathbf{c}^i), \quad \mathbf{x} = (x_1, x_2, \dots, x_d) \quad (10)$$

The centroids are calculated with the expression obtained by the midpoint theorem algorithm:

$$\mathbf{c}^i = \frac{1}{|A|^i} \sum_{\mathbf{x} \in A^i} \mathbf{x} \quad (11)$$

There are many methods for determining the ideal value of k in the K-means algorithm [97–99]. In the present study, two methods were used: the elbow method, which evaluates the decay of the cost function with respect to the number of clusters (n), and PCA. The first algorithm result was applied to the K-means. This search aims to categorize classes of molecules by grouping the cannabinoid and cannabinoid-like molecules based on ES and ADMET results (Algorithm 2).

Algorithm 2 Lloyd's algorithm with similarity matrix as input

$\mathbf{M}_{\text{corr}} \leftarrow$ input data
 Initialization : random centroids $\mathbf{C} = \{\mathbf{c}^1, \dots, \mathbf{c}^k\}$
Repeat for range($\mathbf{i} = 1, \mathbf{i} = 2, \dots, \mathbf{i} = \mathbf{k}$)
 $b(x) = \text{argmin}_i \{\Delta(\mathbf{x}, \mathbf{c}_i)\}, \forall x \in \mathbf{M}_{\text{corr}}$
 Update all clusters centroids in \mathbf{C} :
 $\mathbf{c}^i = \frac{1}{|A|^i} \sum_{\mathbf{x} \in A^i} \mathbf{x}, \forall i \in \{1, 2, \dots, k\}$
Until Mean(cluster centers) stop changing

5. Conclusions

We have identified a set of 30 molecules, out of 253 derived from phytocannabinoids, flavonoids, and terpenoids, as potential new inhibitors against the hAChE (6O4W) enzyme. We employed two machine learning algorithms to categorize the molecules: similarity matrix and K-means. These algorithms were implemented specifically for the data structure to identify patterns within molecules of the same class and between different classes. Principal component analysis was also used to find the most determining molecular descriptors.

From the K-means classification, one of the prenyl bibenzyl molecules, the cyclopropanochromane derivative radulanin J, presented the highest correlation with the AChEI profile of the drug donepezil (0.41, based on AMDET+ES data). The molecules tetrahydrocannabivarin-M2, radunalin H, radunalin J, and trans-cannabitrinol showed positive correlations with the classical AChEIs in the ADMET similarity matrix and electronic structure data. The TCNT-DOP and RDLJ-DNP pairs showed a Pearson correlation coefficient of 0.41, indicating a significant correlation driven by the use of ADMET and electronic structure data.

It is important to note that, despite the promising results, particularly for radulanin J, the present computational framework includes virtual screening, requiring future experimental studies to validate the proposed insights, such as in vitro AChE inhibition assays for direct activity confirmation and IC_{50} determination, cytotoxicity studies using neuronal cell lines to assess safety profiles, and blood–brain barrier permeability assays to evaluate central nervous system penetration potential.

Additionally, while our rigid docking protocol provided consistent ranking criteria for the machine learning pipeline, future studies focusing on the most promising candidates would benefit from more sophisticated computational approaches, including induced-fit

docking, ensemble receptor conformations, explicit water modeling, and enhanced scoring methods, such as the molecular mechanics energies combined with the generalized Born and surface area continuum solvation (MM/GBSA) calculation, to reduce potential docking bias and provide more accurate binding predictions. These validations would provide essential pharmacological data for advancing the developed model.

Moreover, in the current MolSimEx implementation, the ASED metric is effectively equivalent to the L2 norm for clustering, yielding assignments and centroids that are indistinguishable from those of standard K-Means. Accordingly, potential advantages over the classical L2 metric have yet to emerge. Future research that improves centroid estimation and aligns the objective function with ASED may advance MolSimEx, enabling the new norm to exceed L2 in specific cases.

Although these findings are promising, they are based on a relatively narrow dataset, and broader evaluations will be needed to validate the benefits of the metrics. The increasing availability of reliable electronic structure calculations is an open field for drug discovery. Future research should broaden these analyses to include larger and more diverse chemical libraries. This research would enable the identification of even more robust and generalizable molecular similarity patterns for the rational design of new drugs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms262311520/s1>. References [59,100–112] are cited in the supplementary materials.

Author Contributions: Conceptualization, R.C.V., É.C.M.N. and J.B.L.M.; Formal analysis, R.C.V., É.C.M.N. and J.B.L.M.; Investigation, R.C.V.; Methodology, É.C.M.N. and J.B.L.M.; Resources, J.B.L.M.; Supervision, J.B.L.M.; Visualization, R.C.V.; Writing—original draft, R.C.V.; Writing—review and editing, É.C.M.N. and J.B.L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Brazilian funding agencies CNPq (Grant No. 306682/2021-4 and 308419/2025-1) and FAPDF (Grant No. 00193–00000869/2021–31).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The source code (MolSimEx, Molecular Similarity Explorer) of this study is freely available at GitHub (<https://github.com/Rafael-Campos-unb/MolSimEx>) (accessed on 3 November 2025).

Acknowledgments: The authors also acknowledge the support of DPI/BCE/UnB. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xie, Z.; Mi, Y.; Kong, L.; Gao, M.; Chen, S.; Chen, W.; Meng, X.; Sun, W.; Chen, S.; Xu, Z. Cannabis sativa: Origin and history, glandular trichome development, and cannabinoid biosynthesis. *Hortic. Res.* **2023**, *10*, uhad150. [[CrossRef](#)] [[PubMed](#)]
2. Rock, E.M.; Parker, L.A. Constituents of Cannabis sativa. In *Cannabinoids and Neuropsychiatric Disorders*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 1–13. [[CrossRef](#)]
3. Pagano, C.; Navarra, G.; Coppola, L.; Avilia, G.; Bifulco, M.; Laezza, C. Cannabinoids: Therapeutic use in clinical practice. *Int. J. Mol. Sci.* **2022**, *23*, 3344. [[CrossRef](#)]
4. Hanuš, L.O.; Meyer, S.M.; Muñoz, E.; Tagliabue, S.; Scafati, O.; Appendino, G. Phytocannabinoids: A unified critical inventory. *Nat. Prod. Rep.* **2016**, *33*, 1357–1392. [[CrossRef](#)]
5. Gülck, T.; Möller, B.L. Phytocannabinoids: Origins and biosynthesis. *Trends Plant Sci.* **2020**, *25*, 985–1004. [[CrossRef](#)]
6. Finn, D.; Haroutounian, S.; Hohmann, A.; Krane, E.; Soliman, N.; Rice, A. Cannabinoids, the endocannabinoid system, and pain: A review of preclinical studies. *Pain* **2021**, *162*, S5–S25. [[CrossRef](#)]

7. Gojani, E.G.; Wang, B.; Li, D.P.; Kovalchuk, O.; Kovalchuk, I. Anti-Inflammatory Effects of Minor Cannabinoids CBC, THCV, and CBN in Human Macrophages. *Molecules* **2023**, *28*, 6487. [[CrossRef](#)]
8. Kwiecień, E.; Kowalczyk, D. Therapeutic Potential of Minor Cannabinoids in Dermatological Diseases—A Synthetic Review. *Molecules* **2023**, *28*, 6149. [[CrossRef](#)]
9. Cooray, R.; Gupta, V.; Suphioglu, C. Current aspects of the endocannabinoid system and targeted THC and CBD phytocannabinoids as potential therapeutics for Parkinson's and Alzheimer's diseases: A review. *Mol. Neurobiol.* **2020**, *57*, 4878–4890. [[CrossRef](#)]
10. Cassano, T.; Villani, R.; Pace, L.; Carbone, A.; Bukke, V.N.; Orkisz, S.; Avolio, C.; Serviddio, G. From Cannabis sativa to cannabidiol: Promising therapeutic candidate for the treatment of neurodegenerative diseases. *Front. Pharmacol.* **2020**, *11*, 124. [[CrossRef](#)] [[PubMed](#)]
11. Walsh, K.B.; McKinney, A.E.; Holmes, A.E. Minor cannabinoids: Biosynthesis, molecular pharmacology and potential therapeutic uses. *Front. Pharmacol.* **2021**, *12*, 777804. [[CrossRef](#)] [[PubMed](#)]
12. Montero, J.M.S.; Agis-Torres, A.; Solano, D.; Söllhuber, M.; Fernandez, M.; Villaro, W.; Gómez-Cañas, M.; García-Arencibia, M.; Fernández-Ruiz, J.; Egea, J.; et al. Analogues of cannabinoids as multitarget drugs in the treatment of Alzheimer's disease. *Eur. J. Pharmacol.* **2021**, *895*, 173875. [[CrossRef](#)] [[PubMed](#)]
13. Fernández-Ruiz, J. The biomedical challenge of neurodegenerative disorders: An opportunity for cannabinoid-based therapies to improve on the poor current therapeutic outcomes. *Br. J. Pharmacol.* **2019**, *176*, 1370–1383. [[CrossRef](#)]
14. Zou, S.; Kumar, U. Cannabinoid receptors and the endocannabinoid system: Signaling and function in the central nervous system. *Int. J. Mol. Sci.* **2018**, *19*, 833. [[CrossRef](#)]
15. Aymerich, M.S.; Aso, E.; Abellanas, M.A.; Tolon, R.M.; Ramos, J.A.; Ferrer, I.; Romero, J.; Fernández-Ruiz, J. Cannabinoid pharmacology/therapeutics in chronic degenerative disorders affecting the central nervous system. *Biochem. Pharmacol.* **2018**, *157*, 67–84. [[CrossRef](#)]
16. Vasquez, R.; Batista, L.; Cuya, T. Computational Study on the Enzyme–Ligand Relationship between Cannabis Phytochemicals and Human Acetylcholinesterase: Implications in Alzheimer's Disease. *J. Phys. Chem. B* **2023**, *127*, 8780–8795. [[CrossRef](#)]
17. Peña-Guerrero, J.; Nguewa, P.A.; García-Sosa, A.T. Machine learning, artificial intelligence, and data science breaking into drug design and neglected diseases. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11*, e1513. [[CrossRef](#)]
18. McCarthy, M.; Lee, K.L.K. Molecule identification with rotational spectroscopy and probabilistic deep learning. *J. Phys. Chem. A* **2020**, *124*, 3002–3017. [[CrossRef](#)]
19. Walters, W.P.; Barzilay, R. Applications of deep learning in molecule generation and molecular property prediction. *Accounts Chem. Res.* **2020**, *54*, 263–270. [[CrossRef](#)]
20. Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Židek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; et al. Highly accurate protein structure prediction for the human proteome. *Nature* **2021**, *596*, 590–596. [[CrossRef](#)] [[PubMed](#)]
21. Guo, Z.; Wu, S.; Ohno, M.; Yoshida, R. Bayesian algorithm for retrosynthesis. *J. Chem. Inf. Model.* **2020**, *60*, 4474–4486. [[CrossRef](#)] [[PubMed](#)]
22. Dara, S.; Dhamecherla, S.; Jadav, S.S.; Babu, C.M.; Ahsan, M.J. Machine learning in drug discovery: A review. *Artif. Intell. Rev.* **2022**, *55*, 1947–1999. [[CrossRef](#)]
23. Lavecchia, A. Machine-learning approaches in drug discovery: Methods and applications. *Drug Discov. Today* **2015**, *20*, 318–331. [[CrossRef](#)] [[PubMed](#)]
24. Stephenson, N.; Shane, E.; Chase, J.; Rowland, J.; Ries, D.; Justice, N.; Zhang, J.; Chan, L.; Cao, R. Survey of machine learning techniques in drug discovery. *Curr. Drug Metab.* **2019**, *20*, 185–193. [[CrossRef](#)] [[PubMed](#)]
25. Cabello-Solorzano, K.; Ortigosa de Araujo, I.; Peña, M.; Correia, L.; Tallón-Ballesteros, A.J. The Impact of Data Normalization on the Accuracy of Machine Learning Algorithms: A Comparative Analysis. In Proceedings of the International Conference on Soft Computing Models in Industrial and Environmental Applications, Salamanca, Spain, 5–7 September 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 344–353. [[CrossRef](#)]
26. Crismon, M.L. Tacrine: First Drug Approved for Alzheimer's Disease. *Ann. Pharmacother.* **1994**, *28*, 744–751. [[CrossRef](#)] [[PubMed](#)]
27. Li, J.; Wu, H.M.; Zhou, R.L.; Liu, G.J.; Dong, B.R. Huperzine A for Alzheimer's disease. *Cochrane Database Syst. Rev.* **2008**, *2*, CD005592. [[CrossRef](#)]
28. Camps, P.; Cusack, B.; Mallender, W.D.; Achab, R.E.; Morral, J.; Muñoz-Torrero, D.; Rosenberry, T.L. Huprine X is a Novel High-Affinity Inhibitor of Acetylcholinesterase That Is of Interest for Treatment of Alzheimer's Disease. *Mol. Pharmacol.* **2000**, *57*, 409–417. [[CrossRef](#)]
29. Nachon, F.; Carletti, E.; Ronco, C.; Trovaslet, M.; Nicolet, Y.; Jean, L.; Renard, P.Y. Crystal structures of human cholinesterases in complex with huprine W and tacrine: Elements of specificity for anti-Alzheimer's drugs targeting acetyl- and butyrylcholinesterase. *Biochem. J.* **2013**, *453*, 393–399. [[CrossRef](#)]
30. Olin, J.; Schneider, L. Galantamine for dementia due to Alzheimer's disease. *Cochrane Database Syst. Rev.* **2002**, *3*, cd001747 [[CrossRef](#)]

31. Triggle, D.J.; Mitchell, J.M.; Filler, R. The Pharmacology of Physostigmine. *CNS Drug Rev.* **1998**, *4*, 87–136. [[CrossRef](#)]
32. Weinreb, O.; Amit, T.; Bar-Am, O.; Youdim, M.B.H. Ladostigil: A Novel Multimodal Neuroprotective Drug with Cholinesterase and Brain-Selective Monoamine Oxidase Inhibitory Activities for Alzheimers Disease Treatment. *Curr. Drug Targets* **2012**, *13*, 483–494. [[CrossRef](#)]
33. Nascimento, É.C.M.; Martins, J.B. Electronic structure and PCA analysis of covalent and non-covalent acetylcholinesterase inhibitors. *J. Mol. Model.* **2011**, *17*, 1371–1379. [[CrossRef](#)]
34. El Khattabi, M.Z.; El Jai, M.; Lahmadi, Y.; Oughdir, L.; Rahhali, M. Understanding the interplay between metrics, normalization forms, and data distribution in K-means clustering: A comparative simulation study. *Arab. J. Sci. Eng.* **2024**, *49*, 2987–3007. [[CrossRef](#)]
35. Ding, C.; He, X. Cluster Structure of K-means Clustering via Principal Component Analysis. In Proceedings of the Advances in Knowledge Discovery and Data Mining, Sydney, Australia, 26–28 May 2004; Dai, H., Srikant, R., Zhang, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 414–418.
36. Asakawa, Y.; Kondo, K.; Tori, M. Cyclopropanochroman derivatives from the liverwort *Radula javanica*. *Phytochemistry* **1991**, *30*, 325–328. [[CrossRef](#)]
37. Leláková, V.; Béraud-Dufour, S.; Hošek, J.; Šmejkal, K.; Prachyawarakorn, V.; Pailee, P.; Widmann, C.; Václavík, J.; Coppola, T.; Mazella, J.; et al. Therapeutic potential of prenylated stilbenoid macasiamenene F through its anti-inflammatory and cytoprotective effects on LPS-challenged monocytes and microglia. *J. Ethnopharmacol.* **2020**, *263*, 113147. [[CrossRef](#)] [[PubMed](#)]
38. Shen, J.; Zhou, Q.; Li, P.; Wang, Z.; Liu, S.; He, C.; Zhang, C.; Xiao, P. Update on Phytochemistry and Pharmacology of Naturally Occurring Resveratrol Oligomers. *Molecules* **2017**, *22*, 2050. [[CrossRef](#)]
39. Dvir, H.; Silman, I.; Harel, M.; Rosenberry, T.L.; Sussman, J.L. Acetylcholinesterase: From 3D structure to function. *Chem.-Biol. Interact.* **2010**, *187*, 10–22. [[CrossRef](#)] [[PubMed](#)]
40. Silman, I.; Sussman, J.L. Acetylcholinesterase: ‘classical’ and ‘non-classical’ functions and pharmacology. *Curr. Opin. Pharmacol.* **2005**, *5*, 293–302. [[CrossRef](#)]
41. Eubanks, L.M.; Rogers, C.J.; Beuscher, Koob, G.F.; Olson, A.J.; Dickerson, T.J.; Janda, K.D. A Molecular Link between the Active Component of Marijuana and Alzheimer’s Disease Pathology. *Mol. Pharm.* **2006**, *3*, 773–777. [[CrossRef](#)]
42. Gessa, G.L.; Casu, M.A.; Carta, G.; Mascia, M.S. Cannabinoids decrease acetylcholine release in the medial-prefrontal cortex and hippocampus, reversal by SR 141716A. *Eur. J. Pharmacol.* **1998**, *355*, 119–124. [[CrossRef](#)]
43. Acquas, E.; Pisanu, A.; Marrocu, P.; Di Chiara, G. Cannabinoid CB1 receptor agonists increase rat cortical and hippocampal acetylcholine release in vivo. *Eur. J. Pharmacol.* **2000**, *401*, 179–185. [[CrossRef](#)]
44. Toyota, M.; Shimamura, T.; Ishii, H.; Renner, M.; Braggins, J.; Asakawa, Y. New Bibenzyl Cannabinoid from the New Zealand Liverwort *Radula marginata*. *Chem. Pharm. Bull.* **2002**, *50*, 1390–1392. [[CrossRef](#)]
45. Silva, M.A.; Kiametis, A.S.; Treptow, W. Donepezil Inhibits Acetylcholinesterase via Multiple Binding Modes at Room Temperature. *J. Chem. Inf. Model.* **2020**, *60*, 3463–3471. [[CrossRef](#)]
46. Zhao, Q.; Tang, X.C. Effects of huperzine A on acetylcholinesterase isoforms in vitro: Comparison with tacrine, donepezil, rivastigmine and physostigmine. *Eur. J. Pharmacol.* **2002**, *455*, 101–107. [[CrossRef](#)]
47. de Paula, A.A.N.; Martins, J.B.L.; Gargano, R.; dos Santos, M.L.; Romeiro, L.A.S. Electronic structure calculations toward new potentially AChE inhibitors. *Chem. Phys. Lett.* **2007**, *446*, 304–308. [[CrossRef](#)]
48. Nascimento, E.C.M.; Martins, J.B.L.; dos Santos, M.L.; Gargano, R. Theoretical study of classical acetylcholinesterase inhibitors. *Chem. Phys. Lett.* **2008**, *458*, 285–289. [[CrossRef](#)]
49. Nascimento, E.C.M.; Oliva, M.; swiderek, K.; Martins, J.B.L.; Andres, J. Binding Analysis of Some Classical Acetylcholinesterase Inhibitors: Insights for a Rational Design Using Free Energy Perturbation Method Calculations with QM/MM MD Simulations. *J. Chem. Inf. Model.* **2017**, *57*, 958–976. [[CrossRef](#)] [[PubMed](#)]
50. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [[CrossRef](#)]
51. Nascimento, L.A.; Nascimento, É.C.; Martins, J.B. In silico study of tacrine and acetylcholine binding profile with human acetylcholinesterase: Docking and electronic structure. *J. Mol. Model.* **2022**, *28*, 252. [[CrossRef](#)]
52. Colovic, M.B.; Krstic, D.Z.; Lazarevic-Pasti, T.D.; Bondzic, A.M.; Vasic, V.M. Acetylcholinesterase Inhibitors: Pharmacology and Toxicology. *Curr. Neuropharmacol.* **2013**, *11*, 315–335. [[CrossRef](#)] [[PubMed](#)]
53. Kryger, G.; Silman, I.; Sussman, J.L. Structure of acetylcholinesterase complexed with E2020 (Aricept®): Implications for the design of new anti-Alzheimer drugs. *Structure* **1999**, *7*, 297–307. [[CrossRef](#)] [[PubMed](#)]
54. Søndergaard, C.R.; Olsson, M.H.M.; Rostkowski, M.; Jensen, J.H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295. [[CrossRef](#)] [[PubMed](#)]
55. Olsson, M.H.M.; Søndergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKaPredictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537. [[CrossRef](#)] [[PubMed](#)]

56. Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. *Proteins Struct. Funct. Genet.* **1999**, *37*, 228–241. [CrossRef]
57. Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; Forli, S. AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* **2021**, *61*, 3891–3898. [CrossRef]
58. Stone, N.L.; Murphy, A.J.; England, T.J.; O'Sullivan, S.E. A systematic review of minor phytocannabinoids with promising neuroprotective potential. *Br. J. Pharmacol.* **2020**, *177*, 4330–4352. [CrossRef]
59. Hussain, T.; Espley, R.V.; Gertsch, J.; Whare, T.; Stehle, F.; Kayser, O. Demystifying the liverwort *Radula marginata*, a critical review on its taxonomy, genetics, cannabinoid phytochemistry and pharmacology. *Phytochem. Rev.* **2019**, *18*, 953–965. [CrossRef]
60. BIOVIA, Dassault Systèmes, Discovery Studio Visualizer, Version 21.1.0.20298, 2021, San Diego. Available online: <https://www.3ds.com/products/biovia> (accessed on 3 November 2025).
61. Shinbo, Y.; Nakamura, Y.; Altaf-Ul-Amin, M.; Asahi, H.; Kurokawa, K.; Arita, M.; Saito, K.; Ohta, D.; Shibata, D.; Kanaya, S. KNApSACk: A comprehensive species-metabolite relationship database. In *Plant Metabolomics*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 165–181. [CrossRef]
62. Liu, S. (Ed.) *Conceptual Density Functional Theory: Towards a New Chemical Reactivity Theory*; Wiley: Hoboken, NJ, USA, 2022. [CrossRef]
63. Geerlings, P. From Density Functional Theory to Conceptual Density Functional Theory and Biosystems. *Pharmaceuticals* **2022**, *15*, 1112. [CrossRef]
64. Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual Density Functional Theory. *Chem. Rev.* **2003**, *103*, 1793–1874. [CrossRef] [PubMed]
65. Fukui, K.; Yonezawa, T.; Shingu, H. A Molecular Orbital Theory of Reactivity in Aromatic Hydrocarbons. *J. Chem. Phys.* **1952**, *20*, 722–725. [CrossRef]
66. Tsuneda, T. Chemical reaction analyses based on orbitals and orbital energies. *Int. J. Quantum Chem.* **2014**, *115*, 270–282. [CrossRef]
67. Martínez, J. Local reactivity descriptors from degenerate frontier molecular orbitals. *Chem. Phys. Lett.* **2009**, *478*, 310–322. [CrossRef]
68. Klopman, G. Chemical reactivity and the concept of charge- and frontier-controlled reactions. *J. Am. Chem. Soc.* **1968**, *90*, 223–234. [CrossRef]
69. Salem, L. Intermolecular orbital theory of the interaction between conjugated systems. I. General theory. *J. Am. Chem. Soc.* **1968**, *90*, 543–552. [CrossRef]
70. Matysiak, J. Evaluation of electronic, lipophilic and membrane affinity effects on antiproliferative activity of 5-substituted-2-(2,4-dihydroxyphenyl)-1,3,4-thiadiazoles against various human cancer cells. *Eur. J. Med. Chem.* **2007**, *42*, 940–947. [CrossRef]
71. Raya, A.; Barrientos-Salcedo, C.; Rubio-Póo, C.; Soriano-Correa, C. Electronic structure evaluation through quantum chemical descriptors of 17 β -aminoestrogens with an anticoagulant effect. *Eur. J. Med. Chem.* **2011**, *46*, 2463–2468. [CrossRef]
72. Wong, K.Y.; Mercader, A.G.; Saavedra, L.M.; Honarparvar, B.; Romanelli, G.P.; Duchowicz, P.R. QSAR analysis on tacrine-related acetylcholinesterase inhibitors. *J. Biomed. Sci.* **2014**, *21*, 84. [CrossRef]
73. Holikulov, U.; Kazachenko, A.S.; Issaoui, N.; Kazachenko, A.S.; Raja, M.; Al-Dossary, O.M.; Xiang, Z. The molecular structure, vibrational spectra, solvation effect, non-covalent interactions investigations of psilocin. *Spectrochim. Acta Part Mol. Biomol. Spectrosc.* **2024**, *320*, 124600. [CrossRef]
74. e Silva, T.B.; Pereira, M.A.; Malta, V.S.; Bento, E.S.; San-Miguel, M.A.; Zioli, R.L.; Martins, J.B.L.; Sih, A.; Taft, C.A. Study of the structure-activity relationship for theoretical molecular descriptors using density functional theory and chemometric methods in cannabinoid metabolites. *Int. J. Quantum Chem.* **2008**, *108*, 2530–2539. [CrossRef]
75. Becke, A.D. Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *J. Chem. Phys.* **1992**, *96*, 2155–2160. [CrossRef]
76. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785. [CrossRef] [PubMed]
77. Vosko, S.H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211. [CrossRef]
78. Stephens, P.J.; Devlin, F.J.; Chabalowski, C.F.; Frisch, M.J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627. [CrossRef]
79. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [CrossRef]
80. Breneman, C.M.; Wiberg, K.B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **1990**, *11*, 361–373. [CrossRef]

81. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian-16 Revision C.01*; Gaussian Inc.: Wallingford, CT, USA, 2016.
82. Zhurko, G. Chemcraft-Graphical Software for Visualization of Quantum Chemistry. 2005. Available online: <https://chemcraftprog.com> (accessed on 3 November 2025).
83. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. [[CrossRef](#)]
84. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; van der Walt, S., Millman, J., Eds.; pp. 56–61. [[CrossRef](#)]
85. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [[CrossRef](#)]
86. Sander, T.; Freyss, J.; von Korff, M.; Reich, J.R.; Rufener, C. OSIRIS, an entirely in-house developed drug discovery informatics system. *J. Chem. Inf. Model.* **2009**, *49*, 232–246. [[CrossRef](#)]
87. Adeshina, Y.O.; Deeds, E.J.; Karanicolas, J. Machine learning classification can reduce false positives in structure-based virtual screening. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 18477–18488. [[CrossRef](#)]
88. Khan, M.I.; Taehwan, P.; Cho, Y.; Scotti, M.; Priscila Barros de Menezes, R.; Husain, F.M.; Alomar, S.Y.; Baig, M.H.; Dong, J.J. Discovery of novel acetylcholinesterase inhibitors through integration of machine learning with genetic algorithm based in silico screening approaches. *Front. Neurosci.* **2023**, *16*, 1007389. [[CrossRef](#)] [[PubMed](#)]
89. Doring, T.H. Tropmol: A Cloud-Based Web Tool to Predict Acetylcholinesterase Inhibitors by Machine Learning. *SSRN* **2025**. [[CrossRef](#)]
90. Kumar, H.; Datusalia, A.K.; Khatik, G.L. Virtual screening of acetylcholinesterase inhibitors through pharmacophore-based 3D-QSAR modeling, ADMET, molecular docking, and MD simulation studies. *Silico Pharmacol.* **2024**, *12*, 13. [[CrossRef](#)]
91. Ahlgren, P.; Jarneving, B.; Rousseau, R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *J. Am. Soc. Inf. Sci. Technol.* **2003**, *54*, 550–560. [[CrossRef](#)]
92. Di Lena, P.; Fariselli, P.; Margara, L.; Vassura, M.; Casadio, R. Is There an Optimal Substitution Matrix for Contact Prediction with Correlated Mutations? *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2011**, *8*, 1017–1028. [[CrossRef](#)] [[PubMed](#)]
93. Lloyd, S.P. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *20*, 129–137. [[CrossRef](#)]
94. Malhat, M.G.; Mousa, H.M.; El-Sisi, A.B. Clustering of chemical data sets for drug discovery. In Proceedings of the 2014 9th International Conference on Informatics and Systems, St. Petersburg, Russia, 24–27 June 2014; p. DEKM-11. [[CrossRef](#)]
95. Akondi, V.S.; Menon, V.; Baudry, J.; Whittle, J. Novel K-means clustering-based undersampling and feature selection for drug discovery applications. In Proceedings of the 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 18–21 November 2019; pp. 2771–2778. [[CrossRef](#)]
96. Hadipour, H.; Liu, C.; Davis, R.; Cardona, S.T.; Hu, P. Deep clustering of small molecules at large-scale via variational autoencoder embedding and K-means. *BMC Bioinform.* **2022**, *23*, 132. [[CrossRef](#)]
97. Bholowalia, P.; Kumar, A. EBK-means: A clustering technique based on elbow method and k-means in WSN. *Int. J. Comput. Appl.* **2014**, *105*, 8887. [[CrossRef](#)]
98. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
99. Tibshirani, R.; Walther, G.; Hastie, T. Estimating the Number of Clusters in a Data Set Via the Gap Statistic. *J. R. Stat. Soc. Ser. Stat. Methodol.* **2002**, *63*, 411–423. [[CrossRef](#)]
100. Appendino, G.; Giana, A.; Gibbons, S.; Maffei, M.; Gnani, G.; Grassi, G.; Sterner, O. A polar cannabinoid from *Cannabis sativa* var. *Carma*. *Nat. Prod. Commun.* **2008**, *3*, 1934578X0800301207. [[CrossRef](#)]
101. Radwan, M.M.; Chandra, S.; Gul, S.; ElSohly, M.A. Cannabinoids, phenolics, terpenes and alkaloids of cannabis. *Molecules* **2021**, *26*, 2774. [[CrossRef](#)]
102. Gollapudi, S.R.; Telikepalli, H.; Keshavarz-Shokri, A.; Vander Velde, D.; Mitscher, L.A. Glep-dotin C: A minor antimicrobial bibenzyl from *Glycyrrhiza lepidota*. *Phytochemistry* **1989**, *28*, 3556–3557. [[CrossRef](#)]
103. Toyota, M.; Kinugawa, T.; Asakawa, Y. Bibenzyl cannabinoid and bisbibenzyl derivative from the liverwort *Radula perrottetii*. *Phytochemistry* **1994**, *37*, 859–862. [[CrossRef](#)]
104. Tinto, F.; Villano, R.; Kostrzewa, M.; Ligresti, A.; Straker, H.; Manzo, E. Synthesis of the Major Mammalian Metabolites of THC. *J. Nat. Prod.* **2020**, *83*, 2060–2065. [[CrossRef](#)]
105. Pollastro, F.; De Petrocellis, L.; Schiano-Moriello, A.; Chianese, G.; Heyman, H.; Appendino, G.; Tagliatalata-Scafati, O. Amorfrutin-type phytocannabinoids from *Helichrysum umbraculigerum*. *Fitoterapia* **2017**, *123*, 13–17. Correction in Reprint of: Amorfrutin-type phytocannabinoids from *Helichrysum umbraculigerum*. *Fitoterapia* **2018**, *126*, 35–39. [[CrossRef](#)]
106. Chang, J.C.; Lai, Y.H.; Djoko, B.; Wu, P.L.; Liu, C.D.; Liu, Y.W.; Chiou, R.Y.Y. Biosynthesis enhancement and antioxidant and anti-inflammatory activities of peanut (*Arachis hypogaea* L.) arachidin-1, arachidin-3, and isopentadienylresveratrol. *J. Agric. Food Chem.* **2006**, *54*, 10281–10287. [[CrossRef](#)]

107. Muhammad, I.; Ibrahim, M.A.; Kumarihamy, M.; Lambert, J.A.; Zhang, J.; Mohammad, M.H.; Khan, S.I.; Pasco, D.S.; Balachandran, P. Cannabinoid and Opioid Receptor Affinity and Modulation of Cancer-Related Signaling Pathways of Machaeriols and Machaeridiols from *Machaerium Pers.* *Molecules* **2023**, *28*, 4162. [[CrossRef](#)]
108. Asakawa, Y.; Nagashima, F.; Ludwiczuk, A. Distribution of bibenzyls, prenyl bibenzyls, bis-bibenzyls, and terpenoids in the liverwort genus *Radula*. *J. Nat. Prod.* **2020**, *83*, 756–769. [[CrossRef](#)]
109. Serino, E.; Iannotti, F.A.; Al-Hmadi, H.B.; Caprioglio, D.; Moriello, C.; Masi, F.; Hammami, S.; Appendino, G.; Vitale, R.M.; Tagliatela-Scafati, O. PPAR α / γ -targeting amorfrutin phy-tocannabinoids from aerial parts of *Glycyrrhiza foetida*. *J. Nat. Prod.* **2023**, *86*, 2435–2447. [[CrossRef](#)]
110. Al-Hmadi, H.B.; Serino, E.; Pastore, A.; Chianese, G.; Hammami, S.; Stornaiuolo, M.; Tagliatela-Scafati, O. Metabolites from Aerial Parts of *Glycyrrhiza foetida* as Modulators of Targets Related to Metabolic Syndrome. *Biomolecules* **2024**, *14*, 467. [[CrossRef](#)]
111. Weidner, C.; Rousseau, M.; Micikas, R.J.; Fischer, C.; Plauth, A.; Wowro, S.J.; Siems, K.; Het-terling, G.; Kliem, M.; Schroeder, F.C.; et al. Amorfrutin C induces apoptosis and inhibits proliferation in colon cancer cells through targeting mitochondria. *J. Nat. Prod.* **2016**, *79*, 2–12. [[CrossRef](#)]
112. Delle Monache, G.; Botta, B.; Vinciguerra, V.; de Mello, J.; de Andrade Chiappeta, A. Antimicrobial isoflavanones from *Desmodium canum*. *Phytochemistry* **1996**, *41*, 537–544. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.