

## Article

# Cheminformatics Analyses of Tau Ligands Reveal Key Molecular Requirements for the Identification of Potential Drug Candidates against Tauopathies

Luca Pinzi <sup>1</sup>, Annachiara Tinivella <sup>1,2</sup>  and Giulio Rastelli <sup>1,\*</sup> 

<sup>1</sup> Department of Life Sciences, University of Modena and Reggio Emilia, Via G. Campi 103/287, 41125 Modena, Italy; luca.pinzi@unimore.it (L.P.); annachiara.tinivella@unimore.it (A.T.)

<sup>2</sup> Clinical and Experimental Medicine PhD Program, University of Modena and Reggio Emilia, 41125 Modena, Italy

\* Correspondence: giulio.rastelli@unimore.it; Tel.: +39-059-205-8564

**Abstract:** Tau is a highly soluble protein mainly localized at a cytoplasmic level in the neuronal cells, which plays a crucial role in the regulation of microtubule dynamic stability. Recent studies have demonstrated that several factors, such as hyperphosphorylation or alterations of Tau metabolism, may contribute to the pathological accumulation of protein aggregates, which can result in neuronal death and the onset of a number of neurological disorders called Tauopathies. At present, there are no available therapeutic remedies able to reduce Tau aggregation, nor are there any structural clues or guidelines for the rational identification of compounds preventing the accumulation of protein aggregates. To help identify the structural properties required for anti-Tau aggregation activity, we performed extensive cheminformatics analyses on a dataset of Tau ligands reported in ChEMBL. The performed analyses allowed us to identify a set of molecular properties that are in common between known active ligands. Moreover, extensive analyses of the fragment composition of reported ligands led to the identification of chemical moieties and fragment combinations prevalent in the more active compounds. Interestingly, many of these fragments were arranged in recurring frameworks, some of which were clearly present in compounds currently under clinical investigation. This work represents the first in-depth cheminformatics study of the molecular properties, constituting fragments and similarity profiles, of known Tau aggregation inhibitors. The datasets of compounds employed for the analyses, the identified molecular fragments and their combinations are made publicly available as supplementary material.



**Citation:** Pinzi, L.; Tinivella, A.; Rastelli, G. Cheminformatics Analyses of Tau Ligands Reveal Key Molecular Requirements for the Identification of Potential Drug Candidates against Tauopathies. *Molecules* **2021**, *26*, 5039. <https://doi.org/10.3390/molecules26165039>

Academic Editor: Martin Vogt

Received: 5 July 2021

Accepted: 16 August 2021

Published: 20 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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/>).

**Keywords:** Tau; tauopathies; Alzheimer's Disease; cheminformatics; ligand-based; molecular descriptors; molecular fragments; drug repurposing; drug design

## 1. Introduction

Tauopathies are a class of heterogeneous neurodegenerative disorders generally characterized by a progressive decline of cognitive functions, change in personality and memory deficit in patients [1,2]. At present, more than twenty neurodegenerative disorders are classified as tauopathies [3,4], the Alzheimer's Disease (AD) being the most common [5,6]. A common hallmark characterizing these neurological disorders is the abnormal hyperphosphorylation and aggregation of the microtubule-associated protein (MAP) Tau in neurofibrillary tangles (NFTs), especially within the neuronal and *glia* cells [7–9]. Tau is a highly soluble protein, ubiquitously expressed in the human brain in six different isoforms [10], which are normally associated with the microtubules (MTs) [11]. From a structural viewpoint, four different regions have been identified in Tau, which belongs to the class of intrinsically disordered proteins (IDPs) [12]. These include: (i) a N-terminal region (NTR) that extends outwards when Tau binds to the MTs, thus regulating its dynamics; (ii) a microtubule-binding region (MBR) hosting the two well-known hexapeptide

segments (i.e., PHF6 (VQIVYK) and PHF6\* (VQIINK)) involved in the aggregation mechanisms [13,14]; (iii) a proline-rich region (PRR), and; (iv) a C-terminal region (CTR), the latter two being highly conserved in mammals [15]. Under physiological conditions, Tau is involved in the regulation of several signaling processes in cells through the interaction with various binding partners [11,15–17]. Importantly, Tau participates in the regulation of microtubules dynamic equilibrium [18] by stimulating tubulin assembly in neuronal cells [19]. The mechanism by which Tau affects the correct physiological function of cells in the brain has yet to be completely elucidated. However, it has been discovered that its physiological function depends on the fine regulation of Tau phosphorylation and dephosphorylation, which is operated by several different proteins in the brain [20]. Several studies have demonstrated that the abnormal hyperphosphorylation, the presence of mutations, and dysregulations in Tau splicing are among the most common causes of tauopathies [15,21–23]. In particular, these studies have shown that hyperphosphorylated Tau tends to dissociate from MTs, thus triggering a cascade of events that promote self-aggregation into oligomers forming paired helical (PHF) and straight (SF) filaments. In turn, the formation of such aggregates promotes their assembly into the neurofibrillary tangles typically observed in brains of patients affected by tauopathies [21]. Moreover, the accumulation of hyperphosphorylated Tau aggregates also leads to sequestration of normal Tau and other MAP proteins [21], which are co-responsible for cellular toxicity.

Given the physio-pathological role of this protein in neurodegenerative disorders, several efforts have been made to identify compounds able to disrupt Tau aggregation, and thus restoring the normal physiological function of neurons and *glia* cells [24–26]. Notwithstanding, therapeutic remedies currently available for the treatment of tauopathies are limited to amelioration or alleviation of symptoms [2,5]. Based on these premises, the identification of drugs able to prevent or resolve the underlying causes of these neurodegenerative diseases is of utmost importance, especially considering that tauopathies mainly affect elderly people [5,6]. In this respect, several compounds able to modulate Tau aggregation have been reported, most of them acting through non-covalent binding mechanisms [25,27]. Moreover, different approaches have also been pursued to aid in the identification of Tau anti-aggregating agents among natural products [28,29], which represent a rich source of potential drug candidates [30,31]. In addition, inhibitors of protein kinases as GSK-3 $\beta$ , PP<sub>2</sub>A, Fyn and CDK5, which indirectly regulate Tau phosphorylation levels, have been reported [32–34]. Furthermore, efforts for the identification of chemical agents that prevent or abrogate Tau aggregation have also been recently reported [32], aurones [35], diamino-phenothiazines [36], rhodanines [27], phenylthiazolylhydrazides [27,37], N-Phenylamines, benzothiazoles and polyphenols being among the most studied chemical classes [38,39]. However, the majority of the identified ligands resulted from independent screenings, often performed under different experimental conditions. As such, structure-activity relationships (SAR) or pharmacophore requirements useful for the identification of Tau aggregation inhibitors remain largely unknown. In this context, the identification of key structural motifs and molecular properties required to achieve high potency would significantly facilitate drug repurposing and/or the discovery of next generation drug candidates.

Based on these premises, in this work we performed a series of chemoinformatics investigations on ChEMBL [40], which included 2D similarity calculations, molecular fragment analysis, and molecular descriptors evaluation, with the aim of identifying key scaffolds, substructures, and molecular properties responsible for Tau activity. The analyses allowed us to identify a set of chemical fragments, as well as combinations of molecular properties that characterize potent Tau anti-aggregation inhibitors. Such information can be used to assist in the rational design or repositioning of Tau aggregation inhibitors.

## 2. Materials and Methods

### 2.1. Database Preparation

Tau aggregation inhibitors were collected from the ChEMBL database (accessed on 5 June 2020) [40] and filtered to retain only records with activity annotations related to experiments on isolated protein, expressed in terms of Potency. In particular, only activity records obtained through thioflavin T and S fluorescence assays on the human microtubule-associated protein Tau (UniProt ID: P10636) were considered. Activity records and compounds deriving from different experiments or from cell-based assays were removed. Moreover, multiple activity annotations deriving from different experiments on the same compound were also removed, retaining the one with the best value. This procedure allowed us to obtain a total of 49,284 unique ligands, each associated with a single activity annotation. The molecular structures and activity data of the filtered ligands were finally stored as separated files and made available to readers as supplementary material (see Supporting Information). All steps of ligand dataset generation were performed with the KNIME software (version 4.3.2) [41].

### 2.2. 2D similarity Calculations

The 2D similarity degree of the filtered compounds was first assessed by means of different molecular fingerprints, e.g., MACCS and ECFP4 from the OpenEye python toolkits [42], and AtomPairs and TopologicalTorsion from the RDKit libraries [43]. In particular, all-vs-all similarity estimations were performed for compounds with Potency values below 500 nM (905 compounds), for a total of around 820,000 comparisons. Moreover, the similarity profile of these compounds was also evaluated with respect to those having a reported Potency value higher than 1  $\mu$ M (47,246 compounds). The 2D similarity was evaluated in terms of the Tanimoto coefficient ( $T_c$ ), by using default settings [44]. Moreover, the similarity records obtained from the ECFP4 fingerprints-based estimations were further analyzed to evaluate whether one or a selection of active ligands might be considered as representative for the entire population of the most potent Tau aggregation inhibitors considered in the analyses. In particular, extensive analyses were performed to identify a set of compounds among the actives, that were able to retrieve at least 45 ligands with an activity below 500 nM in the curated dataset (corresponding to ~5% of the total), and according to commonly employed similarity thresholds [45]. The selected references were also sequentially combined, up to a maximum of 20 queries *per* combination. This allowed us to identify a set of active ligands, which were able to retrieve the highest number of actives, while keeping the number of inactives low. For each combination, the percentage of retrieved active and inactive ligands was evaluated, and the 20 best performing ones were selected, along with their identified ligands, [45] to be further investigated. Moreover, analysis of the similarity records obtained by ECFP4 fingerprints-based estimations was also performed on the clustered dataset of ligands (see Supporting Information).

### 2.3. Analysis and Comparison of Molecular Descriptors

The molecular properties of the compounds under investigation were first analyzed by using the *QikProp* software available in the Schrödinger suite (release 2020-1) [46], with the default settings. To this aim, the compounds were prepared with *LigPrep* [47] to calculate their ionization states and tautomers potentially present at physiological pH, and to energetically minimize their structure. Then, around 50 molecular properties including drug-likeness and blood-brain-barrier (BBB) permeability were calculated with *QikProp* and compared with the corresponding property ranges of approved drugs.

An especially devised KNIME workflow was also developed to compare molecular properties of the active and inactive compounds in the curated dataset. To this end, 118 different molecular descriptors were first evaluated by using the RDKit nodes implemented in KNIME [43]. Compounds were classified as “active” or “inactive” according to their reported Potency values. In particular, ligands with Potency values below 500 nM were considered as actives, while different activity thresholds (i.e.,  $\geq 1$   $\mu$ M,  $\geq 5$   $\mu$ M,  $\geq 10$   $\mu$ M

and  $\geq 20$   $\mu\text{M}$ ) were considered to classify ligands as “inactive”, these values being very often used in chemoinformatics approaches to define inactivity. Statistical details of compounds classified as inactive according to the various thresholds and types of activity are reported in Table S1 (see Supporting Information). Afterwards, the molecular descriptors of the “active” and “inactive” compounds that provided Pearson Correlation Coefficient (PCC) values higher than 0.95 were removed, and statistical distributions of the others were compared through the use of the one-sided Wilcoxon test [48]. This test has been previously employed to evaluate statistical differences among the properties of populations of ligands in chemoinformatics experiments [49,50], including those that are non-normally distributed. The Bonferroni correction was applied to adjust the significance level ( $p$ -value) of the analyses for multiple comparisons of the statistical tests [51]. Finally, the molecular descriptors that resulted statistically different in the two populations were further evaluated. This allowed us to highlight relevant differences in the molecular properties of the active and inactive ligands at different thresholds of inactivity.

#### 2.4. Molecular Fragment Analyses

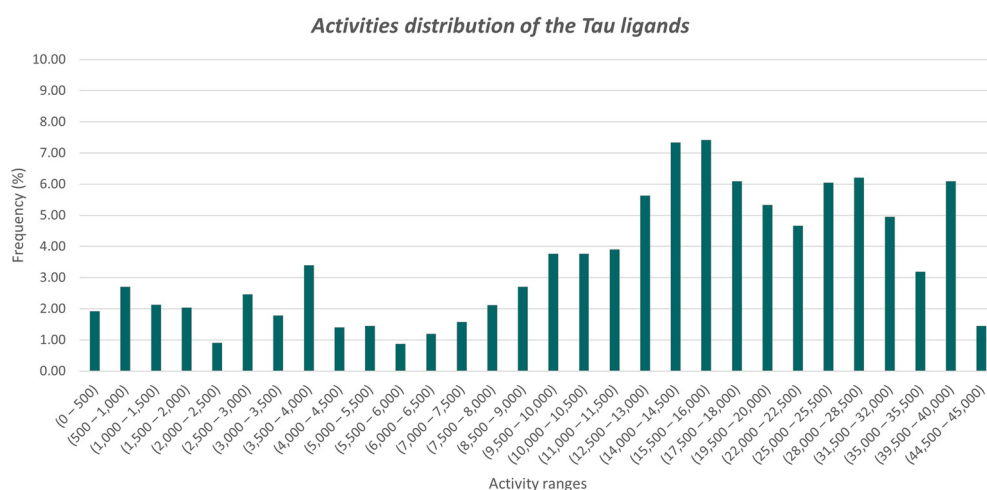
An analysis of the molecular fragments characterizing the compounds in the prepared database was also performed. To this aim, the collected compounds were first fragmented by using an *in house* developed python script implemented with the RDKit libraries [43] and the OpenEye toolkits [42]. Different types of fragmentation algorithms were used in the analyses, including BRICS [52], Bemis-Murcko [53] and Recap [54]. Moreover, the Chomp software (version 3.1.1.2—OpenEye) [55] was also used with default settings, as it allows to generate molecular fragments according to the RLF chemical heuristics seek, which allows to break all non-ring and non-resonance single bonds of a given molecule. Then, fragments with a number of atoms outside the range of 5 to 55, and present in less than three molecules were removed. Duplicate structures derived by fragmentation of different compounds were also removed. The number of unique molecular fragments obtained for the active and inactive Tau ligands is reported in Table S2. Afterwards, a workflow implemented in KNIME (version 4.3.2) was devised to identify molecular fragments and their combinations present only, or in common between, active and inactive Tau ligands. Molecular fragments and their combinations generated in these analyses are made available as supplementary material (see Supporting Information).

### 3. Results and Discussion

#### 3.1. Dataset Preparation for the Analyses

A dataset of Tau ligands was first generated as detailed in Section 2. The selection of Tau aggregation inhibitors was limited to compounds that were assayed under comparable experimental conditions, by selecting compounds tested with Thioflavin T and S fluorescence assays (see Table S3). Although these assays can be performed *in tandem* with other experiments (e.g., circular dichroism spectroscopy or microscopy, and atomic force microscopy) [35,56,57], in order to circumvent potential false positive readouts, data from the Thioflavin T and S fluorescence experiments provide a rich source of information for chemoinformatics analyses and an appropriate statistical representation of the entire population of Tau chemotypes. Indeed, the number of compounds that were not tested with the Thioflavin T and S fluorescence assays was sensibly lower (see Table S1 in the Supporting Information). Activity data distributions showed that more than 95% of the reported ligands have Potency values above 1  $\mu\text{M}$  (Figure 1), while less than 2% had values below 500 nM. This data is of particular interest, considering that more than 100 chemotypes were identified in the 0–500 nM activity range by visual inspection of their chemical structures.

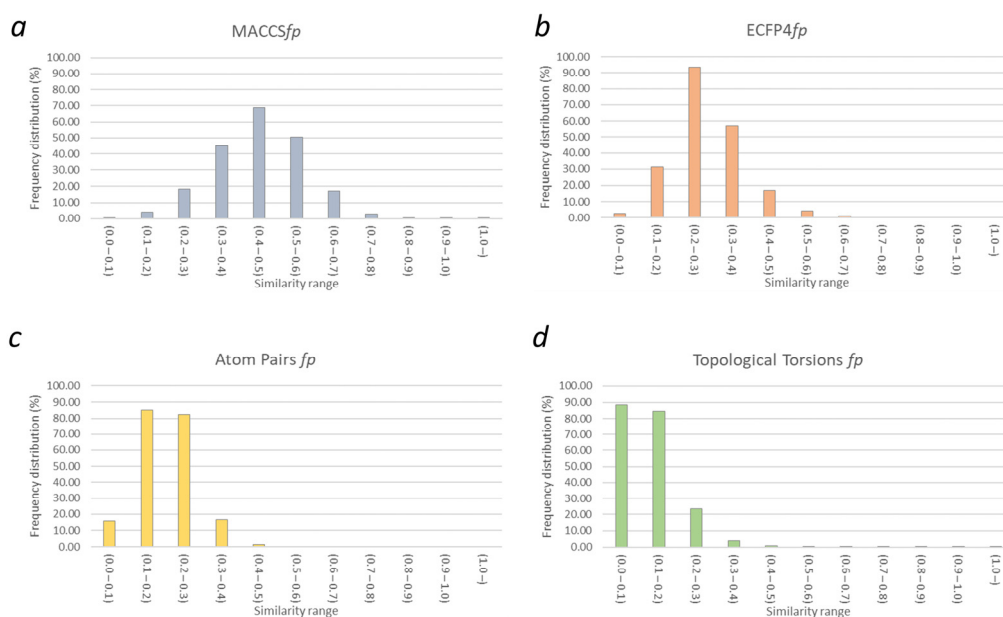
Of note, some of the molecules in the curated dataset have also been tested in different types of assays, providing comparable results (e.g., see CHEMBL140 and CHEMBL191083 in [58,59]).



**Figure 1.** Activity values distribution of known Tau compounds in the curated dataset, each range corresponding to a 500 nM interval. The frequencies are represented by means of their percentage with respect to the total number of activity records (i.e., 49,188). Activity ranges including less than five compounds are not shown.

### 3.2. Similarity Calculations in the Set of Active Compounds

The similarity profile of compounds in the “actives” set (0–500 nM) was evaluated by means of several types of fingerprints, as described above (see Section 2.3). The results, which are shown in Figure 2, highlight an overall low degree of similarity between the compounds of this dataset.



**Figure 2.** Similarity distributions of Tau ligands in the “actives” dataset. Panels (a–d) show results of the 2D similarities calculated with the MACCS, ECFP4, AtomPairs and TopologicalTorsion fingerprints, respectively.

In particular, the similarity analyses showed that the active compounds present high diversity in terms of fragment composition (Figure 2). Indeed, the evaluated Tc values according to MACCSfp were largely below the commonly reported similarity threshold (less than 1% of all comparisons). On the contrary, a higher degree of similarity could be observed in their connectivity tables, as more than 38% of the similarities evaluated with



ECFP4 fingerprints provided Tc values above 0.3 [45]. Table 1 summarizes the number and percentage of similar pairs identified among active compounds.

**Table 1.** Results of the similarity estimations performed in the curated dataset of Tau ligands. The similarity estimations were performed in the “actives” (0–500 nM) and “inactives” ( $\geq 1$   $\mu$ M) datasets.

Type of Fingerprints <sup>1</sup>	Number of Identified Pairs (Actives) <sup>2</sup>	% of Identified Pairs (Actives) <sup>3</sup>	Number of Identified Pairs (Inactives) <sup>2</sup>	% of Identified Pairs (Inactives) <sup>3</sup>
MACCSfp	1859	0.5	27,604	0.1
ECFP4fp	159,125	38.8	13,460,261	30.8
AtomPairsfp	9093	2.2	488,407	1.1
TopologicalTorsionfp	42,477	10.4	2,670,815	6.1

<sup>1</sup> Similarity thresholds were taken from [45] for MACCS and ECFP4, and from [60] for AtomPairs and TopologicalTorsion fingerprints.

<sup>2</sup> The total numbers of active and inactive compounds considered in the analyses are 905 and 47,246, respectively. <sup>3</sup> The total number of comparisons performed by considering only active compounds are 409,965, which became 43,733,220 by including also inactive ligands.

According to the performed similarity estimations, different results were observed by using AtomPairs and TopologicalTorsion fingerprints, which evaluate chemical similarity in terms of atom-based and consecutive non-hydrogen bond environments, respectively [61,62]. Indeed, the percentages of similarity revealed by these fingerprints were 2.2% (AtomPairs) and 10.4% (TopologicalTorsion), these values being calculated according to the thresholds for randomness (95% level) suggested in the RDKit documentation [60]. Overall, this analysis highlighted a low degree of similarity between the curated dataset of the Tau active compounds, which mostly originates from their different fragment and chemotype composition. The adopted fingerprints provided a different performance in identifying pairs of similar ligands, the highest numbers of associations being observed for ECFP4fp and TopologicalTorsionfp also in consensus screenings (Table S4). Of note, ECFP4fp showed the best performance in retrieving active compounds (Table 1). However, ECFP4fp fingerprints yielded the highest number of comparisons above the selected threshold of similarity (30.8% of the total) when inactive ligands were also included in the similarity analyses (Table 1). Consequently, using ECFP4fp in a virtual screening protocol could potentially increase the number of false positives; therefore, the integration with other methods should be carefully evaluated.

The fact that ECFP4fp provided the highest number of similarities suggests that this type of fingerprint was able to more efficiently identify common structural patterns among active ligands. Such a feature might be of interest, for example, for the selection of representative queries in a virtual screening. Indeed, an analysis of the ECFP4fp-based similarity records showed that ChEMBL1558683 was able to retrieve up to 10% of the active ligands, the statistics of the best performing compounds being reported in Table S5. Different results were observed when inactive compounds were also considered in the similarity analyses. In this case, the highest difference in the percentage of retrieved active vs. inactive ligands was obtained for ChEMBL1555206. Although the best performing compounds were able to retrieve a reasonable number of active ligands, their screening performance was not satisfactory. Indeed, the observed enrichment factors (EF), the area under the curve (AUC) and the BEDROC values were, on average, below the performances desirable in virtual screenings [63,64]. The best AUC and BEDROC values were observed for ChEMBL1512606 (AUC  $\sim$ 0.7) and ChEMBL1377126 (BEDROC  $\sim$ 0.38), respectively. To further evaluate whether the common structural patterns observed in the previous ECFP4fp-based analyses might be due to the observed high degree of similarity among small subsets of active ligands, clustering analyses were also performed. In particular, different datasets including active and inactive ligands at different thresholds of inactivity (i.e.,  $\geq 1$   $\mu$ M,  $\geq 5$   $\mu$ M,  $\geq 10$   $\mu$ M and  $\geq 20$   $\mu$ M) were first clustered with Canvas (Schrödinger suite 2020-1) [65], as described in the Supporting Information. Then, the ability of the active compounds in each of the generated clusters in retrieving actives and inactives was evaluated. Interesting results were obtained for clusters identified from the datasets of

the most potent Tau aggregation inhibitors and ligands with a reported Potency value above 10  $\mu\text{M}$  or 20  $\mu\text{M}$ . Indeed, the majority of them presented at least an active compound showing satisfactory AUC and EF values (Table S6), while poorer prediction performances were obtained for clusters defined by activity thresholds of  $\geq 1 \mu\text{M}$  and  $\geq 5 \mu\text{M}$  (Table S6). This result is indicative of the presence of a high number of common structural patterns between small subsets of active ligands, in line with the previously observed chemotype variability.

These results prompted us to also evaluate whether a combination of ligands would be able to more efficiently represent the population of the active compounds in the dataset. To this aim, the similarity data obtained from the ECFP4 $fp$ -based analyses (see Section 2.2) were further processed, combining the records of the best performing compounds, up to a maximum of twenty ligands *per* combination. For each combination we evaluated the percentages of retrieved active and inactive ligands (see Table 2). The list of reference compounds in the identified combinations is reported in Table S7. The identified combinations of references allowed us to retrieve up to one third of the actives in the curated dataset. The best results were obtained for the combination of 20 queries, both in terms of higher number of actives, and greater difference between the percentages of active and inactive ligands retrieved by the similarity records.

**Table 2.** Percentages of active and inactive ligands retrieved by the best performing combinations of reference compounds, according to the ECFP4 similarity estimations.

Number of Queries <i>per</i> Combination	% of Actives ( $\leq 500 \text{ nM}$ ) <sup>1</sup>	% of Inactives ( $\geq 1 \mu\text{M}$ ) <sup>1</sup>	Difference between the % of Actives and Inactives <sup>2</sup>
1	10.8 (98)	3.2 (1568)	7.6 (16)
2	18.1 (164)	5.1 (2465)	13.0 (15)
3	22.4 (203)	6.6 (3189)	15.8 (16)
4	24.2 (219)	6.7 (3238)	17.5 (15)
5	28.3 (256)	9.5 (4591)	18.8 (18)
6	30.1 (272)	10.5 (5074)	19.6 (19)
7	32.7 (296)	12.3 (5944)	20.4 (20)
8	32.2 (291)	12.1 (5847)	20.1 (20)
9	33.7 (305)	13.6 (6572)	20.1 (22)
10	33.8 (306)	13.9 (6717)	19.9 (22)
11	32.9 (298)	12.9 (6234)	20.0 (21)
12	33.7 (305)	13.3 (6427)	20.4 (21)
13	35.0 (317)	14.3 (6910)	20.7 (22)
14	35.0 (317)	14.4 (6959)	20.6 (22)
15	35.1 (318)	14.5 (7007)	20.6 (22)
16	35.7 (323)	14.8 (7152)	20.9 (22)
17	36.8 (333)	15.7 (7587)	21.1 (23)
18	36.8 (333)	15.7 (7587)	21.1 (23)
19	36.8 (333)	15.6 (7539)	21.2 (23)
20	37.1 (336)	15.9 (7684)	21.2 (23)

<sup>1</sup> The number of compounds retrieved by the combination of references is reported in round brackets. <sup>2</sup> The number of inactives *per* active compound is reported in round brackets. The total numbers of active and inactive compounds are 905 and 47,246, respectively.

As shown in Table 2 and Figure S1, the rate of retrieved actives per number of queries grows higher than that of the inactives, for combinations of one to seven references. In contrast, the number of actives retrieved by combinations of ten or more queries remained more or less stable. However, results reported in Table S7 showed that the use of combinations of queries still provided poor discriminating performances according to the AUC and BEDROC indexes. Overall, although the use of combinations of queries increased the number of retrieved active compounds, their prediction performances were still not satisfactory. Altogether, results of the similarity analyses did not allow the identification of compounds (or combinations thereof) to be used as representatives of the entire popula-

tion of actives, or able to discriminate actives from inactives. Instead, results implied the presence of a number of small groups of highly structurally similar ligands in the analyzed dataset, suggesting that 2D similarity estimations alone might not be an ideal method to perform, for example, virtual screening on this target.

### 3.3. Analysis and Comparison of Molecular Descriptors

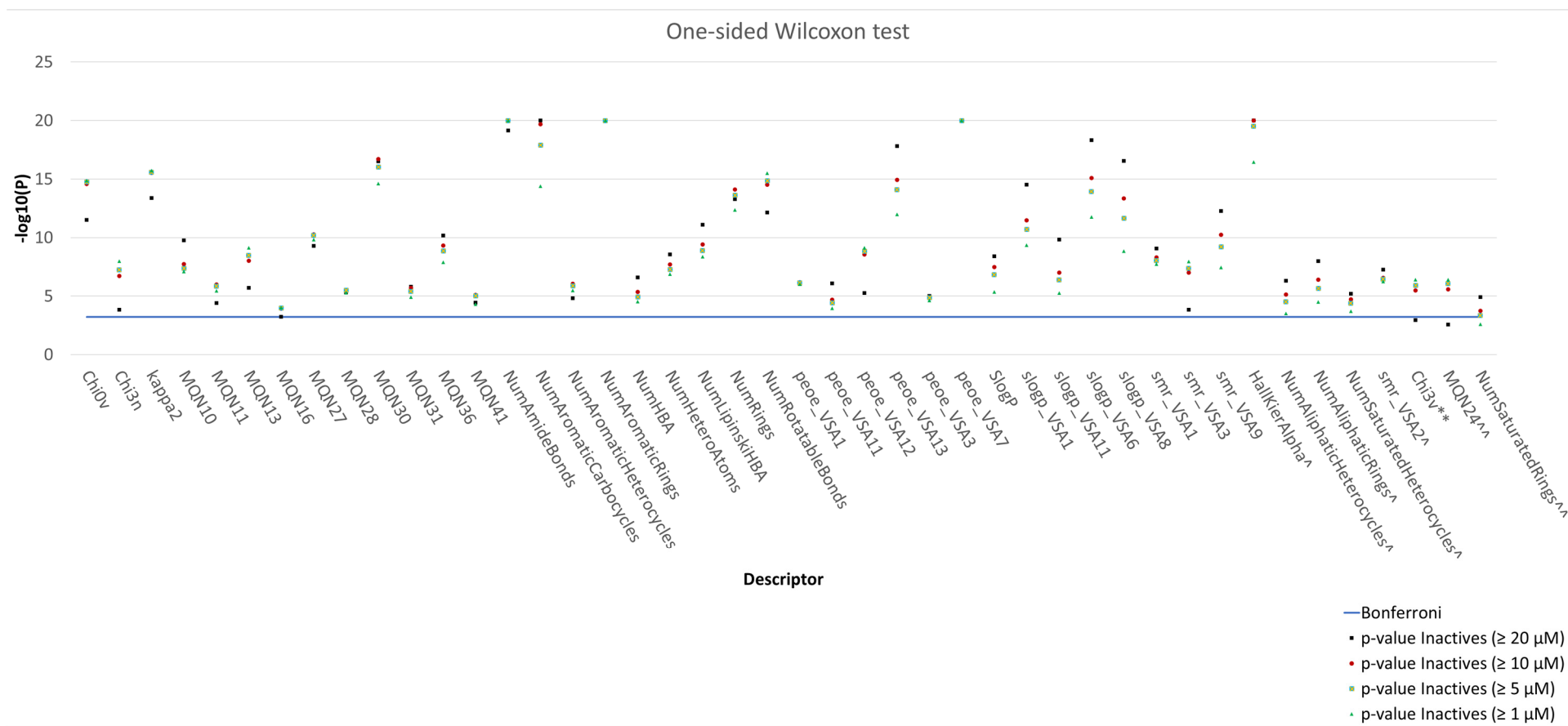
The compounds were then evaluated for their drug-like properties and blood-brain-barrier (BBB) permeability. This analysis was performed by means of the *QikProp* software available in the Schrödinger suite (release 2020-1) [46,47]. The results showed that the majority of the compounds present good drug-like properties and are able to cross the BBB, such prediction being especially true for compounds with activity below 500 nM. The results of this analysis are reported in Table S8.

Extensive *in silico* analyses were then performed to evaluate whether the active and inactive compounds present statistically different molecular properties. The analyses were performed on different populations of inactive ligands, i.e., by considering  $\geq 1$   $\mu\text{M}$ ,  $\geq 5$   $\mu\text{M}$ ,  $\geq 10$   $\mu\text{M}$  and  $\geq 20$   $\mu\text{M}$  inactivity thresholds, and by using the 118 molecular descriptors implemented in the “RDKit Descriptor Calculation” node available in KNIME. Molecular descriptors with  $\text{PCC} \geq 0.95$  were removed to avoid potential inter-correlation issues. Then, a statistical comparison of the remaining molecular descriptors was performed through a one-sided Wilcoxon test, to evaluate whether the mean values in the actives and inactives datasets were statistically different. The results, shown in Figure 3 and Figure S2, demonstrate that 41 of the calculated molecular descriptors provided statistically different results (see Table S9), although none of them alone was able to clearly separate actives from inactives (see Table S10). These results are likely due to the high molecular diversity observed in the dataset.

Interestingly, these analyses allowed us to delineate a set of molecular properties that should be present in active compounds.

In particular, the introduction of aryl-substituted H-bond acceptors is generally favored. Indeed, the *NumAmideBonds* and *NumHBA* descriptors were on average higher for the active compounds. Moreover, active compounds had a higher number of aromatic and heteroaromatic rings compared to the inactives (on average from 2 to 4 and up to 2, respectively). Also, the presence of aliphatic moieties is associated with inactivity, as highlighted by the lower values of the *NumAliphaticHeterocycles*, *NumAliphaticRings* and *NumSaturatedHeterocycles* molecular descriptors in the active compounds. Consistently, the *NumSaturatedRings* of inactive compounds was lower at 5  $\mu\text{M}$ , 10  $\mu\text{M}$  and 20  $\mu\text{M}$  inactivity thresholds. Together, these results suggest that an increase of the aliphatic character of the molecules result in a decrease of activity. Moreover, it also suggests that aromatic compounds able to establish  $\pi$ - $\pi$  interactions may more efficiently interact with Tau. The active compounds had, on average, higher values of *logP* (Table S10). Active and inactive compounds also differ for their atomic valence connectivity indexes, which tend to be higher for the former class (e.g., *Chi0v*, *Chi3n* and *Kappa 2*) [66]. Moreover, higher values of MOE-type descriptors related to the surface area with partial charges (e.g., *PEOE\_VSA<sub>1,3,7,11,12,13</sub>*), molar refractivity (e.g., *SMR\_VSA<sub>1,3,9</sub>*) and *LogP* (e.g., *SlogP\_VSA<sub>1,6,8,11</sub>*) were also observed for the more active compounds [66,67]. Finally, topological descriptors based on the count of simple structural features, i.e., cyclic divalent nodes (MQN30) and 6-membered rings (MQN36) [68], were higher for the active compounds. In conclusion, the analysis of molecular descriptors provided some hints on molecular properties that active compounds should possess, which could be useful for drug design.



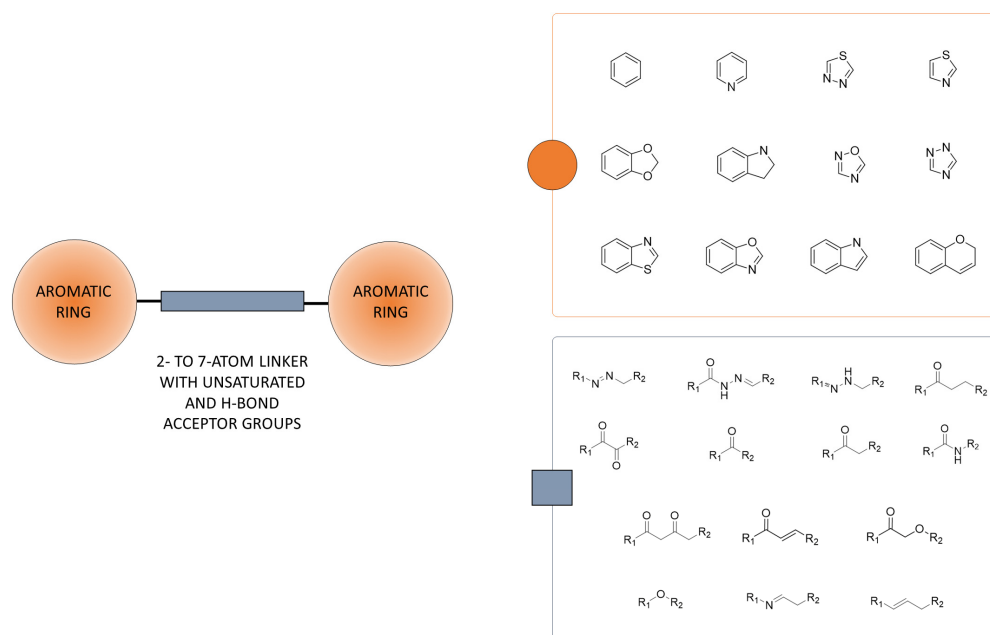


**Figure 3.** Distribution plot of statistically different molecular descriptors. Each distribution is represented by means of its negative logarithmic  $p$ -value of the one-sided Wilcoxon test. The Bonferroni threshold, which is displayed as a blue line, highlights the adjusted significance level ( $p$ -value) for multiple comparisons at statistical tests. Molecular descriptors marked with “~” are those whose values are on average lower for the inactive compounds with respect to the actives. Molecular descriptors marked as “\*\*\*” present values that are, on average, higher for the active compounds, with respect to inactives at 1  $\mu$ M, 5  $\mu$ M and 10  $\mu$ M activity thresholds. Molecular descriptors that are marked with “^^” present values that are, on average, higher for the active compounds, with respect to inactives at 5  $\mu$ M, 10  $\mu$ M and 20  $\mu$ M activity thresholds.

### 3.4. Analysis of Molecular Fragments

An analysis of the molecular fragment composition for compounds in the curated dataset was also performed. Fragmentation was performed by means of five different algorithms. This allowed us to exhaustively identify chemical moieties and functional groups present in Tau active and/or inactive compounds at different thresholds of inactivity. In particular, we were able to identify 38 chemical moieties exclusively present in the active ligands (Table S11), and 70,330, 63,678, 57,407 and 25,637 molecular fragments present only in the inactives at 1  $\mu$ M, 5  $\mu$ M, 10  $\mu$ M and 20  $\mu$ M thresholds of inactivity, respectively. Statistics related to the identified fragments are reported in Table S12. Visual inspection of the molecular fragments of active compounds revealed the presence of at least one hydrogen bond acceptor group. Moreover, around 50% of them had also one or more aromatic or heteroaromatic ring(s), often bonded together or through carbonyl and/or linear alkenes linkers. A high number of fragments in common between the active and inactive sets have also been detected. However, also in this case, aromatic and heteroaromatic groups and fragments bearing H-bond acceptors were generally prevalent in the active compounds (Table S12). On the contrary, molecular fragments with hydrophilic rings, such as piperazine and morpholine, or saturated rings, were more present in the inactive compounds. A list of the more frequently observed molecular fragments is reported in Table S13, along with their percentages of occurrence in the active and inactive sets. Altogether, these results suggest that aromatic/heteroaromatic fragments and H-bond acceptors are expected to contribute favorably to activity, as opposed to hydrophilic or saturated chemical moieties. The complete list of molecular fragments identified in the analysis is provided as supplementary material (see Supporting Information). Interestingly, we also found that certain combinations of molecular fragments are especially present in the active dataset (Table S14). These include the combination of 2 or 3 molecular fragments, for example the phenyl and carbonyl groups with fragments bearing other H-bond acceptors, which were present in more than 32% and 16% of the active compounds, respectively (see Table S14, sections A and B). The combinations of molecular fragments identified in active compounds are provided as supplementary material (see Supporting Information). Of note, visual inspection of the more frequent fragment combinations revealed that active compounds are often organized into molecular frameworks composed by two aromatic or heteroaromatic rings separated by different 2- to 7-atom linkers. This latter moiety is very often decorated with H-bond acceptors or may consist of rigid linear unsaturated functional groups (Figure 4).

Notably, the identified framework is present in some of the already reported Tau-based drug candidates currently undergoing clinical trials (e.g., xalsalate and curcumin) (Figure S3) [69,70], as well as in the natural compounds xanthohumol and licochalcone A that act as Tau aggregation inhibitors [70,71] (Figure S3). The latter chemotype is currently under evaluation on several targets related to neurodegenerative diseases, including monoaminoxidase B,  $\alpha\beta$ -amyloid and  $\alpha$ -sinuclein [72–74]. These findings further strengthen the potential use of molecular fragment combinations based on this framework as starting points for the development of novel Tau aggregation inhibitors.



**Figure 4.** General molecular framework identified by visual inspection of the more frequently observed combinations of fragments in the active compounds. A list of the top-100 combinations observed in the more active ligands is reported in Sections A and B of Table S15.

#### 4. Conclusions

In this work, we have reported the first extensive cheminformatics study on Tau ligands that systematically analyzed the similarity profiles, molecular properties and chemical fragments of a large dataset of compounds with known activity on Tau aggregation. In particular, similarity estimations were performed by means of different types of fingerprints, identifying a high degree of structural diversity. Comparison of the molecular properties of active and inactive compounds confirmed the presence of significant structural diversity and allowed us to outline a set of molecular properties that compounds should possess in order to display high potency. A comparative analysis of the molecular fragments of Tau ligands corroborated the results obtained with the molecular descriptors and brought attention to the need for aromatic/heteroaromatic rings and hydrogen bond acceptors, as opposed to hydrophilic or saturated chemical moieties. Notably, an analysis of fragment combinations showed that active compounds are often organized into molecular frameworks composed by two aromatic or heteroaromatic rings separated by different 2- to 7-atom linkers, the latter moiety being very often decorated with H-bond acceptors or possibly consisting of rigid linear unsaturated functional groups. Altogether, the results of this study showed the possibility to identify a set of molecular properties and fragment combinations that can be useful for *de novo* drug design or drug repurposing campaigns.

**Supplementary Materials:** The following are available online, Figure S1: Percentages of active and inactive compounds retrieved by means of the selected query combinations.; Figure S2: Distribution plots of the molecular descriptors compared in the analyses.; Figure S3: Chemical structures of the xalsalate, curcumin, xanthohumol and licochalcone A known Tau aggregation inhibitors.; Table S1: Number of ligands according to different activity thresholds and type.; Table S2: Number of unique molecular fragments identified from the analyses of the active and inactive datasets.; Table S3: Assay types and description related to the compounds considered in this work.; Table S4: Statistics related to the combinations of the different fingerprints employed in the similarity estimations.; Table S5: Percentages of actives and inactive ligands retrieved by the best performing queries, according to the ECFP4fp similarity estimations.; Table S6: Best-three performing queries identified for each cluster in the different datasets of ligands.; Table S7: Results of the ECFP4fp-based similarity screenings obtained

for the best performing set of queries per combination.; Table S8: Percentages of Tau aggregation inhibitors that present good drug-like properties according to commonly recommended values, as evaluated with QikProp (Schrödinger 2020-1).; Table S9: Mean ( $\pm$ Standard deviation), and 10th and 90th percentiles evaluated for the molecular descriptors that emerged as different for the active and inactive compounds.; Table S10: Percentage of compounds with molecular descriptors within the 10th and 90th percentile thresholds of the active compounds.; Table S11: Molecular fragments observed exclusively in the active compounds.; Table S12: Statistics related to molecular fragments identified in the active and inactive datasets.; Table S13: Molecular fragments that provided the larger difference in their prevalence across the active and inactive classes.; Table S14: Combinations of fragments more frequently observed in the active ligands, according to percentages of occurrence.; Ligands clustering; The datasets of compounds employed for the analyses, the molecular fragments and their combinations identified are provided as supplementary material.

**Author Contributions:** Conceptualization, G.R.; data curation, analysis and visualization, L.P.; methodology, L.P. and A.T.; original draft preparation, L.P.; review and editing, L.P., A.T. and G.R.; funding acquisition, G.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by FAR—Fondo di Ateneo per la Ricerca 2019, grant number 166835, and by Banca Popolare dell’Emilia Romagna (BPER). The APC was funded by MDPI.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank OpenEye Scientific Software, Inc., for a free academic license of the OpenEye Toolkit.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Sample Availability:** Samples are available from the authors.

## References

1. Chi, H.; Sang, T.-K.; Chang, H.-Y. Tauopathy. In *Cognitive Disorders*; Foyaca Sibat, H., Ed.; IntechOpen: Rijeka, Croatia, 2019; pp. 33–63.
2. Josephs, K.A. Current Understanding of Neurodegenerative Diseases Associated With the Protein Tau. *Mayo Clinic Proc.* **2017**, *92*, 1291–1303. [[CrossRef](#)] [[PubMed](#)]
3. Arendt, T.; Stieler, J.T.; Holzer, M. Tau and Tauopathies. *Brain Res. Bull.* **2016**, *126*, 238–292. [[CrossRef](#)] [[PubMed](#)]
4. Williams, D.R. Tauopathies: Classification and Clinical Update on Neurodegenerative Diseases Associated with Microtubule-Associated Protein Tau. *Intern. Med. J.* **2006**, *36*, 652–660. [[CrossRef](#)]
5. Orr, M.E.; Sullivan, A.C.; Frost, B. A Brief Overview of Tauopathy: Causes, Consequences, and Therapeutic Strategies. *Trends Pharmacol. Sci.* **2017**, *38*, 637–648. [[CrossRef](#)] [[PubMed](#)]
6. Alzheimer’s Association. 2019 Alzheimer’s Disease Facts and Figures. *Alzheimer’s Dement.* **2019**, *15*, 321–387. [[CrossRef](#)]
7. Sergeant, N.; Delacourte, A.; Buée, L. Tau Protein as a Differential Biomarker of Tauopathies. *Biochim. Biophys. Acta (BBA) Mol. Basis Dis.* **2005**, *1739*, 179–197. [[CrossRef](#)]
8. Ittner, L.M.; Götz, J. Amyloid- $\beta$  and Tau—A Toxic Pas de Deux in Alzheimer’s Disease. *Nat. Rev. Neurosci.* **2011**, *12*, 65–72. [[CrossRef](#)]
9. Spillantini, M.G.; Goedert, M. Tau Pathology and Neurodegeneration. *Lancet Neurol.* **2013**, *12*, 609–622. [[CrossRef](#)]
10. Bakota, L.; Ussif, A.; Jeserich, G.; Brandt, R. Systemic and Network Functions of the Microtubule-Associated Protein Tau: Implications for Tau-Based Therapies. *Mol. Cell. Neurosci.* **2017**, *84*, 132–141. [[CrossRef](#)]
11. Morris, M.; Maeda, S.; Vossel, K.; Mucke, L. The Many Faces of Tau. *Neuron* **2011**, *70*, 410–426. [[CrossRef](#)]
12. Skrabana, R.; Skrabanova, M.; Csokova, N.; Sevcik, J.; Novak, M. Intrinsically Disordered Tau Protein in Alzheimer’s Tangles: A Coincidence or a Rule? *Bratisl. Lek. Listy* **2006**, *107*, 354–358.
13. von Bergen, M.; Barghorn, S.; Li, L.; Marx, A.; Biernat, J.; Mandelkow, E.-M.; Mandelkow, E. Mutations of Tau Protein in Frontotemporal Dementia Promote Aggregation of Paired Helical Filaments by Enhancing Local  $\beta$ -Structure. *J. Biol. Chem.* **2001**, *276*, 48165–48174. [[CrossRef](#)]
14. von Bergen, M.; Friedhoff, P.; Biernat, J.; Heberle, J.; Mandelkow, E.-M.; Mandelkow, E. Assembly of  $\tau$  Protein into Alzheimer Paired Helical Filaments Depends on a Local Sequence Motif ( $^{306}$ VQIVYK $^{311}$ ) Forming  $\beta$  Structure. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5129–5134. [[CrossRef](#)] [[PubMed](#)]

15. Trushina, N.I.; Bakota, L.; Mulikidjanian, A.Y.; Brandt, R. The Evolution of Tau Phosphorylation and Interactions. *Front. Aging Neurosci.* **2019**, *11*, 256. [[CrossRef](#)] [[PubMed](#)]
16. Brandt, R.; Trushina, N.I.; Bakota, L. Much More Than a Cytoskeletal Protein: Physiological and Pathological Functions of the Non-Microtubule Binding Region of Tau. *Front. Neurol.* **2020**, *11*, 590059. [[CrossRef](#)] [[PubMed](#)]
17. Uversky, V.N. Intrinsically Disordered Proteins and Their (Disordered) Proteomes in Neurodegenerative Disorders. *Front. Aging Neurosci.* **2015**, *7*, 18. [[CrossRef](#)] [[PubMed](#)]
18. Janning, D.; Igaev, M.; Sündermann, F.; Brühmann, J.; Beutel, O.; Heinisch, J.J.; Bakota, L.; Piehler, J.; Junge, W.; Brandt, R. Single-Molecule Tracking of Tau Reveals Fast Kiss-and-Hop Interaction with Microtubules in Living Neurons. *Mol. Biol. Cell* **2014**, *25*, 3541–3551. [[CrossRef](#)]
19. Weingarten, M.D.; Lockwood, A.H.; Hwo, S.Y.; Kirschner, M.W. A Protein Factor Essential for Microtubule Assembly. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1858–1862. [[CrossRef](#)] [[PubMed](#)]
20. Wang, J.-Z.; Grundke-Iqbal, I.; Iqbal, K. Kinases and Phosphatases and Tau Sites Involved in Alzheimer Neurofibrillary Degeneration. *Eur. J. Neurosci.* **2007**, *25*, 59–68. [[CrossRef](#)] [[PubMed](#)]
21. Gong, C.-X.; Iqbal, K. Hyperphosphorylation of Microtubule-Associated Protein Tau: A Promising Therapeutic Target for Alzheimer Disease. *Curr. Med. Chem.* **2008**, *15*, 2321–2328. [[CrossRef](#)]
22. Schoch, K.M.; DeVos, S.L.; Miller, R.L.; Chun, S.J.; Norrbom, M.; Wozniak, D.F.; Dawson, H.N.; Bennett, C.F.; Rigo, F.; Miller, T.M. Increased 4R-Tau Induces Pathological Changes in a Human-Tau Mouse Model. *Neuron* **2016**, *90*, 941–947. [[CrossRef](#)]
23. Kent, S.A.; Spires-Jones, T.L.; Durrant, C.S. The Physiological Roles of Tau and A $\beta$ : Implications for Alzheimer’s Disease Pathology and Therapeutics. *Acta Neuropathol.* **2020**, *140*, 417–447. [[CrossRef](#)]
24. Dominguez-Mejide, A.; Vasili, E.; Outeiro, T.F. Pharmacological Modulators of Tau Aggregation and Spreading. *Brain Sci.* **2020**, *10*, 858. [[CrossRef](#)]
25. Cisek, K.; Cooper, G.L.; Huseby, C.J.; Kuret, J. Structure and Mechanism of Action of Tau Aggregation Inhibitors. *Curr. Alzheimer Res.* **2014**, *11*, 918–927. [[CrossRef](#)]
26. Pradeepkiran, J.A.; Reddy, P.H. Structure Based Design and Molecular Docking Studies for Phosphorylated Tau Inhibitors in Alzheimer’s Disease. *Cells* **2019**, *8*, 260. [[CrossRef](#)] [[PubMed](#)]
27. Bulic, B.; Pickhardt, M.; Mandelkow, E. Progress and Developments in Tau Aggregation Inhibitors for Alzheimer Disease. *J. Med. Chem.* **2013**, *56*, 4135–4155. [[CrossRef](#)] [[PubMed](#)]
28. Calcul, L.; Zhang, B.; Jinwal, U.K.; Dickey, C.A.; Baker, B.J. Natural Products as a Rich Source of Tau-Targeting Drugs for Alzheimer’s Disease. *Future Med. Chem.* **2012**, *4*, 1751–1761. [[CrossRef](#)] [[PubMed](#)]
29. Deshpande, P.; Gogia, N.; Singh, A. Exploring the Efficacy of Natural Products in Alleviating Alzheimer’s Disease. *Neural Regen. Res.* **2019**, *14*, 1321–1329. [[CrossRef](#)]
30. Rastelli, G.; Pellati, F.; Pinzi, L.; Gamberini, M.C. Repositioning Natural Products in Drug Discovery. *Molecules* **2020**, *25*, 1154. [[CrossRef](#)] [[PubMed](#)]
31. Atanasov, A.G.; Zotchev, S.B.; Dirsch, V.M.; Orhan, I.E.; Banach, M.; Rollinger, J.M.; Barreca, D.; Weckwerth, W.; Bauer, R.; Bayer, E.A.; et al. Natural Products in Drug Discovery: Advances and Opportunities. *Nat. Rev. Drug Discov.* **2021**, *20*, 200–216. [[CrossRef](#)]
32. Jadhav, S.; Avila, J.; Schöll, M.; Kovacs, G.G.; Kövari, E.; Skrabana, R.; Evans, L.D.; Kontsekova, E.; Malawska, B.; de Silva, R.; et al. A Walk through Tau Therapeutic Strategies. *Acta Neuropathol. Commun.* **2019**, *7*, 22. [[CrossRef](#)]
33. Li, C.; Götz, J. Tau-Based Therapies in Neurodegeneration: Opportunities and Challenges. *Nat. Rev. Drug Discov.* **2017**, *16*, 863–883. [[CrossRef](#)]
34. Soeda, Y.; Takashima, A. New Insights Into Drug Discovery Targeting Tau Protein. *Front. Mol. Neurosci.* **2020**, *13*, 590896. [[CrossRef](#)]
35. Lunven, L.; Bonnet, H.; Yahiaoui, S.; Yi, W.; Da Costa, L.; Peuchmaur, M.; Boumendjel, A.; Chierici, S. Disruption of Fibers from the Tau Model AcPHF6 by Naturally Occurring Aurones and Synthetic Analogues. *ACS Chem. Neurosci.* **2016**, *7*, 995–1003. [[CrossRef](#)]
36. Wischik, C.M.; Edwards, P.C.; Lai, R.Y.; Roth, M.; Harrington, C.R. Selective Inhibition of Alzheimer Disease-like Tau Aggregation by Phenothiazines. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11213–11218. [[CrossRef](#)]
37. Larbig, G.; Pickhardt, M.; Lloyd, D.G.; Schmidt, B.; Mandelkow, E. Screening for Inhibitors of Tau Protein Aggregation into Alzheimer Paired Helical Filaments: A Ligand Based Approach Results in Successful Scaffold Hopping. *Curr. Alzheimer Res.* **2007**, *4*, 315–323. [[CrossRef](#)]
38. Necula, M.; Chirita, C.N.; Kuret, J. Cyanine Dye N744 Inhibits Tau Fibrillization by Blocking Filament Extension: Implications for the Treatment of Tauopathic Neurodegenerative Diseases. *Biochemistry* **2005**, *44*, 10227–10237. [[CrossRef](#)]
39. Necula, M.; Kaye, R.; Milton, S.; Glabe, C.G. Small Molecule Inhibitors of Aggregation Indicate That Amyloid Beta Oligomerization and Fibrillization Pathways Are Independent and Distinct. *J. Biol. Chem.* **2007**, *282*, 10311–10324. [[CrossRef](#)] [[PubMed](#)]
40. Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A.P.; Chambers, J.; Mendez, D.; Motow, P.; Atkinson, F.; Bellis, L.J.; Cibrián-Uhalte, E.; et al. The ChEMBL Database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954. [[CrossRef](#)] [[PubMed](#)]



41. Berthold, M.R.; Cebon, N.; Dill, F.; Gabriel, T.R.; Kötter, T.; Meinel, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Data Analysis, Machine Learning and Applications, Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V.*; Breisgau, Germany, 7–9 March 2007; Preisach, C., Burkhardt, H., Schmidt-Thieme, L., Decker, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 319–326.
42. OpenEye Toolkits 2020.2.2 OpenEye Scientific Software, Santa Fe, NM. Available online: <http://www.eyesopen.com> (accessed on 20 January 2021).
43. RDKit: Open-Source Cheminformatics. Available online: <http://www.rdkit.org> (accessed on 5 June 2020).
44. Willett, P. Searching Techniques for Databases of Two- and Three-Dimensional Chemical Structures. *J. Med. Chem.* **2005**, *48*, 4183–4199. [[CrossRef](#)] [[PubMed](#)]
45. Jasial, S.; Hu, Y.; Vogt, M.; Bajorath, J. Activity-Relevant Similarity Values for Fingerprints and Implications for Similarity Searching. *F1000Res* **2016**, *5*. [[CrossRef](#)]
46. *Schrödinger Release 2020-1: QikProp*; Schrödinger, LLC: New York, NY, USA, 2020.
47. *Schrödinger Release 2020-1: LigPrep*; Schrödinger, LLC: New York, NY, USA, 2020.
48. Bauer, D.F. Constructing Confidence Sets Using Rank Statistics. *J. Am. Stat. Assoc.* **1972**, *67*, 687–690. [[CrossRef](#)]
49. Kuenemann, M.A.; Labbé, C.M.; Cerdan, A.H.; Sperandio, O. Imbalance in Chemical Space: How to Facilitate the Identification of Protein-Protein Interaction Inhibitors. *Sci. Rep.* **2016**, *6*, 23815. [[CrossRef](#)] [[PubMed](#)]
50. Blake, J.F. Examination of the Computed Molecular Properties of Compounds Selected for Clinical Development. *Biotechniques* **2003**, *8*, 16–20. [[CrossRef](#)]
51. Armstrong, R.A. When to Use the Bonferroni Correction. *Ophthalmic Physiol. Opt.* **2014**, *34*, 502–508. [[CrossRef](#)]
52. Degen, J.; Wegscheid-Gerlach, C.; Zaliani, A.; Rarey, M. On the Art of Compiling and Using “Drug-Like” Chemical Fragment Spaces. *ChemMedChem* **2008**, *3*, 1503–1507. [[CrossRef](#)]
53. Bemis, G.W.; Murcko, M.A. The Properties of Known Drugs. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887–2893. [[CrossRef](#)]
54. Lewell, X.Q.; Judd, D.B.; Watson, S.P.; Hann, M.M. RECAP—Retrosynthetic Combinatorial Analysis Procedure: A Powerful New Technique for Identifying Privileged Molecular Fragments with Useful Applications in Combinatorial Chemistry. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 511–522. [[CrossRef](#)]
55. CHOMP 3.1.1.2: OpenEye Scientific Software, Santa Fe, NM. Available online: <http://www.eyesopen.com> (accessed on 20 January 2021).
56. Jameson, L.P.; Smith, N.W.; Dzyuba, S.V. Dye-Binding Assays for Evaluation of the Effects of Small Molecule Inhibitors on Amyloid (A $\beta$ ) Self-Assembly. *ACS Chem. Neurosci.* **2012**, *3*, 807–819. [[CrossRef](#)]
57. Munishkina, L.A.; Fink, A.L. Fluorescence as a Method to Reveal Structures and Membrane-Interactions of Amyloidogenic Proteins. *Biochim. Biophys. Acta (BBA) Biomembr.* **2007**, *1768*, 1862–1885. [[CrossRef](#)]
58. Okuda, M.; Hijikuro, I.; Fujita, Y.; Teruya, T.; Kawakami, H.; Takahashi, T.; Sugimoto, H. Design and Synthesis of Curcumin Derivatives as Tau and Amyloid  $\beta$  Dual Aggregation Inhibitors. *Bioorganic Med. Chem. Lett.* **2016**, *26*, 5024–5028. [[CrossRef](#)]
59. Fuse, S.; Matsumura, K.; Fujita, Y.; Sugimoto, H.; Takahashi, T. Development of Dual Targeting Inhibitors against Aggregations of Amyloid- $\beta$  and Tau Protein. *Eur. J. Med. Chem.* **2014**, *85*, 228–234. [[CrossRef](#)]
60. Available online: <http://rdkit.blogspot.com/2013/10/fingerprint-thresholds.html> (accessed on 5 June 2020).
61. Carhart, R.E.; Smith, D.H.; Venkataraghavan, R. Atom Pairs as Molecular Features in Structure-Activity Studies: Definition and Applications. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 64–73. [[CrossRef](#)]
62. Nilakantan, R.; Bauman, N.; Dixon, J.S.; Venkataraghavan, R. Topological Torsion: A New Molecular Descriptor for SAR Applications. Comparison with Other Descriptors. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 82–85. [[CrossRef](#)]
63. Jain, A.N.; Nicholls, A. Recommendations for Evaluation of Computational Methods. *J. Comput. Aided Mol. Des.* **2008**, *22*, 133–139. [[CrossRef](#)]
64. Truchon, J.-F.; Bayly, C.I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *J. Chem. Inf. Model.* **2007**, *47*, 488–508. [[CrossRef](#)]
65. *Schrödinger Release 2020-1: Canvas*; Schrödinger, LLC: New York, NY, USA, 2020.
66. Hall, L.H.; Kier, L.B. The Molecular Connectivity Chi Indexes and Kappa Shape Indexes in Structure-Property Modeling. In *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 1991; pp. 367–422.
67. Hall, L.H.; Kier, L.B. The Nature of Structure-Activity Relationships and Their Relation to Molecular Connectivity. *Eur. J. Med. Chem.* **1977**, *12*, 307–312.
68. Nguyen, K.T.; Blum, L.C.; van Deursen, R.; Reymond, J.-L. Classification of Organic Molecules by Molecular Quantum Numbers. *ChemMedChem* **2009**, *4*, 1803–1805. [[CrossRef](#)] [[PubMed](#)]
69. Min, S.-W.; Cho, S.-H.; Zhou, Y.; Schroeder, S.; Haroutunian, V.; Seeley, W.W.; Huang, E.J.; Shen, Y.; Masliah, E.; Mukherjee, C.; et al. Acetylation of Tau Inhibits Its Degradation and Contributes to Tauopathy. *Neuron* **2010**, *67*, 953–966. [[CrossRef](#)] [[PubMed](#)]
70. Thapa, P.; Upadhyay, S.P.; Suo, W.Z.; Singh, V.; Gurung, P.; Lee, E.S.; Sharma, R.; Sharma, M. Chalcone and Its Analogs: Therapeutic and Diagnostic Applications in Alzheimer’s Disease. *Bioorganic Chem.* **2021**, *108*, 104681. [[CrossRef](#)] [[PubMed](#)]
71. Lin, T.-H.; Chiu, Y.-J.; Lin, C.-H.; Lin, C.-Y.; Chao, C.-Y.; Chen, Y.-C.; Yang, S.-M.; Lin, W.; Mei Hsieh-Li, H.; Wu, Y.-R.; et al. Exploration of Multi-Target Effects of 3-Benzoyl-5-Hydroxychromen-2-One in Alzheimer’s Disease Cell and Mouse Models. *Aging Cell* **2020**, *19*, e13169. [[CrossRef](#)] [[PubMed](#)]

72. Iacovino, L.G.; Pinzi, L.; Facchetti, G.; Bortolini, B.; Christodoulou, M.S.; Binda, C.; Rastelli, G.; Rimoldi, I.; Passarella, D.; Di Paolo, M.L.; et al. Promising Non-Cytotoxic Monosubstituted Chalcones to Target Monoamine Oxidase-B. *ACS Med. Chem. Lett.* **2021**. [[CrossRef](#)] [[PubMed](#)]
73. Hsieh, C.-J.; Xu, K.; Lee, I.; Graham, T.J.A.; Tu, Z.; Dhavale, D.; Kotzbauer, P.; Mach, R.H. Chalcones and Five-Membered Heterocyclic Isosteres Bind to Alpha Synuclein Fibrils in Vitro. *ACS Omega* **2018**, *3*, 4486–4493. [[CrossRef](#)]
74. Jeon, K.-H.; Lee, E.; Jun, K.-Y.; Eom, J.-E.; Kwak, S.Y.; Na, Y.; Kwon, Y. Neuroprotective Effect of Synthetic Chalcone Derivatives as Competitive Dual Inhibitors against  $\mu$ -Calpain and Cathepsin B through the Downregulation of Tau Phosphorylation and Insoluble A $\beta$  Peptide Formation. *Eur. J. Med. Chem.* **2016**, *121*, 433–444. [[CrossRef](#)] [[PubMed](#)]