Journal of Medicinal Chemistry

pubs.acs.org/jmc

Article

Systematic Investigation of Dual-Target-Directed Ligands

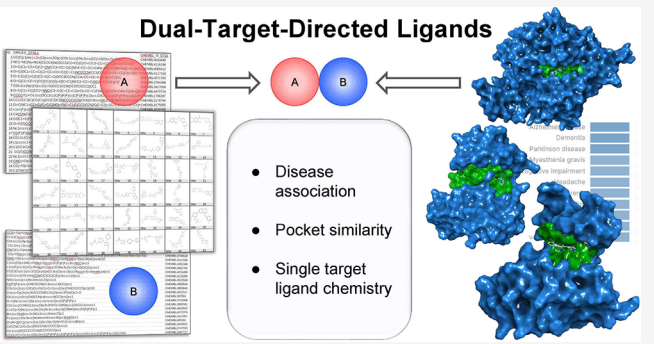
Vittorio Lembo and Giovanni Bottegoni*

Cite This: J. Med. Chem. 2024, 67, 10374−10385

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Multitarget-directed ligands (MTDLs) are compounds rationally designed to affect multiple targets, aiming for a better therapeutic profile. For over 20 years, MTDLs have garnered increasing attention, the idea being that their full potential would have been achieved, thanks to unprecedented target combinations and advanced medicinal chemistry strategies. This study presents a literature mining effort resulting in a data set of dual-target-directed ligands (DTDLs), the fundamental example of MTDLs. We used this data set to evaluate the rationale behind target selection and the chemical novelty of DTDLs targeting specific protein combinations. Our analysis focused on DTDL targets in terms of biological function, disease association, structure, and chemogenomic traits. We also compared DTDLs with single-target compounds. We found that well-known target pathology associations often guide DTDL design, leveraging existing chemical scaffolds and binding pocket similarities. These findings highlight the current state of the field and suggest substantial untapped potential for rational polypharmacology.



INTRODUCTION

Simultaneously modulating multiple targets has shown promise in enhancing the effectiveness of drugs and overcoming resistance, especially in the framework of multifactorial diseases.[1] Multitarget-directed ligands are molecules rationally designed to exert activity on multiple targets and have emerged as a well-established strategy in drug design. Developing MTDLs entails two key steps: (i) identifying a group of targets that, modulated together, offer improved therapeutic outcomes compared to single target therapies, and (ii) designing a single chemical structure capable of interacting with all selected targets. A singular molecular entity offers, in fact, several advantages with respect to drug cocktails.[2] Several strategies for the identification of optimal target combinations have been proposed. The most intuitive and early adopted approach relies on developing molecules with multiple activities at targets that have been previously and independently validated for the same pathology.[3] Ideally, this target combination has already been addressed by a drug cocktail. For example, several MDTLs developed for the treatment of complex conditions of the central nervous system, such as Alzheimer's disease (AD), Parkinson's disease, schizophrenia, and mood disorders, addressed a set of targets previously engaged by combination therapy.[4−6] Strategies that rely on computational techniques for the identification of target combinations have also been reported.[7] These approaches propose utilizing machine learning to comprehensively analyze omics data and prospectively identify combinations of targets. This is what has been described as "increasing the opportunity space for druggable targets".[8]

In terms of actual chemical design, two different strategies can be conceived: (i) starting from a molecule active at one of the targets of interest and engineering other activities into it, or (ii) searching for entirely new scaffolds capable of binding all sought targets.[9] In the first approach, the degree of overlap that is obtained can vary and MTDLs are often classified accordingly.[1] Here, a key challenge is to balance the gain of new activities by introducing novel features on a given scaffold with preserving as much as possible the starting pharmacophore, thus retaining activity at the original target. Intuitively, it is more straightforward to design MTDLs acting at structurally similar targets (e.g., belonging to the same protein family) or that share similar or identical substrates (e.g., targeting enzymes that belong to the same metabolic pathway).[10,11] Overall, while effective, this strategy could potentially yield molecules of limited chemical novelty. With novelty in mind, one can turn to a screening approach, sequentially testing a library of compounds at multiple targets. However, this approach combines a resource-intensive nature with an overall

Received: April 8, 2024
Revised: May 17, 2024
Accepted: May 20, 2024
Published: June 6, 2024

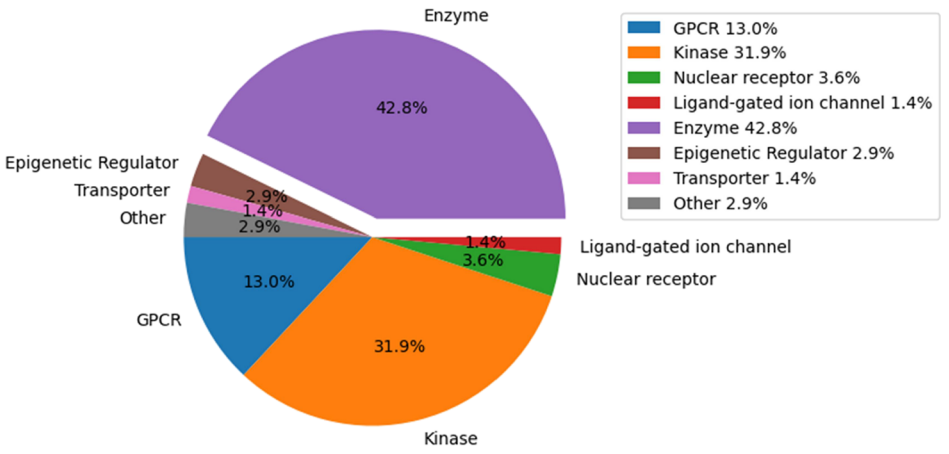


Figure 1. Distribution of individual target classes in the DTDL data set.

limited chance of identifying compounds with the desired array of activities.[12]

In this study, we decided to systematically investigate these critical points: (i) what is the rationale more often adopted in selecting target combinations? (ii) what is the degree of chemical novelty that can be found in MTDLs with respect to single target compounds already reported for the MTDL targets? We compiled a list of compounds that are multitarget by design, thus excluding molecules whose polypharmacological profile was only serendipitously discovered. Furthermore, to collect and analyze data as homogeneous as possible, we only focused on the most basic form of MTDLs: molecules with activity at only two targets (DTDLs).

RESULTS AND DISCUSSION

Our literature search returned 941 articles. We retained 158 articles focused on rationally conceived DTDLs, while 783 articles were excluded (see Experimental Section). A summary of the literature search is included as a spreadsheet in the Supporting Information. A representative DTDL from each study was included in the data set described here (see Table S1), which eventually encompassed 158 entries. Each molecule was annotated in terms of molecular ChEMBL ID and ChEMBL ID of each target from the target pair on which the molecule is active (henceforth simply referred to as the target pair). The complete data set, including canonical SMILES notation and UniProt ID of each target, is available as a spreadsheet in the Supporting Information.

We systematically analyzed the target pairs in the set. The total number of targets was 126. Targets were grouped according to their biological function (Figure 1). Kinases were reported as a separate group, considering the large number of DTDLs active in this protein family.

The most populated groups were enzymes (42.8%) and kinases (31.9%). GPCRs accounted for 13%, nuclear receptors for 3.6%, ligand-gated ion channels for 1.4%, epigenetic regulators for 2.9%, and transporters for 1.4%. The NOD-like receptor 3 (NLRP3), a component of the inflammasome, the induced myeloid leukemia cell differentiation protein Mcl-1 (MCL-1), a protein involved in apoptosis, the advanced glycosylation end product-specific receptor (RAGE), and the heat shock protein HSP 90-alpha (HSP90) were not assigned to any specific group ("Other" in Figure 1). The more frequently recurring targets were acetylcholinesterase (AChE) (15), epidermal growth factor receptor (EGFR) (14), monoamine oxidase B (MAO-B), and ALK tyrosine kinase receptor (ALK) (12 each). In terms of targets, the collected DTDLs covered a diverse set representing therapeutically relevant protein families. We also analyzed the distribution of target pairs (see Table S2). The most populated groups were enzyme−enzyme (34%) and kinase−kinase (32.1%). We could also find examples of enzymes paired with kinases, GPCRs, and nuclear receptors. There are also examples of GPCR—GPCR target pairs (3.7%). All other combinations were either populated by three examples or less or not represented. The most frequently recurring target pair was AChE—MAO-B (7 times).

We also annotated individual targets in terms of how they are associated with diseases and characterized each target pair accordingly. To this aim, we used the Open Targets Platform, a web-based open-source platform that, integrating multiple data sources, returns an estimate, called the Overall Association Score (OAS, ranging from 0 to 1), of the association of a protein with a disease.[24] The analysis was performed on 155 of the 158 target pairs since we only considered targets with available data in the Open Targets Platform. Comparing the OAS to all listed diseases of each target pair from the set, we observed that all pairs (100%) shared an association with at least one common disease. In 152 of 155 pairs (98.1%), both targets had an association with the disease reported as the focus of the synthetic campaign. In the remaining 3 pairs, only one target presented some level of association with the investigated disease. Furthermore, out of 310 targets, 196 (63.2%) were ranked in the 95th OAS percentile or above among the targets associated with the investigated disease. In 70 instances, both targets of the DTDL were ranked in the 95th OAS percentile or above for the investigated disease. Network biology proposes that targets instrumental to a sought therapeutic effect can be identified, thanks to the application of network theory to omics data. In particular, since the concurrent targeting of multiple nodes is considered fundamental to overcoming the redundancy that is typically associated with biological networks, pioneering work in the field put forward the idea that a strong interplay would have existed between the rational design of MTDLs and identification of unprecedented target combinations, thanks to network biology.[8,13] However, even when considering the relatively simple case of DTDLs, data on the collected combinations pointed toward previous and independent validation within a certain pathology as one of the key drivers

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

for target selection. In only three of the selected protein combinations, an association with the investigated disease was entirely missing for at least one of the targets. In this regard, it should also be pointed out that OAS is largely affected by the availability of data.[14] A target selection largely driven by pre-existing knowledge of target-disease association was also confirmed by how DTDLs aimed at Alzheimer's disease (21; 13.3%) and cancer (86; 54.4%) were largely represented in the data set. Both these conditions are multifactorial in nature with several independently validated targets already reported, and they have been the focus of MTDL discovery campaigns since the early days of the discipline.[4,15]

We then moved toward a chemogenomic approach, collecting single target compounds with reported activity at each member of the target pair and investigating the existence of an overlap in chemical space between the two sets. To avoid biasing our analysis, all rationally designed MTDLs with activity at the target pair (e.g., compounds that belong to the same series as the selected representative) and ligands published after the selected DTDL were removed from each set. For each target pair, a dissimilarity matrix was generated assessing Tanimoto distance (TD) between the fingerprint of each molecule in one set with respect to the corresponding representation of all molecules from the other one. In this way, we were able to determine the minimum TD between any two molecules individually active at each target in the pair and annotate each compound accordingly. The existence of at least one pair of molecules from the respective sets exhibiting a TD between 0 and 0.1 was considered indicative of an overlap. In Figure 2, the distribution of this minimum TD is reported.

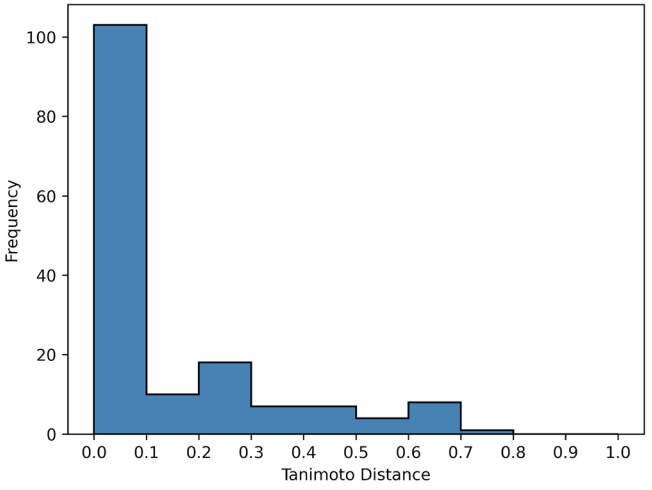


Figure 2. Distribution of the minimum TD values from the DTDL data set.

Most target pairs (102 out of 158; 64.6%) presented a pre-existing overlap in chemical space with at least one identical or almost identical molecule in the set (TD between 0 and 0.1). In 92 out of 158 cases (58.2%), the two sets had at least one molecule in common (TD equal 0), that, while likely not rationally developed as a DTDL (see Experimental Section), was annotated for activity on both targets. In only 13 cases, the minimum distance was over 0.5. TD, and, in the entire set, it never exceeded 0.71. We then investigated each target pair aiming at discriminating cases in which the common space between the targets was extended from cases where only one or a few molecules accounted for the low TD minimum value. The TD distributions of two target pairs, namely, EGFR and receptor tyrosine-protein kinase erbB-2 (HER2) (blue in Figure 3) and butyrylcholinesterase (BuChE) and cannabinoid

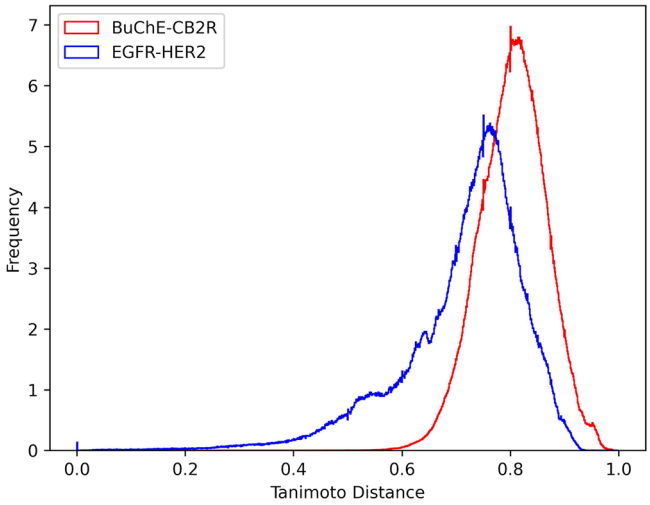


Figure 3. Comparison of TD distribution between the ligands of two target pairs. EGFR-HER2 in blue and BuChE-CB2R in red. The chart was generated by normalizing the distributions so that the total area under each histogram is equal to 1. The height of each bin represents the probability density in that data interval.

receptor 2 (CB2R) (red in Figure 3) will be compared to illustrate this point. Both pairs return a minimum TD between single target compounds equal to 0. Overall, both distributions look similar, with the EGFR-HER2 distribution marginally shifted toward the left. However, when focusing on the leftmost part of the plot (TD ranging from 0 to 0.5), the overlap in chemical space between EGFR and HER2 appears widespread, while the overlap between BuChE and CB2R is only due to a few individual compounds.

In the data set, 53 target pairs (33.5%) shared over 100 molecules with a minimum TD lower than 0.1. Sixteen target pairs (10.1%) shared over 1000. Out of these 53 target pairs: 34 consisted of two kinases, 2 of Class A GPCRs, 8 pairs were proteins belonging to the same biosynthetic pathway, and in 7 cases the two targets shared similar endogenous ligands. Only 2 target pairs encompassed unrelated proteins: Polo-like kinase 1 (PLK-1)—bromodomain-containing protein 4 (BRD4) (165 ligand pairs) and BRD4—transcription initiation factor TFIID subunit 1 (TAF1) (119 ligand pairs). Structurally or functionally related proteins are often modulated by chemically similar ligands enhancing the possibility to rationally design compounds with activities at the pair building upon this overlap.[10] Furthermore, 73 pairs (46.2%) only shared less than 10 molecules showing TDs between 0 and 0.1. We also investigated overlap in chemical space by employing a clustering approach. For each target pair, both single target sets were merged and partitioned by means of a modified nearest neighbor algorithm (see Experimental Section). Clustering enabled us to confirm the presence of overlap in chemical space between single target sets of molecules, excluding potential noise generated by one or a few overrepresented scaffolds. In particular, we were interested in assessing the existence of hybrid clusters, namely, clusters containing at least one molecule from each single target set. For 111 target pairs (70.3%), we observed the presence of one

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

or more hybrid clusters. Moreover, 68 cases (43%) presented different areas of overlap, as indicated by the identification of 10 hybrid clusters or more. In eight target pairs [serotonin transporter (SERT)—norepinephrine transporter (NET), phosphatidylinositol-4,5-bisphosphate 3-kinase α (PI3K-α)—mammalian target of rapamycin (mTOR), and stem cell growth factor receptor (c-KIT)—platelet-derived growth factor receptor α (PDGFR-α)], over 20% of all clusters were hybrid ones, highlighting an overlap that extended beyond a few specific molecules. A spreadsheet that summarizes the results of the clustering exercise is available in the Supporting Information. To summarize, access to an already mapped overlap in chemical space, even if limited to a small number of compounds, was confirmed as another possible driver for identifying target combinations for DTDLs. A limitation of this chemogenomic approach was that the TD was not adjusted for either activity[16] or promiscuity, the latter referring to the total number of targets a molecule is active at. Even if we considered only relatively potent molecules with a pChEMBL value of 5 or greater, the affinities for each target varied widely, spanning between 4 and 6 orders of magnitude. If we consider affinity as a measure of a molecule's relevance in defining the chemical space of a target, it could be theoretically possible to assign different weights to the same TD value based on the affinity of each compared molecule for that target. Similarly, a compound acting on a few specific targets might contribute differently to defining a shared chemical space compared to one that binds aspecifically. We intend to delve deeper into these concepts in future research.

We also compared the binding pockets of each target pair. We employed DeeplyTough, an algorithm based on a convolutional neural network that expresses the dissimilarity between pockets as a non-normalized Euclidean distance. The analysis was performed on 121 of the 158 target pairs since we only considered targets with available crystal structures. Based on the results obtained from DeeplyTough, we generated the distribution reported in Figure 4.

The leftmost peak (Figure 4) accounts for 47 protein pairs whose pockets were separated by distances between 0 and 0.2. It highlights the presence of several DTDLs targeting conserved pockets, namely, pairs of pockets that the algorithm is trained to consider a match. Out of these 47 target pairs, 37

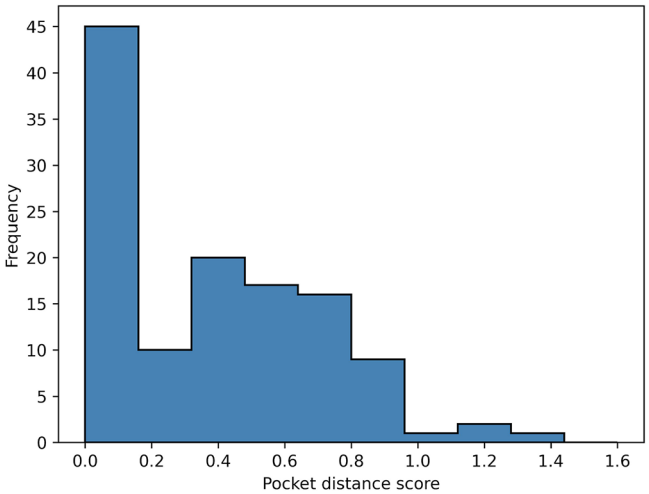


Figure 4. Distribution of pocket distances calculated using DeeplyTough.

consisted of two kinases, thus enzymes with highly conserved binding sites. Two pairs were formed by proteins from the same metabolic pathway: phospholipase A2 (PLA2) and leukotriene-A4-hydrolase (LTA4H)[17] and microsomal prostaglandin E synthase-1 (mPGES-1) and 5-lipoxygenase activating protein (FLAP),[18] respectively. Four pocket pairs were formed by indoleamine-pyrrole 2,3-dioxygenase-1 (IDO-1) and tryptophan 2,3-dioxygenase (TDO), two enzymes both involved in tryptophan metabolism.[19,20] Two pairs were formed by two epigenetic regulators [Bromodomain and PHD finger-containing protein 1 (BRPF1) and transcription intermediary factor 1-alpha (TRIM24)][21] and two class A aminergic GPCR [dopamine D2 receptor (D2) and serotonin 2a receptor (5-HT2a)].[22] Last, two pairs were formed by unrelated proteins: ALK—HSP90[23] and EGFR—carbonic anhydrase IX (CA-IX).[24] Only a few cases of structurally unrelated proteins (e.g., rAChE—*N*-methyl-D-aspartate receptor (NMDAR);[25] BuChE—CB2R;[26] estrogen receptor α (ER-α)—Histone deacetylase 6 (HDAC6);[27] HDAC6—Apoptosis regulator Bcl-2 (BCL-2)),[28] consistently mapped apart in pocket space, could be identified. Binding site similarity could thus be described as another driver that strongly influenced the identification of target combinations. Indeed, several authors have suggested how targeting proteins that share very similar binding pockets improves the feasibility of designing polypharmacological drugs.[11] Overlap in single target chemical space and pocket similarity are, of course, correlated (Figure 5). For target pairs with similar binding sites, there frequently

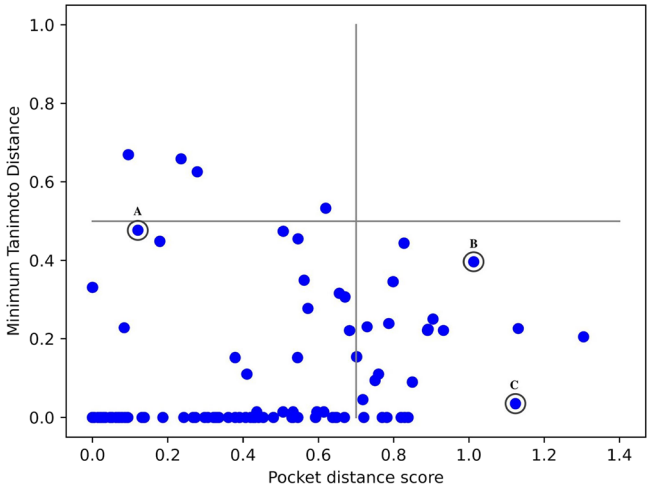


Figure 5. Pocket distance and minimum TD values of the most similar ligand pair for the target combinations. Discussed examples are labeled explicitly.

were one or more pre-existing molecules able to modulate activity at both proteins. Conversely, as the distance between pockets increased, there was an increase in minimum TD between single target compounds.

There are examples departing from this general trend to be discussed in detail. The PLA2/LTA4H target pair (labeled A in Figure 5) displayed a pocket distance of 0.12, hinting at a conserved binding site, while the most similar single target molecules exhibit a TD value of 0.48.[17] However, single target compounds ChEMBL1644549 (**1**, active at PLA2) and ChEMBL71853 (**2**, active at LTA4H) shared the same (2S)-4-methyl-2-(2-oxoacetamido)pentanoic acid group (Figure 6). TD expresses a score generated by the entire fingerprints and

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

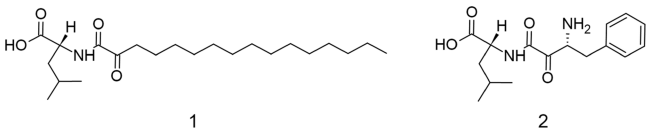


Figure 6. Structures of DTDL compounds **1** and **2** sharing a common substructure.

this value could fail at highlighting the presence of a common core. Thus, we also compared DTDLs and the single target compounds active at the involved targets in terms of common substructures (see below).

In the second case (B in Figure 5), the rAChE/NMDAR target pair displayed a pocket distance of 1.01 with the most similar single target molecules for this pair, exhibiting TD equal to 0.4. Linked DTDLs are a straightforward way to overcome the absence of pocket similarity in DTDL design. In this specific case, Simoni and colleagues successfully designed memagal (ChEMBL2178784, **3**, Figure 7), a dual rAChE/

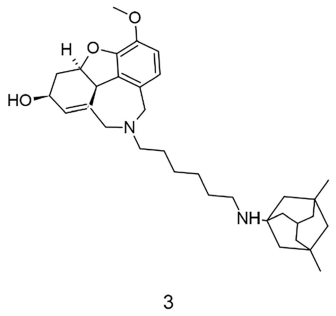


Figure 7. Chemical structure of **3**.

NMDAR inhibitor conceived for the treatment of AD, consisting of the combination of galantamine (AChE inhibitor) and memantine (NMDAR antagonist) via a six-methylene unit spacer.[25]

In the third example (C in Figure 5), the BuChE/CB2R target pair was associated with the remarkable pocket distance (1.12), but the most similar ligands (ChEMBL1095500, **4**, active at AChE and ChEMBL73572, **5**, active at CB2R; Figure 8) returned a minimum TD value equal to 0.03.[26] Having

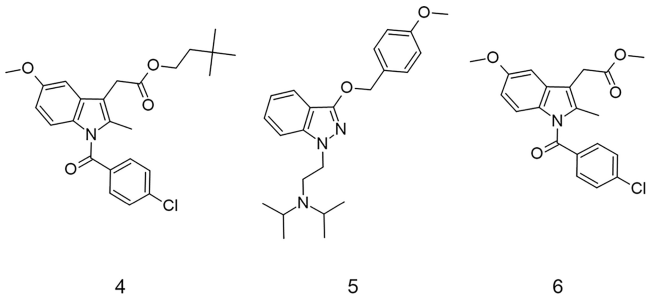


Figure 8. Relevant compounds in the development of BuChE-CB2R DTDLs.

observed how some members of an indole-bearing series displayed activity at AChE, González-Naranjo and co-workers were able to devise a cleverly merged DTDL, ChEMBL3116280 (**6**, Figure 8). This highlights how, in specific cases, it is possible to design a fused or merged DTDL exploiting an overlap in chemical space to overcome or, at least, to compensate, for limited pocket similarity.

Shifting our perspective from targets to the actual DTDLs, we were able to group the compounds in our data set according to the framework originally proposed by Morphy and Rankovic.[1] Based on the level of overlap between pharmacophoric elements associated with activity at one or both targets, compounds were labeled as merged, fused, and linked DTDLs. A linked MTDL consists of distinct molecules, known to interact at different targets, connected by a linker. This strategy offers a way to combine two chemical entities that modulate different targets, without the need for massive structural optimization. MTDLs obtained exploiting partially overlapping pharmacophores are classified as fused MTDLs. In this approach, the pharmacophores of the starting molecules lack largely overlapping areas. However, it is still possible to act on the structure of a starting compound by introducing a local modification that, while subtle, can endow the molecule with activity at another target (while retaining activity at the original one). Last, pharmacophoric features can be seamlessly combined into a unified framework. When this is possible, the obtained structures are usually simpler and with lower molecular weight as compared to linked or fused multitarget compounds. In principle, the physicochemical profile of a merged DTDL should not significantly differ from that of a single target compound. Despite containing an element of arbitrariness, due to cases exhibiting intermediate characteristics that cannot be univocally assigned to a specific group, this classification is widely adopted in MTDL literature and was thus adopted here for consistency. The largest group in the data set was that of molecules classified as merged (91 out of 158; 57.6%). The second most populated group was the one of fused DTDLs (56 out of 158; 35.4%). Only 11 compounds (out of 158; 7%) were classified as linked DTDLs. Then, we analyzed the physicochemical profile of the collected molecules. We calculated MW, LogP, and topological polar surface area (TPSA) for each DTDL and compared these values with those extracted from corresponding single target ligand sets. Eighty-seven DTDLs (55.1%) presented a MW value above the average MW of both sets of single target ligands. Eighty-six DTDLs (54.4%) presented a LogP value above the average and 81 DTDLs (51.3%) presented a TPSA value below the average of both sets. Conversely, 34 DTDLs (21.5%) presented a MW value below the average of both sets of single target ligands, 46 (29.1%) a LogP value below and 51 (32.3%) a TPSA value above the average of both sets, respectively. Despite the prevalence of merged compounds, DTDLs exhibited consistently higher MW and lipophilicity with respect to single target analogs. While the restricted size of this data set poses a challenge in extrapolating broadly applicable trends, data gathered here confirm the intuitions originally proposed by several authors, and later episodically reported also by others, on how MTDLs tend to grow heavier and more lipophilic with respect to single target compounds.[29,30]

We also assessed the structural difference between each DTDL and single target compounds active at the target pair in terms of TD.

Let us assume that each DTDL is active at two targets, target A and target B. We retrieved from ChEMBL molecules with high affinity at target A (set A) and target B (set B) according to the procedure reported in the Experimental Section. We measured the TD of the DTDL from all molecules from set A

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

and all molecules from set B. In the scatter plot in Figure 9, each DTDL (blue dot) is assigned coordinates based on the

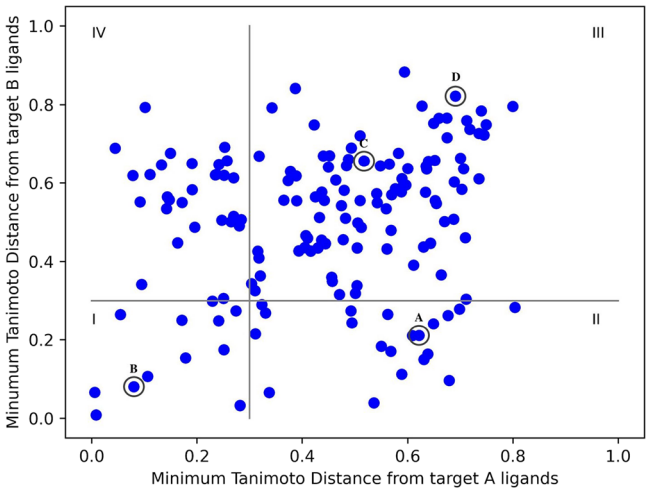
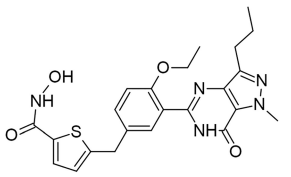


Figure 9. Minimum TDs of each DTDL from compounds with reported activity at each member of the target pair. Discussed examples are labeled explicitly.

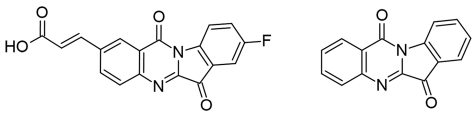
minimum TD calculated from all molecules in set A (horizontal axis) or from those in set B (vertical axis). The generated plot can be divided into 4 quadrants (labeled I, II, III, and IV in Figure 9). Quadrant I is populated by 12 (7.6%) DTDLs with clear similarity (minimum TD < 0.3) to at least one molecule from set A and one from set B. Quadrants II and IV are populated by 48 DTDLs (30.4%) that closely resembled at least one compound from one of the sets (quadrant II, set A, and quadrant IV, set B, respectively) while displaying low or no similarity toward molecules active at the other set. Last, 99 DTDLs (62.7%) are found in quadrant III, which encompasses DTDLs that, in terms of TD, were different from molecules from both sets. DTDLs distribution is thus strongly asymmetric. ChEMBL4212186 (**7**, A in Figure 9) is a merged HDAC6—phosphodiesterase 5A (PDE5A) inhibitor for the potential treatment of AD.[31] Despite a pocket distance of 0.89 and the absence of obvious overlapping areas in the single target chemical space, Rabal et al. were able to design **7** by inserting a Zn^{2+} chelating group, which introduced activity at HDAC6, into the 5-(2-ethoxyphenyl)-1-methyl-3-propyl-1,6-dihydro-7Hpyrazolo[4,3-*d*]pyrimidin-7-one scaffold of sildenafil (a PDE5 inhibitor) via a thiophene ring linker (Figure 10).

ChEMBL4645973 (**8**, B in Figure 9) is an IDO-1—TDO dual inhibitor developed by Li and co-workers for the treatment of cancer.[20] This molecule closely resembles tryptanthrin (**9**), a natural product known for its activity against both IDO-1 and TDO (Figure 11). This molecule can



7

Figure 10. Structure of **7**.



8 9

Figure 11. Structures of **8** and **9**.

thus be considered an example of optimization from a common precursor that already possesses dual activity. In these cases, a key hurdle to overcome in designing D/MTDLs is limiting the activity to only selected members of a protein superfamily or of a metabolic pathway, while retaining selectivity toward other closely related proteins.

Molecules populating quadrant III (Figure 9) were usually obtained either (i) by a pharmacophore merging strategy with equivalent contributions from both targets or (ii) by optimization of a common core associated with activity at both proteins. The latter strategy can be exemplified by farnesoid X receptor (FXR) modulator—soluble epoxide hydrolase (sEH) inhibitor ChEMBL4066332 (**10**, C in Figure 9), developed by Schmidt et al.,[32] for the treatment of nonalcoholic fatty liver disease. As reported in Figure 12, the most similar single target compounds from each target of the activity pair were the sEH inhibitor ChEMBL3763549 (**11**) (TD 0.55) and the FXR modulator ChEMBL3613439 (**12**) (TD 0.64). Despite TD values for both targets, these compounds both bore an aryl amide group. Once again, the design of compound **10** was inspired by the structure of previously reported ligands that individually showed activity on the sought targets. In fact, the identification of a shared framework among the chemical structures of known ligands facilitated the development of a merged-type DTDL despite the pocket distance (0.64).

The dual fatty acid amide hydrolase (FAAH) and monoacylglycerol lipase (MAGL) inhibitor ChEMBL2376224 (**13**, D in Figure 9), reported by Holtfrerich and co-workers,[33] presented another interesting case, which highlighted the limits of an analysis solely based on TD. Compound **13** exhibits a chemical structure with nontrivial overlap with known single target compounds. In TD space (Figure 13), the closest molecules to compound **13** are the FAAH inhibitor ChEMBL521038 (**14**) (TD 0.69) and the MAGL inhibitor ChEMBL2159777 (**15**) (TD 0.82). The similarity with single target compounds clearly emerges only when a three-dimensional pharmacophore match is considered.

In addition, it is worth noting how the physicochemical profile of **13** with lower MW and LogP values and a TPSA is slightly above the average values displayed by single target compounds active at FAAH and MAGL. Despite the similarity of the target pair pockets (rFAAH/MAGL pocket distance is equal to 0.32) and more than 100 molecule pairs from single target sets presenting a TD below 0.1, **13** introduces an element of novelty with respect to already reported single target ligands. The trends extrapolated from the distribution reported in Figure 9 were also confirmed by cluster analysis when a library of compounds encompassing the DTDL and the single target ligands from the target pair were processed using the same nearest neighbor clustering procedure previously described. In 135 instances (85.4%), DTDLs clustered as singletons. Three DTDLs (active as inhibitors of PI3K-α/mTOR, EGFR/HER2, and IDO-1/TDO target pairs) were clustered with compounds active at both targets. Finally, 20

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

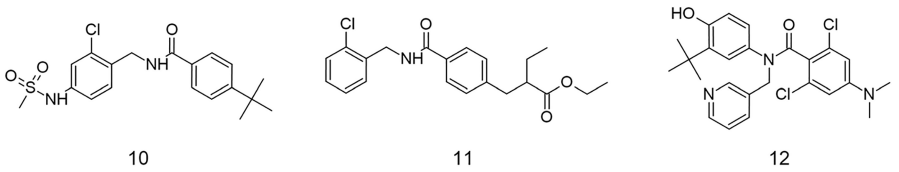


Figure 12. Relevant compounds in the development of sEH-FXR DTDL.

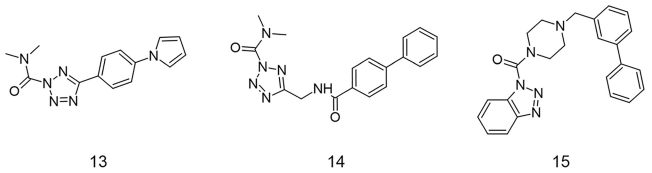


Figure 13. Relevant compounds in the development of FAAH-MAGL DTDL.

DTDLs were clustered with compounds active at only one of the targets involved. In general, the structure of DTDLs appeared largely based on the structure of single target compounds. This connection is straightforward, and somewhat expected, for molecules active at closely related targets or for molecules whose dual activity profile is generated by the introduction of a local functional group on a pre-existing single target compound. This relationship is often less intuitive for molecules populating quadrant III (Figure 9), and it might escape an analysis solely based on TD. For example, TD likely fails at capturing how part of a linked DTDL is, by design, identical to a single target molecule. To overcome this limitation, we also compared molecules in terms of maximum common substructure, calculating an MCES (Maximum Common Edge Substructure) relative score with the Rascal MCES algorithm. This approach was useful for identifying DTDLs that were globally different according to TD value, often clustered as singletons but locally shared a common motif with a single target compound. MCES between each DTDL and the molecules active on the targets from the activity pair was identified and both Johnson similarity and MCES relative score were calculated (see Experimental Section). The analysis was conducted considering a Johnson-predicted similarity threshold equal to 0.7 to avoid isolating small and not particularly informative fragments. Due to this threshold, an MCES was investigated for only 135 out of 158 DTDLs.

Most DTDLs populated the upper-right corner of the four-quadrant plot reported in Figure 14. Indeed, 114 DTDLs (84.4%) presented an MCES relative score of 70% or above for at least one molecule present in each set of single target ligands. The dual AChE inhibitor/cannabinoid receptor 1 (CBR1) antagonist ChEMBL570159 (**16**, A in Figure 14)[34] presents an MCES relative score of 83.3% to ChEMBL478666 (**17**) (AChE inhibitor) and of 86.5% to ChEMBL175291 (**18**) (CBR1 antagonist). As reported in Figure 15, the tacrine motif of **17**—responsible for AChE inhibition—and the almost entire structure of **18** are substructures of compound **16**.

Moreover, considering how the introduction of novel ring systems is regarded as an indicator of chemical novelty,[35] we compared rings and fused ring systems from each DTDL with those extracted from single target compounds active at the relevant target pair. Our analysis revealed that in 137 out of 158 cases (86.7%), each DTDL ring was represented among the rings extracted from the single target sets. It is worth

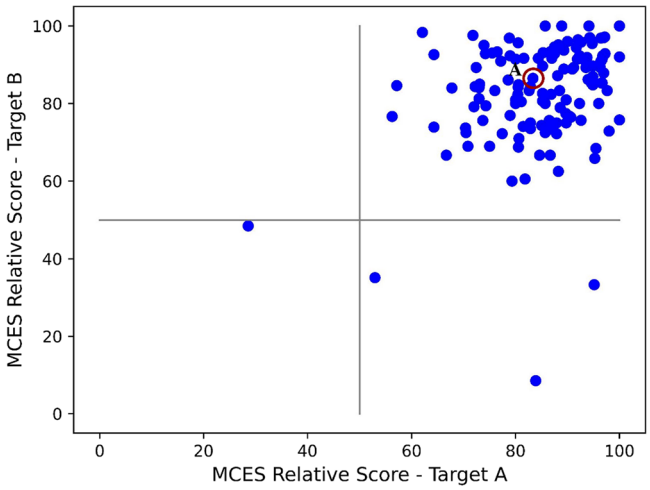


Figure 14. Highest MCES relative scores between the compounds active at target A (horizontal axis) and at target B (vertical axis) to each DTDL. The discussed example is labeled explicitly.

noticing that this analysis is particularly influenced by the size of single target ligand sets. Of the 21 cases where a new ring or ring system was found, 4 of them [xanthine dehydrogenase (XOD)—NLRP3 target pair, DNA gyrase subunit A from *Escherichia coli* (eGYRA)—topoisomerase IV subunit A from *Escherichia coli* (eTOPO4), DNA gyrase subunit B from *Staphylococcus aureus* (GYRB) - DNA topoisomerase 4 subunit B from *Staphylococcus aureus* (sTOPO4), and BCL-2—MCL-1 target pair] were extracted from DTDLs targeting proteins associated with a limited number of single target ligands. The DTDL rings and ring systems are reported in Figure S1. MCES and ring analysis confirmed what the systematic comparison of TD from single target compounds and cluster analysis already suggested: a polypharmacological profile is obtained by either introducing conservative modifications on pre-existing structures or rearranging privileged fragments.

CONCLUSIONS

This study describes a systematic analysis of DTDLs, the fundamental MTDL form. A set of 158 rationally conceived DTDLs reported in the literature was extracted from comprehensive and curated public data available through ChEMBL. Several parameters were adopted to ensure the right balance between data set cleanliness and inclusiveness. Indeed, we only considered molecules reported in medicinal chemistry journals regularly monitored by the curators of ChEMBL. We focused on DTDLs with a pChEMBL value equal to or greater than 5 (corresponding to a broadly defined activity of at least 10 μM), associated with the interaction with specific targets, thus excluding generic or phenotypic effects. Moreover, we only considered data associated with an assay confidence score equal to or greater than 8. This data set played a crucial role in investigating two main aspects: understanding target combi-

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

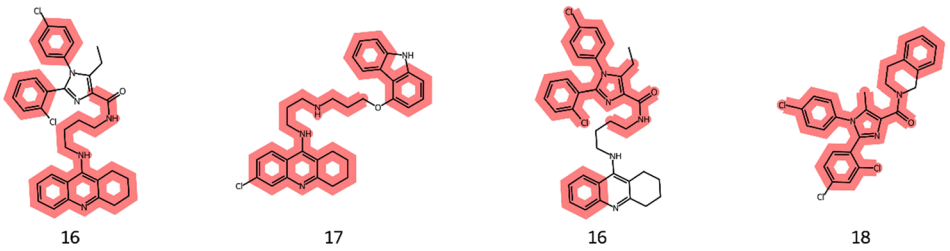


Figure 15. MCES (red) between **16** and **17** (on the left), and between **16** and **18** (on the right).

nations and assessing the chemical strategies used in designing DTDLs. Concerning target combinations, despite the ever-increasing wealth of available data from electronic clinical records, large-scale genomic sequencing, proteomics, and high-throughput screening, a strong pre-existing association between investigated targets and the targeted pathology was what mainly influenced the selection process. The analysis also revealed the importance of binding pocket similarity and the existence of a common chemical space between single target ligands. Regarding the design of DTDLs, almost all examples found were based, to some extent, on preexisting structures. If the targets shared highly similar binding pockets or exhibited significant overlap in chemical space, the resulting DTDLs usually resembled ligands already active on both targets. When designing DTDLs to interact with targets lacking pocket similarity or overlapping chemical space, identifying a tolerant region in a known single target compound played a key role in introducing activity at the second target. In cases where no tolerant region could be identified, designing linked DTDLs, with an increase in LogP and molecular weight, was often the adopted strategy. While relying on an established target-disease association or on the existing understanding of the structure–activity relationship for a scaffold are sound strategies for drug discovery campaigns, this investigation revealed that a considerable portion of DTDL therapeutic potential remains untapped. This potential lies in (i) expanding the array of proteins investigated in campaigns for the development of DTDLs, (ii) unprecedented target combinations, and (iii) innovative structures specifically designed for rational polypharmacology. The application of machine learning in the rational design of multitarget drugs and the exploitation of ultralarge compound libraries show promise and may lead to breakthroughs in the field.[36−38]

EXPERIMENTAL SECTION

Data Set Compilation. We retrieved from PubMed articles describing rationally designed dual-activity compounds. The following keywords were used: multitarget, MTDL, polypharmacology, and dual inhibitor. For each PubMed query, the overall number of retrieved articles, the number of retained articles, and the number of excluded articles were retained. Discarded articles were categorized as (i) not actually about small molecules (e.g., the DTDL is a peptide); (ii) not active at specific targets, article in which the pharmacological activity of the DTDL was not associated with a specific protein; (iii) reported compounds are active at more than two targets (MTDLs); (iv) actually about single target compounds, no DTDL reported; (v) multitarget concept, article that discussed or mentioned polypharmacology without reporting new molecules; (vi) active at two or more members of the same family, the reported DTDL was generically active at two or more members of the same protein family (e.g., aspecific MAO inhibitor); (vii) no activity data on ChEMBL;[39] (viii) pChEMBL value less than 5, articles in which no DTDL presented a pChEMBL value equal to or greater than 5; (ix) no corresponding molecule found in ChEMBL; (x) correction to previous articles; (xi) editorial; (xii) viewpoint article; (xiii) perspective; and (xiv) retracted article, based on the reasons for their exclusion. Since our planned analysis relied on molecules recorded in the ChEMBL database, we exclusively considered articles from the six core medicinal chemistry journals that contribute data to ChEMBL: Journal of Medicinal Chemistry (ACS), ACS Medicinal Chemistry Letters (ACS), European Journal of Medicinal Chemistry (Elsevier), Bioorganic and Medicinal Chemistry (Elsevier), Bioorganic and Medicinal Chemistry Letters (Elsevier), and MedChemComm (RSC). Articles describing molecules active at more than two targets were discarded, including cases in which one of the two targets is defined as encompassing multiple isoforms or members of the same family, e.g., histone deacetylase (HDAC) or cholinesterase (ChE) families. For instance, several articles describing a compound or a series of compounds generically annotated as monoamine oxidase (MAO) and/or ChE inhibitors were discarded. Furthermore, we only selected molecules whose mechanism of action was related to a specific protein target rather than to a more generic effect (e.g., antioxidant). Only the molecule the authors identified as the lead (frontrunner) for each compound series was retained. When a single molecule could not be unequivocally identified, we selected the one with the lowest molecular weight. We only retained molecules consistently displaying at both targets: a pChEMBL activity value equal to or greater than 5 and an assay confidence score equal to or greater than 8. Each molecule was annotated according to canonical SMILES notation,[40,41] univocal ID assigned in the ChEMBL database, year of publication, UniProt name, and ChEMBL target ID of each target.[42]

Single Target Compound Data Sets. Ligands from ChEMBL annotated for activity at each of the targets in the pair (pChEMBL value equal to or greater than 5 and assay confidence score equal to or greater than 8) were retrieved using Python scripts purposely written to interact with the database's API. Canonical SMILES notation for each compound was generated using the ligand-associated molfile reported in ChEMBL, employing the MolFromMolBlock function from the Chem module of RDKit[43] (version 2023-09-05). Subsequently, nonunivocal SMILES were removed from the set. To ensure that each set represented a snapshot of the molecules active on those targets before the publication of the investigated DTDL, we removed all molecules published after the reference DTDL. The year of publication for each DTDL retrieved from ChEMBL was collected and used as a filter to remove ligands within each associated single target data set that was published later. We also removed from each set other rationally conceived multitarget compounds with activity at the same targets, including molecules from the same series of the analyzed DTDL. We identified compounds with activity at both proteins and retrieved all original publications discussing these dual-activity molecules. Each molecule from these articles was manually investigated: (i) rationally devised DTDLs were removed; (ii) compounds that were engineered to have a multitarget profile but, for any reason, did not display activity at one or more of the investigated targets (failed DTDLs) were removed; (iii) single target molecules simply used as standard or as a starting point in the development of DTDLs (i.e., when a known single target compound is tested on the second target to confirm that there is no pre-existing cross activity and the reported multitarget profile is obtained through rational design)[44] were retained.

Similarity Assessment. The similarity between chemical structures was expressed in terms of Tanimoto Distance (TD),

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

which ranges from 0, when two molecules are encoded by two identical fingerprints, to 1, when the fingerprints do not have any bit in common. Fingerprints employed in this work were generated by the RDKFingerprint function as implemented in the Chem module of RDKit (v. 2023-09-05). This fingerprint format, also known as RDKit-specific fingerprint,[45] can be considered a variant of the standard Daylight fingerprint. The fingerprint size was 4096 bits.

Similarity between Single Target Data Sets. TD was used to assess the presence of overlapping areas in the chemical space explored by single target compounds with reported activity at the DTDL targets. TD was computed between the fingerprint of every ligand active at one target against those of each ligand active at the other protein. In this way, we were able to identify compounds returning the minimum TD for each target pair. Furthermore, to reduce the effect of noise introduced by redundancy, cluster analysis was performed. Ligands from individual target sets were merged in a combined set while retaining the original target association. Then, ligands were clustered using Butina's algorithm[46] (as implemented in RDKit v. 2023-09-05), adopting a threshold of 0.2 TD.

Similarity between DTDLs and Single Target Compounds. TD between fingerprints was also used to express the similarity between each DTDL and single target compounds annotated for activity at the target pair. In this case, all single target compounds with molecular weight (MW) lower than 198 or greater than 800 Da were excluded from single target sets. MW upper and lower bounds were chosen based on MW of tacrine (MW = 198.27) and ChEMBL370181 (MW = 793.79). In addition to TD, the Rascal MCES (Maximum Common Edge Substructure) score, which was calculated according to the procedure implemented in RDKit (v. 2023-09-05), was also employed. This metric determines the maximum common substructure based on bonds (edges). The resulting maximum common substructure from Rascal MCES may be represented by disconnected fragments. The procedure that computes the Rascal MCES also returns the Johnson similarity between the investigated molecules, which is calculated as the squared sum of the number of atoms and bonds in the MCES divided by the product of the sums of the numbers of atoms and bonds in the input molecules. This value ranges between 0 and 1, indicating the extent of the MCES. Rascal MCES analysis was performed as a multistep process. First, the maximum value for the Johnson similarity between all investigated compound pairs was predicted. Subsequently, the actual MCES analysis was only performed between molecule pairs with a predicted Johnson similarity equal to or greater than 0.7. The MCES relative score was computed for each DTDL—single target ligand pair. This score, defined as the ratio of the size of the MCES to the total number of bonds in the single target ligand, ranges between 0 and 1.

Ring Annotation. Ring systems were extracted from each molecule by means of a purposely written Python script. Rings were collected after fragmenting each molecule on single, nonring bonds. In this way, the procedure could correctly distinguish between rings such as benzodiazepine and benzodiazepinone that only differ because of the presence of a carbonyl group. Stereoisomers were considered as distinct entries.

Pocket Similarity. When available, we retrieved the crystallographic structure of each target from RCSB PDB.[47] We only considered holo-structures in complex with ligands with activity at that target as reported in ChEMBL. For acetylcholinesterase from *Mus musculus* (mAChE), acetylcholinesterase from *Rattus norvegicus* (rAChE), 5-lipoxygenase from *Rattus norvegicus* (r5-LOX), cyclooxygenase-2 from *Rattus norvegicus* (rCOX-2), angiotensin-converting enzyme from *Rattus norvegicus* (rACE), and phosphatidylinositol-4,5-bisphosphate 3-kinase α from *Mus musculus* (mPI3K-α) human orthologs were considered. When multiple structures were available, the one cocrystallized with the ligand displaying the lowest TD from the DTDL was considered. For each crystal structure, additional chains that do not describe the binding site, additional ligands, solvents (including water), and ions were removed. The actual pocket was defined considering all protein residues with at least one nonhydrogen atom within 5 Å from the cocrystallized ligand. Distances between pockets were calculated using DeeplyTough. This software is based on a convolutional neural network and compares, in an alignment-free fashion, feature vectors encoding the three-dimensional coordinates of pocket residues.[48] DeeplyTough was trained on the Vertex data set, which encompasses 6029 protein structures.[49]

Software. All purposely developed scripts were written in Python3.8, using RDKit (v. 2023-09-05) and ChEMBL API.[39] DeeplyThough was employed for the pocket similarity analysis. All scripts are available from the authors upon request and will soon be made available in a public repository.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.4c00838.

Complete DTDL data set (CSV)
Results of PubMed queries (CSV)
Complete results of the analysis of overlapping areas in single target ligands chemical space performed by clustering (CSV)
DTDL data set, target pair distribution, and chemical structure of DTDLs rings and ring systems displayed by a DTDL and not by its target pair (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Giovanni Bottegoni − *Department of Biomolecular Sciences, Università degli Studi di Urbino Carlo Bo, 61029 Urbino, Italy; Institute of Clinical Sciences, University of Birmingham, B15 2TT Birmingham, U.K.;* orcid.org/0000-0003-1251-583X; Email: giovanni.bottegoni@uniurb.it

Author

Vittorio Lembo − *Department of Biomolecular Sciences, Università degli Studi di Urbino Carlo Bo, 61029 Urbino, Italy; Computational and Chemical Biology, Istituto Italiano di Tecnologia, 16163 Genova, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.4c00838

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been funded by the European Union—NextGenerationEU under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem grant ECS00000041—VITALITY—CUP H33C22000430006. The authors would like to express their gratitude to the anonymous reviewers, which greatly contributed to improving the quality and clarity of this manuscript. Their thorough review and insightful comments were instrumental in shaping the final version of the paper.

■ ABBREVIATIONS

5-HT2a	serotonin 2a receptor
AChE	acetylcholinesterase
AD	Alzheimer's disease
ALK	ALK tyrosine kinase receptor

BCL-2	apoptosis regulator Bcl-2
BRD4	bromodomain-containing protein 4
BRPF1	bromodomain and PHD finger-containing protein 1
BuChE	butyrylcholinesterase
CA-IX	carbonic anhydrase IX
CBR1	cannabinoid receptor 1
CBR2	cannabinoid receptor 2
ChE	cholinesterase
c-KIT	stem cell growth factor receptor
D2	dopamine D2 receptor
DTDLs	dual-target-directed Ligands
EGFR	epidermal growth factor receptor
eGYRA	DNA gyrase subunit A from *Escherichia coli*
ER-α	estrogen receptor α
eTOPO4	topoisomerase IV subunit A from *Escherichia coli*
FAAH	fatty acid amide hydrolase
FLAP	5-lipoxygenase activating protein
FXR	farnesoid X receptor
GYRB	DNA gyrase subunit B from *Staphylococcus aureus*
HDAC	histone deacetylase
HDAC6	histone deacetylase 6
HER2	receptor tyrosine-protein kinase erbB-2
HSP90	heat shock protein HSP 90-α
IDO-1	indoleamine-pyrrole 2,3-dioxygenase-1
LTA4H	leukotriene-A4-hydrolase
mAChE	acetylcholinesterase from *Mus musculus*
MAGL	monoacylglycerol lipase
MAO	monoamine oxidase
MAO-B	monoamine oxidase B
MCES	maximum common edge substructure
MCL-1	induced myeloid leukemia cell differentiation protein Mcl-1
mPGES-1	microsomal prostaglandin E synthase-1
mPI3K-α	phosphatidylinositol-4,5-bisphosphate 3-kinase α from *Mus musculus*
MTDLs	multitarget-directed ligands
mTOR	mammalian target of rapamycin
MW	molecular weight
NET	norepinephrine transporter
NLRP3	NOD-like receptor 3
NMDAR	*N*-methyl-D-aspartate receptor
OAS	overall association score
PDE5A	phosphodiesterase 5A
PDGFR- α	platelet-derived growth factor receptor α
PI3K-α	phosphatidylinositol-4,5-bisphosphate 3-kinase α
PLA2	phospholipase A2
PLK-1	polo-like kinase 1
r5-LOX	5-lipoxygenase from *Rattus norvegicus*
rACE	angiotensin-converting enzyme from *Rattus norvegicus*
rAChE	acetylcholinesterase from *Rattus norvegicus*
RAGE	advanced glycosylation end product-specific receptor
rCOX-2	cyclooxygenase-2 from *Rattus norvegicus*
sEH	soluble epoxide hydrolase
SERT	serotonin transporter
sTOPO4	DNA topoisomerase 4 subunit B from *Staphylococcus aureus*
TAF1	transcription initiation factor TFIID subunit 1
TD	tanimoto distance
TDO	tryptophan 2,3-dioxygenase
TPSA	topological polar surface area
TRIM24	transcription intermediary factor 1-alpha
XOD	xanthine dehydrogenase

REFERENCES

(1) Morphy, R.; Rankovic, Z. Designed multiple ligands. An emerging drug discovery paradigm. In. *J. Med. Chem.* **2005**, *48*, 6523−6543.

(2) Bolognesi, M. L. Polypharmacology in a Single Drug: Multitarget Drugs. *Curr. Med. Chem.* **2013**, *20* (13), 1639 DOI: 10.2174/0929867311320130004.

(3) Bottegoni, G.; Favia, A. D.; Recanatini, M.; Cavalli, A. The role of fragment-based and computational methods in polypharmacology. *Drug Discovery Today* **2012**, *17*, 23 DOI: 10.1016/j.drudis.2011.08.002.

(4) Roth, B. L.; Sheffer, D. J.; Kroeze, W. K. Magic shotguns versus magic bullets: Selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discovery* **2004**, *3*, 353 DOI: 10.1038/nrd1346.

(5) Kampen, S.; Duy Vo, D.; Zhang, X.; Panel, N.; Yang, Y.; Jaiteh, M.; Matricon, P.; Svenningsson, P.; Brea, J.; Loza, M. I.; Kihlberg, J.; Carlsson, J. Structure-Guided Design of G-Protein-Coupled Receptor Polypharmacology. *Angew. Chem., Int. Ed.* **2021**, *60* (33), 18022 DOI: 10.1002/anie.202101478.

(6) Cavalli, A.; Bolognesi, M. L.; Mìnarini, A.; Rosini, M.; Tumiatti, V.; Recanatini, M.; Melchiorre, C. Multi-target-directed ligands to combat neurodegenerative diseases. *J. Med. Chem.* **2008**, *51*, 347 DOI: 10.1021/jm7009364.

(7) Chua, H. E.; Bhowmick, S. S.; Tucker-Kellogg, L. Synergistic target combination prediction from curated signaling networks: Machine learning meets systems biology and pharmacology. *Methods* **2017**, *129*, 60−80, DOI: 10.1016/j.ymeth.2017.05.015.

(8) Hopkins, A. L. Network Pharmacology: The Next Paradigm in Drug Discovery. *Nat. Chem. Biol.* **2008**, *4*, 682−690, DOI: 10.1038/nchembio.118.

(9) Bolognesi, M. L. Harnessing Polypharmacology with Medicinal Chemistry. *ACS Med. Chem. Lett.* **2019**, *10* (3), 273−275, DOI: 10.1021/acsmedchemlett.9b00039.

(10) Peters, J. U. Polypharmacology—Foe or friend? *J. Med. Chem.* **2013**, *56*, 8955−8971, DOI: 10.1021/jm400856t.

(11) Morphy, R. Selectively nonselective kinase inhibition: striking the right balance. *J. Med. Chem.* **2010**, *53*, 1413−1437, DOI: 10.1021/jm901132v.

(12) Morphy, R.; Kay, C.; Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discovery Today* **2004**, *9*, 641 DOI: 10.1016/S1359-6446(04)03163-0.

(13) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Towards Drugs Targeting Multiple Proteins in a Systems Biology Approach. *Curr. Top. Med. Chem.* **2007**, *7* (10), 943 DOI: 10.2174/156802607780906690.

(14) Ochoa, D.; Hercules, A.; Carmona, M.; Suveges, D.; Baker, J.; Malangone, C.; Lopez, I.; Miranda, A.; Cruz-Castillo, C.; Fumis, L.; Bernal-Llinares, M.; Tsukanov, K.; Cornu, H.; Tsirigos, K.; Razuvayevskaya, O.; Buniello, A.; Schwartzentruber, J.; Karim, M.; Ariano, B.; Martinez Osorio, R. E.; Ferrer, J.; Ge, X.; Machlitt-Northen, S.; Gonzalez-Uriarte, A.; Saha, S.; Tirunagari, S.; Mehta, C.; Roldán-Romero, J. M.; Horswell, S.; Young, S.; Ghoussaini, M.; Hulcoop, D. G.; Dunham, I.; McDonagh, E. M. The next-generation Open Targets Platform: reimagined, redesigned, rebuilt. *Nucleic Acids Res.* **2023**, *51* (D1), D1353 DOI: 10.1093/nar/gkac1046.

(15) Knight, Z. A.; Lin, H.; Shokat, K. M. Targeting the cancer kinome through polypharmacology. *Nat. Rev. Cancer* **2010**, *10*, 130 DOI: 10.1038/nrc2787.

(16) Jasial, S.; Hu, Y.; Vogt, M.; Bajorath, J. Activity-relevant similarity values for fingerprints and implications for similarity searching. *F1000Research* **2016**, *5*, 591 DOI: 10.12688/f1000research.8357.2.

(17) Meng, H.; Liu, Y.; Zhai, Y.; Lai, L. Optimization of 5-hydroxytryptamines as dual function inhibitors targeting phospholi-

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385

pase A2 and leukotriene A4 hydrolase. *Eur. J. Med. Chem.* **2013**, *59*, 160−167.
(18) Gür, Z. T.; Çalışkan, B.; Garscha, U.; Olgaç, A.; Schubert, U. S.; Gerstmeier, J.; Werz, O.; Banoglu, E. Identification of multi-target inhibitors of leukotriene and prostaglandin E2 biosynthesis by structural tuning of the FLAP inhibitor BRP-7. *Eur. J. Med. Chem.* **2018**, *150*, 876−899.
(19) Yang, L.; Chen, Y.; He, J.; Njoya, E. M.; Chen, J.; Liu, S.; Xie, C.; Huang, W.; Wang, F.; Wang, Z.; Li, Y.; Qian, S. 4,6-Substituted-1H-Indazoles as potent IDO1/TDO dual inhibitors. *Bioorg. Med. Chem.* **2019**, *27* (6), 1087−1098.
(20) Li, Y.; Zhang, S.; Wang, R.; Cui, M.; Liu, W.; Yang, Q.; Kuang, C. Synthesis of novel tryptanthrin derivatives as dual inhibitors of indoleamine 2,3-dioxygenase 1 and tryptophan 2,3-dioxygenase. *Bioorg. Med. Chem. Lett.* **2020**, *30* (11), 127159−127159.
(21) Bennett, J.; Fedorov, O.; Tallant, C.; Monteiro, O.; Meier, J.; Gamble, V.; Savitsky, P.; Nunez-Alonso, G. A.; Haendler, B.; Rogers, C.; Brennan, P. E.; Müller, S.; Knapp, S. Discovery of a Chemical Tool Inhibitor Targeting the Bromodomains of TRIM24 and BRPF. *J. Med. Chem.* **2016**, *59* (4), 1642 DOI: 10.1021/acs.jmedchem.5b00458.
(22) Aranda, R.; Villalba, K.; Raviña, E.; Masaguer, C. F.; Brea, J.; Areias, F.; Domínguez, E.; Selent, J.; López, L.; Sanz, F.; Pastor, M.; Loza, M. I. Synthesis, binding affinity, and molecular docking analysis of new benzofuranone derivatives as potential antipsychotics. *J. Med. Chem.* **2008**, *51* (19), 6085 DOI: 10.1021/jm800602w.
(23) Geng, K.; Liu, H.; Song, Z.; Zhang, C.; Zhang, M.; Yang, H.; Cao, J.; Geng, M.; Shen, A.; Zhang, A. Design, synthesis and pharmacological evaluation of ALK and Hsp90 dual inhibitors bearing resorcinol and 2,4-diaminopyrimidine motifs. *Eur. J. Med. Chem.* **2018**, *152*, 76 DOI: 10.1016/j.ejmech.2018.04.019.
(24) Zhang, B.; Liu, Z.; Xia, S.; Liu, Q.; Gou, S. Design, synthesis and biological evaluation of sulfamoylphenyl-quinazoline derivatives as potential EGFR/CAIX dual inhibitors. *Eur. J. Med. Chem.* **2021**, *216*, 11330 DOI: 10.1016/j.ejmech.2021.113300.
(25) Simoni, E.; Daniele, S.; Bottegoni, G.; Pizzirani, D.; Trincavelli, M. L.; Goldoni, L.; Tarozzo, G.; Reggiani, A.; Martini, C.; Piomelli, D.; Melchiorre, C.; Rosini, M.; Cavalli, A. Combining Galantamine and Memantine in Multitargeted, New Chemical Entities Potentially Useful in Alzheimer's Disease. *J. Med. Chem.* **2012**, *55* (22), 9708−9721.
(26) González-Naranjo, P.; Pérez-Macias, N.; Campillo, N. E.; Pérez, C.; Arán, V. J.; Girón, R.; Sánchez-Robles, E.; Martín, M. I.; Gómez-Cañas, M.; García-Arencibia, M.; Fernández-Ruiz, J.; Páez, J. A. Cannabinoid agonists showing BuChE inhibition as potential therapeutic agents for Alzheimer's disease. *Eur. J. Med. Chem.* **2014**, *73*, 56−72.
(27) Tang, C.; Li, C.; Zhang, S.; Hu, Z.; Wu, J.; Dong, C.; Huang, J.; Zhou, H.-B. Novel Bioactive Hybrid Compound Dual Targeting Estrogen Receptor and Histone Deacetylase for the Treatment of Breast Cancer. *J. Med. Chem.* **2015**, *58* (11), 4550−4572.
(28) Zhou, R.; Fang, S.; Zhang, M.; Zhang, Q.; Hu, J.; Wang, M.; Wang, C.; Zhu, J.; Shen, A.; Chen, X.; Zheng, C. Design, synthesis, and bioactivity evaluation of novel Bcl-2/HDAC dual-target inhibitors for the treatment of multiple myeloma. *Bioorg. Med. Chem. Lett.* **2019**, *29* (3), 349 DOI: 10.1016/j.bmcl.2018.12.052.
(29) Zhou, J.; Jiang, X.; He, S.; Jiang, H.; Feng, F.; Liu, W.; Qu, W.; Sun, H. Rational Design of Multitarget-Directed Ligands: Strategies and Emerging Paradigms. *J. Med. Chem.* **2019**, *62* (20), 8881 DOI: 10.1021/acs.jmedchem.9b00017.
(30) Bickerton, G. R.; Paolini, G. V.; Besnard, J.; Muresan, S.; Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nat. Chem.* **2012**, *4* (2), 90 DOI: 10.1038/nchem.1243.
(31) Rabal, O.; Sánchez-Arias, J. A.; Cuadrado-Tejedor, M.; de Miguel, I.; Pérez-González, M.; García-Barroso, C.; Ugarte, A.; Estella-Hermoso de Mendoza, A.; Sáez, E.; Espelosin, M.; Ursua, S.; Haizhong, T.; Wei, W.; Musheng, X.; Garcia-Osta, A.; Oyarzabal, J. Design, synthesis, biological evaluation and in vivo testing of dual phosphodiesterase 5 (PDE5) and histone deacetylase 6 (HDAC6)-selective inhibitors for the treatment of Alzheimer's disease. *Eur. J. Med. Chem.* **2018**, *150*, 506−524.
(32) Schmidt, J.; Rotter, M.; Weiser, T.; Wittmann, S.; Weizel, L.; Kaiser, A.; Heering, J.; Goebel, T.; Angioni, C.; Wurglics, M.; Paulke, A.; Geisslinger, G.; Kahnt, A.; Steinhilber, D.; Proschak, E.; Merk, D. A Dual Modulator of Farnesoid X Receptor and Soluble Epoxide Hydrolase To Counter Nonalcoholic Steatohepatitis. *J. Med. Chem.* **2017**, *60* (18), 7703−7724.
(33) Holtfrerich, A.; Hanekamp, W.; Lehr, M. 4-Phenoxyphenyl)-tetrazolecarboxamides and related compounds as dual inhibitors of fatty acid amide hydrolase (FAAH) and monoacylglycerol lipase (MAGL. *Eur. J. Med. Chem.* **2013**, *63*, 64−75.
(34) Lange, J. H. M.; Coolen, H. K. A. C.; van der Neut, M. A. W.; Borst, A. J. M.; Stork, B.; Verveer, P. C.; Kruse, C. G. Design, Synthesis, Biological Properties, and Molecular Modeling Investigations of Novel Tacrine Derivatives with a Combination of Acetylcholinesterase Inhibition and Cannabinoid CB1 Receptor Antagonism. *J. Med. Chem.* **2010**, *53* (3), 1338−1346.
(35) Taylor, R. D.; Maccoss, M.; Lawson, A. D. G. Rings in drugs. *J. Med. Chem.* **2014**, *57*, 5845 DOI: 10.1021/jm4017625.
(36) Schneider, P.; Walters, W. P.; Plowright, A. T.; Sieroka, N.; Listgarten, J.; Goodnow, R. A.; Fisher, J.; Jansen, J. M.; Duca, J. S.; Rush, T. S.; Zentgraf, M.; Hill, J. E.; Krutoholow, E.; Kohler, M.; Blaney, J.; Funatsu, K.; Luebkemann, C.; Schneider, G. Rethinking drug design in the artificial intelligence era. *Nat. Rev. Drug Discovery* **2020**, *19*, 353 DOI: 10.1038/s41573-019-0050-3.
(37) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; Tolmachev, A. A.; Shoichet, B. K.; Roth, B. L.; Irwin, J. J. Ultra-large library docking for discovering new chemotypes. *Nature* **2019**, *566* (7743), 224 DOI: 10.1038/s41586-019-0917-9.
(38) Besnard, J.; Ruda, G. F.; Setola, V.; Abecassis, K.; Rodriguiz, R. M.; Huang, X. P.; Norval, S.; Sassano, M. F.; Shin, A. I.; Webster, L. A.; Simeons, F. R. C.; Stojanovski, L.; Prat, A.; Seidah, N. G.; Constam, D. B.; Bickerton, G. R.; Read, K. D.; Wetsel, W. C.; Gilbert, I. H.; Roth, B. L.; Hopkins, A. L. Automated design of ligands to polypharmacological profiles. *Nature* **2012**, *492* (7428), 215 DOI: 10.1038/nature11691.
(39) Davies, M.; Nowotka, M.; Papadatos, G.; Dedman, N.; Gaulton, A.; Atkinson, F.; Bellis, L.; Overington, J. P. ChEMBL web services: Streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* **2015**, *43* (W1), W612 DOI: 10.1093/nar/gkv352.
(40) Weininger, D. SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31 DOI: 10.1021/ci00057a005.
(41) Daylight Theory: SMILES.
(42) Bateman, A.; Martin, M. J.; Orchard, S.; Magrane, M.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E. H.; Britto, R.; Bye-A-Jee, H.; Cukura, A.; Denny, P.; Dogan, T.; Ebenezer, T. G.; Fan, J.; Garmiri, P.; da Costa Gonzales, L. J.; Hatton-Ellis, E.; Hussein, A.; Ignatchenko, A.; Insana, G.; Ishtiaq, R.; Joshi, V.; Jyothi, D.; Kandasaamy, S.; Lock, A.; Luciani, A.; Lugaric, M.; Luo, J.; Lussi, Y.; MacDougall, A.; Madeira, F.; Mahmoudy, M.; Mishra, A.; Moulang, K.; Nightingale, A.; Pundir, S.; Qi, G.; Raj, S.; Raposo, P.; Rice, D. L.; Saidi, R.; Santos, R.; Speretta, E.; Stephenson, J.; Totoo, P.; Turner, E.; Tyagi, N.; Vasudev, P.; Warner, K.; Watkins, X.; Zaru, R.; Zellner, H.; Bridge, A. J.; Aimo, L.; Argoud-Puy, G.; Auchincloss, A. H.; Axelsen, K. B.; Bansal, P.; Baratin, D.; Batista Neto, T. M.; Blatter, M. C.; Bolleman, J. T.; Boutet, E.; Breuza, L.; Gil, B. C.; Casals-Casas, C.; Echioukh, K. C.; Coudert, E.; Cuche, B.; de Castro, E.; Estreicher, A.; Famiglietti, M. L.; Feuermann, M.; Gasteiger, E.; Gaudet, P.; Gehant, S.; Gerritsen, V.; Gos, A.; Gruaz, N.; Hulo, C.; Hyka-Nouspikel, N.; Jungo, F.; Kerhornou, A.; Le Mercier, P.; Lieberherr, D.; Masson, P.; Morgat, A.; Muthukrishnan, V.; Paesano, S.; Pedruzzi, I.; Pilbout, S.; Pourcel, L.; Poux, S.; Pozzato, M.; Pruess, M.; Redaschi, N.; Rivoire, C.; Sigrist, C. J. A.; Sonesson, K.; Sundaram, S.; Wu, C. H.; Arighi, C. N.; Arminski, L.; Chen, C.; Chen, Y.; Huang, H.; Laiho, K.; McGarvey, P.; Natale, D. A.; Ross, K.; Vinayaka, C. R.; Wang, Q.; Wang, Y.; Zhang, J. UniProt: the Universal Protein Knowledgebase in

2023. *Nucleic Acids Res.* **2023**, *51* (D1), D523 DOI: 10.1093/nar/gkac1052.
(43) Landrum, G. *RDKit: Open-Source Cheminformatics*. In Http://Www.Rdkit.Org/, 2006; Vol. 3.
(44) De Simone, A.; Ruda, G. F.; Albani, C.; Tarozzo, G.; Piomelli, D.; Cavalli, A.; Bottegoni, G. Applying a multitarget rational drug design strategy: the first set of modulators with potent and balanced activity toward dopamine D3 receptor and fatty acid amide hydrolase. *Chem. Commun.* **2014**, *50* (38), 4904 DOI: 10.1039/c4cc00967c.
(45) Daylight Theory: Fingerprints.
(46) Butina, D. Unsupervised data base clustering based on daylight's fingerprint and Tanimoto similarity: A fast and automated way to cluster small and large data sets. *J. Chem. Inf. Comput. Sci.* **1999**, *39* (4), 747 DOI: 10.1021/ci9803381.
(47) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *In Nucleic Acids Research*, **2000**; Vol. *28*.
(48) Simonovsky, M.; Meyers, J., Deeply Tough: Learning Structural Comparison of Protein Binding Sites. *J. Chem. Inf. Model.* **2020**, *60* (4).
(49) Chen, Y. C.; Tolbert, R.; Aronov, A. M.; McGaughey, G.; Walters, W. P.; Meireles, L., Prediction of Protein Pairs Sharing Common Active Ligands Using Protein Sequence, Structure, and Ligand Similarity. *J. Chem. Inf. Model.* **2016**, *56* (9).

https://doi.org/10.1021/acs.jmedchem.4c00838
J. Med. Chem. 2024, 67, 10374−10385