

University of Groningen

## In Silico Design and Selection of New Tetrahydroisoquinoline-Based CD44 Antagonist Candidates

Ruiz-Moreno, Angel J; Reyes-Romero, Atilio; Dömling, Alexander; Velasco-Velázquez, Marco A

*Published in:*  
Molecules

*DOI:*  
[10.3390/molecules26071877](https://doi.org/10.3390/molecules26071877)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2021

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Ruiz-Moreno, A. J., Reyes-Romero, A., Dömling, A., & Velasco-Velázquez, M. A. (2021). In Silico Design and Selection of New Tetrahydroisoquinoline-Based CD44 Antagonist Candidates. *Molecules*, 26(7), [1877]. <https://doi.org/10.3390/molecules26071877>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



## Article

# In Silico Design and Selection of New Tetrahydroisoquinoline-Based CD44 Antagonist Candidates

Angel J. Ruiz-Moreno <sup>1,2,3,4</sup> , Atilio Reyes-Romero <sup>4</sup> , Alexander Dömling <sup>4,\*</sup> and Marco A. Velasco-Velázquez <sup>1,2,\*</sup> 

- <sup>1</sup> Departamento de Farmacología, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), Ciudad de México 04510, México; angel\_ruiz\_m@comunidad.unam.mx
  - <sup>2</sup> Unidad Periférica de Investigación en Biomedicina Translacional, Facultad de Medicina, Universidad Nacional Autónoma de México (UNAM), Félix Cuevas 540, Ciudad de México 03229, México
  - <sup>3</sup> Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM), Ciudad de México 04510, México
  - <sup>4</sup> Drug Design Group, Department of Pharmacy, University of Groningen, 9700 AD Groningen, The Netherlands; atilioreyesromero@gmail.com
- \* Correspondence: a.s.s.domling@rug.nl (A.D.); marcovelasco@unam.mx (M.A.V.-V.); Tel.: +31-50-363-330 (A.D.); +52-55-5623-2282 (M.A.V.-V.)

**Abstract:** CD44 promotes metastasis, chemoresistance, and stemness in different types of cancer and is a target for the development of new anti-cancer therapies. All CD44 isoforms share a common N-terminal domain that binds to hyaluronic acid (HA). Herein, we used a computational approach to design new potential CD44 antagonists and evaluate their target-binding ability. By analyzing 30 crystal structures of the HA-binding domain (CD44HAbd), we characterized a subdomain that binds to 1,2,3,4-tetrahydroisoquinoline (THQ)-containing compounds and is adjacent to residues essential for HA interaction. By computational combinatorial chemistry (CCC), we designed 168,190 molecules and compared their conformers to a pharmacophore containing the key features of the crystallographic THQ binding mode. Approximately 0.01% of the compounds matched the pharmacophore and were analyzed by computational docking and molecular dynamics (MD). We identified two compounds, Can125 and Can159, that bound to human CD44HAbd (hCD44HAbd) in explicit-solvent MD simulations and therefore may elicit CD44 blockage. These compounds can be easily synthesized by multicomponent reactions for activity testing and their binding mode, reported here, could be helpful in the design of more potent CD44 antagonists.

**Keywords:** tetrahydroisoquinoline; CD44; computational combinatorial chemistry; pharmacophore; molecular dynamics



**Citation:** Ruiz-Moreno, A.J.; Reyes-Romero, A.; Dömling, A.; Velasco-Velázquez, M.A. In Silico Design and Selection of New Tetrahydroisoquinoline-Based CD44 Antagonist Candidates. *Molecules* **2021**, *26*, 1877. <https://doi.org/10.3390/molecules26071877>

Academic Editor: Rainer Riedel

Received: 27 January 2021

Accepted: 14 March 2021

Published: 26 March 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/>).

## 1. Introduction

CD44 is a transmembrane glycoprotein that functions as a receptor for the glycosaminoglycan hyaluronic acid (HA), an integral component of the extracellular matrix [1,2]. CD44 is expressed on multiple cells, including embryonic stem cells and differentiated cells, mediating cellular functions such as adhesion, homing, migration, and extravasation [1,2]. CD44 transcript can undergo alternative splicing, generating multiple isoforms of CD44, but all of them conserve intact the HA-binding domain (HAbd) and, therefore, can be activated by HA [3].

CD44 expression correlates with unfavorable clinical outcomes in multiple types of cancer [4–8]. CD44 activation by HA in cancer cells induces transcriptional and epigenetic changes that stimulate signaling pathways controlling invasiveness and metastasis, chemoresistance, and stemness [9–11]. For instance, in breast cancer cells, HA binding to CD44 induces epithelial–mesenchymal transition, which increases cell migration and invasive capacity [12], and promotes survival under detached conditions during the development of metastasis [13]. Moreover, CD44 is expressed in cancer stem cells that survive



chemotherapy in models of glioblastoma [14], breast [15], pancreatic [16], colorectal [17], and prostate [18] cancer. Consistent with its key role in cancer progression, CD44 silencing impairs chemoresistance, clonogenicity, tumorigenicity, and/or metastasis [19–21]. Therefore, blockage of HA-binding to CD44 has been proposed as a potential therapeutic strategy for cancer.

The CD44HAbd is located in the N-terminal end of the extracellular region of the receptor. Structural analysis of murine CD44HAbd crystals showed that only 13 residues along a shallow groove mediate HA-binding [22]. The residues Arg41, Tyr42, Arg78, Tyr79 in hCD44HAbd (Arg45, Tyr46, Arg82, Tyr83 in mCD44HAbd) have been previously described as essential for HA-binding by directed mutagenesis experiments or crystal analysis [23,24]. Given the lack of an obvious druggable pocket in the HA-binding site, small molecule inhibitors that interact with allosteric sites within the CD44HAbd have been developed [24–27]. However, those compounds bind to CD44HAbd in the high micromolar or even low millimolar range, limiting further applications. Therefore, there is a need for new CD44 antagonists with improved affinity, efficacy, and physicochemical properties for future effective translation to the clinic.

Herein we designed and evaluated the binding of new potential CD44 antagonists using an *in silico* strategy. We identified that small molecules sharing a 1,2,3,4-tetrahydroisoquinoline (THQ) motif are frequently co-crystallized with CD44HAbd in a subdomain adjacent to the residues that are essential for HA-binding. By computational combinatorial chemistry (CCC), we generated libraries including more than 168,000 THQ-containing molecules. The new molecules (i) could be easily synthesized by multicomponent reactions, (ii) are diverse, and (iii) display drug-like physicochemical properties. We selected a subset of 163 candidates matching the key features of the reported THQ binding mode for further analysis by computational docking. The nine candidates with the highest frequency of poses reproducing the reported THQ binding mode were analyzed by molecular dynamics (MD). Our results allowed the identification of two compounds predicted to stably bind to hCD44HAbd in an aqueous solution. Those compounds may be useful as CD44 antagonists, and the information of their binding mode can be employed as the basis for the design of new bioactive molecules that target CD44.

## 2. Results

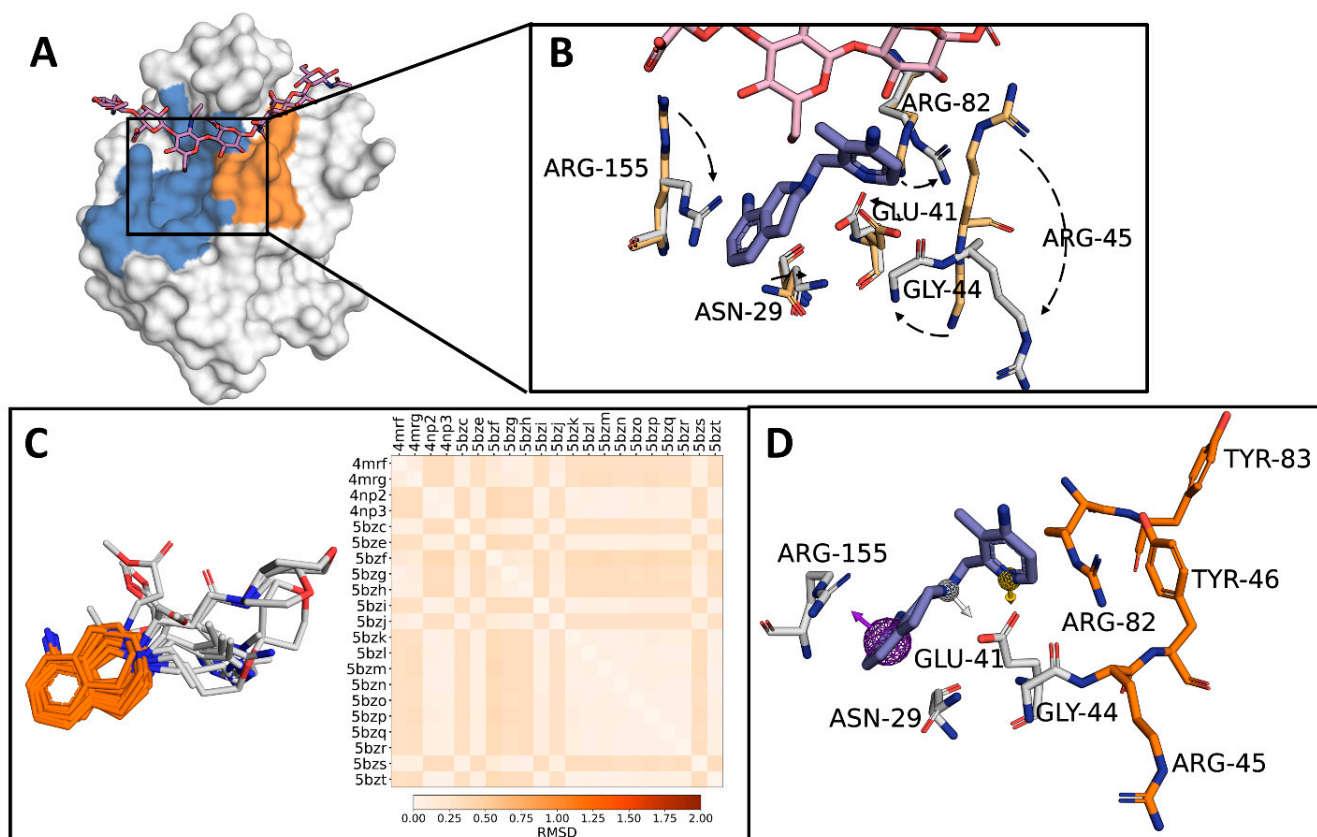
### 2.1. Identification of a Target Subdomain within the CD44HAbd and Generation of a THQ-Based Pharmacophore

Aiming to identify relevant regions for drug design, we compared the 30 crystal structures available in Protein Data Bank (PDB) that comprise the HA-binding domain of human (three structures) or mouse (27 structures) CD44 (Table S1). The three human structures correspond to the apo form of CD44HAbd. For mCD44HAbd, 2JCP represents the apo-CD44HAbd, three structures are co-crystallized with HA (2JCQ, 2JCR, and 4MRD), and the rest are co-crystallized with molecules weighting 100–250 Da. Within the structures containing small molecules, 21 of them are co-crystallized with compounds containing the THQ motif. Sequence identity analysis showed 100% identity among all hCD44HAbd, 99–100% among mCD44HAbd, and 86–88% between hCD44HAbd and mCD44HAbd (Figure S1A). Due to the high identity among CD44HAbd structures, we compared all of them in a structural analysis. The root mean square deviation (RMSD) profiles for alpha-carbons and full atoms showed the higher deviations on some residues previously reported as essential for HA-binding by direct mutagenesis experiments (Arg41, Tyr42, Arg78, and Tyr79) in hCD44HAbd [23] (Figure S1B).

Analysis of the 21 structures co-crystallized with THQ-containing small molecules identified that all these ligands bind to a pocket contiguous to the HA-binding domain (Figure 1A). Their binding drives a shift in multiple residues of mCD44HAbd, including some of the key residues participating in HA-binding, namely, Arg45, Arg82, and Arg155 (Figure 1B). Alignment of THQ-composing atoms showed that the binding mode is highly conserved among the 21 crystals analyzed (Figure 1C). Thus, we model a pharmacophore from the co-crystallized molecules containing the THQ substructure. The generated model



included four pharmacophoric descriptors—aromatic, two hydrogen bond donors, and a positively charged ion (Figure 1D).

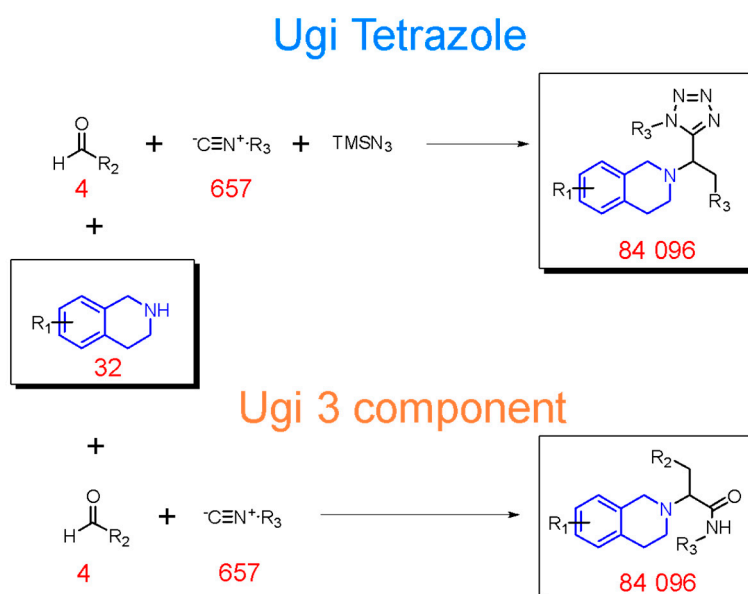


**Figure 1.** Generation of a 1,2,3,4-tetrahydroisoquinoline (THQ)-based pharmacophore. (A) CD44HAbd (white surface) has a subdomain where the THQ-containing molecules can bind (blue) and partially overlaps with the region containing the key residues for CD44–HA binding (Arg45, Tyr46, Arg82, and Tyr83-orange-). HA is shown in pink. (B) Spatial positions of the lateral chain of residues Asn29, Glu41, Gly44, Arg45, Arg82, and Arg155 from mCD44HAbd in the HA-bound form (residues shown as beige sticks; HA in pink) vs. the form bound to a THQ-containing molecule (residues shown as grey sticks; ligand in blue). Dashes represent the residue shifts between both states. (C) Structural alignment and root mean square deviation (RMSD) comparison of the THQ scaffold (orange) of the molecules co-crystallized with mCD44HAbd. (D) Pharmacophore model generated using THQ-containing molecules as a template. Purple sphere: aromatic; yellow sphere: hydrogen bond donor; white sphere: merged hydrogen bond donor and positively charged ion. A molecule with the THQ substructure is included (blue) to show the interacting residues in CD44 (grey sticks) and its proximity to residues essential for HA-binding (orange sticks).

## 2.2. Generation of THQ-Containing Libraries

To identify new compounds with the potential capability to interfere with the CD44–HA binding, we employed CCC to generate two libraries of compounds that include the THQ scaffold. To facilitate the synthesis of our compounds in subsequent research, we decided to use multicomponent reactions (MCR) synthesis routes. Considering the characteristics of the THQ-containing small molecules co-crystallized with mCD44HAbd, we implemented the Ugi four-component tetrazole synthesis [28,29] and Ugi three-component reaction [30] for our CCC experiments. For each MCR route, we obtained 84,096 different compounds (Scheme 1).





**Scheme 1.** CCC strategy for the generation of the test libraries.

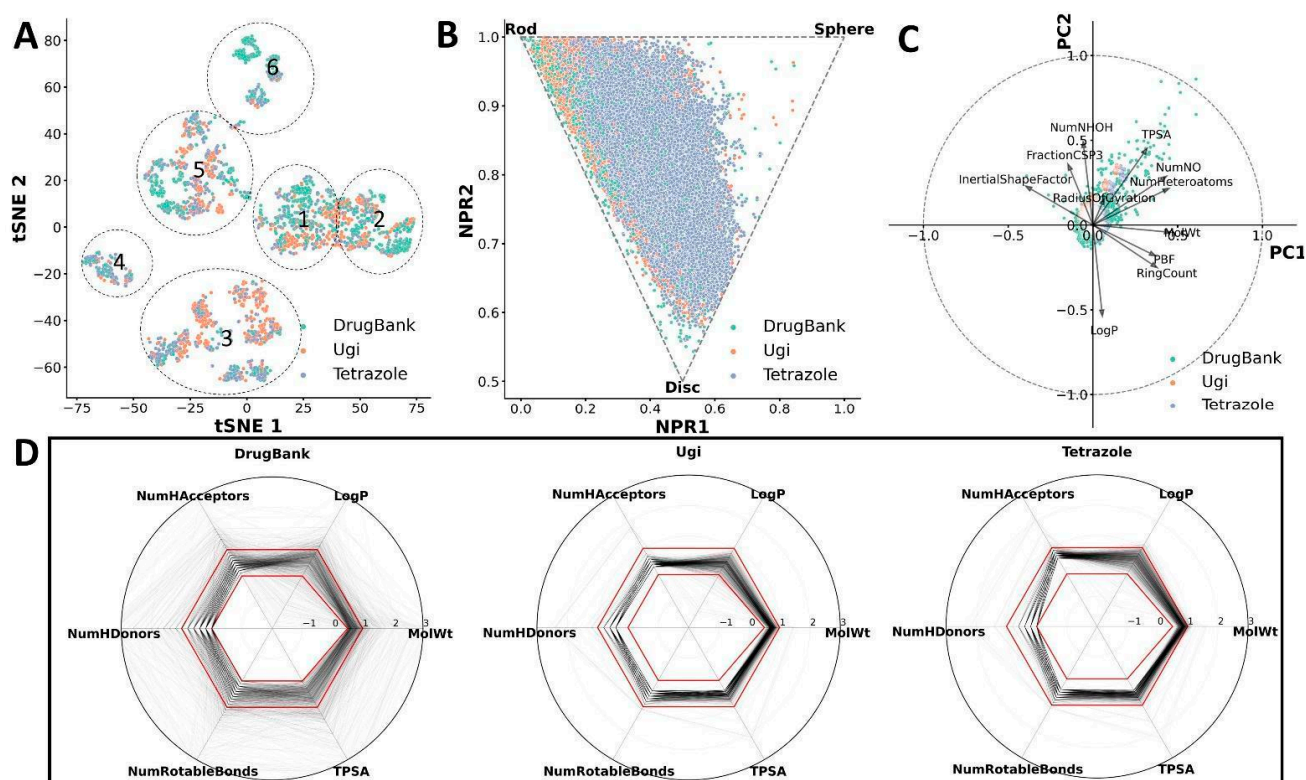
### 2.3. Cheminformatic Analysis of CCC Compounds

In order to explore the chemical diversity and physicochemical properties of the designed compounds, we employed a series of cheminformatic analyses, calculating 30 different 2D/3D-shape and physicochemical molecular descriptors. For comparison, we also studied the compounds within DrugBank Database 5.0.10, a library of 1542 FDA-approved small molecules [31]. The diversity analysis of 1500 randomly sampled compounds of each library, using t-distributed stochastic neighbor embedding (tSNE) employing Molecular ACCess System (MACCS) keys [32], showed that the compounds from the tetrazole and Ugi libraries display similar structural diversity to the compounds inside DrugBank. The K-means clustering showed that compounds from the CCC and DrugBank libraries distribute similarly on five out of six clusters, whereas the sixth cluster was enriched in DrugBank small molecules (Figure 2A).

A normalized principal moments ratio (NPR) analysis was conducted to assess the molecular shape distribution of compounds. The results showed that the minimum energy conformers of the compounds from the three libraries presented similar 3D shapes, predominantly rod- and disk-shaped, with only a few compounds displaying a spherical shape (Figure 2B). We also performed a principal components analysis (PCA) employing non-redundant molecular descriptors selected by their correlation (Figure S2). PCA showed that the compounds in CCC-generated libraries possess similar molecular descriptors and physicochemical properties to those of the DrugBank Database, displaying a dense accumulation of the compounds at the origin of the principal components (PCs), PC 1 and PC2. Topological polar surface area (TPSA) and logarithmic partition coefficient (LogP) were the primary descriptors, correlating positively with PC1 and PC2 and negatively with PC2, respectively (Figure 2C).

Finally, the extended Lipinski's rule of five (Ro5) analysis showed that most newly designed compounds comply with the physicochemical properties required for oral use [33]. Interestingly, the CCC-generated libraries displayed a more homogeneous distribution inside the extended Ro5 than the group of compounds in DrugBank (Figure 2D).





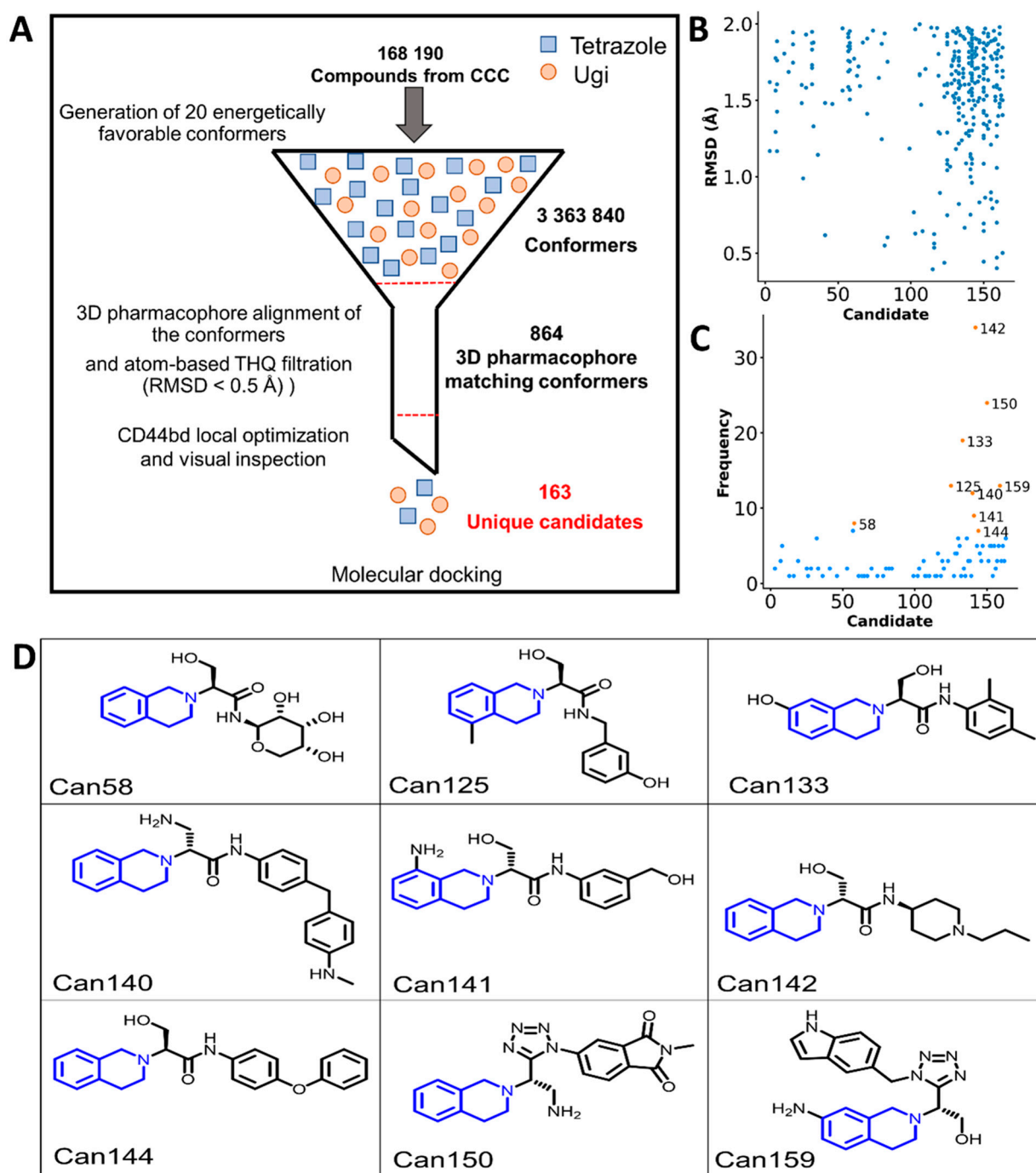
**Figure 2.** Characteristics of the generated libraries. (A) t-distributed stochastic neighbor embedding (tSNE) chart of structural diversity analysis for the compounds generated by Ugi three-component reaction (Ugi) or Ugi four-component tetrazole synthesis (Tetrazole). For comparison, a database containing FDA-approved molecules was included (DrugBank). (B) Normalized principal moments ratio (NPR) analysis. (C) Principal components analysis (PCA) for molecular and physicochemical descriptors. (D) Lipinski's rule of five (Ro5) analysis, which included molecular weight (MolWt), logarithmic partition coefficient (LogP), number of hydrogen bond donors (NumHDonors) and acceptors (NumHAcceptors), and the topological polar surface area (TPSA). A gray line represents each compound, and the density indicates the frequency of compounds.

## 2.4. Virtual Screening

To identify new compounds with the theoretical ability to bind CD44, we generated expanded libraries containing 20 energetically favorable conformers for each compound within the CCC-generated libraries, retrieving 3,363,840 conformers. The expanded libraries were screened by alignment to the pharmacophore, followed by local optimization in CD44HAbd and visual inspection. We identified 864 conformers from 163 unique compounds that matched the selection criteria (Figure 3A). Only those molecules were employed for the subsequent experiments.

The docking protocol employed for filtering was validated by docking the 21 THQ-containing molecules co-crystallized with mCD44HAbd into protein 5BZM. We observed that the crystal pose was more frequently reproduced in compounds with smaller substituents on the THQ motif (Figure S3A). Thus, additional analysis of candidates considered only the position of the THQ atoms.





**Figure 3.** Virtual screening of THQ-containing molecules as potential CD44 antagonists. (A) The compounds within the CCC-generated libraries were sequentially filtered using the depicted strategy. (B) RMSD analysis comparing the THQ position in crystals vs. the poses obtained by docking of 163 unique compounds to human or mouse CD44HAbd. (C) Analysis of the frequency of poses with RMSD < 2 Å allowed the selection of nine candidates (orange dots). (D) Structure of the nine candidates selected by virtual screening with the THQ motif highlighted in blue.

The 163 candidates were docked against hCD44HAbd and mCD44HAbd for comparison. For each candidate/receptor pair, the docking scores of 25 poses were analyzed (Figure S4). The docking poses were compared to the coordinates of THQ crystalized on mCD44HAbd since none of the available human crystal structures contained THQ-derived compounds. We focused on the compounds showing docking poses matching the crystallized THQ atoms with RMSD < 2 Å (Figure 3B). We selected the compounds with the



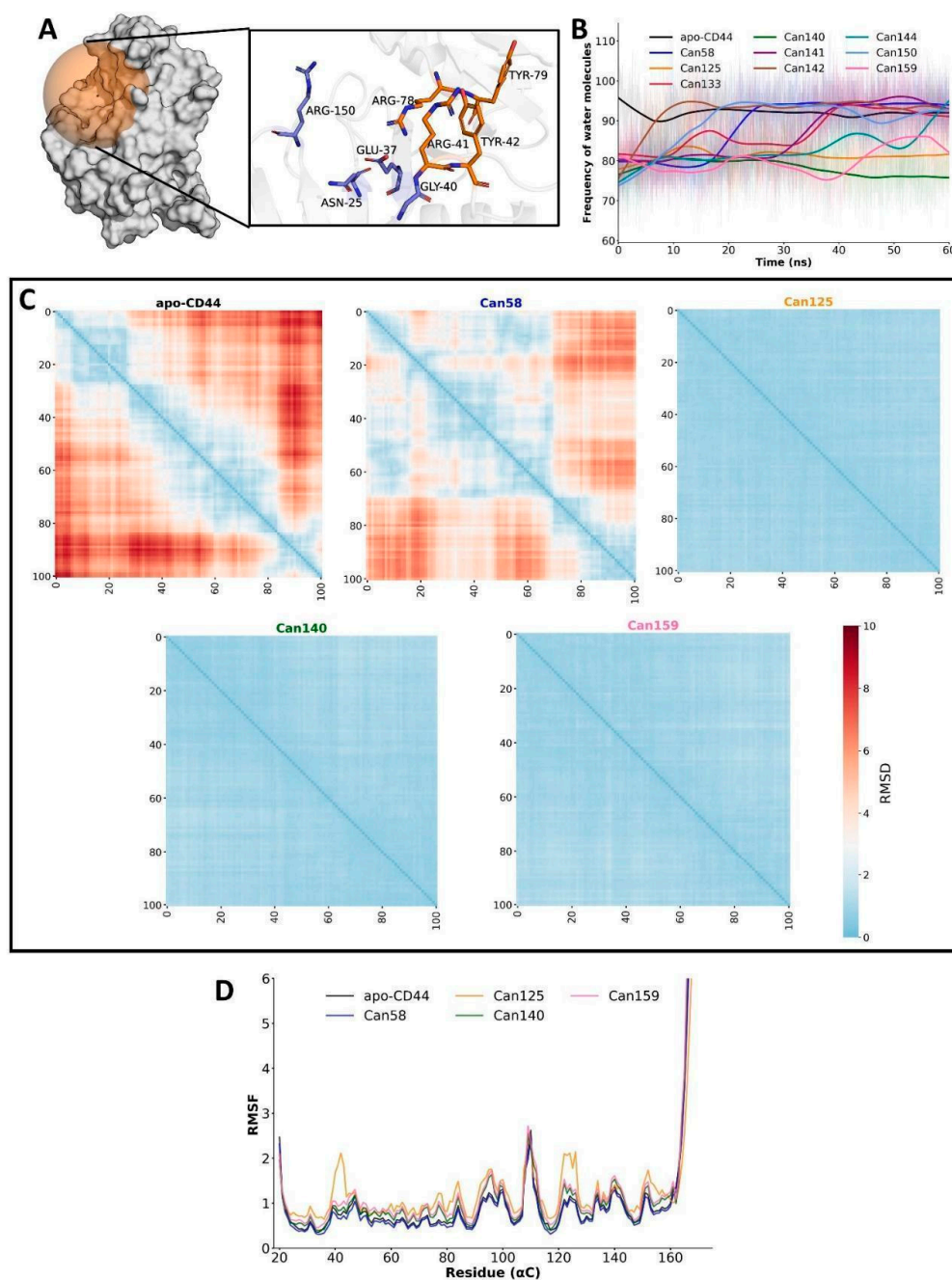
highest frequency of matching poses, ranging from 8/50 to 34/50 (Figure 3C). For the nine selected molecules (Figure 3D and Table S2), we assessed the THQ-binding site selectivity by docking the compounds into four/five additional pockets using three relevant forms of the receptor: apo-hCD44HAbd, HA-bound mCD44HAbd, and THQ-containing molecule mCD44HAbd. With the exception of the candidate (Can) 142, all molecules were predicted to bind the region of interest with better or similar affinity than other pockets (Figure S5). We then performed pose clustering analysis (Figure S6) to identify the best pose for molecular dynamics (MD) simulations.

### 2.5. Molecular Dynamics and Free Energy Calculation

We performed solvent explicit MD simulations to characterize the binding of the selected candidates to hCDHA44. The apo-hCD44HAbd was included as a control. Our analysis focused on the THQ binding site reported for mCD44HAbd (Figure 4A). By quantifying the water molecules displacement in the selected region, we identified that candidate Can58, Can133, Can141, Can142, and Can150 left the cavity during the simulation. Furthermore, Can58, Can142, and Can150 moved from the binding site at the early steps of the MD simulation. On the other hand, three compounds, (Can125, Can140, and Can159) remained on the binding site during the 100 ns of simulation and maintained a constant number of local water molecules (Figure 4B). Interestingly, the global backbone RMSD analysis of the systems with those three candidates showed a different profile than the one generated by apo-CD44HAbd (Figure 4C) or by systems with other candidates (Figure S7). Additionally, a shift on some residues was observed on the root mean square fluctuation (RMSF) profile generated for Can125 (Figure 4D).

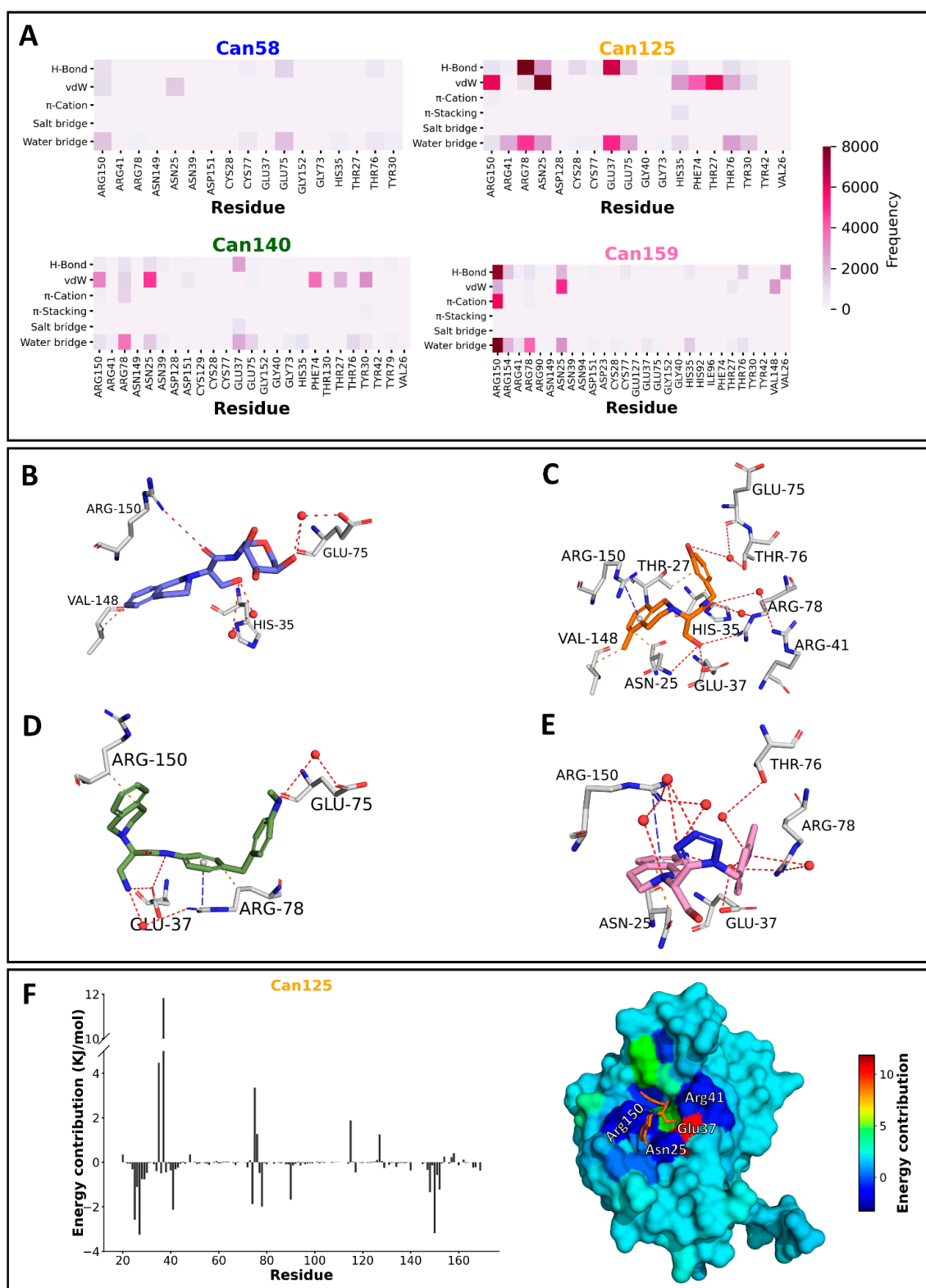
The frequencies of molecular interactions generated by Can125, Can140, and Can159 were studied during the whole simulation; Can58 was considered in the analysis for comparison (Figure S8). As expected from the water analysis, Can58 showed a low frequency of interactions in the THQ binding site, which included water bridges and hydrogen bonds with Arg150 and Glu75 (Figure 5A,B). Of the candidates studied by MD, Can125 formed the highest number of interactions, the most frequent were hydrogen bonds with Arg41 and Glu37, Van der Waals interactions with Arg150, Arg78, and Thr27, and water bridges with Arg78 and Glu37 (Figure 5A,C). Can140 showed Van der Waals interactions with residues Arg150, Asn25, and Phe74 predominantly, but it was also able to form water bridges with Arg78 and Glu37 (Figure 5A,D). Can159 showed a high frequency of  $\pi$ -cation, and hydrogen bonds interactions with Arg150, and water bridges with Arg78, and Arg150. Moreover, the Can159 also displayed the frequent formation of Van der Waals interactions with Asn25 (Figure 5A,E).





**Figure 4.** Molecular dynamics (MD) analysis identified Can 125 as a potential CD44 antagonist. (A) Spatial representation of the analyzed pocket in hCDHAbd. Inset shows the lateral chains of residues reported as essential for HA-binding (orange sticks) or those that mediate the interaction with the THQ-containing compounds employed for pharmacophore modeling (blue sticks). (B) Water molecules displacement analysis for the nine candidates shown in Figure 3D. (C) Pairwise backbone RMSD matrix along 100 ns of MD simulation from systems including candidates with stable binding to hCD44HAbd. The unliganded protein (apo-CD44) and the system with Can58 are included for comparison. (D) Alpha-carbon RMSF analysis for the systems presented in C.





**Figure 5.** Identification and importance of residues mediating compound binding. **(A)** Types of interactions that support stable candidate/hCD44HAbd interactions and their frequency. Note that Can58, presented here for comparison, left the binding site during the simulation. **(B–E)** The 3D molecular interactions of a representative frame from MD simulation of the complexes between Can58 **(B)**, Can125 **(C)**, Can140 **(D)**, and Can159 **(E)** with hCD44HAbd (gray sticks). **(F)** Per-residue energy decomposition for Can125. Most positive values correspond to His35 and Glu37, and most negative values to Asn25, Thr27, Arg41, and Arg150.



Calculation of the binding free energy (Table 1) showed that Can140 has lower binding energy than Can125 and Can159, with the electrostatic energy as the component that contributes most to these differences. Nevertheless, Can140 showed large energy fluctuations during the simulations (Figure S9), suggesting a possible rearrangement of the binding pose during the experiment. In contrast, Can125 and Can159 showed stable energetic profiles (Figure S9). Per-residue energetic decomposition (Figure 5F and Figure S10) revealed that Can125 binding to hCD44HAbd is supported by energetically favorable interactions with residues essential for HA-binding (Arg41 and Arg78) or residues selected in the pharmacophore modeling (Arg150). Together, these observations suggest that Can125 is the most promising compound for biological evaluation.

**Table 1.** Average free energy calculated from MD simulations (mean  $\pm$  standard deviation).

Compound	Van der Waals Energy	Electrostatic Energy	Polar Solvation Energy	SASA Energy	Binding Energy
Can58	$-33.871 \pm 35.190$	$-18.447 \pm 29.906$	$47.592 \pm 66.887$	$-4.667 \pm 4.853$	$-9.393 \pm 41.581$
Can125	$-116.512 \pm 15.988$	$-52.356 \pm 23.296$	$123.555 \pm 27.579$	$-13.962 \pm 1.252$	$-59.274 \pm 17.744$
Can140	$-96.931 \pm 36.947$	$-128.983 \pm 66.843$	$106.522 \pm 90.455$	$-12.078 \pm 4.022$	$-131.470 \pm 41.310$
Can159	$-99.395 \pm 18.056$	$-34.928 \pm 25.315$	$112.733 \pm 40.928$	$-11.318 \pm 1.938$	$-32.908 \pm 17.750$

### 3. Discussion

Due to the essential physiological and pathological roles of CD44, several crystal structures of its HAbd have been solved, either in the unligated form (apo) or in complex with HA. Our structural analysis of the crystals available at PDB corroborated the previously identified shifts in residues participating in HA-binding, including Arg41, Tyr42, Arg78, and Tyr79 in hCD44HAbd, and Arg45, Tyr46, Arg82, and Tyr83 in mCD44HAbd [22–24,34]. We found the shifts in Arg41 and Arg78 as particularly important for drug design because (i) the shift in Arg41 has been identified as a trigger for the conversion of high (active) to low (inactive) affinity conformations of CD44HAbd [22] and (ii) both residues are close to a pocket that binds to small molecules with the THQ motif and suffer conformation changes induced by ligand binding.

The THQ-binding pocket has been employed for the development of molecules that display a similar affinity for human or mouse CD44HAbd (ranging from 0.4  $\mu$ M to 6.9  $\mu$ M for hCD44HAbd, and 0.5  $\mu$ M to 11.2  $\mu$ M mCD44HAbd) [24]. Thus, we used 21 murine crystals containing THQ-based molecules for developing a pharmacophore that contained the key interactions mediating the binding of those compounds to mCD44HAbd and hypothesized that the model could be used in the identification of new antagonists for the human version of the receptor.

By using MCR-based CCC, we also generated libraries of easily synthesizable compounds that contain the THQ substructure. MCR are one-pot reactions in which two or more starting materials are used simultaneously; thus, most of the atoms from the initial building blocks are incorporated into the final product of the reaction [35,36]. The Ugi four-component tetrazole synthesis [28,29] and Ugi three-component reaction [30] are well described, easy to perform, and have been suggested as synthesis methods for diverse drug-like molecules [36]. Additionally, we selected building blocks that are commercially available at low-cost, which will allow compound synthesis and activity evaluation in further studies. Chemoinformatic characterization of the generated databases showed that the compounds are highly diverse but contain similar structural and physicochemical characteristics to those of marketable molecules. The DrugBank Database drugs frequently adopt rod- and disc-shapes [37] and comply with Lipinski's Ro5 [33,38]. The compounds within our CCC-generated libraries predominantly displayed these shapes, due to the high predominance of non-cyclic molecules [37,39], and have drug-like physicochemical characteristics. Molecules with the THQ substructure have been identified as nicotinic [40] or



muscarinic [41] receptor antagonists, in addition to inhibitors of the angiotensin-converting enzyme [42]. Thus, the databases reported here may be useful starting points for identifying new compounds with those activities.

A robust exploration of the conformational space allowed the selection of 163 unique candidates that matched with the pharmacophoric model. To overcome the lack of structural information regarding the binding mode of THQ-containing molecules to hCD44HAbd, we investigated the binding mode of the candidates in hCD44HAbd and mCD44HAbd by a docking protocol that was able to reproduce the crystallographic binding mode of most co-crystallized molecules with an all-atoms RMSD threshold  $< 3\text{\AA}$ . We identified nine candidates that reproduced the THQ crystallographic pose with an RMSD  $< 2\text{\AA}$  and high frequency. A similar strategy, using a THQ-based pharmacophore for screening identified potential anticonvulsant compounds [43].

MD in explicit solvent further characterized the binding capability of the best nine candidates from virtual screening. We applied this method considering that the effects of solvation play a key role in forming molecular interactions in ligand–protein complexes. Thus, simulations employing explicit solvent allow the study of the most realistic and detailed level of physical chemistry of solvation [44]. We found that only three candidates (Can125, Can140, and Can159) remained bound to the THQ binding site during the entire MD simulation. In contrast, the Can142 left the THQ binding site at an early stage of the MD simulations; this observation might correlate with the fact that this candidate also showed higher docking scores for other pockets in CD44HAbd than the THQ-binding pocket (Figure S5).

Moreover, the candidates that remained bound during all simulations induced drastic decreases in the RMSD values of the hCD44HAbd backbone compared with those of the apo structure. The ligand-induced transition to a less flexible conformation of the protein can modulate its activation and improve both the compound's affinity and residence time [45,46]. Hence, a reduction in the target's conformational dynamics is a desirable characteristic of a drug-like molecule.

Although Can125, Can140, and Can159 displayed molecular interactions with residues involved in the HA-binding, including some reported as essential, only Can125 and Can159 reproduced the interactions predicted by the pharmacophore. Per-residue energetic decomposition corroborated that residues at the THQ-binding pocket support the binding of these two candidates to hCD44HAbd. However, only Can125 originated an RMSF profile that diverged from the unligated hCD44HAbd; specifically, it induced fluctuations in residues Arg41 and 120–126. The ligand-induced shift on Arg41 was different from the one identified on the crystal structures containing HA, which is considered essential for the transition from inactive to the active state in CD44 [22]. The changes in residues 120–126, which comprise a loop adjacent to the THQ-binding pocket, may participate in target constraint since they do not contribute to the ligand-binding energy. We hypothesize that the conformational changes induced by Can125, especially on Arg41, may impede the binding of the HA to hCD44HAbd.

On the other hand, Glu37 contributed negatively to the binding energy of Can125, which may be caused by the method employed for energy calculations. Calculations performed in implicit solvent offer a fast approach for binding energy assessment but are not yet well parameterized for complex problems that consider the presence of all solvent molecules [44]. Thus, our calculations may be underestimating the energetic contribution of water bridges and hydrogen bonds formed between the side chain of Glu37 and the methyl alcohol of Can125. Moreover, it is also possible that the proximity between opposite hydrogen bond acceptors (the carboxylate of Glu37 and the oxygen in the acetamide group of Can125) represents an unfavorable energetic contribution. This finding represents an opportunity to improve the chemical features of potential antagonists of CD44 to be proposed in future studies.

We did not assess the possible effect of the best candidates in the binding of other CD44 ligands, as aggrecan, osteopontin, collagen, or CD74 [3,47,48], given the lack of



corresponding structural data. However, we found that the THQ-binding site was able to allocate the best-ranked poses of Can125 and Can159 among all CD44HAbd pockets, suggesting a better affinity for this site over other regions of CD44HAbd. Thus, we speculate that these compounds may elicit competitive inhibition only for ligands with binding sites overlapping with that of HA, such as aggrecan [49]. On the other hand, we do not have evidence to propose that binding sites outside the CD44HAbd could be affected by the candidates, although the compounds restricted the conformational dynamics of CD44HAbd.

#### 4. Materials and Methods

##### 4.1. Sequence and Structural Alignments

Human and murine CD44 binding-domain (hCD44HAbd and mCD44HAbd, respectively) crystal structures were retrieved from the Protein Data Bank (PDB). All structures were aligned using UCSF Chimera 1.14 [50] and the A chain of the entry 1UUH as reference. RMSD was calculated for alpha-carbons, backbone, and all atoms. Sequence alignments were performed using the pairwise2 module of Biopython 1.78 [51]. For small-molecule atom-based alignment and comparison, a python tailored-made script was created employing the Cheminformatics Toolbox RDKit (<http://www.rdkit.org>, accessed date 25 February 2021) and the Pymol 2.4 API (PyMOL Molecular Graphics System, Schrödinger, LLC). The script is available at [https://github.com/AngelRuizMoreno/CD44\\_antagonist](https://github.com/AngelRuizMoreno/CD44_antagonist) (accessed date 25 February 2021).

##### 4.2. The 3D Pharmacophore Modeling

The 3D pharmacophoric model was generated by using 21 mCD44HAbd crystal structures containing small molecules with a THQ motif (Table S1) and the Pharmit server [52]. The most relevant molecular features were chosen by visual inspection, and their 3D coordinates in the mCD44HAbd pocket were set using the threshold of RMSD < 0.5 Å among THQ atoms.

##### 4.3. Combinatorial Computational Chemistry

The creation of compound libraries by the CCC approach was carried out using Reactor 20.17 from ChemAxon (<http://www.chemaxon.com>, accessed date 25 February 2021). The Ugi Tetrazole and Ugi 3 component reactions were selected as synthesis routes to generate compounds synthetically accessible by MCR. For the CCC experiments, libraries of building blocks were made by searching highly diverse and low-cost commercially available starting materials using the sci-finder platform (<https://scifinder.cas.org>). The building blocks library consisted of 32 substituted 1,2,3,4-THQ, 4 aldehydes, and 657 isocyanides.

##### 4.4. Cheminformatic Analysis

The resulting compounds from CCC were stored in two different libraries according to their synthesis origin. For comparison, the DrugBank dataset was included in the cheminformatic analysis. For each library, we computed tSNE, NPR, PCA, and Ro5. The tSNE analysis was computed using the Tanimoto similarity among the MACCS keys for each molecule using RDKit; then, the tSNE calculation was conducted using Scikit learn 0.23.1 [53], implementing two components and a perplexity value of 50. After tSNE analysis, a silhouette-based k-means clustering was performed using Scikit learn 0.23.1 [53]. Similarly, the PCA was performed using Scikit learn and employing ten selected non-redundant molecular descriptors of 30 different 2D and 3D molecular descriptors, which were computed using RDKit. The NPR analysis was carried out by calculating the NPR1 and NPR2 descriptors for the molecules. Finally, the Ro5 analysis was performed by calculating the molecular weight, LogP, number of hydrogen bond donors and acceptors, and TPSA.



#### 4.5. The 3D Pharmacophoric Matching

We generated 20 energetically favorable conformers for each compound within our working libraries by using the Mcnf module of Moloc [54]. All the conformers were aligned against the THQ-atoms and pharmacophore using the Pharmit server [52]. The best conformers were selected using an RMSD threshold of 0.5 Å against the pharmacophoric model descriptors. Finally, a local optimization was performed using the mCD44HAbd surface using Moloc [54,55], followed by visual inspection. The molecules that after local optimization kept the 3D pharmacophoric matching (RMSD < 0.5 Å) were selected for further studies.

#### 4.6. Molecular Docking

According to the identity and RMSD values, we found a very high similarity between all hCD44bd. Therefore, we choose the first hCD44bd crystal reported (1UUH). Regarding the mCD44HAbd our data also indicated a very high similarity among all available structures. Nonetheless, for docking experiments, we focused on the mCD44HAbd containing THQ-molecules, and we choose the crystal 5BZM, which contains a THQ-molecule displaying all molecular features according to our pharmacophoric model.

For the validation of the molecular docking protocol, the crystal structure of mCD44HAbd (5BZM) was used as the receptor and 21 co-crystallized THQ-containing molecules as ligands. For virtual screening, proteins hCD44HAbd (1UUH) and mCD44HAbd (5BZM) were employed as receptors, and the 163 molecules selected by pharmacophore filtering as ligands. Secondary dockings were carried out into additional pockets identified Fpocket v3.0 [56] in ligand-free hCD44HAbd (1UUH), HA-bound mCD44HAbd (2JCR), or THQ-containing molecule mCD44HAbd (5BZM).

Proteins were prepared by removing co-crystallized waters, solvent molecules, and adding charges and hydrogens using Chimera 1.14 [50]. Ligands were prepared by adding explicit hydrogens and tautomeric states at pH 7.4. and generating 3D coordinates with Standardized 19.20.0 (<http://www.chemaxon.com>, accessed date 25 February 2021). For virtual screening, docking was performed into hCD44HAbd and mCD44HAbd within a sphere with 8 Å of radius and center in the Thr27 or Thr31, respectively. Secondary dockings were carried out using as reference the coordinates of each additional pocket identified. For each ligand, 50 runs of the genetic algorithm were performed for the conformational search. Each pose was evaluated employing the PLP Chemscore scoring function established in the GOLD software from the Cambridge Crystallographic Data Center (CCDC) [57]. For each compound, the best 25 poses were saved for analysis. Finally, hierarchical clustering analysis of the poses was performed using Scipy 1.5.2 [58].

#### 4.7. Molecular Dynamics

The MD simulations were carried out using Gromacs 5.0.4 [59]. Selected candidates and hCD44HAbd (1UUH) were parameterized using the CGenFF and CHARMM36 force field, respectively, through the CHARMM-GUI (<http://www.charmm-gui.org/>, accessed date 25 February 2021) [60]. The systems were built by adding TIP3P water molecules to the ligand-hCD44HAbd complexes, neutralizing ions, and establishing periodic boundary conditions (PBC) by using the multicomponent assembler of the CHARMM-GUI. Before production, the systems were minimized and then equilibrated under an NVT assembly. During the production phase, an NPT assembly was performed at 310.15 K for 100 ns saving velocities, energy, and positions every ten ps. Water molecules displacement was computed by quantifying the number of water molecules within a 6 Å radius sphere covering the THQ-binding pocket, every frame of the simulation. Analysis of ligand-target interactions was carried out by a python tailored-made script ([https://github.com/AngelRuizMoreno/CD44\\_antagonist](https://github.com/AngelRuizMoreno/CD44_antagonist), accessed date 25 February 2021) implementing MDAnalysis [61] and PLIP [62].



#### 4.8. Interactions Free Energy Calculations

Full-length candidate-hCD44HAbd trajectories were employed for free energy calculations using the molecular mechanics energies combined with the Poisson–Boltzmann surface area continuum solvation (MM/PBSA) [63] by g\_mmpbsa v1.6 package [64]. Computation of the potential energy in vacuum, polar solvation energy, and non-polar solvation energy were performed to calculate the average binding energy. Per-residue energetic decomposing and maps were created to show the contribution of each residue to the binding energy.

## 5. Conclusions

Our experiments demonstrate that the compounds Can125 [3-hydroxy-N-(3-hydroxy benzyl)-2-(5-methyl-3,4-dihydroisoquinolin-2(1H)-yl)propenamide] and Can159 [2-(1-((1H-indol-5-yl)methyl)-1H-tetrazol-5-yl)-2-(7-amino-3,4-dihydroisoquinolin-2(1H)-yl)ethan-1-ol] bind with high theoretical affinity to the murine and human structures of CD44HAbd and stabilize the conformational dynamics of the protein. Therefore, those compounds may elicit a blocking effect on HA-binding.

**Supplementary Materials:** The following are available online: Table S1: mCD44HAbd crystal structures employed for 3D pharmacophoric modeling, Figure S1: Sequence and structural analysis of CD44HAbd crystals, Figure S2: Correlation matrix of the molecular descriptors from CCC-libraries, Figure S3: Validation of the docking protocol, Figure S4: Ligand efficiency from docking scores for the 163 unique candidates, Table S2: SMILES codes and formal names of the nine selected candidates, Figure S5: Comparison of theoretical binding of the best nine candidates among multiple pockets within CD44HAbd, Figure S6: Docking pose clustering for the best nine candidates, Figure S7: Pairwise backbone RMSD matrix along 100 ns of MD simulation and alpha-carbon RMSF analysis, Figure S8: Frequency analysis of molecular interactions, Figure S9: Energy calculations generated from MD simulations, Figure S10: Per-residue energy decomposition for the MD-simulated binding of candidates.

**Author Contributions:** Conceptualization, A.J.R.-M., M.A.V.-V. and A.D.; methodology, A.J.R.-M. and A.R.-R.; software, A.J.R.-M.; formal analysis, A.J.R.-M. and M.A.V.-V.; writing—original draft preparation, A.J.R.-M. and M.A.V.-V.; writing—review and editing, A.J.R.-M., M.A.V.-V., A.R.-R. and A.D.; funding acquisition, M.A.V.-V. and A.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by PAPIIT UNAM IN219719, PAPIIT UNAM IV200121, CONACYT A1-S-18285, LANCAD-UNAM-DGTIC-364 2018–2019 (M.A.V.-V.). Research in the lab of A.D. was supported by an ITN “AcceleratedEarly stage drug dIScovery” (AEGIS, grant agreement No. 675555) and COFUND ALERT (grant agreement No. 665250), and a KWF Kankerbestrijding grant (grant agreement No. 10504).

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data, including the in silico libraries, are available from the authors upon request.

**Acknowledgments:** A.J.R.-M. is a student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received fellowship 584534 from CONACYT. We thank Alfredo Torres-Larios for his helpful comments and for reviewing the final version of the manuscript and Diego Prada-Gracia for the advisory for strategies and analysis workflows.

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

**Sample Availability:** Not applicable.



## References

- Ponta, H.; Sherman, L.; Herrlich, P.A. CD44: From adhesion molecules to signalling regulators. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 33–45. [\[CrossRef\]](#)
- Zöller, M. CD44: Can a cancer-initiating cell profit from an abundantly expressed molecule? *Nat. Rev. Cancer* **2011**, *11*, 254–267. [\[CrossRef\]](#)
- Senbanjo, L.T.; Chellaiah, M.A. CD44: A multifunctional cell surface adhesion receptor is a regulator of progression and metastasis of cancer cells. *Front. Cell Dev. Biol.* **2017**, *5*. [\[CrossRef\]](#)
- Wu, K.; Xu, H.; Yuan, X.; Tian, Y.; Liu, Y.; Liu, Q.; Wu, H.; Wu, G.S. Enrichment of CD44 in basal-type breast cancer correlates with EMT, cancer stem cell gene profile, and prognosis. *OTT* **2016**, *9*, 431–444. [\[CrossRef\]](#)
- Si, D.; Yin, F.; Peng, J.; Zhang, G. High Expression of CD44 Predicts a Poor Prognosis in Glioblastomas. *Cancer Manag. Res.* **2020**, *12*, 769–775. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wu, G.; Song, X.; Liu, J.; Li, S.; Gao, W.; Qiu, M.; Yang, C.; Ma, Y.; Chen, Y. Expression of CD44 and the survival in glioma: A meta-analysis. *Biosci. Rep.* **2020**, *40*, BSR20200520. [\[CrossRef\]](#)
- Chen, J.; Zhou, J.; Lu, J.; Xiong, H.; Shi, X.; Gong, L. Significance of CD44 expression in head and neck cancer: A systemic review and meta-analysis. *BMC Cancer* **2014**, *14*, 15. [\[CrossRef\]](#) [\[PubMed\]](#)
- Papadaki, C.; Manolakou, S.; Lagoudaki, E.; Pontikakis, S.; Ierodiakonou, D.; Vogiatzoglou, K.; Messaritakis, I.; Trypaki, M.; Giannikaki, L.; Sfakianaki, M.; et al. Correlation of PKM2 and CD44 protein expression with poor prognosis in platinum-treated epithelial ovarian cancer: A retrospective study. *Cancers (Basel)* **2020**, *12*, 1013. [\[CrossRef\]](#)
- Bourguignon, L.Y.W.; Spevak, C.C.; Wong, G.; Xia, W.; Gilad, E. Hyaluronan-CD44 interaction with protein kinase C $\epsilon$  promotes oncogenic signaling by the stem cell marker nanog and the production of microRNA-21, leading to down-regulation of the tumor suppressor protein PDCD4, anti-apoptosis, and chemotherapy resistance. *J. Biol. Chem.* **2009**, *284*, 26533–26546. [\[CrossRef\]](#)
- Bourguignon, L.Y.W. Matrix Hyaluronan-CD44 Interaction Activates MicroRNA and LncRNA Signaling Associated With Chemoresistance, Invasion, and Tumor Progression. *Front. Oncol.* **2019**, *9*, 492. [\[CrossRef\]](#)
- Bourguignon, L.Y.W.; Wong, G.; Shiina, M. Up-regulation of histone methyltransferase, DOT1L, by matrix hyaluronan promotes microRNA-10 expression leading to tumor cell invasion and chemoresistance in cancer stem cells from head and neck squamous cell carcinoma. *J. Biol. Chem.* **2016**, *291*, 10571–10585. [\[CrossRef\]](#)
- Hill, A.; McFarlane, S.; Mulligan, K.; Gillespie, H.; Draffin, J.E.; Trimble, A.; Ouhtit, A.; Johnston, P.G.; Harkin, D.P.; McCormick, D.; et al. Cortactin underpins CD44-promoted invasion and adhesion of breast cancer cells to bone marrow endothelial cells. *Oncogene* **2006**, *25*, 6079–6091. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cieply, B.; Koontz, C.; Frisch, S.M. CD44S-hyaluronan interactions protect cells resulting from EMT against anoikis. *Matrix Biol.* **2015**, *48*, 55–65. [\[CrossRef\]](#)
- Gudbergsson, J.M.; Christensen, E.; Kostikov, S.; Moos, T.; Duroux, M.; Kjær, A.; Johnsen, K.B.; Andresen, T.L. Conventional Treatment of Glioblastoma Reveals Persistent CD44+ Subpopulations. *Mol. Neurobiol.* **2020**, *57*, 3943–3955. [\[CrossRef\]](#)
- Yu, F.; Yao, H.; Zhu, P.; Zhang, X.; Pan, Q.; Gong, C.; Huang, Y.; Hu, X.; Su, F.; Lieberman, J.; et al. Let-7 regulates self renewal and tumorigenicity of breast cancer cells. *Cell* **2007**, *131*, 1109–1123. [\[CrossRef\]](#)
- Hong, S.P.; Wen, J.; Bang, S.; Park, S.; Song, S.Y. CD44-positive cells are responsible for gemcitabine resistance in pancreatic cancer cells. *Int. J. Cancer* **2009**, *125*, 2323–2331. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dylla, S.J.; Beviglia, L.; Park, I.-K.; Chartier, C.; Raval, J.; Ngan, L.; Pickell, K.; Aguilar, J.; Lazetic, S.; Smith-Berdan, S.; et al. Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy. *PLoS ONE* **2008**, *3*, e2428. [\[CrossRef\]](#)
- Wang, L.; Huang, X.; Zheng, X.; Wang, X.; Li, S.; Zhang, L.; Yang, Z.; Xia, Z. Enrichment of prostate cancer stem-like cells from human prostate cancer cell lines by culture in serum-free medium and chemoradiotherapy. *Int. J. Biol. Sci.* **2013**, *9*, 472–479. [\[CrossRef\]](#)
- Hiraga, T.; Ito, S.; Nakamura, H. Cancer stem-like cell marker CD44 promotes bone metastases by enhancing tumorigenicity, cell motility, and hyaluronan production. *Cancer Res.* **2013**, *73*, 4112–4122. [\[CrossRef\]](#)
- Su, Y.-J.; Lai, H.-M.; Chang, Y.-W.; Chen, G.-Y.; Lee, J.-L. Direct reprogramming of stem cell properties in colon cancer cells by CD44. *EMBO J.* **2011**, *30*, 3186–3199. [\[CrossRef\]](#)
- Gao, Y.; Foster, R.; Yang, X.; Feng, Y.; Shen, J.K.; Mankin, H.J.; Hornicek, F.J.; Amiji, M.M.; Duan, Z. Up-regulation of CD44 in the development of metastasis, recurrence and drug resistance of ovarian cancer. *Oncotarget* **2015**, *6*, 9313–9326. [\[CrossRef\]](#)
- Banerji, S.; Wright, A.J.; Noble, M.; Mahoney, D.J.; Campbell, I.D.; Day, A.J.; Jackson, D.G. Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction. *Nat. Struct. Mol. Biol.* **2007**, *14*, 234–239. [\[CrossRef\]](#)
- Bajorath, J.; Greenfield, B.; Munro, S.B.; Day, A.J.; Aruffo, A. Identification of CD44 residues important for hyaluronan binding and delineation of the binding site. *J. Biol. Chem.* **1998**, *273*, 338–343. [\[CrossRef\]](#)
- Liu, L.K.; Finzel, B.C. Fragment-based identification of an inducible binding site on cell surface receptor CD44 for the design of protein-carbohydrate interaction inhibitors. *J. Med. Chem.* **2014**, *57*, 2714–2725. [\[CrossRef\]](#)
- Baggio, C.; Barile, E.; Di Sorbo, G.; Kipps, T.J.; Pellecchia, M. The cell surface receptor CD44: NMR-based characterization of putative ligands. *ChemMedChem* **2016**, *11*, 1097–1106. [\[CrossRef\]](#)
- Aguirre-Alvarado, C.; Segura-Cabrera, A.; Velázquez-Quesada, I.; Hernández-Esquivel, M.A.; García-Pérez, C.A.; Guerrero-Rodríguez, S.L.; Ruiz-Moreno, A.J.; Rodríguez-Moreno, A.; Pérez-Tapia, S.M.; Velasco-Velázquez, M.A. Virtual screening-driven repositioning of etoposide as CD44 antagonist in breast cancer cells. *Oncotarget* **2016**, *7*, 23772–23784. [\[CrossRef\]](#)



27. Pustuła, M.; Czub, M.; Łabuzek, B.; Surmiak, E.; Tomala, M.; Twarda-Clapa, A.; Guzik, K.; Popowicz, G.M.; Holak, T.A. NMR fragment-based screening for development of the CD44-binding small molecules. *Bioorg Chem.* **2019**, *82*, 284–289. [\[CrossRef\]](#)
28. Zhang, J.; Patil, P.; Kurpiewskab, K.; Kalinowska, T.; Iuścikb, J.; Dömling, A. Hydrazine in the Ugi Tetrazole Reaction Synthesis. *Synthesis (Stuttg)* **2016**, *48*, 1122–1130.
29. Neochoritis, C.G.; Zhao, T.; Dömling, A. Tetrazoles via multicomponent reactions. *Chem. Rev.* **2019**, *119*, 1970–2042. [\[CrossRef\]](#)
30. Tripolitsiotis, N.P.; Thomaidi, M.; Neochoritis, C.G. The Ugi three-component reaction; a valuable tool in modern organic synthesis. *Eur. J. Org. Chem.* **2020**, *2020*, 6525–6554. [\[CrossRef\]](#)
31. Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **2006**, *34*, D668–D672. [\[CrossRef\]](#)
32. Polton, D.J. Installation and operational experiences with MACCS (Molecular Access System). *Online Rev.* **1982**, *6*, 8. [\[CrossRef\]](#)
33. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26. [\[CrossRef\]](#)
34. Liu, L.K.; Finzel, B. High-resolution crystal structures of alternate forms of the human CD44 hyaluronan-binding domain reveal a site for protein interaction. *Acta Crystallogr. Sect. Struct. Biol. Commun.* **2014**, *70*, 1155–1161. [\[CrossRef\]](#)
35. Zarganes-Tzitzikas, T.; Chandgude, A.L.; Dömling, A. Multicomponent Reactions, Union of MCRs and beyond. *Chem. Rec.* **2015**, *15*, 981–996. [\[CrossRef\]](#)
36. Dömling, A.; Wang, W.; Wang, K. Chemistry and biology of multicomponent reactions. *Chem. Rev.* **2012**, *112*, 3083–3135. [\[CrossRef\]](#)
37. Awale, M.; Reymond, J.-L. Web-based 3D-visualization of the DrugBank chemical space. *J. Cheminform.* **2016**, *8*, 25. [\[CrossRef\]](#)
38. Doak, B.C.; Zheng, J.; Dobritsch, D.; Kihlberg, J. How beyond rule of 5 drugs and clinical candidates bind to their targets. *J. Med. Chem.* **2016**, *59*, 2312–2327. [\[CrossRef\]](#)
39. Li, J.; Di Lorenzo, V.; Patil, P.; Ruiz-Moreno, A.J.; Kurpiewska, K.; Kalinowska-Tłuścik, J.; Velasco-Velázquez, M.A.; Dömling, A. Scaffolding-Induced Property Modulation of Chemical Space. *ACS Comb. Sci.* **2020**, *22*, 356–360. [\[CrossRef\]](#)
40. Bowman, W.C. Neuromuscular block. *Br. J. Pharmacol.* **2006**, *147* (Suppl 1), S277–S286. [\[CrossRef\]](#)
41. Doroshenko, O.; Fuhr, U. Clinical pharmacokinetics and pharmacodynamics of solifenacin. *Clin. Pharm.* **2009**, *48*, 281–302. [\[CrossRef\]](#)
42. Asmar, R.; Sayegh, F.; Tracz, W.; Hlawaty, M.; Olszowska, M.; Jourde, M.; Vincent, M.; Goujoun, B.; Maldonado, J. Reversal of left ventricular hypertrophy with the ACE inhibitor moexipril in patients with essential hypertension. *Acta Cardiol.* **2002**, *57*, 31–32.
43. De Luca, L.; Gitto, R.; Barreca, M.L.; Caruso, R.; Quartarone, S.; Citraro, R.; De Sarro, G.; Chimirri, A. 3D pharmacophore models for 1,2,3,4-tetrahydroisoquinoline derivatives acting as anticonvulsant agents. *Arch. Pharm. Chem. Life Sci.* **2006**, *339*, 388–400. [\[CrossRef\]](#)
44. Zhang, L.Y.; Gallicchio, E.; Friesner, R.A.; Levy, R.M. Solvent models for protein–ligand binding: Comparison of implicit solvent poisson and surface generalized born models with explicit solvent simulations. *J. Comput. Chem.* **2001**, *22*, 591–607. [\[CrossRef\]](#)
45. Amaral, M.; Kokh, D.B.; Bomke, J.; Wegener, A.; Buchstaller, H.P.; Eggenweiler, H.M.; Matias, P.; Sirrenberg, C.; Wade, R.C.; Frech, M. Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat. Commun.* **2017**, *8*, 2276. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Ferina, J.; Daggett, V. Visualizing protein folding and unfolding. *J. Mol. Biol.* **2019**, *431*, 1540–1564. [\[CrossRef\]](#)
47. Naor, D.; Sionov, R.V.; Ish-Shalom, D. CD44: Structure, function and association with the malignant process. In *Advances in Cancer Research*; Vande Woude, G.F., Klein, G., Eds.; Academic Press: Cambridge, MA, USA, 1997; Volume 71, pp. 241–319.
48. Shi, X.; Leng, L.; Wang, T.; Wang, W.; Du, X.; Li, J.; McDonald, C.; Chen, Z.; Murphy, J.W.; Lolis, E.; et al. CD44 is the signaling component of the macrophage migration inhibitory factor-CD74 receptor complex. *Immunity* **2006**, *25*, 595–606. [\[CrossRef\]](#)
49. Fujimoto, T.; Kawashima, H.; Tanaka, T.; Hirose, M.; Toyama-Sorimachi, N.; Matsuzawa, Y.; Miyasaka, M. CD44 binds a chondroitin sulfate proteoglycan, aggrecan. *Int. Immunol.* **2001**, *13*, 359–366. [\[CrossRef\]](#)
50. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [\[CrossRef\]](#)
51. Cock, P.J.A.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423. [\[CrossRef\]](#)
52. Sunseri, J.; Koes, D.R. Pharmit: Interactive exploration of chemical space. *Nucleic Acids Res.* **2016**, *44*, W442–W448. [\[CrossRef\]](#)
53. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
54. Gerber, P.R.; Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput. Aided. Mol. Des.* **1995**, *9*, 251–268. [\[CrossRef\]](#)
55. Gerber, P.R. Topological pharmacophore description of chemical structures using MAB-Force-Field-Derived data and corresponding similarity measures. In *Fundamentals of Molecular Similarity*; Carbó-Dorca, R., Gironés, X., Mezey, P.G., Eds.; Mathematical and Computational Chemistry; Springer US: Boston, MA, USA, 2001; pp. 67–81. ISBN 978-1-4757-3273-3.
56. Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinform.* **2009**, *10*, 168. [\[CrossRef\]](#)



- 
57. Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748. [[CrossRef](#)]
  58. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)]
  59. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [[CrossRef](#)]
  60. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [[CrossRef](#)]
  61. Gowers, R.J.; Linke, M.; Barnoud, J.; Reddy, T.J.E.; Melo, M.N.; Seyler, S.L.; Domański, J.; Dotson, D.L.; Buchoux, S.; Kenney, I.M.; et al. MDAnalysis: A python package for the rapid analysis of molecular dynamics simulations. In Proceedings of the 15th Python in Science Conference (SciPy 2016), Austin, TX, USA, 11–17 July 2016; pp. 98–105.
  62. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447. [[CrossRef](#)]
  63. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [[CrossRef](#)]
  64. Kumari, R.; Kumar, R.; Lynn, A. g\_mmpbsa—A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962. [[CrossRef](#)] [[PubMed](#)]