



OPEN

Identification of non-classical hCA XII inhibitors using combination of computational approaches for drug design and discovery

Mohammad M. Al-Sanea^{1,9}, Garri Chilingaryan^{2,3,9}✉, Narek Abelyan^{3,4,9}, Grigor Arakelov², Harutyun Sahakyan^{2,4}, Vahram G. Arakelov², Karen Nazaryan², Shaimaa Hussein⁵, Gharam M. Alazmi¹, Haifa E. Alsharari¹, Waad M. Al-faraj¹, Faten S. Alruwaili¹, Nouf O. Albilasi¹, Tahani S. Alsharari¹, Abdulaziz A. S. Alsaleh¹, Turki M. Alazmi¹, Atiah H. Almalki^{7,8}, Nasser H. Alotaibi⁶ & Mohamed A. Abdelgawad¹

Human carbonic anhydrase XII (hCA XII) isozyme is of high therapeutic value as a pharmacological target and biomarker for different types of cancer. The hCA XII is one of the crucial effectors that regulates extracellular and intracellular pH and affects cancer cell proliferation, invasion, growth and metastasis. Despite the fact that interaction features of hCAs inhibitors with the catalytic site of the enzyme are well described, lack in the selectivity of the traditional hCA inhibitors based on the sulfonamide group or related motifs is an urgent issue. Moreover, drugs containing sulfanomides can cause sulfa allergies. Thus, identification of novel non-classical inhibitors of hCA XII is of high priority and is currently the subject of a vast field of study. This study was devoted to the identification of novel potential hCA XII inhibitors using comprehensive set of computational approaches for drug design discovery: generation and validation of structure- and ligand-based pharmacophore models, molecular docking, re-scoring of virtual screening results with MMGBSA, molecular dynamics simulations, etc. As the results of the study several compounds with alternative to classical inhibitors chemical scaffolds, in particular one of coumarin derivative, have been identified and are of high interest as potential non-classical hCA XII inhibitors.

Carbonic anhydrases (CAs, EC 4.2.1.1) are zinc metalloenzymes that catalyze the reversible hydration of carbon dioxide into bicarbonate and a proton¹. Sixteen CA isoforms are found in humans and all vary in kinetic properties, subcellular localization and distribution to the tissues. Human carbonic anhydrase XII (hCA XII) is induced by hypoxia and highly expressed within the hypoxic core of many solid tumor types². The hCA XII regulates extracellular and intracellular pH homeostasis of the cancer cells, thus, mediating cancer cells invasion, proliferation, metastasis, progression and tumor growth^{3–5}.

In general, CA inhibitors (CAIs) are compounds equipped with an effective zinc-binding group (ZBG) capable of chelating the prosthetic zinc ion placed inside the hCA binding site that is essential for these enzymes' catalytic action. Sulfonamide moieties or related structural motifs represent the most common ZBGs shared by typical CAIs (such as sulfamides and sulfamates) and are the most significant groups of CAIs with multiple ligands with high inhibitory potency have been recorded to date⁶. These groups are especially effective in endowing high-affinity small-molecule ligands with hCAs, as they enable not only the proper coordination of the catalytic zinc ion, but also the forming of associations of H-bonds with key protein residues located in the zinc-binding cavity

¹Pharmaceutical Chemistry Department, College of Pharmacy, Jouf University, Sakaka 72341, Aljouf, Saudi Arabia. ²Institute of Molecular Biology of NAS RA, 0014 Yerevan, Armenia. ³Institute of Biomedicine and Pharmacy, Russian-Armenian University, 0051 Yerevan, Armenia. ⁴Foundation for Armenian Science and Technology, 0033 Yerevan, Armenia. ⁵Department of Pharmacology, College of Pharmacy, Jouf University, Sakaka 72341, Aljouf, Saudi Arabia. ⁶Department of Clinical Pharmacy, College of Pharmacy, Jouf University, Sakaka 72341, Aljouf, Saudi Arabia. ⁷Department of Pharmaceutical Chemistry, College of Pharmacy, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia. ⁸Addiction and Neuroscience Research Unit, Health Science Campus, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia. ⁹These authors contributed equally: Mohammad M. Al-Sanea, Garri Chilingaryan and Narek Abelyan. ✉email: g_chilingaryan@mb.sci.am

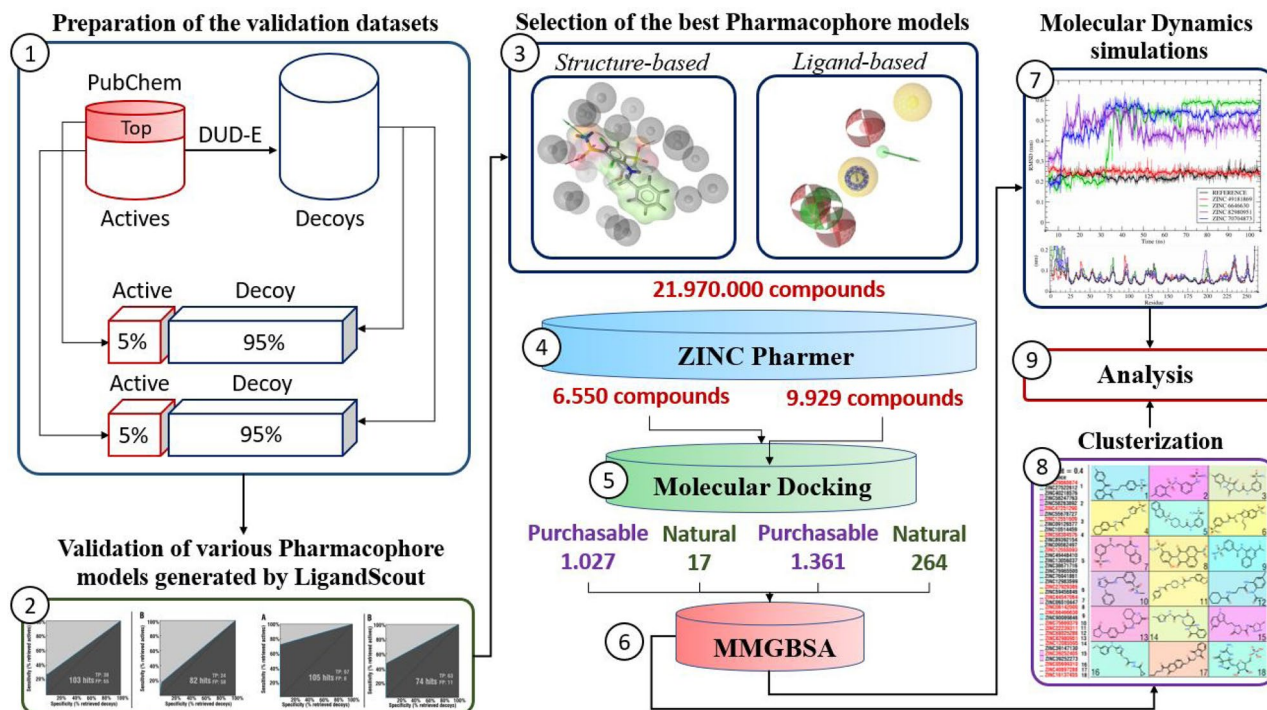


Figure 1. Schematic representation of general methodology and approaches applied in the study. Schematic illustration was drawn using Adobe Illustrator 2018 (www.adobe.com/products/illustrator).

area. A great amount of effort has been put in drug design and discovery of hCA XII potential inhibitors^{7–15}. Currently, SLC-011, one of the most promising compounds that is also based on the ureido-substituted benzene sulphonamide core (USB) has reached phase Ib/II of clinical trials as hCA IX/XII inhibitor^{16,17}. However, because of the high amino acid conservation found at the level of their catalytic site and adjacent regions in the various hCA isoforms, most of these ligands are insufficiently selective against particular hCAs, including hCA XII which is currently the subject of a vast field of study¹⁸. In addition, a small but significant percentage of the general population cannot be treated with sulfonamide-based compounds due to a sulfa allergy^{19,20}. Therefore, novel CAIs should be not only isoform specific, but also non-classical, i.e. not based on sulfonamides, sulfamates, or sulfamides²⁰. Non-classical CAIs are an especially useful resource in this sense for finding isoform specificity and avoiding possible off-target events, side effects and adverse reactions (such as sulfur allergies) associated with the use of sulfurized ligands²⁰. Notwithstanding, the development of hCA IX/XII selectivity inhibitors over hCA I/II, which are ubiquitously distributed and involved in key physiological processes, is still a difficult challenge, although some examples of selective ligands have been reported^{21,22}.

Computer-aided drug design (CADD) comprises a broad range of theoretical and computational approaches that are part of modern drug discovery^{23,24}. CADD methods have made key contributions to the development of drugs that are in clinical use or in clinical trials^{25,26}. Such methods have emerged and evolved along with experimental approaches used in drug design. This study was devoted to identification of novel potential non-classical hCA XII inhibitors using the combination of computational approaches, including pharmacophore modeling (structure-, and ligand-based), molecular docking, MMGBSA re-scoring and molecular dynamics simulations.

Materials and methods

Combination of the several computer-aided drug design and discovery methods, tools and approaches were applied in this study (Fig. 1). Structure- and ligand-based pharmacophores modelling was used for initial filtration of ZINC database of chemical compounds. Molecular docking and additional re-scoring with MMGBSA were used for further filtration and identification of potential hit compounds. Molecular dynamics simulations were used for assessment of interaction stability between hCA XII isozyme and top identified chemical compounds. Additionally, chemical scaffolds of identified compounds were analyzed and compared to the reference ligand (classical hCA inhibitor).

Structure-based pharmacophore generation. Structure-based pharmacophore model was generated using LigandScout v4.4²⁷. 22 crystal structures of the hCA XII in complexes with inhibitors (X-ray resolution <2 Å) were downloaded from Protein Data Bank (PDB)²⁸. Several approaches for the structure-based pharmacophore model generation were tested: (1) based on the separate crystal structures, (2) based on the combination of different crystal structures, (3) merged and (4) shared pharmacophores. Best pharmacophore model was generated using crystal structure of the catalytic domain of human carbonic anhydrase isozyme XII with benzene sulfonamide derivative (PDB ID: 4QJW, resolution 1.55 Å, chain C). All features that characterize

interactions with the water molecules and hydrophobic interactions were removed from the obtained pharmacophore model. Additionally, exclusion volumes coat feature of LigandScout v4.4 was applied.

Ligand-based pharmacophore generation. Ligand-based pharmacophore model was also generated using LigandScout v4.4²⁷. 135 compounds active against hCA XII enzyme were downloaded from PubChem²⁹ BioAssay dataset. These 135 active compounds were clustered using pharmacophore alignment score as a similarity measure and the average method for cluster distance calculations. Maximum number of the conformations and cluster distance values were set to 3 and 0.4, respectively. As a result, 14 clusters were identified and representative compounds of these clusters were used for the creation of ligand-based pharmacophore models (Table S1). The final best ligand-based pharmacophore model was obtained based on the biggest cluster. “H bond donor” feature was removed from the selected ligand-based pharmacophore model.

Validation of generated structure- and ligand-based pharmacophore models. 2500 compounds with $K_i < 10$ against hCA XII enzyme were downloaded from PubChem BioAssay and ChEMBL³⁰ databases. 5000 decoy compounds were generated using a Database of Useful Decoys: Enhanced (DUD-E)³¹ and were not included in the validation datasets in order to avoid bias. 3D structures of decoys were generated by Open Babel v3.1.1³² software based on the SMILES obtained by using DUD-E. Two separate validation datasets were prepared and each of them consisted of 5% active compounds and 95% decoys. First validation dataset included top 135 active compounds based on the K_i values. Second dataset included 135 randomly selected active compounds. Hydrogens were added to all compounds using Open Babel v3.1.1 software. The iConBest method of LigandScout v4.4 was used for the generation of the conformations for active and decoy compounds.

For the validation of the generated pharmacophore models several widely established metrics were used. The area value under the receiver operating characteristic (ROC AUC) ranges from 0 to 1, where 1 is perfect classification, while values below 0.5 indicate random classification. ROC is widely used to evaluate virtual screening and pharmacophore modeling methods^{33,34} and defined as a graphical representation of the test sensitivity in relation to its specificity or false-positive rate. The AUC is the probability of active compounds being ranked earlier than decoy compounds. The classifier “precision” represents the share of true positives (TP) among all hits ($TP/(TP + FP)$, where FP is false positive compounds). The classifier “specificity” represents the ratio of the active compounds found in the list of compounds identified as “true positives” (TP/A , where A is active compounds). The “Specificity” classifier is the ratio of true negative compounds to all compounds in the database, excluding active ones. Enrichment Factor (EF) measures the fraction of active compounds found in a specific percentage, solving the problem of comparing the results for datasets with different active/inactive compound ratios³⁵. The EF for 1, 5, 10 and 100% was calculated for the share of true positives among the molecules identified as hit compounds using generated pharmacophore models.

Virtual screening. ZINC “purchasable” (21,777,093 compounds) and ZINC “natural and derivatives” (197,488 compounds) datasets were used for the selection of library of compounds. ZINCPharmer³⁶ was used for the filtration of these datasets using generated structure-based and ligand-based pharmacophore models. Identified hit compounds were used for the molecular docking against the active site of the hCA XII enzyme (PDB ID 4QJW). AutoDock Tools³⁷ was used for the estimation of the grid box ($< 27 \text{ \AA}$), calculation of the protein's and compounds' charges and the addition of the polar hydrogens. Virtual screening was performed using AutoDock Vina software³⁸, which has been regarded as highly efficient software for molecular docking and virtual screening procedures based on the recent benchmark studies among both academic and commercial software^{39,40}. AutoDock Vina uses “Iterated Local Search global optimizer” similar to that by Abagyan et al. and Broyden-Fletcher-Goldfarb-Shanno (BFGS), which is quasi-Newton method for the local optimization, as a search algorithm and hybrid scoring function (empirical + knowledge-based function) inspired in the X-Score function³⁸. Standard, recommended by the developers, parameters were used for the virtual screening procedure.

Re-scoring using MMGBSA method. The algorithm for the molecular mechanics-generalized Born surface area (MMGBSA) calculations is based on a freely-available AmberTools suite. The algorithm can be described in three stages, (1) receptor and ligands parametrization, (2) minimization and (3) MM/GBSA and MM/PBSA calculations. At the first stage, the ff14SB force field⁴¹ is used to describe protein parameters, and General Amber Force Field (GAFF)⁴² with AM1-BCC charge model⁴³ is used for small molecule parametrization. Next, the algorithm prepares necessary input files (coordinates and topologies) with mbondi3 radii using LEaP. Minimization is performed in generalized Born implicit solvent models ($igb = 8$) using the sander engine. Finally, for the free energy calculations, the algorithm uses the MMPBSA.py program⁴⁴ for the MMGBSA calculations. The algorithm⁴⁵ is implemented in a bash script and can be run in parallel in most Linux distributions without any additional libraries for parallelization. The full code is available at the following link: <https://github.com/sahakyankh/iPBSA>.

Molecular dynamics simulations. The molecular dynamics simulations were carried out using AMBER20⁴⁶ molecular dynamics package. The ff14SB force field was used for protein parametrization and GAFF for the ligand parameterization with AM1-BCC charge model. Minimized conformations of complexes of hCA XII with bound compounds were taken from the previous step (re-scoring of docked complexes with MMGBSA) and used as starting positions for corresponding simulations. The complexes were solvated in TIP3P water model and Na^+/Cl^- ions at 150 mM concentration⁴⁷. The Monte Carlo barostat⁴⁸ with reference pressure at 1 bar and Langevin thermostat⁴⁹ with collision frequency (γ_{ln}) 2 ps^{-1} were used to keep the temperature at 310.15 K. The Particle Mesh Ewald (PME) method with 1.0 nm cutoff was used for the long-range electro-

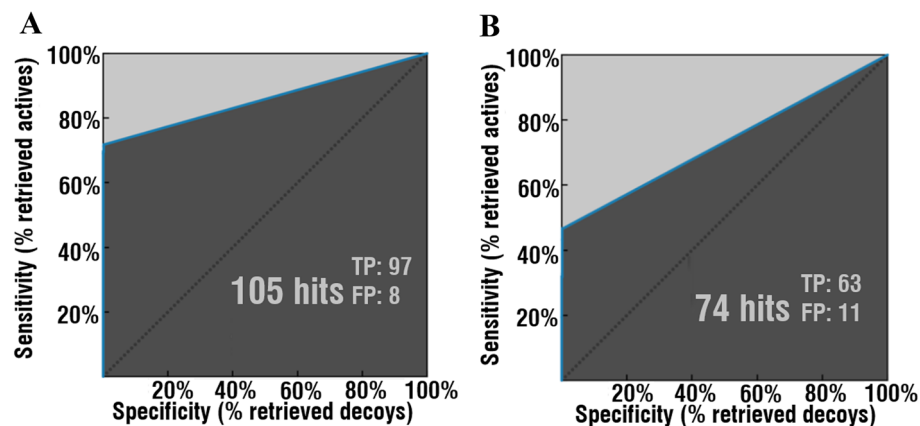


Figure 2. Validation of the generated ligand-based pharmacophore on the two datasets (A,B) Identified hit compounds out of 2635 total compounds (135 actives, 2500 decoys). Figure was obtained using LigandScout v4.4²⁷ (www.inteligand.com/ligandscout/). Schematic illustration was drawn using Adobe Illustrator 2018 (www.adobe.com/products/illustrator).

static interactions. Bonds involving hydrogen were constrained using the SHAKE algorithm with 2 fs integration step⁵⁰. Each simulation consisted of 5 ns of system minimization and equilibration and 100 ns of conventional molecular dynamics simulation. Finally, for every simulation, binding free energies were re-calculated using the same MMGBSA method and MMPBSA.py program, using 250 snapshots with equal intervals collected from the last 20 ns of simulation. RMSD and SASA was calculated as indicators of stability of studied complexes during simulation. Besides, RMSF analysis was performed to measure the average atomic flexibility of the Ca atoms of the docked complexes. Radius of gyration was calculated as an indicator of protein structure compactness during simulation. Hydrogen bonding analysis was performed to identify similarities and differences in interaction patterns between studied compounds and amino acid residues of hCA XII binding site.

Results and discussion

Ligand-based pharmacophore generation. As a result of the clusterization of 135 compounds active against hCA XII, 14 clusters have been obtained. The best ligand-based pharmacophore model was obtained based on the biggest cluster (Supplementary Figure S1). The final ligand-based pharmacophore model included three “H-bond (acceptor)” features, two “H-bond (donor)” features, one “hydrophobic” and “aromatic” features of the LigandScout v4.4. Validation tests of the generated ligand-based pharmacophore model demonstrated high results on two separate datasets. First dataset (Fig. 2A): 105 hits (97 TP, 8 FP), AUC 1, 5, 10, 100–1, 1, 1, 0.86; EF 1, 5, 10, 100–19.8, 18.3, 18.3, 18.3. Second dataset (Fig. 2B): 74 hit compounds (63 TP, 11 FP) AUC 1, 5, 10, 100–1, 1, 1, 0.73; EF 1, 5, 10, 100–19.8, 16.9, 16.9, 16.9.

Structure-based pharmacophore generation. Structure-based pharmacophore model generated based on the crystal structure of catalytic domain of hCA isozyme XII with benzenesulfonamide derivative (PDB ID: 4QJW) demonstrated LigandScout’s binding affinity score of – 17.18, which includes interaction and desolvation energies. The selected structure-based pharmacophore model included two “H-bond (acceptor)” features that are important for the interaction of the ligand with ASN64 and THR198 of the active site of hCA XII enzyme, and two “H-bond (donor)” feature that are important for interaction with PRO200 and GLU104 (Fig. 3).

Validation of structure-based pharmacophore model in the case of the first dataset (Fig. 4A) demonstrated following values: 103 hit compounds (38 TP, 65 FP), AUC 1, 5, 10, 100–0.72, 0.93, 0.95, 0.63; EF 1, 5, 10, 100–7.6, 7.3, 7.3, 7.3. In the case of the second validation dataset (Fig. 4B): 82 hit compounds (24 TP, 58 FP), AUC 1, 5, 10, 100–0.97, 0.92, 0.93, 0.58; EF 1, 5, 10, 100–3.8, 5.8, 5.8, 5.8 (Fig. 4).

Filtration of ZINC database using generated pharmacophore models. Both selected ligand-based and structure-based pharmacophore models were applied to filter ZINC “purchasable” and “natural and derivatives” compounds datasets (total of 21.970.000 compounds) using ZINCPharmer, separately. 6357 and 193 hit compounds were identified from “purchasable” and “natural and derivatives” datasets, respectively, using a structure-based pharmacophore model. Other 8415 and 1514 hit compounds were identified using a ligand-based pharmacophore model. All identified hit compounds were used for molecular docking against the active site of the hCA XII enzyme.

Molecular docking. 1361 “purchasable” and 264 “natural and derivatives” compounds with similar or higher docking scores compared to the reference molecule (co-crystallized ligand of 4QJW, benzenesulfonamide derivative: – 8 kcal/mol) were identified as the result of molecular docking of hit compounds obtained by application of the ligand-based pharmacophore (Table S2). In case of hit compounds identified by structure-based pharmacophore, 1027 “purchasable” and 17 “natural and derivatives” compounds were identified as a result of

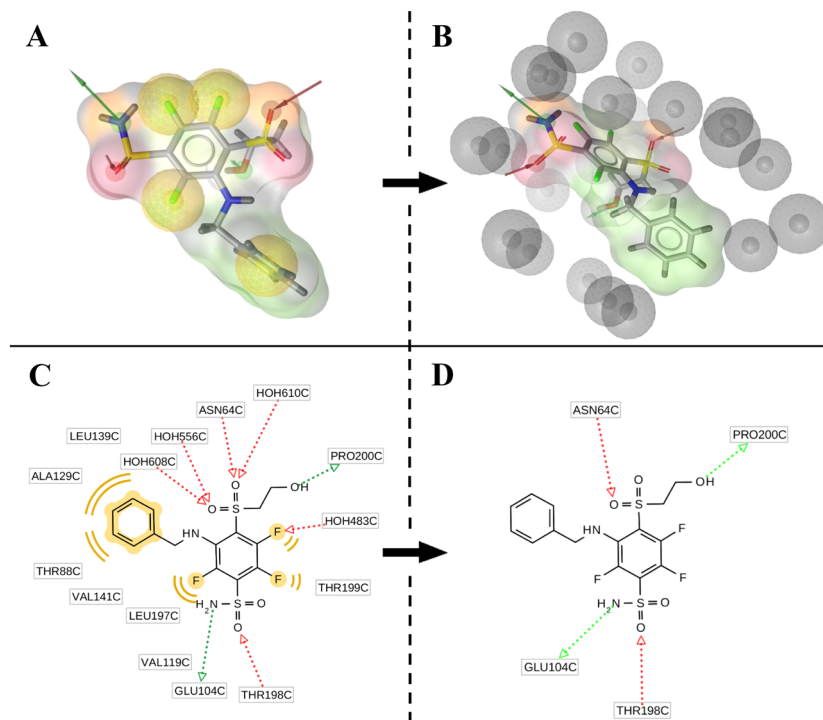


Figure 3. The best generated structure-based pharmacophore model. 3D (A) and 2D (C) representations of the pharmacophore model with all initial features. Final pharmacophore models after removing features responsible for the interactions with water molecules and hydrophobic interactions (B,D) and addition of the exclusion volumes coat (B). Figures of structure-based pharmacophore model was obtained using LigandScout v4.4²⁷ (www.inteligand.com/ligandscout/). Schematic illustration was drawn using Adobe Illustrator 2018 (www.adobe.com/products/illustrator).

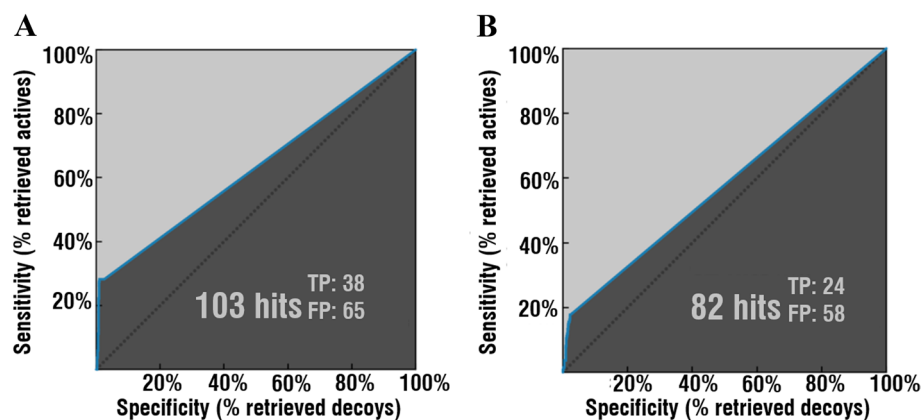


Figure 4. Validation of the generated structure-based pharmacophore on the two datasets (A,B). Identified hit compounds out of 2635 total compounds (135 actives, 2500 decoys). Figure was obtained using LigandScout v4.4²⁷ (www.inteligand.com/ligandscout/). Schematic illustration was drawn using Adobe Illustrator 2018 (www.adobe.com/products/illustrator).

molecular docking (Table S3). All of the compounds with similar or higher docking scores in comparison to the reference ligand were passed to the additional stages of minimization and re-scoring using MMGBSA method.

MMGBSA re-scoring. Only 38 compounds from the “purchasable” (Table 1) and 2 from “natural and derivatives” (Table 2) datasets demonstrated similar or higher binding energy values than reference ligand as the result of MMGBSA re-scoring. 24 out of aforementioned 38 compounds were obtained with the use of structure-based pharmacophore model and the rest 15 with the use of the ligand-based pharmacophore model.

Structure-based approach				Ligand-based approach			
ZINC ID	Energy (kJ/mol)	ZINC ID	Energy (kJ/mol)	ZINC ID	Energy (kJ/mol)	ZINC ID	Energy (kJ/mol)
ZINC66466630	-157,83	ZINC76941861	-129,56	ZINC82980951	-159,25	ZINC21761334	-133,43
ZINC16137455	-153,97	ZINC39147130	-128,51	ZINC68025286	-147,68	ZINC10514459	-130,47
ZINC90089846	-152,20	ZINC75669379	-127,26	ZINC27522612	-144,36	ZINC38671716	-129,54
ZINC76965500	-146,49	ZINC39252273	-127,24	ZINC12555593	-141,68	ZINC55678727	-128,52
ZINC05699310	-139,63	ZINC12551609	-126,40	ZINC29565674	-139,19	ZINC27929386	-128,16
ZINC09126577	-139,14	ZINC89392154	-126,25	ZINC59456846	-137,50	ZINC13056037	-127,18
ZINC44547064	-138,05	ZINC12085595	-125,75	ZINC49448410	-136,54	ZINC40897288	-124,45
ZINC09562497	-135,91	ZINC39252405	-125,39	ZINC06510447	-136,40	ligand_4qjw	-124,20
ZINC22239311	-134,62	ZINC12983599	-124,99				
ZINC47251290	-134,41	ZINC58304576	-124,67				
ZINC58247763	-132,89	ZINC06142500	-124,23				
ZINC40218576	-131,14	ligand_4qjw	-124,20				
ZINC58263892	-129,77						

Table 1. Binding energies of the top “purchasable” compounds after re-scoring using MM-GBSA.

Structure-based approach		Ligand-based approach	
ZINC ID	Energy (kJ/mol)	ZINC ID	Energy (kJ/mol)
ZINC70704873	-105,68	ZINC49181869	-139,62
ZINC04221765	-93,03	ZINC49181861	-133,30
ZINC02131655	-85,34	ZINC49181866	-120,74
ZINC70699917	-77,17	ZINC08829478	-111,70
ZINC15958674	-76,88	ZINC08792367	-109,86

Table 2. Binding energies of the top “natural and derivatives” compounds after re-scoring using MMGBSA.

Both compounds identified from “natural and derivatives” dataset were obtained with the use of a ligand-based pharmacophore model.

In order to analyze chemical diversity of identified compounds, an agglomerative hierarchical clustering method based on the “ECFP” fingerprint implemented in ICM-Pro software⁵¹ was performed (Fig. 5). As a result of the clusterization of the identified 38 compounds, 18 clusters were obtained. As expected, most of the identified compounds (Fig. 5, clusters 1–10) have sulfonamide moieties or related structural motifs that are widely known to promote small-molecules interaction with hCAs active site. However, several of the identified compounds (Fig. 5, clusters 11–18) have different from sulfonamide chemical moieties. These compounds include cyclohexanecarboxiamide, propanamide, acetamide, cyclopropanecarbohydrazide, thiazolidine, cyclopropane, carbonylhydrazide, chromen-7, tetrahydrofuran derivatives and analogues.

Remarkably, one of these compounds, identified with the use of ligand-based pharmacophore (ZINC82980951), showed the lowest binding energy among all compounds as the result of MMGBSA re-scoring (Table 1). Compounds that are chemically different from the traditional hCA XII inhibitors (which include sulfonamide chemical group) are of special interest for further analysis and investigation as potential alternative hCA XII inhibitors.

In the case of the “natural and derivatives” compounds only two compounds (ZINC49181869 and ZINC49181861, Table 2), which are stereoisomers of the same coumarin derivative, showed higher binding energies than reference ligand.

Molecular dynamics. From the list of compounds, identified from the “purchasable” dataset (Table 1), two compounds with lowest binding energy (one identified using structure-based pharmacophore—ZINC66466630 and another one using ligand-based pharmacophore model—ZINC82980951) were selected for additional MD simulations. In the case of the “natural and derivatives” dataset (Table 2), two compounds with the lowest binding energies, identified using structure-based (ZINC70704873) and ligand-based pharmacophore (ZINC49181869) models, were also selected for MD simulations.

The reference compound (co-crystallized ligand of 4QJW structure) was stable during the whole simulation and has maintained conformation close to its initial crystal state (fluctuations around 0.05 nm, Fig. 6).

From the four tested compounds only ZINC49181869 was stable during the whole simulation and also maintained conformation close to the one predicted by molecular docking fluctuations around 0.05 nm. Compound ZINC70704873 stabilized after ~ 55 ns of the simulation and maintains its conformation with fluctuations around ~ 0.05 nm. ZINC66466630 stabilized after ~ 70 ns and maintains its conformation with fluctuations

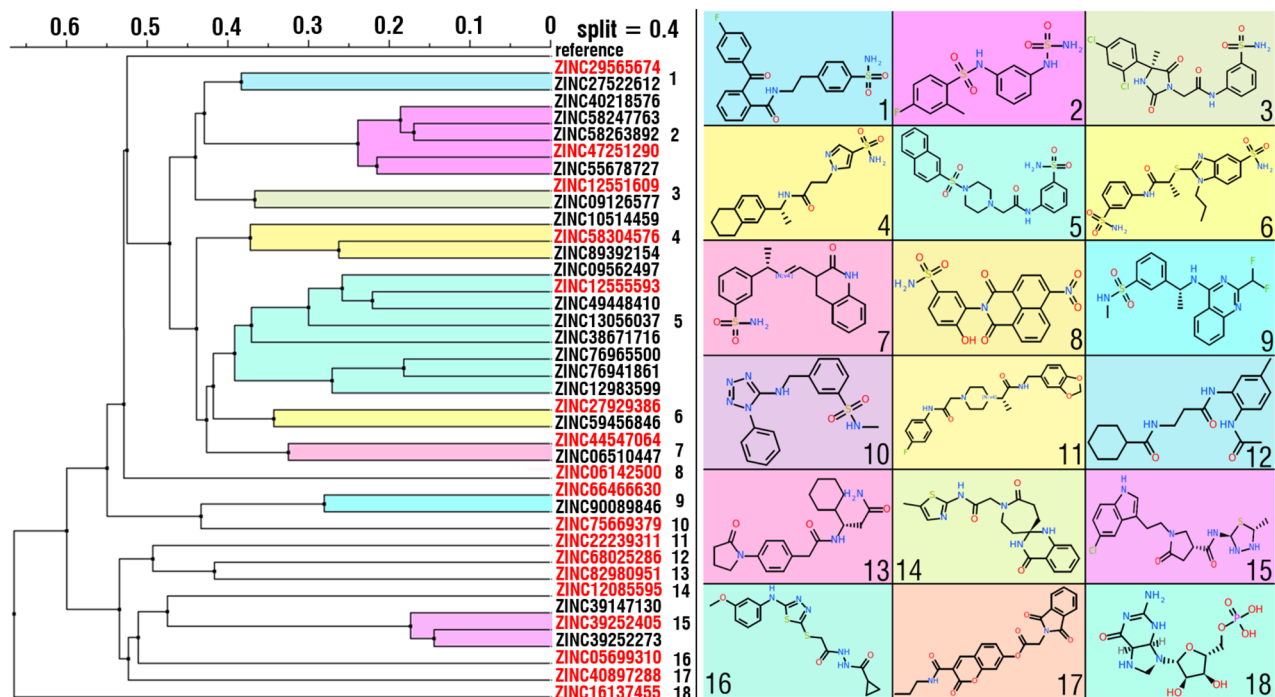


Figure 5. Results of the clusterization of the identified compounds as the result of MMGBSA re-scoring procedure. 2D structures of representative are presented for all 18 clusters. IDs of representative compounds of clusters with more than one compound are colored in red. Clusterization dendrogram and figures of chemical structures were obtained using ICM-PRO⁵¹ (http://www.molsoft.com/icm_pro.html). Schematic illustration was drawn using Adobe Illustrator 2018 (www.adobe.com/products/illustrator).

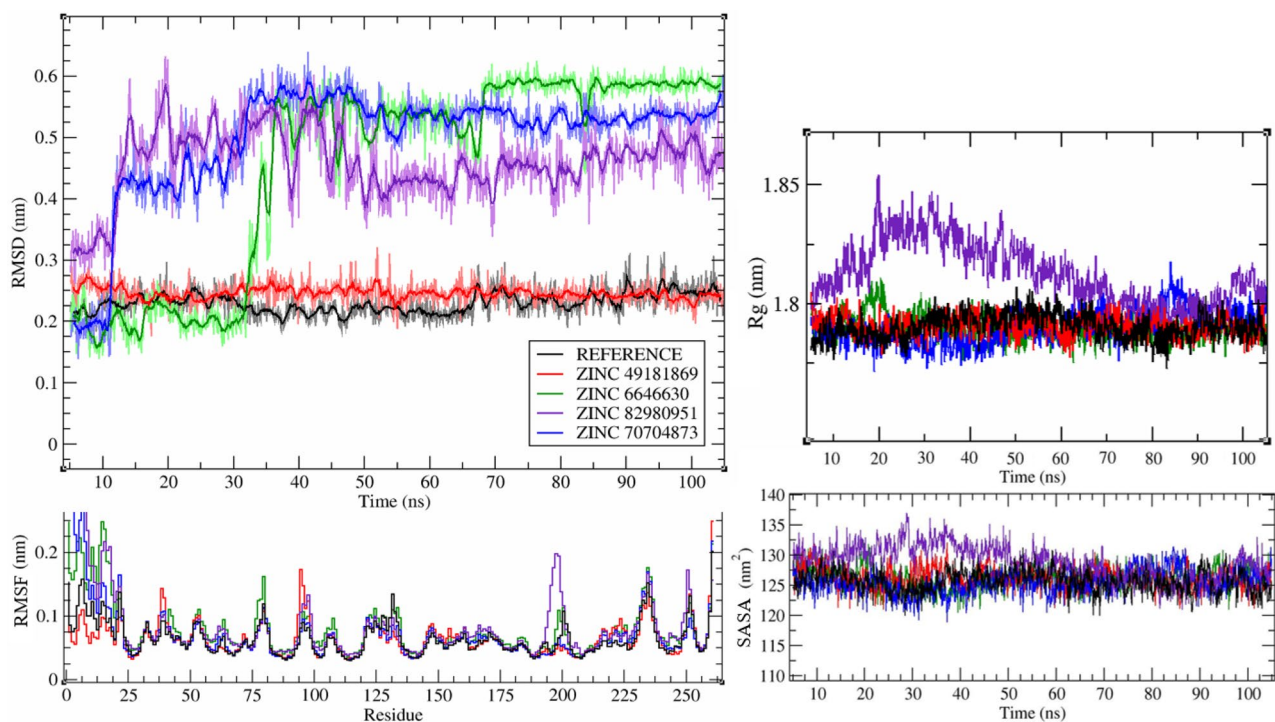


Figure 6. RMSD values of studied compounds and Rg, RMSF and SASA values of hCA XII during performed MD simulations. First 5 ns represent equilibration stage of performed MD simulations.

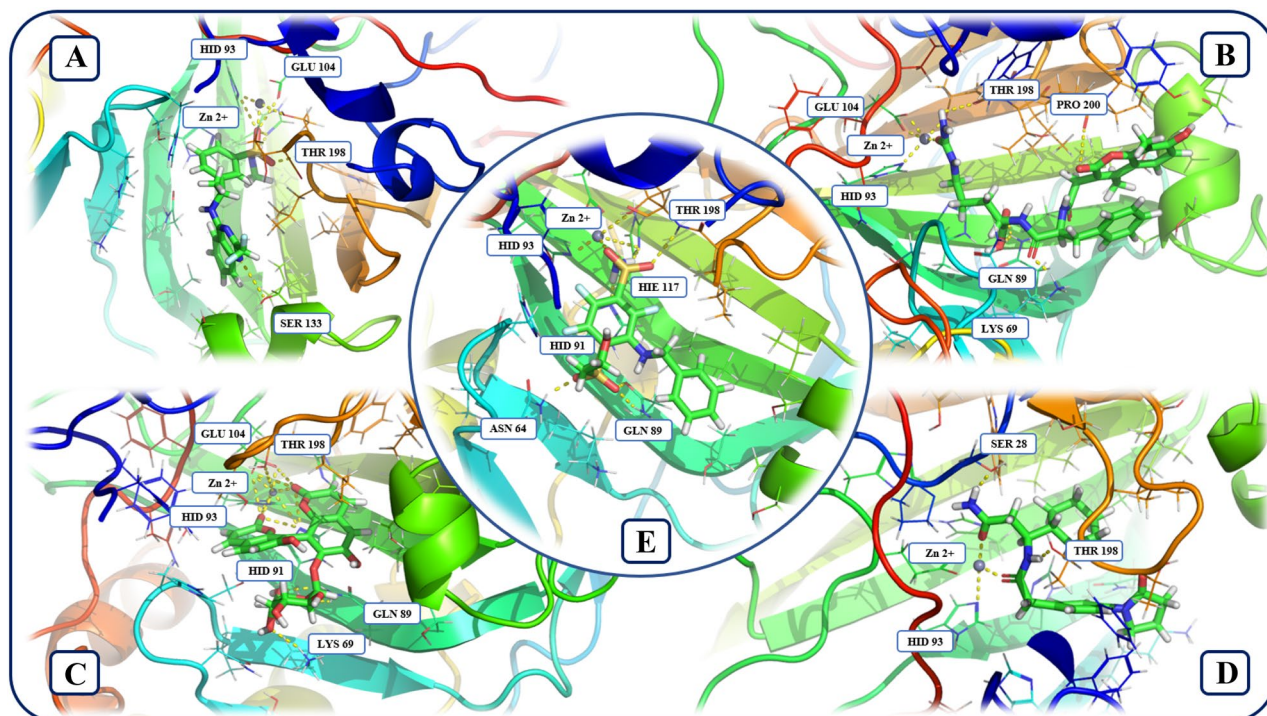


Figure 7. Interaction of the selected compounds with amino acid residues and Zn^{2+} ion in the binding site of the hCA XII isozyme. (A) ZINC66466630, (B) ZINC70704873, (C) ZINC49181869, (D) ZINC82980951, (E) Reference. Figures of complexes were obtained using PyMOL v. 2.3.2 (<https://pymol.org>). Schematic illustration was drawn using Adobe Illustrator 2018 (www.adobe.com/products/illustrator).

around ~ 0.025 nm. ZINC82980951 stabilized after ~ 50 ns and maintains its conformation with fluctuations around ~ 0.1 nm. Based on the obtained RMSD, RMSE, Rg and SASA values all studied complexes stabilize within performed simulations.

In its stable conformation, during molecular dynamics simulation, the reference ligand had five hydrogen bonds with the following amino acid residues of active site of the hCA XII enzyme: HIE 117, HID 91, GLN 89, ASN 64, THR 198. In the presence of the reference ligand, HID 93, HIE 117 and THR 198 amino acid residues coordinate Zn^{2+} ion (Fig. 7).

Similarities and differences in interaction patterns of the amino acid residues of hCA XII binding site with compounds and reference ligand is of particular interest for the evaluation of the studied compounds as potential alternative inhibitors of the hCA XII. ZINC66466630 and ZINC82980951 compounds and reference ligand have hydrogen bond with the THR 198 amino acid residue of the hCA XII active site. ZINC70704873, as the reference ligand, has hydrogen bonds with THR 198 and GLN 89 residues. ZINC49181869 and reference ligand both have hydrogen bonds with GLN 89 and HID 91 residues.

At the same time, all studied compounds have unique hydrogen bonds with other amino acid residues of the hCA XII isozyme. ZINC66466630 has hydrogen bonds with HID 93, GLU 104 and SER 133 residues. ZINC70704873 has hydrogen bonds with PRO 200 and LYS 69. ZINC49181869 has hydrogen bonds with HID 93, GLU 104 and LYS 71 residues. Finally, ZINC82980951 has hydrogen bond with SER 30 amino acid residue.

Information on the involvement of amino acid residues of the hCA XII binding site in coordination of Zn^{2+} ion in the presence of different compounds is also of high value for drug design and discovery of potential hCA inhibitors. Common amino acid residues that interact with the Zn^{2+} in the presence of tested compounds and reference ligand: HID 93, THR 198 (ZINC66466630, ZINC70704873, ZINC49181869), HID 93 (ZINC82980951). Unique, in comparison to the reference ligand, amino acid residues that interact with the Zn in the presence of the tested compounds: GLU 104 (ZINC70704873), HID 91 and GLU 104 (ZINC49181869), HID 93 (ZINC82980951).

Additionally, comparison of the shapes and chemical structures of the selected compounds to the reference ligand (classical inhibitor of classical inhibitor of hCA XII) was performed using ROCS_report utility tool (Fig. 8).

Only one out of four studied compounds (ZINC66466630) has sulfonamide group, which is inherent in classical hCA inhibitors. Three other studied compounds have alternative to classical inhibitors chemical structures.

One of these compounds—ZINC49181869 (coumarin derivative) is of higher interest as potential alternative inhibitor of the hCA XII, since it demonstrated relatively good parameters based on various indicators. This compound showed exceptional stability during the whole length of molecular dynamics simulation. It also has high affinity to the hCA XII isozyme ($-139,62$ kJ/mol) based on the MMGBSA calculations. Besides, it is a natural compound that has various good ADMET properties predicted by SwissADME⁵³ (consensus Log P of 0.58; Log S of around -3.5 and is not inhibitor of CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4). Remarkably coumarins are regarded as a promising new class of non-classical inhibitors of CAs^{20,54}.

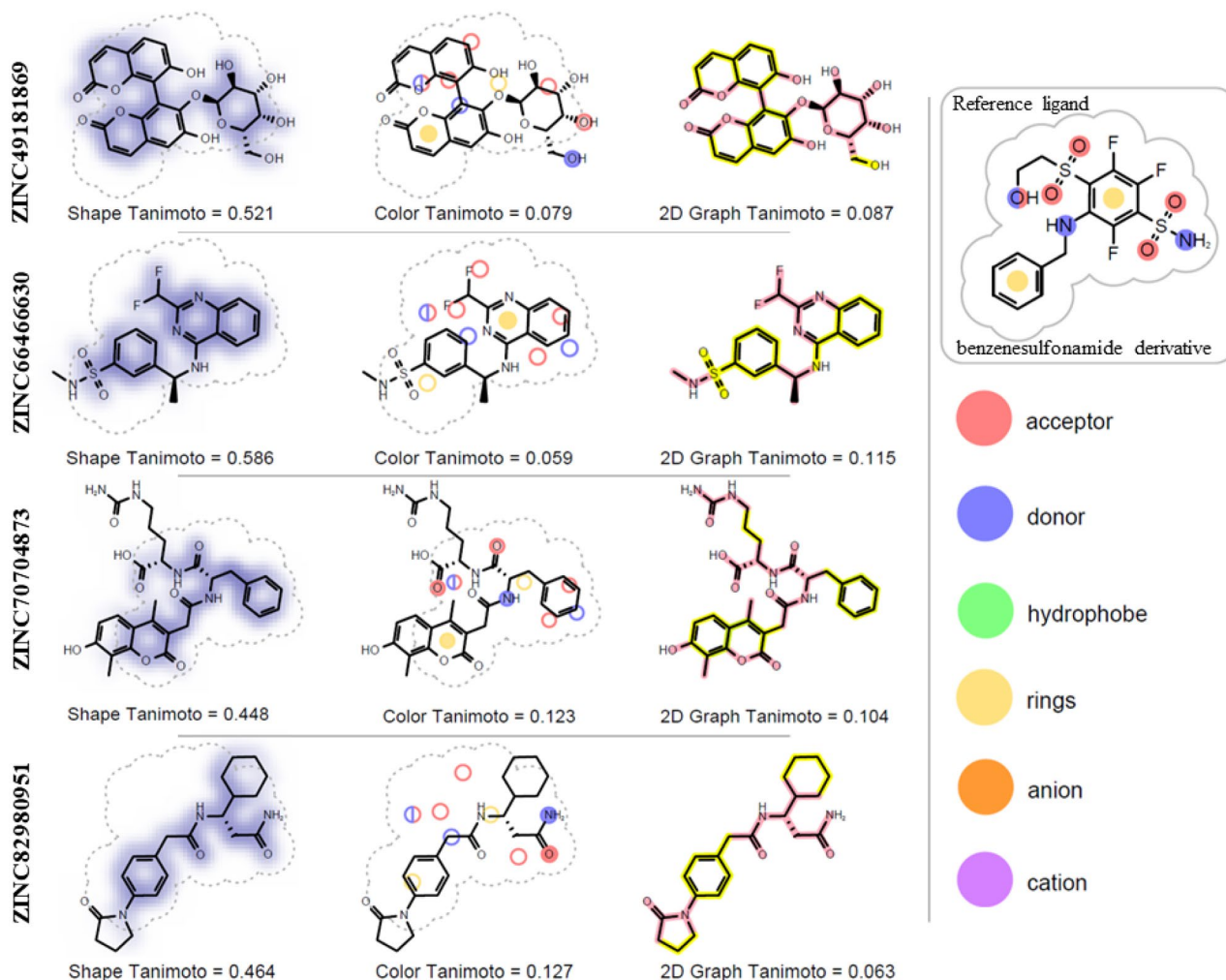


Figure 8. Comparison of the shape and chemical structures of identified compounds with the benzenesulfonamide derivative (representative classical inhibitor of hCA XII). Figure was obtained using OpenEye ROCS⁵² ROCSReport utility (<https://www.eyesopen.com/rocs>).

Conclusion

Despite the fact that interaction features of hCAs inhibitors with the catalytic site of the hCA enzyme are well described, classical inhibitors that have sulfonamides group or related motifs are lacking selectivity to the particular hCA isoforms and can also cause sulfa allergies in patients. Novel inhibitors of hCA XII should be both, isoform specific and non-classical, i.e. not based on sulfonamides, sulfamates, or sulfamides. This study was devoted to the identification of novel potential hCA XII inhibitors using comprehensive set of computational approaches for drug design discovery. As the results of the study several compounds with alternative to classical inhibitors chemical scaffolds have been identified and are of high interest as potential non-classical hCA XII inhibitors. The most promising out of identified compounds is coumarin derivative (ZINC49181869) that demonstrated relatively better indicators and properties based on the performed calculations and analysis. Remarkably, coumarin derivatives were recognized as a new class of non-classical inhibitors of CAs based on several relatively recent studies. The results of our study signify the potency of coumarins as non-classical inhibitor of hCA XII isozyme.

Received: 4 April 2021; Accepted: 12 July 2021

Published online: 30 July 2021

References

- Supuran, C. T. Structure and function of carbonic anhydrases. *Biochem. J.* **473**, 2023–2032 (2016).
- Wykoff, C. C. *et al.* Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Res.* **60**, 7075–7083 (2000).
- Chiche, J., Brahimi-Horn, M. C. & Pouyssegur, J. Tumour hypoxia induces a metabolic shift causing acidosis: A common feature in cancer. *J. Cell Mol. Med.* **14**, 771–794 (2010).
- Mboge, M. Y., Mahon, B. P., McKenna, R. & Frost, S. C. Carbonic anhydrases: Role in pH control and cancer. *Metabolites* **8**, 20 (2018).

5. Türeci, O. *et al.* Human carbonic anhydrase XII: cDNA cloning, expression, and chromosomal localization of a carbonic anhydrase gene that is overexpressed in some renal cell cancers. *Proc. Natl. Acad. Sci. USA* **95**, 7608–7613 (1998).
6. Carta, F., Supuran, C. T. & Scozzafava, A. Sulfonamides and their isosters as carbonic anhydrase inhibitors. *Future Med. Chem.* **6**, 1149–1165 (2014).
7. Supuran, C. T. Advances in structure-based drug discovery of carbonic anhydrase inhibitors. *Expert Opin. Drug Discov.* **12**, 61–88 (2017).
8. Bonardi, A. *et al.* Structural investigations on coumarins leading to chromeno[4,3-c]pyrazol-4-ones and pyrano[4,3-c]pyrazol-4-ones: New scaffolds for the design of the tumor-associated carbonic anhydrase isoforms IX and XII. *Eur. J. Med. Chem.* **146**, 47–59 (2018).
9. Bozdag, M. *et al.* Discovery of 4-hydroxy-3-(3-(phenylureido)benzenesulfonamides as SLC-0111 analogues for the treatment of hypoxic tumors overexpressing carbonic anhydrase IX. *J. Med. Chem.* **61**, 6328–6338 (2018).
10. Cornelio, B. *et al.* 5-Arylisothiazol-3(2H)-one-1, (1)-(di)oxides: A new class of selective tumor-associated carbonic anhydrases (hCA IX and XII) inhibitors. *Eur. J. Med. Chem.* **175**, 40–48 (2019).
11. Thacker, P. S. *et al.* Design, synthesis and biological evaluation of coumarin linked 1,2,4-oxadiazoles as selective carbonic anhydrase IX and XII inhibitors. *Bioorg. Chem.* **98**, 103739 (2020).
12. Said, M. A. *et al.* Synthesis, biological and molecular dynamics investigations with a series of triazolopyrimidine/triazole-based benzenesulfonamides as novel carbonic anhydrase inhibitors. *Eur. J. Med. Chem.* **185**, 111843 (2020).
13. Abdel-Aziz, A. A.-M. *et al.* Design, synthesis, and carbonic anhydrase inhibition activity of benzenesulfonamide-linked novel pyrazoline derivatives. *Bioorg. Chem.* **87**, 425–431 (2019).
14. Ali, M. *et al.* Benzylaminoethureido-tailed benzenesulfonamides: Design, synthesis, kinetic and X-ray investigations on human carbonic anhydrases. *Int. J. Mol. Sci.* **21**, 20 (2020).
15. Abdelrahman, M. A. *et al.* Novel benzofuran-based sulphonamides as selective carbonic anhydrases IX and XII inhibitors: Synthesis and in vitro biological evaluation. *J. Enzyme Inhib. Med. Chem.* **35**, 298–305 (2020).
16. Mboge, M. Y. *et al.* Structure activity study of carbonic anhydrase IX: Selective inhibition with ureido-substituted benzenesulfonamides. *Eur. J. Med. Chem.* **132**, 184–191 (2017).
17. Mboge, M. Y. *et al.* Selective inhibition of carbonic anhydrase IX over carbonic anhydrase XII in breast cancer cells using benzene sulfonamides: Disconnect between activity and growth inhibition. *PLoS One* **13**, e0207417 (2018).
18. Lomelino, C. & McKenna, R. Carbonic anhydrase inhibitors: A review on the progress of patent literature (2011–2016). *Expert Opin. Ther. Pat.* **26**, 947–956 (2016).
19. Kelly, T. E. & Hackett, P. H. Acetazolamide and sulfonamide allergy: A not so simple story. *High Alt. Med. Biol.* **11**, 319–323 (2010).
20. Lomelino, C. L., Supuran, C. T. & McKenna, R. Non-classical inhibition of carbonic anhydrase. *Int. J. Mol. Sci.* **17**, 20 (2016).
21. Krasavin, M. *et al.* Continued exploration of 1,2,4-oxadiazole periphery for carbonic anhydrase-targeting primary arene sulfonamides: Discovery of subnanomolar inhibitors of membrane-bound hCA IX isoform that selectively kill cancer cells in hypoxic environment. *Eur. J. Med. Chem.* **164**, 92–105 (2019).
22. Andreucci, E. *et al.* The carbonic anhydrase IX inhibitor SLC-0111 sensitises cancer cells to conventional chemotherapy. *J. Enzyme Inhib. Med. Chem.* **34**, 117–123 (2019).
23. Baig, M. H. *et al.* Computer aided drug design: Success and limitations. *Curr. Pharm. Des.* **22**, 572–581 (2016).
24. Talevi, A. Computer-aided drug design: An overview. *Methods Mol. Biol.* **1762**, 1–19 (2018).
25. Macalino, S. J. Y., Gosu, V., Hong, S. & Choi, S. Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **38**, 1686–1701 (2015).
26. Yu, W. & MacKerell, A. D. Computer-aided drug design methods. *Methods Mol. Biol.* **1520**, 85–106 (2017).
27. Wolber, G. & Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* **45**, 160–169 (2005).
28. Berman, H. M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
29. Kim, S. *et al.* PubChem substance and compound databases. *Nucleic Acids Res.* **44**, D1202–D1213 (2016).
30. Gaulton, A. *et al.* ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–D1107 (2012).
31. Mysinger, M. M., Carchia, M., Irwin, J. & Shoichet, B. K. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.* **55**, 6582–6594 (2012).
32. O’Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J. Cheminform.* **3**, 33 (2011).
33. Empereur-mot, C. *et al.* Predictiveness curves in virtual screening. *J. Cheminform.* **7**, 20 (2015).
34. Pascual, R., Almansa, C., Plata-Salamán, C. & Vela, J. M. A new pharmacophore model for the design of sigma-1 ligands validated on a large experimental dataset. *Front. Pharmacol.* **10**, 20 (2019).
35. Krüger, D. M. & Evers, A. Comparison of structure- and ligand-based virtual screening protocols considering hit list complementarity and enrichment factors. *ChemMedChem* **5**, 148–158 (2010).
36. Koes, D. R. & Camacho, C. J. ZINCPharmer: Pharmacophore search of the ZINC database. *Nucleic Acids Res.* **40**, W409–W414 (2012).
37. Morris, G. M. *et al.* AutoDock4 and autoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
38. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
39. Wang, Z. *et al.* Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: The prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys.* **18**, 12964–12975 (2016).
40. Chilingaryan, G. *et al.* Combination of consensus and ensemble docking strategies for the discovery of human dihydroorotate dehydrogenase inhibitors. *Sci. Rep.* **11**, 11417 (2021).
41. Maier, J. A. *et al.* ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
42. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
43. Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II Parameterization and validation. *J. Comput. Chem.* **23**, 1623–1641 (2002).
44. Miller, B. R. *et al.* MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).
45. Sahakyan, H. Improving virtual screening results with MM/GBSA and MM/PBSA rescoring. *J. Comput. Aided Mol. Des.* <https://doi.org/10.1007/s10822-021-00389-3> (2021).
46. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).
47. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
48. Oliver, D. S., Cunha, L. B. & Reynolds, A. C. Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Math. Geol.* **29**, 61–91 (1997).
49. Wu, X. & Brooks, B. R. Self-guided Langevin dynamics simulation method. *Chem. Phys. Lett.* **381**, 512–518 (2003).

50. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
51. Abagyan, R., Totrov, M. & Kuznetsov, D. ICM—a new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **15**, 488–506 (1994).
52. Hawkins, P. C. D., Skillman, A. G. & Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **50**, 74–82 (2007).
53. Daina, A., Michielin, O. & Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 42717 (2017).
54. Touisni, N. *et al.* Glycosyl coumarin carbonic anhydrase IX and XII inhibitors strongly attenuate the growth of primary breast tumors. *J. Med. Chem.* **54**, 8271–8277 (2011).

Acknowledgements

The authors would like to extend their sincere appreciation to the central laboratory at Jouf University for the support of this study.

Author contributions

M.M.A.: conceptualization, project management, fund acquisition and data analysis.; G.C. and N.A.: conceptualization, experimental design, molecular modeling, simulations and data analysis; G.A., H.S. and V.A.: data collection, analysis and validation.; K.N., G.M.A., H.E.A., W.M.A., F.S.A., N.Q.A., T.S.A., A.A.S.A., T.M.A., A.H.A., N.H.A. and M.A.A.: discussion of the results, writing, review, editing, plotting the graphical results and figures preparation. All authors revised and approved the final manuscript.

Funding

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this work through the project number "375213500".

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94809-x>.

Correspondence and requests for materials should be addressed to G.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021