*Chem Biol Drug Des 2014*; **83:** 247–258

Research Article

# Discovering Isozyme-Selective Inhibitor Scaffolds of Human Carbonic Anhydrases Using Structural Alignment and De novo Drug Design Approaches

**Fu Xiang[1], Jun Xiang[1,2], Yuanping Fang[1], Mingju Zhang[1,*] and Maoteng Li[2,*]**

*[1]Hubei Key Laboratory of Economic Forest Germplasm Improvement and Resources Comprehensive Utilization, Huanggang Normal University, Huanggang 438000, China*
*[2]College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*
**Corresponding authors: Mingju Zhang, lc_xiangfu@sohu.com; Maoteng Li, limaoteng426@hust.edu.cn*

**The development of isozyme-selective carbonic anhydrase inhibitors is currently still a great challenge. In the present study, protein–ligand complex structures were obtained by AutoDock Vina with SBR ((*R*)-*N*-(3-indol-1-yl-2-methyl-propyl)-4-sulfamoyl-benzamide) as the only inhibitor docked into the binding pockets of human isozymes CA I, II, IV, VI, IX, XII, and XIII. To make the spatial structures of complexes comparable, the co-ordinates for CA domains were reassigned based on structural alignments. With preferred docking poses of SBR been reduced to seed structures, the LigBuilder was used to build up inhibitor molecules. The results suggested that sulfonamide derivatives with naphthalene, fluorene, and acridan as the scaffold structures can be the potential isozyme-selective CAIs, especially for isozymes CA II, IV, and IX.**

**Key words:** carbonic anhydrase, drug design, molecular docking, selective inhibitor, structural alignment

**Abbreviations:** CA, carbonic anhydrase; CAIs, carbonic anhydrase inhibitors; SBR, (*R*)-*N*-(3-indol-1-yl-2-methyl-propyl)-4-sulfamoyl-benzamide.

Received 19 July 2013, revised 24 August 2013 and accepted for publication 15 September 2013

Carbonic anhydrases (CAs, EC 4.2.1.1) are present in humans in at least 15 different isoforms, 12 of which are catalytically active (CA I–IV, CA VA, VB, VI, VII, IX, and CA XII–XIV) (1–3). For their efficiently catalyzation on the chemical interconversion of carbon dioxide to bicarbonate ions, CAs are involved not only in respiration, pH changes, regulation of $CO_2$ fixation, and transport of $CO_2$/bicarbonate, also in electrolyte secretion in a variety of tissues and in many biosynthetic reactions, such as lipogenesis, ureagenesis, gluconeogenesis, tumorigenicity, bone resorption, calcification, and many other physiological or pathological processes (3,4). Consequently, many of CAs are well-established drug targets for the design of novel pharmacological agents useful in the treatment or prevention of many disorders, such as altitude sickness, epilepsy, glaucoma, obesity, tumor, acid–base disequilibria, and other neuromuscular diseases (2–7).

The catalytic and inhibition mechanisms of CAs are understood in great detail, and this favored the molecular design of potent inhibitors, and some of which possess important clinical applications (2). However, due to indiscriminate inhibition of isoforms other than the target one, most of the clinically available carbonic anhydrase inhibitors (CAIs) known to date are associated with undesired side-effects, including numbness and tingling of extremities, metallic taste, depression, fatigue, malaise, weight loss, decreased libido, gastrointestinal irritation, metabolic acidosis, renal calculi, and transient myopia (3). For the prospects of overcoming the systemic side-effects and of achieving the desired effect at lower dosages of a drug, selectivity of inhibitors against closely related enzyme isoforms is very attractive to researchers.

One of the major classes of CAIs is the unsubstituted sulfonamides and their bioisosteres, that is, the sulfamates, sulfamides, and related compounds (3). In fact, the importance of an unsubstituted sulfonamide group for CA inhibition has been confirmed by previously published reports on structure–activity relationships for this class of inhibitors (8). Sulfonamides disturb the catalytic cycle by binding in their deprotonated form ($RSO_2NH^-$) to the Zn(II) ion in carbonic anhydrase (1,8,9). Figure 1 shows the structural elements required for sulfonamide inhibitors of CAs (9,10), and the side chain might possess a hydrophilic link able to interact with the hydrophilic part of the active site and a hydrophobic moiety which can interact in the hydrophobic pocket. Several selective CAIs against pharmacologically relevant isozymes have been found by rational design approaches based on X-ray crystal structures of enzyme–inhibitor complexes (9), 3D QSAR selectivity models (11), tail approach (9,12), and dynamic combinatorial libraries

 doi: 10.1111/cbdd.12234

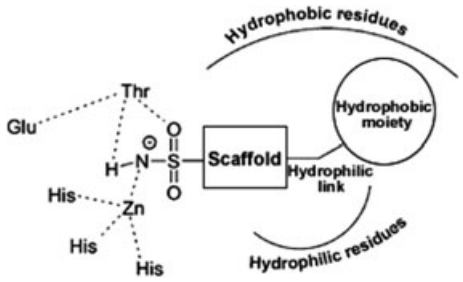



**Figure 1:** Structure elements of sulfonamide inhibitors in the carbonic anhydrase active site. A sulfonamide attached to the scaffold binds the Zn(II) ion and interacts with threonine by H bond. The hydrophilic link and hydrophobic moiety of inhibitors can interact with the hydrophilic and hydrophobic parts of the active site, respectively.

(13), etc. However, it is still remains challenging to extract selectivity discriminating features as a guideline to synthesis of CAIs.

In fact, the lack of structural information about the nature of the interactions between CAs and inhibitors has made it a difficult task to discover a lead compound through structure-based drug design and docking studies. More than that, the prediction of the bound conformation of a ligand in the active site of a protein is a major challenge in current structure-based drug design. In this work, we will predict protein–ligand complex structures for human CA isoforms with known structures using the docking program AutoDock Vina (14) to place the same inhibitor in the binding pockets of isoforms. Our interest in the automatic docking was stimulated by its significantly improving the accuracy of the binding mode predictions. To make the spatial structures of complexes comparable, the three-dimensional co-ordinates for CA isoforms were reassigned to share the same co-ordinate space based on structural alignments. Based on the complex structures with correct binding mode from AutoDock Vina, the de novo design method will be used to build ligand molecules within the binding pockets of human CAs. By so doing, novel isozyme-selective inhibitor scaffolds could be expected to be discovered.

## Experimental Procedures

Multiple structural alignment of human CAs was computed using program STAMP implemented in MultiSeq, which is a part of VMD (15,16). The prediction of protein–ligand complex structures was performed using AutoDock Vina (14) version 1.0.3. The de novo design of selective inhibitors was performed using LigBuilder (17) version 1.2.

### *Data set and structural alignment*

Nine human CAs of known structure up to date were selected from Protein Data Bank (PDB) for their high resolution and non-mutant: (CA I: *2foy*; CA II: *1if7*; CA III: *1z93*; CA IV: *1znc*; CA VI: *3fe4*; CA VIII: *2w2j*; CA IX: *3iai*; CA XII: *1jd0*; CA XIII: *3czv*). The crystal structure of CA II complexed with SBR ((*R*)-*N*-(3-indol-1-yl-2-methyl-propyl)-4-sulfamoyl-benzamide), that is, *1if7*, was used as a reference protein. For its high affinity for CA II, $K_d$ approximately 30 pM (18), SBR was used as the only inhibitor docked into the binding pockets of 9 human CAs by program AutoDock Vina (14) to obtain protein–ligand complex structures for each isoform.

With metal ions being left and water molecular and bound ligand being eliminated, all 9 isozymes were aligned using program STAMP implemented in the software package VMD (15,16) version 1.8.7. For structure comparisons, we confine our attention to the protein domains of human CAs. Based on the multidimensional *QR* factorization (19,20) implemented in VMD, the non-redundant domain data set of 9 human CA isozymes, corresponding to the PDB code mentioned above, was composed of d2foya, d1if7a, d1z93a, d1znca, d3fe4b, d2w2ja, d3iaib, d1jd0b, and d3czva (the 'd' means domain, the 'a' and 'b' mean A chain and B chain, respectively). With d1if7a as the reference structure, the co-ordinates for domains were reassigned to share the same co-ordinate space after structural alignment (shown in Figure 2), which can not only make all domains share the same 'search space' (i.e. grid box) required by AutoDock Vina program, but also help to identify the potent complexes of SBR with 9 isozymes predicted by AutoDock Vina (14).

Structural similarity measures $Q_H$ and $Q_{res}$ are derived from the structural identity measure $Q$, which is used in protein folding to compare the pair distances in a protein conformation to the native on (21). $Q_{res}$ is designed to calculate structural similarity of each residue in a set of aligned structures, and $Q_H$ to measure the similarity between two structures (15). They ranges from 0 to 1, lower scores indicate low similarity and higher scores high similarity.

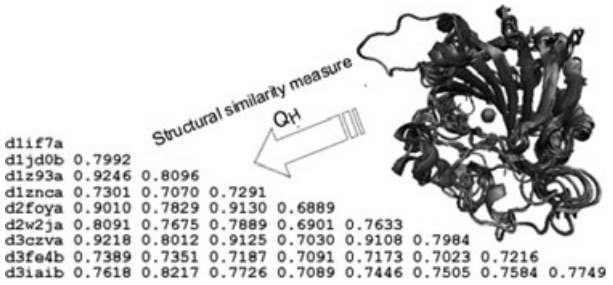



**Figure 2:** Structural superposition and similarity ($Q_H$) of nine human carbonic anhydrase domains. The ball presents the superposition of Mg(II) in d3fe4b (CA VI) and Zn(II)s in d2foya (CA I), d1if7a (CA II), d1z93a (CA III), d1znca (CA IV), d3iaib (CA IX), d1jd0b (CA XII), and d3czva (CA XIII). There is no metal ion in d2w2ja (CA VIII). $Q_H$ is designed to measure the similarity between two structures. $Q_H > 0.7$ means strongly structural homology.

### Docking simulation

The ligand, that is, SBR, was extracted directly from the complex structure *1if7*. AutoDock Vina (14) was used for automatic placement of SBR in the binding pockets of 9 isozyme domains mentioned above. The correct prediction of the binding mode was validated based on the extremely similarity between the poses of SBR in the binding pockets of d1if7a and *lif7* (Figure 3A). Receptor and ligand preparations were carried out using the AutoDockTools (22) version 1.5.4, including repairing missing atoms, adding polar hydrogens, and assigning charges and solvation parameters, etc. A grid of 36 × 40 × 36 points in the *x*, *y*, and *z*-axis direction was built with a grid spacing of 1 Å, and co-ordinates of the center was: $x = 0.074$, $y = 4.526$, and $z = 12.46$. In all cases, the grid box was large enough to guarantee independence of the docking results from small variations of the binding-site definitions. To look for an optimal binding mode, each domain was computed 100 times using AutoDock Vina (14) with num_modes = 20.

### De novo design

With the docked complex structure for each CA domain being used as the starting point, the software package LigBuilder (17) was used in the structure-based design of potential selective CAIs. The first step was to find the key interaction residues in the active site using the POCKET module with the minimal distance between any two pharmacophore features being 3.00. The next step was to construct molecules for the target protein by applying the genetic algorithm as implemented in the GROW module. The rough procedure of GROW is like this: (i) generate the initial population based on the seed structure; (ii) choosing 'parents' molecules from the current population into the mating pool; (iii) copy the 'elite' of the current population into the new population; (iv) fill out the new population by performing structural manipulations on the molecules in the mating pool; and (v) go to Step 2 until the preset number of generations has been reached. The numbers of the population and the 'parents' molecules to be selected into the mating pool were set to 8000 and 500, respectively. In addition, both the linking manipulation, which will result in intramolecular rings whenever it is possible, and the 'smart mutation' manipulation employed by GROW were forbidden. Due to the stochastic nature of genetic algorithm, we have run the program five times to sample the solution space adequately. According to the Lipinski rules (23), 200 top candidates were selected with maximal molecular weight and logP value being 500 and 5, respectively.

Table 1 presents the chemical structures and distributions of CAI scaffolds derived from top-scored 200 molecules.

### Cross-docking

Four molecules with multiple rings, which came from the top hit compounds designed for CA II, IV, VI, and IX, respectively, were selected to do cross-docking with CA I (*2foy*), CA II (*1if7*), CA IV (*1znc*), CA IX (*3iai*), CA XII (*1jd0*), and CA XIII (*3czv*). CA VI was excluded for its Mg(II) in d3fe4b. The four molecules **B1**, **B2**, **B3**, and **B4** are naphthalenesulfonamide, *N*-sulfonamide-phthalimide, 3-sulfonamide-acridan, and fluorenesulfonamide derivatives, respectively (Table 2). The parameters for docking are the same as above. Each domain was computed 10 times using AutoDock Vina (14) with num_modes = 20.

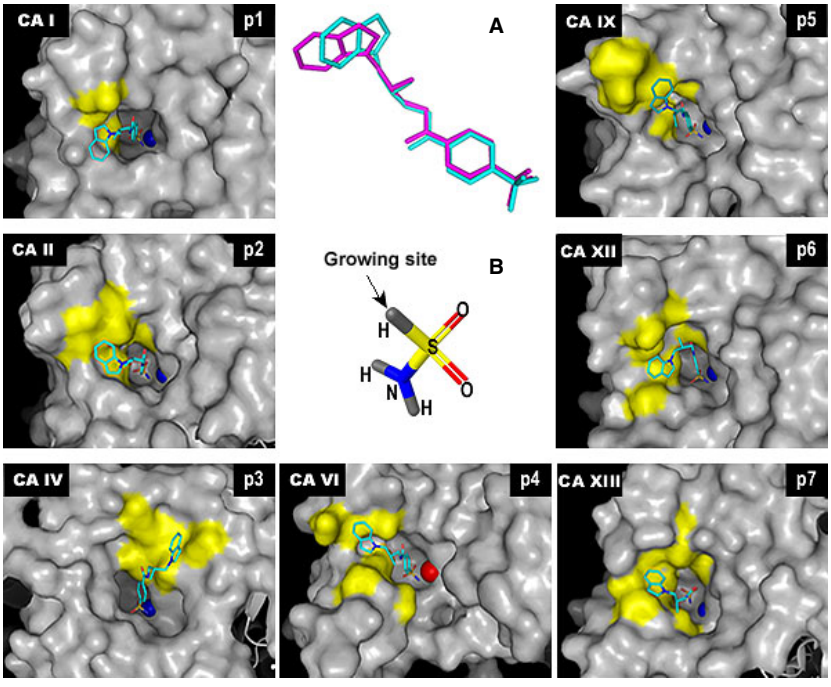



**Figure 3:** The poses of SBR in the active sites of carbonic anhydrase (CA) I, II, IV, VI, IX, XII, and XIII predicted by program AutoDock Vina, named as p1, p2, p3, p4, p5, p6, and p7, respectively. The hydrophobic half of the active site cleft is shown in yellow, while Zn(II) in blue and Mg (II) in red. (A) The superposition of SBR (in magenta) and p2 (in cyan). (B) The seed structure derived from p1-7 with the hydrogen attached to sulfur as the growing site. Graphics were generated with PyMOL software (DeLano Scientific LLC, Palo Alto, CA, USA).

## Results and Discussion

### *Structural alignment*

All nine CA domains in non-redundant set are readily structurally aligned by the STAMP algorithm and display clear structural homology. In Figure 2, all nine CA domains are superposed, and the Mg(II) in d3fe4b and Zn(II)s in d2foya, d1if7a, d1z93a, d1znca, d3iaib, d1jd0b, and d3czva are overlapped well with rmsd from 0.7791 Å to 1.9211 Å. A striking feature is that the structural core is well conserved, with $Q_H$ from 0.6889 to 0.9218. In general, $Q \geq 0.4$ corresponds to structures with 5 Å root-mean-square deviations or less, and this value indicates visible structural similarity (21), and $Q_H > 0.7$ means strongly structural homology. The high structure conservation contributes to the non-selective inhibition of CAs. After alignment, the co-ordinates for CA domains in non-redundant set were reassigned with d1if7a as the reference structure. In so doing, the failure of the scoring function implemented in docking program AutoDock Vina can be avoided and the receptor–ligand complexes predicted with correct binding mode can be manually screened out without difficulty. More than that, the search space required by AutoDock Vina can be shared the same one after the reassignment of co-ordinates for all nine CA domains in non-redundant set.

**Table 1:** The chemical structures and distributions of CAI scaffolds from top 200 molecules

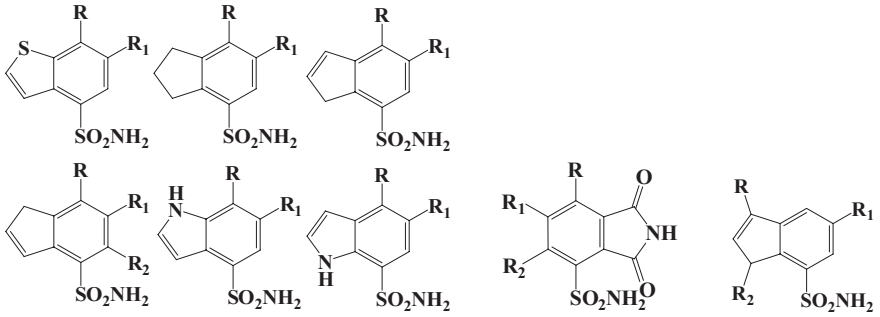
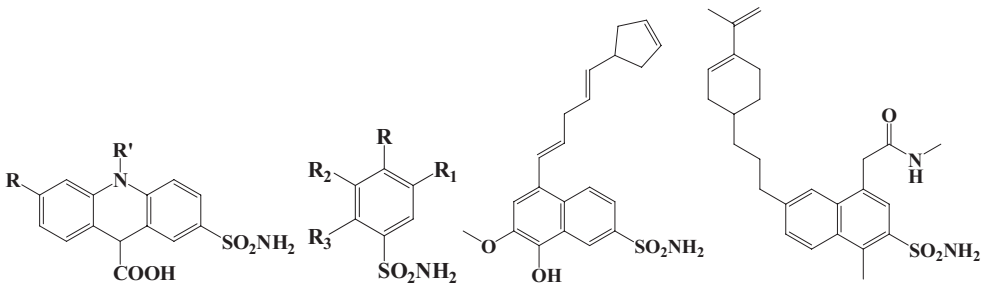
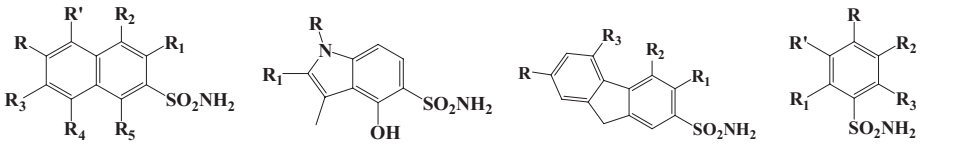
| Isozyme | The chemical structures and distributions of CAI scaffolds[a] |
|---|---|
| CA I |  **A1** (126/63%) **A2** (59/29.5%) **A3** (6/3%) |
| |  **A4** (4/2%) **A5** (3/1.5%) **A6** (1/0.5%) **A7** (1/0.5%) |
| CA II |  **A8** (189/94.5%) **A9** (6/3%) **A10** (3/1.5%) **A11** (2/1%) |

**Table 1:** continued

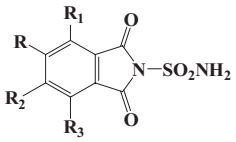
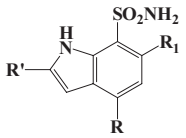
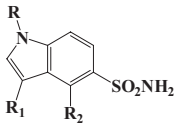
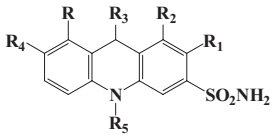
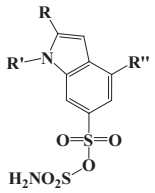
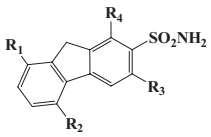
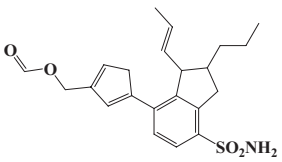
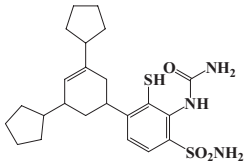
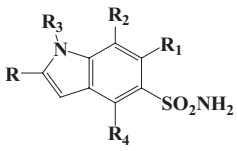
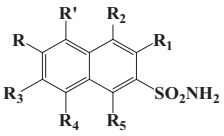
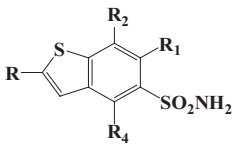
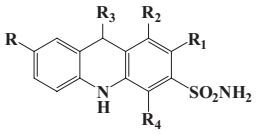
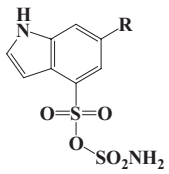
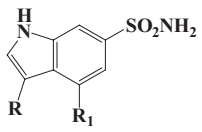
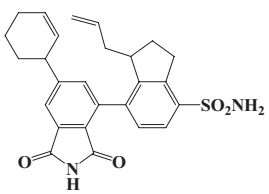
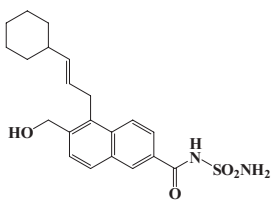
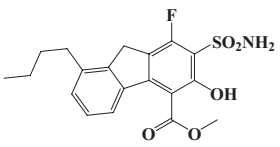
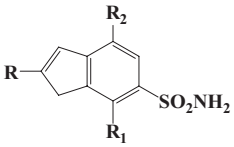
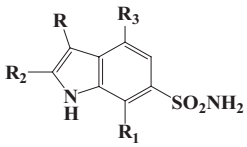
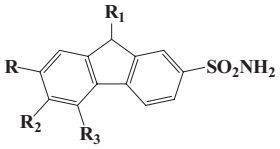
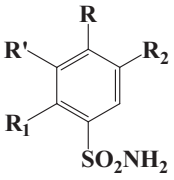
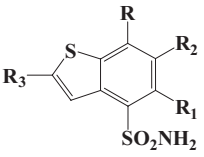
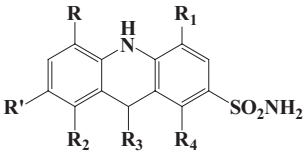
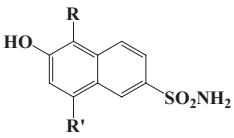
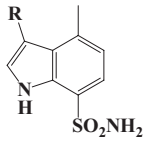
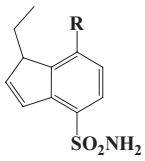
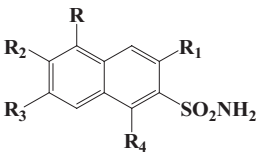
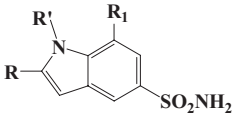
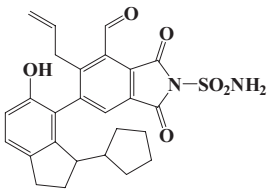
| Isozyme | The chemical structures and distributions of CAI scaffolds[a] | | | |
|---|---|---|---|---|
| CA IV |  **A12** (172/86%) |  **A13** (9/4.5%) |  **A14** (6/3%) |  **A15** (6/3%) |
| |  **A16** (3/1.5%) |  **A17** (2/1%) |  **A18** (1/0.5%) |  **A19** (1/0.5%) |
| |  **A20** (80/40%) |  **A21** (48/24%) |  **A22** (37/18.5%) |  **A23** (20/10%) |
| CA VI |  **A24** (8/4%) |  **A25** (2/1%) |  **A26** (2/1%) | |
| |  **A27** (1/0.5%) |  **A28** (1/0.5%) |  **A29** (1/0.5%) | |

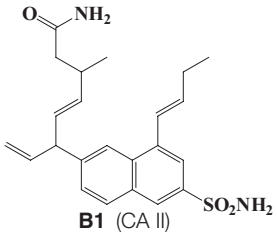
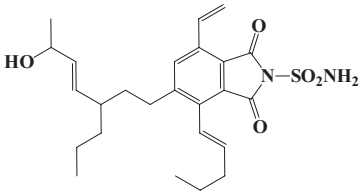
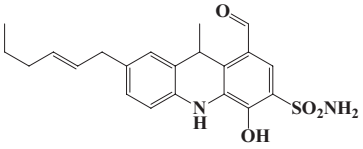
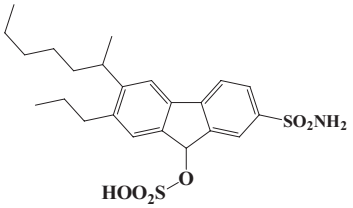
**Table 1:** continued

| Isozyme | The chemical structures and distributions of CAI scaffolds[a] | | |
|---|---|---|---|
| CA IX |  |  |  |
| | **A30** (134/67%) | **A31** (63/31.5%) | **A32** (3/1.5%) |
| CA XII |  |  |  |
| | **A33** (179/89.5%) | **A34** (11/5.5%) | **A35** (6/3%) |
| |  |  |  |
| | **A36** (2/1%) | **A37** (1/0.5%) | **A38** (1/0.5%) |
| CA XIII |  |  |  |
| | **A39** (190/95%) | **A40** (9/4.5%) | **A41** (1/0.5%) |

[a]The substitutions *R* and *R′* means necessary, while $R_n$ ($n$ = 1, 2, 3, 4, 5) substitutes occasionally.

Figure 4 presents the sequence alignment corresponding to the structural alignment of active site residues of nine CA domains in non-redundant set. Thirty-six active site residues previously defined as forming the cavity of the enzyme (24) were aligned with a combination of + (active-site hydrogen bond network) and Z (zinc-liganded histidine) signs. As shown in Figure 4, CA VI and CA IV present thirty-five residues for missing a residue in position 7 and 131, respectively. The residues with great variations in sequence are signed by boxes in Figure 4. It is worthy of note that the residues in positions indicated by boxes have high structure similarity with $Q_{res} > 0.6$, which means the selectivity profile against various isozymes may contribute to the small variation in the structure of CAIs. In fact, several amino acids located in the active site had been reported to be involved in inhibitor recognition and to have a role in the selectivity of various inhibitors toward different isozymes (25–28). The amino acid in position 65, which is Ala for CA II; Thr for CA III and CA VI; and Ser for CA I, CA IV, CA IX, CA XII, and CA XIII, has been reported to have an important role in discriminating inhibitor binding affinity toward different isozymes (26). The Asn 67 of CA II

**Table 2:** Cross-docking of de novo designed molecules against CA isozymes

| Inhibitor | Isozyme[a] | | | | | |
|---|---|---|---|---|---|---|
| | CA I | CA II | CA IV | CA IX | CA XII | CA XIII |
|  **B1** (CA II) | – | + | – | – | – | – |
|  **B2** (CA IV) | – | – | + | + | + | + |
|  **B3** (CA VI) | – | – | – | + | – | – |
|  **B4** (CA IX) | – | – | + | – | – | – |

[a]– shows interaction without sulfonamide binding the Zn(II) ion in active site; + shows interaction with sulfonamide binding the Zn(II) ion.

was shown to be involved in the binding of many classes of inhibitors (25,29). It is worth mentioning that the amino acid in position 67 is His for CA I, Arg for CA III, Gln for CA VI and IX, Met for CA IV, and Lys for CA XII. In particular, a difference among various CA isozymes in both sequence and structure is the amino acid in position 131, which is Phe for CA II, III, and XIII; Leu for CA I; Tyr for CA VI; Val for CA IX; and Ala for CA XII (see Figure 4). In fact, Phe 131 is known to be very important for the binding of sulfonamide inhibitors to CAs (25): In the case of CA II, the bulky side chain of this Phe residue not only increases steric hindrance for the inhibitors, also may participate in stacking interactions with groups present in it. Except for Tyr in CA VI, other residues, that is, Leu, Val, and Ala, are less bulky than Phe in CA II, which make the active sites of CA I, IX, and XII as a consequence being larger than the CA II active site. Correspondingly, CA IV active site is larger than CA II active site for missing a residue in position 131 (Figure 4). So, bulky inhibitors can be more accommodated within the active sites of CA I, IV, IX, and XII than those of CA II and VI.

## Docking simulation

Molecular docking aims to predict and rank the structure (s) arising from the association between a given ligand and a target protein of known 3D structure, and it thus can be used to try to predict the bound conformation of known binders, when the experimental structures are unavailable (30). The accuracy of the docked binding poses and the

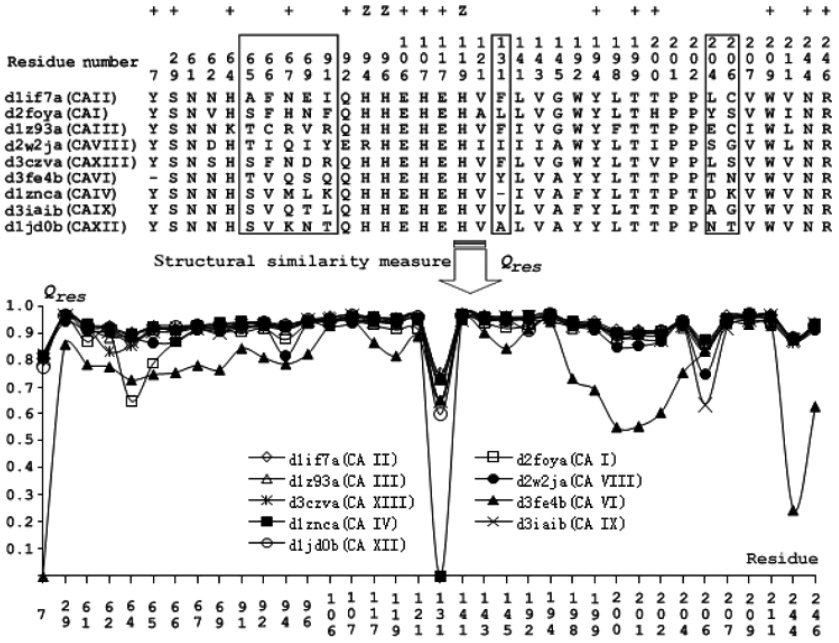



**Figure 4:** Sequence alignment resulting from the structural alignment of active site residues of nine human carbonic anhydrase domains. Thirty-six active site residues forming the active site are aligned. Symbols: +, means active site hydrogen bond network; Z, zinc-liganded histidine. Amino acid residues poorly conserved are indicated by boxes. Residue numbers are based on the CA II sequences. Qres is designed to measure structural similarity of each residue in a set of aligned structures. Qres > 0.7 means highly structural conservation.

correct predictions of the binding affinities are the two main problems in docking. Especially, the reliable scoring function remains inaccessible and is a key weakness with today's docking and scoring algorithms (31). Based on strongly structural conservation, the SBR from complex *1if7* can be used as the reference to select manually the docking poses of SBR in 9 CA domains after co-ordinates being reassigned, which can remedy the defects of the scoring function implemented in docking software AutoDock Vina.

For each isozyme, 2000 docking poses of SBR were generated by AutoDock Vina software. AutoDock Vina is a new program for drug discovery, molecular docking and virtual screening, offering multicore capability, high performance, and enhanced accuracy (14). The most preferred docking poses of SBR in CA I, II, IV, VI, IX, XII, and XIII are presented in Figure 3 and named as **p1**, **p2**, **p3**, **p4**, **p5**, **p6**, and **p7**, respectively. When aligned in PyMOL the rmsd between SBR and **p2** is 1.23 Å with 1 atom rejected and 0.78 Å with 4 atoms rejected (Figure 3A), indicating a high degree of similarity between the two poses. Reproducing the docked conformation of SBR in CA II with a reasonable rmsd is the basis for our motivation to use AutoDock Vina software to predict the most preferred docking poses of SBR in human CA domains. With SBR as a reference, no proper docking poses were found in CA III and VIII. In fact, CA VIII is non-catalytically active (1–3) for the amino acid in position 94 being Arg not His (Figure 4). CA III, unlike CA VIII, has a weak binding affinity to sulfonamides. The Phe in position 198 (Figure 4) was identified as a major contributor to such a unique feature of CA III (32).

According to Figure 3, p1, p2, p4, p5, p6, and p7 show a similar spatial arrangement with the indole group oriented toward the hydrophobic part of the active site cleft, making numerous van der Waals interactions with residues placed in this region. Figure 5 presents the active site regions in the CA I-**p1**, CA II-**p2**, CA IV-**p3**, and CA VI-**p4** complexes showing the residues involved in the hydrophobic interactions. These residues, in position 121, 131, 135, 198, 200, 202, and 204 (Figures 4 and 5), have been already reported to establish important interactions with the organic scaffold of other CA inhibitors (33), especially contributing to stabilize the inhibitor within the active site cleft. CA IV loses a residue in position 131 and forms a loop not helix at the hydrophobic part (Figures 2 and 4), which lead to the entrance of active site cleft of CA IV become more planar than those of other isozymes (Figure 3). Consequently, the indole group of **p3** pointed toward the hydrophobic pocket comprised of Leu 69, Lys 91, and Val 121 (Figures 3 and 5). In particular, the ionized sulfamate $NH^-$ group donates a hydrogen bond to the hydroxyl group of Thr 199 (Thr 220 in CA VI-**p4**), which has been observed in all adducts of sulfonamides/sulfamates with various CA isozymes (33–35), while one of the two sulfamate oxygens accepts a hydrogen bond from the backbone NH group of Thr 199.

### *Isozyme-selective scaffolds of CAIs*

With the structural constraints of human CA domains, the de novo drug design program LigBuilder 1.2 was used to build up inhibitor molecules iteratively using a library of organic fragments (17). Programs of this type normally require the user to define an initial 'seed' in a binding site, after which the computer builds a series of molecules by adding to the growing structures the most suitable fragments taken from a given library (36). So, choosing a proper seed structure is very important because all the

resultant molecules are the derivatives of the seed structure. To build seed structure, the preferred pose of SBR in each docked complex (i.e. **p1**-**7** shown in Figure 3) was reduced to the form H*-$SO_2NH_2$ with the H* as the growing site (Figure 3B). With H*-$SO_2NH_2$ used as the starting structure of de novo design, fragments will be added to the growing site to generate new molecules. All the molecules are developed and evolved with a genetic algorithm procedure. To score the derivatives according to the relative binding affinity, the program LigBuilder 1.2 employs the empirical binding free energy function including van der Waals, hydrogen bond, hydrophobic, and entropic terms (37). In addition, the bioavailability rules were also applied to screen the derivatives with good physicochemical properties as a drug (23).

For each isozyme, the top-scored 200 molecules were then selected to discover CAI scaffolds with isozyme preferences. Except for a single-ring structure, that is, benzene ring, most of the scaffolds are multiple-ring structures as shown in Table 1. In particular, the predominant structures of the inhibitor molecules designed by program LigBuilder 1.2 for CA I, II, IV, VI, IX, and XIII are multiple rings, while that for CA XII is benzene ring (**A1**, **A8**, **A12**, **A20**, **A30**, **A33**, and **A39**).

As for CA XII, the molecules with benzene ring scaffold account for 89.5%, while only 3 and 2 benzenesulfonamide derivatives for CA I and II, respectively. It should be noted that the substitutions at the para and/or meta positions of benzenesulfonamide are necessary (**A5**, **A11**, and **A33**). For the preference for benzene ring scaffold, the benzenesulfonamides with substitutions at the para and meta positions can be the potential isozyme-selective inhibitors of CA XII. In fact, CA XII is a sulfonamide-avid isoform, being strongly inhibited by many types of sulfonamides, and the most selective human CA XII over CA II inhibitors are benzenesulfonamide derivatives (38).

As for CA I, the compounds with indole, indene, indan, and thionaphthene rings used as the scaffolds reach 66% of 200 molecules (**A1** and **A3**). Interestingly, these multiple rings, especially the indole ring, have been designed for each isozyme as shown in Table 1. Thereinto, the compounds with indole ring as the scaffold structure account for 3%, 7.5%, 41%, 31.5%, and 4.5% of 200 top molecules designed for isozymes CA II, IV, VI, IX, and X III, respectively (**A9**, **A13**, **A14**, **A20**, **A26**, **A31**, and **A40**). For CA VI, there is a compound with indan ring as the scaffold (**A27**), and the compounds with thionaphthene ring as scaffolds account for 22.5% (**A22** and **A24**). The compounds with indene ring as the scaffold structure reach 67% for CA IX (**A30)**. With benzene ring as the predominant scaffold structure, 5.5%, 0.5%, and 0.5% of 200 molecules designed for CA XII possess the thionaphthene, indole, and indene scaffold structures, respectively (**A34**, **A37**, and **A38**). Consequently, the sulfonamide derivatives with multiple-ring structures such as indole, indene, indan, and thionaphthene as the scaffolds lack the preference for CA isozymes.

The predominant structures of inhibitor molecules designed for CA II and XIII are naphthalene ring, which account for

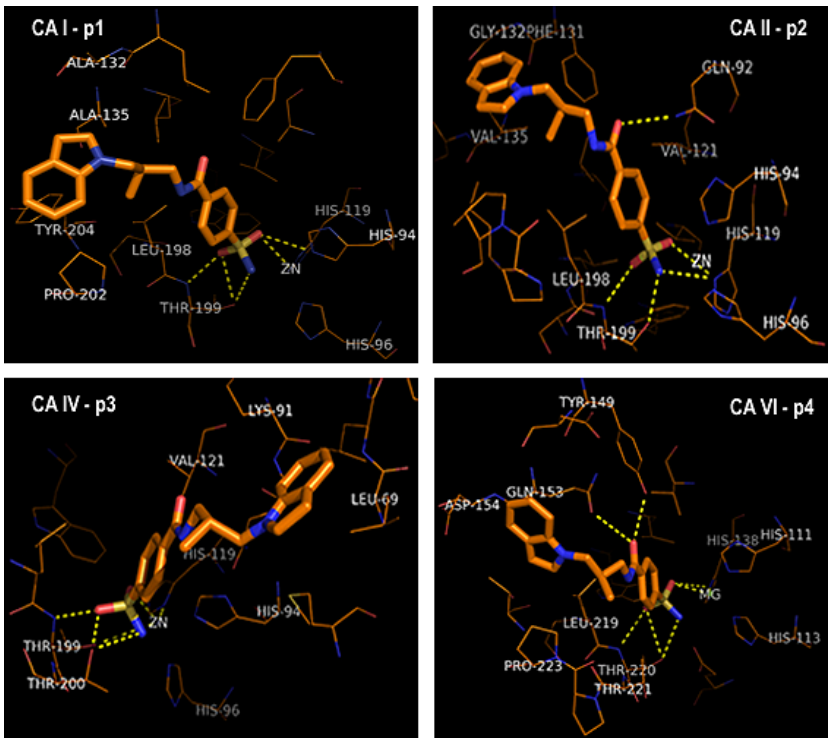



**Figure 5:** Stereo view of the active site regions in the CA I-p1, CA II-p2, CA IV-p3, and CA VI-p4 complexes, showing the residues participating in recognition of the inhibitor molecule. Hydrogen bonds are shown as dotted lines in yellow. Graphics were generated with PyMOL software (DeLano Scientific LLC, Palo Alto, CA, USA).

94.5% and 95% of top 200 molecules, respectively (**A8** and **A39**). In addition, there are 2, 48, and 2 naphthalenesulfonamide derivatives designed for CA I, VI, and XII, respectively (**A6**, **A7**, **A21**, and **A36**). It is worth noting that all the molecules are 2-naphthalenesulfonamide derivatives, and the substitutions at 5 and 6 positions of naphthalene ring are necessary for compounds **A8** and **A21**. These results implied that the sulfonamide derivatives with naphthalene ring scaffold can be the potential isozyme-selective CAIs, especially for isozymes CA II and XIII. In fact, it has been demonstrated that compounds of the $RSO_2NH_2$ type (where *R* is a benzene, naphthalene, or pyridine ring) specifically inhibit carbonic anhydrase in very small concentrations (39). As seen from Table 1, the preferences of benzenesulfonamide and naphthalenesulfonamide derivatives for CA isozymes depend on the substitutions in the benzene and naphthalene rings, respectively.

The phthalimide scaffold accounts for 86% of top 200 molecules and becomes the predominant structures of inhibitors designed for CA IV (**A12**). Such scaffold structures are observed in one and fifty-nine molecules designed for CA XIII (**A41**) and CA I (**A2**), respectively. As seen from Table 1, the important difference between **A2** and **A12** is the substitution of the sulfonamide group. Like as naphthalenesulfonamide derivatives, the *N*-sulfonamide-phthalimide derivatives can be the potential isozyme-selective CAIs, especially for isozyme CA IV.

The sulfonamide derivatives with multiple rings, such as acridan and fluorene, as scaffold structures have been observed. The compounds with acridan as the scaffold structure account for 2%, 3%, and 3% of 200 top molecules designed for isozymes CA I, IV, and XII, respectively (**A4**, **A15**, and **A35**). In particular, the 3-sulfonamide-acridan derivatives reach 10% of 200 molecules designed for CA VI (**A23**). It is worth noting that the compounds designed for CA I and XII are 2-sulfonamide-acridan derivatives (**A4** and **A35**) and those for CA IV and VI are 3-sulfonamide-acridan derivatives (**A15** and **A23**). More than that, there are different substitutions among **A4**, **A15**, **A23**, and **A35**. Based on these observations, one can conclude that the acridan can be the isozyme-selective scaffold of CAIs. Compare with acridan-sulfonamide derivatives, less fluorenesulfonamide derivatives were designed for CA isozymes. As seen from Table 1, there are 3 fluorenesulfonamide derivatives designed for CA II and IX (**A10** and **A32**), two for CA IV (**A17**), and only one for CA VI (**A29**). Such results implied that the fluorene can be the isozyme-selective scaffold of inhibitor molecules against CA II, IV, VI, and IX.

### *Cross-docking*

To ascertain the isozyme-selective inhibitor scaffolds, four sulfonamide derivatives with naphthalene, phthalimide, acridan, and fluorene as multiple-ring scaffolds, respectively, were selected to perform cross-docking experiment. A comparison of the cross-docking result is illustrated in Table 2. Not unexpectedly, de novo designed molecule **B1**, which was designed for CA II, interacts selectively with CA II. By contrast, **B2**, which was designed for CA IV, has less selectivity against CA IV, IX, XII, and XIII. **B3** and **B4**, which were designed for CA VI and IX, respectively, were selected to do cross-docking for their novel scaffold structures. As with **B1**, they have selectivity against CA IX and IV, respectively. It is worth noting that the cross-docking experiments with **B2**, **B3,** and **B4** are not in agreement with de novo molecular design. In fact, the de novo design builds inhibitor molecules based on an initial seed structure in an active site by adding to the growing structures, the most suitable fragments taken from a library of organic fragments. The designed molecules then were constrained to the inner structure of isozyme active site. In other words, even though the designed molecules could be appropriate for the inner structure of active site, it can be too bulky to enter into the active site. It is the reason why **B4** does not dock to CA IX in cross-docking experiments, even though that is the isozyme for which it was designed. However, regardless of the difference between cross-docking and de novo design, cross-docking results presented in Table 2 indicated that sulfonamide derivatives with naphthalene, fluorene, and acridan as the scaffold structures can be the potential isozyme-selective CAIs, especially for isozymes CA II, IV, and IX.

## Conclusions

In the present study, the complexes of CA isozymes with SBR were predicted using structural alignment and molecular docking approaches. With the structural constraints of human CA domains, de novo drug design produced preferred molecular structures for each isozymes based on complex structures. The results from cross-docking and de novo design suggested that sulfonamide derivatives with naphthalene, fluorene, and acridan as the scaffold structures can be the potential isozyme-selective inhibitors against CA II, IV, and IX.

Finally, it is the authors' opinion that the approaches used in this study can be combined with tail approaches reported previously to find novel virtual sulfonamide compounds, which can be a potent approach to develop isozyme-selective CAIs.

## Acknowledgments

This work was supported by Doctoral Foundation of Huanggang Normal University (2011CD220), Natural Science Funds for Distinguished Young Scholars of Hubei Province of China (2010CDA097) and New Century Talents Support Program of the Ministry of Education of China (NCET110172).

## References


1. Supuran C.T., Scozzafava A. (2002) Applications of carbonic anhydrase inhibitors and activators in therapy. Expert Opin Ther Pat;12:217–242.
2. Pastorekova S., Parkkila S., Pastorek J., Supuran C.T. (2004) Carbonic anhydrases: current state of the art, therapeutic applications and future prospects. J Enzym Inhib Med Chem;19:199–299.
3. Supuran C.T. (2008) Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. Nat Rev Drug Discov;7:168–181.
4. Supuran C.T., Scozzafava A. (2007) Carbonic anhydrases as targets for medicinal chemistry. Bioorgan Med Chem;15:4336–4350.
5. Stiti M., Cecchi A., Rami M., Abdaoui M., Barragan-Montero V., Scozzafava A., Guari Y., Winum J.Y., Supuran C.T. (2008) Carbonic anhydrase inhibitor coated gold nanoparticles selectively inhibit the tumor-associated isoform IX over the cytosolic isozymes I and II. J Am Chem Soc;130:16130–16131.
6. Winum J.Y., Scozzafava A., Montero J.L., Supuran C.T. (2009) Inhibition of carbonic anhydrase IX: a new strategy against cancer. Anticancer Agents Med Chem;9:693–702.
7. Pastorekova S., Zatovicova M., Pastorek J. (2008) Cancer-associated carbonic anhydrases and their inhibition. Curr Pharm Design;14:685–698.
8. Maren T.H. (1976) Relations between structure and biological activity of sulfonamides. Annu Rev Pharmacol;16:309–327.
9. Supuran C.T., Scozzafava A., Casini A. (2003) Carbonic anhydrase inhibitors. Med Res Rev;23:146–189.
10. Thiry A., Ledecq M., Cecchi A., Dogne J., Wouters J., Supuran C.T., Masereel B. (2006) Indanesulfonamides as carbonic anhydrase inhibitors. Toward structure-based design of selective inhibitors of the tumor-associated isozyme CA IX. J Med Chem;49:2743–2749.
11. Weber A., Bohm M., Supuran C.T., Scozzafava A., Sotriffer C.A., Klebe G. (2006) 3D QSAR selectivity analyses of carbonic anhydrase inhibitors: insights for the design of isozyme selective inhibitors. J Chem Infor Model;46:2737–2760.
12. Turkmen H., Durgun M., Yilmaztekin S., Emul M., Innocenti A., Vullo D., Scozzafava A., Supuran C.T. (2005) Carbonic anhydrase inhibitors. Novel sulfanilamide/acetazolamide derivatives obtained by the tail approach and their interaction with the cytosolic isozymes I and II, and the tumor-associated isozyme IX. Bioorg Med Chem Lett;15:367–372.
13. Nasr G., Petit E., Vullo D., Winum J., Supuran C.T., Barboiu M. (2009) Carbonic anhydrase-encoded dynamic constitutional libraries: toward the discovery of isozyme-specific inhibitors. J Med Chem;52:4853–4859.
14. Trott O., Olson A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem;31:455–461.
15. Roberts E., Eargle J., Wright D., Luthey-Schulten Z. (2006) MultiSeq: unifying sequence and structure data for evolutionary analysis. BMC Bioinformatics;7:382.
16. Humphrey W., Dalke A., Schulten K. (1996) VMD-Visual Molecular Dynamics. J Mol Graph;14:33–38.
17. Wang R., Gao Y., Lai L. (2000) LigBuilder: a multi-purpose program for structure-based drug design. J Mol Model;6:498–516.
18. Grzybowski B.A., Ishchenko A.V., Kim C., Topalov G., Chapman R., Christianson D.W., Whitesides G.M., Shakhnovich E.I. (2002) Combinatorial computational method gives new picomolar ligands for a known enzyme. PNAS;99:1270–1273.
19. O'Donoghue P., Luthey-Schulten Z. (2003) On the Evolution of Structure in Aminoacyl-tRNA Synthetases. Microbiol Mol Biol R;67:550–573.
20. O'Donoghue P., Luthey-Schulten Z. (2005) Evolutionary Profiles Derived from the QR Factorization of Multiple Structural Alignments Gives an Economy of Information. J Mol Biol;346:875–894.
21. Eastwood M.P., Hardin C., Luthey-Schulten Z., Wolynes P.G. (2001) Evaluating protein structure prediction schemes using energy landscape theory. IBM J Res Dev;45:475–497.
22. Sanner M.F. (1999) Python: a programming language for software integration and development. J Mol Graph Model;17:57–61.
23. Lipinski C.A., Lombardo F., Dominy B.W., Feeney P.J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliver Rev;23:3–25.
24. Tashian R.E. (1992) Genetics of the mammalian carbonic anhydrases. Adv Genet;30:321–356.
25. Menchise V., De Simone G., Alterio V., Di Fiore A., Pedone C., Scozzafava A., Supuran C.T. (2005) Carbonic Anhydrase Inhibitors: stacking with Phe131 Determines Active Site Binding Region of Inhibitors As Exemplified by the X-ray Crystal Structure of a Membrane-Impermeant Antitumor Sulfonamide Complexed with Isozyme II. J Med Chem;48:5721–5727.
26. Winum J.Y., Temperini C., El Cheikh K., Innocenti A., Vullo D., Ciattini S., Montero J.L., Scozzafava A., Supuran C.T. (2006) Carbonic anhydrase inhibitors: clash with Ala65 as a means for designing inhibitors with low affinity for the ubiquitous isozyme II, exemplified by the crystal structure of the topiramate sulfamide analogue. J Med Chem;49:7024–7031.
27. Guzel O., Temperini C., Innocenti A., Scozzafava A., Salman A., Supuran C.T. (2008) Carbonic anhydrase inhibitors. Interaction of 2-(hydrazinocarbonyl)-3-phenyl-1H-indole-5-sulfonamide with 12 mammalian isoforms: kinetic and X-ray crystallographic studies. Bioorg Med Chem Lett;18:152–158.
28. Di Fiore A., Pedone C., D'Ambrosio K., Scozzafava A., De Simone G., Supuran C.T. (2006) Carbonic anhydrase

inhibitors: valdecoxib binds to a different active site region of the human isoform II as compared to the structurally related cyclooxygenase II "selective" inhibitor celecoxib. Bioorg Med Chem Lett;16:437–442.
29. Alterio V., Vitale R.M., Monti S.M., Pedone C., Scozzafava A., Cecchi A., De Simone G., Supuran C.T. (2006) Carbonic Anhydrase Inhibitors: X-ray and Molecular Modeling Study for the Interaction of a Fluorescent Antitumor Sulfonamide with Isozyme II and IX. J Am Chem Soc;128:8329–8335.
30. Sousa S.F., Fernandes P.A., Ramos M.J. (2006) Protein-ligand docking: current status and future challenges. Proteins: Struct, Funct, Bioinf;65:15–26.
31. Leach A.R., Shoichet B.K., Peishoff C.E. (2006) Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. J Med Chem;49:5851–5855.
32. LoGrasso P.V., Tu C., Jewell D.A., Wynns G.C., Laipis P.J., Silverman D.N. (1991) Catalytic enhancement of human carbonic anhydrase III by replacement of phenylalanine-198 with leucine. Biochemistry;30:8463–8470.
33. De Simone G., Di Fiore A., Menchise V., Pedone C., Antel J., Casini A., Scozzafava A., Wurl M., Supuran C.T. (2005) Carbonic anhydrase inhibitors. Zonisamide is an effective inhibitor of the cytosolic isozyme II and mitochondrial isozyme V: solution and X-ray crystallographic studies. Bioorg Med Chem Lett;15:2315–2320.
34. Abbate F., Casini A., Owa T., Scozzafava A., Supuran C.T. (2004) Carbonic anhydrase inhibitors: E7070, a sulfonamide anticancer agent, potently inhibits cytosolic isozymes I and II, and transmembrane, tumor-associated isozyme IX. Bioorg Med Chem Lett;14:217–223.
35. Vitale R.M., Alterio V., Innocenti A., Winum J., Monti S.M., De Simone G., Supuran C.T. (2009) Carbonic anhydrase inhibitors. Comparison of Aliphatic Sulfamate/Bis-sulfamate Adducts with Isozymes II and IX as a Platform for Designing Tight-Binding, More Isoform-Selective Inhibitors. J Med Chem;52:5990–5998.
36. Schneider G., Fechner U. (2005) Computer-based de novo design of drug-like molecules. Nat Rev Drug Discov;4:649–663.
37. Wang R., Liu L., Lai L., Tang Y. (1998) SCORE: a new empirical method for estimating the binding affinity of a protein-ligand complex. J Mol Model;4:379–394.
38. Vullo D., Innocenti A., Nishimori I., Pastorek J., Scozzafava A., Pastorekova S., Supuran C.T. (2005) Carbonic anhydrase inhibitors. Inhibition of the transmembrane isozyme XII with sulfonamides: a new target for the design of antitumor and antiglaucoma drugs? Bioorg Med Chem Lett;15:963–969.
39. Mann T., Keilin D. (1940) Sulphanilamide as a specific inhibitor of carbonic anhydrase. Nature;146:164–165.