

Successful Virtual Screening for Novel Inhibitors of Human Carbonic Anhydrase: Strategy and Experimental Confirmation

Sven Grüneberg,[†] Milton T. Stubbs, and Gerhard Klebe*

Institute of Pharmaceutical Chemistry, University of Marburg, Marbacher Weg 6, D-35032 Marburg, Germany

Received November 26, 2001

Virtual screening of compound libraries is an alternative and complementary approach to high-throughput screening in the lead discovery process. A new strategy is described to search for possible leads of human carbonic anhydrase II, applying a protocol of several consecutive hierarchical filters involving a preselection based on functional group requirements and fast pharmacophore matching. A suitable pharmacophore is derived by a sophisticated “hot spot” analysis of the binding site to detect regions favorable for protein–ligand interactions. In subsequent steps, molecular similarity with known reference ligands is used to rerank the hits from the pharmacophore matching. Finally the best scored candidates are docked flexibly into the protein binding pocket. After examination of the affinity predictions, 13 compounds were selected for experimental testing. Of these 13, three could be shown to be subnanomolar, one is nanomolar, while a further seven are micromolar inhibitors. The binding mode of two hits could be confirmed by crystal structure analysis. The novelty of the discovered leads is best supported by the fact that a search in the patent literature showed the newly discovered subnanomolar compounds to comprise scaffolds not yet covered by existing patents.

Introduction

The discovery of novel innovative leads is the key element and starting point for any new drug discovery project. Genomics seeks to provide us with the sequences of all genes coding for proteins that make up the biology of a particular organism. Although this compilation does not directly reveal the functions of proteins that should be modulated in order to treat a specific disease, it is hoped that by means of proteomics and bioinformatics the most appropriate levels for drug intervention can be detected.¹ Accordingly, the number of possible targets for drug therapy is expected to increase from ca. 500 at present to 5000–10000.² To match this dramatic rate of drug target identification, it is necessary to develop new strategies to speed the discovery of novel putative leads.

At present, two alternative, yet complementary techniques are employed: experimental high-throughput screening (HTS) of large compound libraries and computational methods for virtual screening (VS) and de novo design. Enormous effort has been put into the large-scale automation of HTS.³ Together with combinatorial chemistry, which is able to increase compound synthesis by some orders of magnitude, this spelled the end of any rational and knowledge-based approach. However, initial euphoria that designated this technique as a universal lead generator has subsided as a result of the considerable costs involved and disappointingly low hit rates.⁴

The low hit rates are frequently due to inadequacies in quality and quantity of the compound libraries used

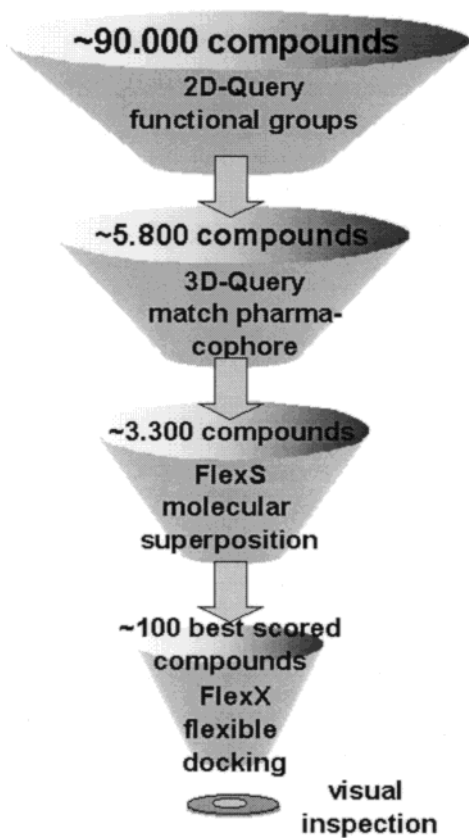
for testing. In addition, problems occur concerning false-positives and negative binding of test candidates. Quite depressing rates have been reported for the translation of apparently active hits from HTS into leads suitable for a subsequent optimization into drug candidates.⁴ Virtual screening represents an alternative, allowing selection of possible hits computationally.⁵ The requirements are opposite to those for HTS; the latter is technology-driven, largely ignoring the structural properties of the target, while VS depends on prior information about factors necessary for binding to the target. Thus, it is knowledge-driven. As a starting point, either the 3D structure of the target protein is available from crystallography, NMR, or homology modeling or a rigid reference ligand is known in its bioactive conformation, allowing development of a sophisticated pharmacophore model. In VS, compound screening does not consume valuable substance material synthesized in advance, yet even more importantly, the screening of libraries of virtual compounds can postpone synthesis of the most promising hits to a later stage. Experimental deficiencies such as limited solubility and other effects that can interfere with the assay conditions are not important on the first run. However, VS requires reliable methods for flexible docking or molecular superpositioning to define reasonable binding modes that predict binding affinities. The computed affinities are most critical and define the criteria for the final selection. To investigate the scope of VS for lead discovery, we selected human carbonic anhydrase II (hCAII) as a suitable target to develop a strategy for selecting possible new leads. As a proof-of-concept for a limited set of compounds, experimental enzyme inhibition was tested and the crystal structures of two potential hits were determined.

We have chosen hCAII for our case study for several reasons: (1) it was and still is a valid target in the

* To whom correspondence should be addressed. Phone: +49-6421-2821313. Fax: +49-6421-2828994. E-mail: klebe@mail.uni-marburg.de.

[†] Present address: Aventis Pharma Deutschland GmbH, DI&A Lead Generation, Computational Chemistry/Molecular Modeling, D-65926 Frankfurt am Main, Germany.

Scheme 1



treatment of glaucoma,⁶ and CAs are involved in a large variety of physiological processes that make them potential drug targets; (2) the crystal structure of hCAII has been solved to high resolution and several protein–ligand complexes have been determined; (3) when Relibase⁷ is used, a comparative analysis of 24 high-resolution complex crystal structures showed that the binding site is rather constant. In the present contribution, we describe a strategy to search for possible lead structures, applying a protocol of several consecutive hierarchical filters. The initial database comprised about 90 000 entries (Scheme 1). In the first selection step, putative ligands were chosen that bear functional groups to anchor zinc and agree with a protein-derived pharmacophore model. Subsequently molecular similarity with known reference ligands and flexible docking to the target protein served as sequential filters to reduce the initial set to about 100 prospective entries. On the basis of the resulting scoring and ranking of the selected hits and a final visual inspection, 13 compounds were selected for testing. To control the performance of our strategy, 35 known hCAII inhibitors were added to the search sample whose appearance in the hit list served to calibrate and validate the approach at each stage.

Several VS studies have already been described.⁸ However, most of them are postmodeling sessions, assessing the success of the approach by analyzing the achieved enrichment factors of known actives in a set of candidate molecules. Although some studies have assayed the suggested hits, final proof-of-concept by experimental means requires the crystallographic confirmation of the suggested binding modes. The issue of using enrichment factors as some kind of figure-of-merit

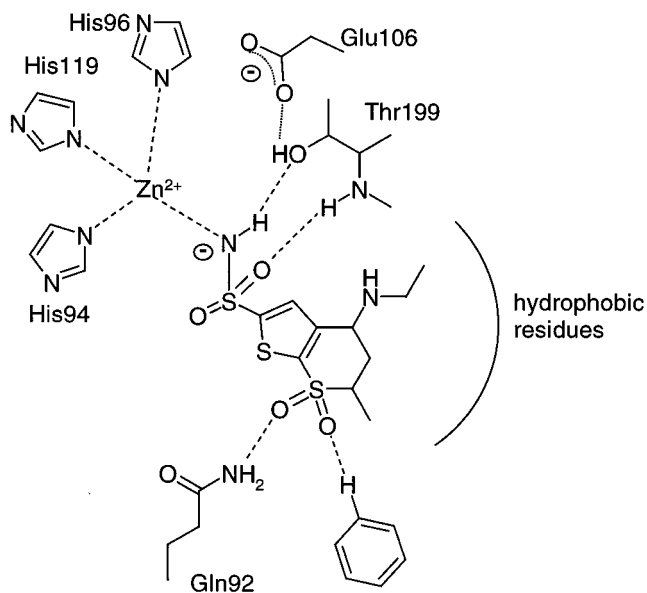


Figure 1. Schematic drawing of the binding mode of dorzolamide to hCAII

in VS has been controversial.^{9,10} Usually, for this purpose, rather potent compounds (nanomolar binders) are selected to compose a test sample. The performance of a VS run is then assessed with respect to the success in retrieving these test compounds as the most prominent VS hits. It has been debated whether the retrieval of potent known hits among a large background of weak binders or nonbinders is an indicative test.¹⁰ Furthermore, and the present study highlights this dilemma, the ratio of the number of prospective hits found by VS should increase from filter step to filter step in a strategy involving a consecutive hierarchical filtering. Thus, the enrichment factor must drop gradually because the known potent compounds in the original test sample must compete at later stages with the successfully retrieved VS hits for the best ranks on a hit list. Accordingly, enrichment factors should only be applied on early filtering steps of the retrieval; too strong focus on the recovery of known actives at later stages can be counterproductive for the discovery of VS hits.

Data Analysis and Results

Binding Site of the Target Protein. Human carbonic anhydrase II (hCAII, EC 4.2.1.1) is a metalloenzyme that catalyzes the reversible hydration of CO₂ to HCO₃⁻. This single domain protein is composed of 260 amino acids with a central stranded sheet. The catalytic zinc is coordinated by the imidazole moieties of three histidines located deep in a conical amphiphilic binding pocket (Figure 1). In the ligand-free state, water molecules fill the active site. One of them, presumably present as OH⁻, occupies the fourth coordination site at zinc.¹¹ It is also hydrogen-bonded to Thr199, which in turn is assumed to form a hydrogen bond to the substrate CO₂. An additional water, W338 (the so-called “deep water”), is found in the ligand-free state and replaced in the enzymatic reaction by CO₂.¹²

Presently, 14 different isoenzymes of human carbonic anhydrases are known, localized in different compartments of the organism.¹³ One isoform of hCAII is concentrated in the ciliary body of the eye, where it is

responsible for aqueous production. Glaucoma leads to reduced removal of ocular water, so the intraocular pressure increases and thereby damage of the optic nerve can occur. Potent inhibitors of hCAII thus offer a possible means to reduce intraocular pressure.

The enzyme is weakly inhibited by anions such as HS^- , CN^- , NCO^- , N_3^- , HSO_3^- , I^- , and HCOO^- .¹⁴ However, the most potent inhibitors all have in common a terminal sulfonamide group connected to an aromatic or heterocyclic portion (Figure 1). The likely deprotonated RSO_2NH^- group binds via its nitrogen to the tetrahedrally coordinated zinc, replacing the previously bound OH^- with the remaining proton at the sulfonamide group, forming a hydrogen bond to Thr199. One of the sulfonamide oxygens interacts with the backbone NH of Thr199, whereas the second oxygen replaces deep water W338. The acidity of the sulfonamide group has predominate influence on binding. Whereas $\text{CH}_3\text{SO}_2\text{NH}_2$ ($\text{p}K_a = 10.5$) inhibits hCAII at 3 mM, the more acidic $\text{CF}_3\text{SO}_2\text{NH}_2$ ($\text{p}K_a = 5.8$) inhibits at nanomolar levels.^{15,16} Inhibitors with an extended hydrophobic moiety form hydrophobic contacts to Val121, Phe131, Val143, Leu198, and Trp209. The nearby Gln92 serves as an additional anchor for several inhibitors, fixing the ligand via hydrogen bonding. Crystallographically, the binding modes of acetazolamide,¹⁷ dorzolamide,¹⁸ and brinzolamide¹⁹ have been determined. In addition to sulfonamides, only hydroxamates²⁰ have been identified as a zinc-coordinating group in inhibitors, with reduced binding affinity for hCAII. In total, 24 crystal structures of hCAII complexes deposited in the PDB²¹ served for our initial data analysis. The binding pocket is rather rigid, with only His64 exhibiting alternative conformations. Mobility of this residue is linked to its functional role in the enzyme proton relay mechanism.

Upon ligand binding, several water molecules are displaced from the binding site. On the other hand, a group of water molecules are found recurrently in most of the considered crystal structures at the same positions. A direct comparison of complex and apo structures using the water database in Relibase²² identifies four water molecules (1cil:²³ W301, W325, W393, W394) as the most likely to remain in the binding site upon ligand binding. This analysis is based on criteria such as B factors, local hydrophobicity, and number of contacting hydrogen-bonding partners in the protein.²² W393 interacts with the backbone NH of Asn62 and Asn67, W301 interacts with the OH of Tyr7 and the backbone carbonyl of His64, W394 is bound to Glu69, and W395 mediates an interaction between W393 and W301. Interestingly enough, none of these water molecules are involved in hydrogen bonding to ligands in any of the crystallographically resolved complex structures. To further back up and confirm the assumed water hydrogen bond network, an analysis using the MAB/Moloc²⁴ force field was performed. As a result of this analysis, these favorably placed water molecules were included in our final docking, in particular to add to the steric restriction of the binding pocket. Of course, this consideration is only valid as long as the ligands all occupy a comparable volume of the binding site. This aspect is important during later stages of ligand docking in a filtering step.

Preprocessing of the Ligand Data. All ligands considered in this study have been transformed from 2D to 3D via the program CORINA.²⁵ Protonation states have been assumed in the standard setting as suggested by CORINA (e.g., carboxylates deprotonated, amines protonated). Partial charges have been attributed to the atoms according to the method of Gasteiger and Marsili²⁶ as implemented in SYBYL.²⁷ Finally the data were converted in the UNITY system using the conversion routines supplied by Tripos.²⁸

Selection of the Test Set and Initial Filtering To Satisfy the Lipinski Rules. A set of 35 characterized hCAII inhibitors were included in the search sample at each filtering step. At the early steps of our search, the appearance of these compounds in the list of prospective hits served to calibrate and control our methods. The members of the test set are listed in Table 1. Among them, 13 inhibitors were considered whose binding mode had been determined crystallographically. Another 22 high-affinity ligands (IC_{50} , K_i , K_d values of 40–1.2 nM) were included to increase the structural diversity of the reference set. Furthermore, two micromolar hydroxamate inhibitors were included to incorporate this compound class. The molecular weight within the test sample covered a range between 75 and 478 Da; the computed MlogP²⁹ varied between -1.46 and 1.90. The molecules comprised between 2 and 16 rotatable bonds and the number of H-bond donors or acceptors was between 1–4 and 2–7, respectively. Furthermore, Tanimoto similarity indices^{30,31} (scaled to values between 0 and 100) were computed, taking either dorzolamide or acetazolamide as reference (Table 1).

Search entries for our analysis were taken from the Maybridge (61,186 entries)³² and the LeadQuest (37,841 entries)³³ databases. Application of the criteria defined by the Lipinski rules³⁴ (≥ 5 H-bond donors (no. of OH and NH groups), ≥ 10 H-bond acceptors (no. of O or N atoms), $\text{MW} \leq 500$ Da, $\text{MlogP} \leq 5$) reduced this initial set to 98 850 entries. The minimal reduction showed that both databases have already been preselected with respect to these criteria.

Preliminary Filtering for Zinc-Binding Anchor Groups. Coordinating to zinc appears to be of predominant importance for binding to hCAII. Accordingly, we preselected the initial data sample for those entries containing one of the terminal functional groups listed in Table 2. These groups had already been described as putative zinc binders in other zinc proteases. The extraction was performed using UNITY,²⁸ resulting in a reduction to 5904 entries.

Protein-Derived Pharmacophore. As a next step in our hierarchical filtering, the ligands of the sample set had to match a predefined pharmacophore criterion. Such a precondition was directly extracted from the requirements imposed by the binding site. To highlight those areas of the binding site where a putative ligand can favorably interact with the protein, we applied LUDI,³⁵ GRID,³⁶ SuperStar,³⁷ and DrugScore³⁸ as alternative methods. LUDI places potential interaction sites for hydrogen-bond donors and acceptors and hydrophobic groups within the binding pocket according to a set of rules derived from composite crystal-field environments observed in the crystal packing of small organic molecules.^{39,40} SuperStar likewise identifies

Table 1. Compounds Used in the Test Set, with Chemical Formula, Name, PDB Code If Crystal Structure Available, Affinity Data, Tanimoto Index with Respect to Dorzolamide (First Line) or Acetazolamide (Second Line), Calculated MlogP, and Literature Reference

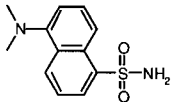
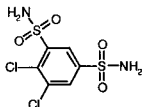
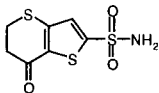
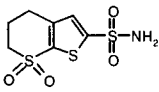
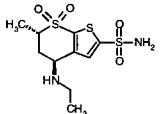
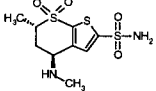
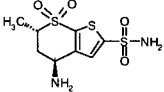
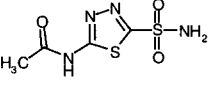
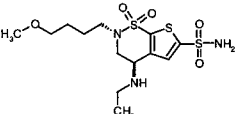
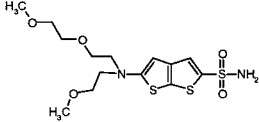
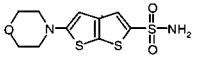
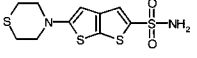
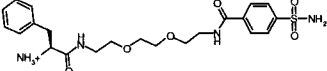
Compound	name	PDB	$K_i / K_d / IC_{50}$	Tanimoto Dorzolamide Acetazolamide	MlogP	lit.
	Dansylamide MNS	1okl	$K_d = 0.93 \mu\text{M}$	22.2 17.4	1.60	[56]
	Dichlorophenamide _DCPA	—	$IC_{50} = 40.0 \text{ nM}$	21.0 19.5	0.19	[57]
	_VTS	—	$K_i = 1.15 \text{ nM}$ $IC_{50} = 3.5 \text{ nM}$	45.6 21.4	0.11	[58]
	_STS	—	$K_i = 3.3 \text{ nM}$ $IC_{50} = 4.5 \text{ nM}$	77.4 22.3	-1.02	[58]
	Dorzolamide Trusopt™ MK-507 ETS	1cil	$K_i = 0.372 \text{ nM}$ ($IC_{50} = 0.18 \text{ nM}^{28}$)	100 25.2	-0.84	[59] [18]
	MTS	1cim	$K_i = 1.88 \text{ nM}$	98.9 25.0	-1.14	[18]
	PTS	1cin	$K_i = 1.52 \text{ nM}$	93.8 24.4	-1.46	[18]
	Diamox Acetazolamide AZM	1yda	$K_i = 6.0 \text{ nM}$ ($K_i = 22.0 \text{ nM}^{35}$) ($IC_{50} = 3.4 \text{ nM}^{28}$)	25.2 100.0	-1.58	[61]
	Brinzolamide Azopt AL-4862 _Azopt	1a42	$IC_{50} = 3.2 \text{ nM}$ ($K_d = 0.13 \text{ nM}$) Boriack-Sjodin, 1998 Prot. Science)	63.2 26.7	-0.96	[62]
	No. 55 _TIO55	—	$K_i = 1.58 \text{ nM}$ $IC_{50} = 2.8 \text{ nM}$	39.3 27.4	-0.22	[63]
	No. 51 _TIO51	—	$K_i = 1.1 \text{ nM}$ $IC_{50} = 3.0 \text{ nM}$	39.4 26.9	0.19	[63]
	No. 52 _TIO52	—	$K_i = 0.44 \text{ nM}$ $IC_{50} = 2.6 \text{ nM}$	40.3 26.8	0.98	[63]
	EG ₃ PheNH ₃ ⁺ EG3	1cny	$K_d = 14.0 \text{ nM}$	36.7 24.3	0.07	[64]

Table 1. (Continued)

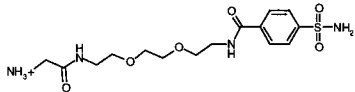
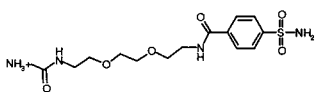
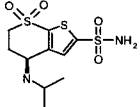
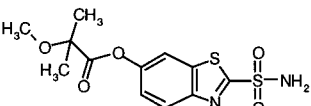
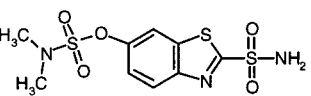
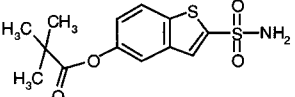
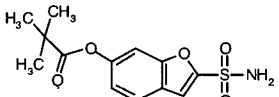
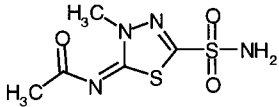
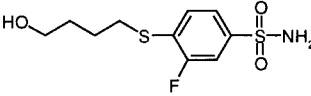
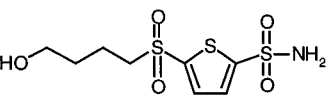
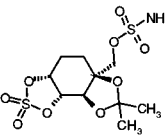
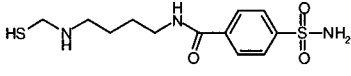
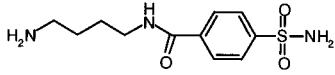
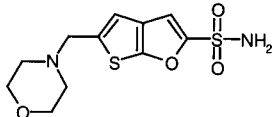
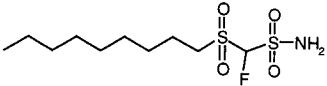
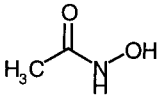
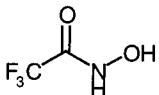
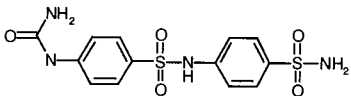
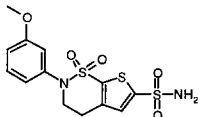
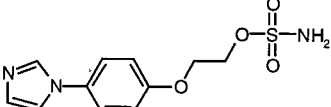
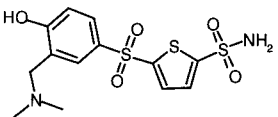
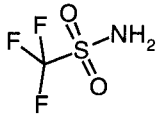
Compound	name	PDB	K_i / K_d / IC_{50}	Tanimoto Dorzolamide Acetazolamide	MlogP	lit.
	EG ₃ GlyNH ₃ ⁺ EG1	1cnw	$K_d = 19.0$ nM	30.6 26.5	-1.33	[64]
	EG ₃ NH ₃ ⁺ EG2	1cnx	$K_d = 43.0$ nM	30.8 26.7	-0.70	[64]
	(MK-927) MK-417 (S) _MK-927	–	$K_i = 0.635$ nM $IC_{50} = 0.54$ nM	94.5 26.2	-0.84	[60]
	No.39 _BTZ39	–	$IC_{50} = 1.2$ nM	30.0 28.7	-0.10	[65]
	No.49 _BTZ49	–	$IC_{50} = 1.2$ nM	29.0 30.3	0.22	[65]
	No.25 _BTZ25	–	$IC_{50} = 2.6$ nM	38.3 22.8	1.69	[57]
	_BFU	–	$IC_{50} = 3.0$ nM	28.3 18.0	0.92	[66]
	Methazolamid MZA _MZA	–	$IC_{50} = 21.0$ nM	23.7 23.4	-1.70	[57]
	_SSF	–	$IC_{50} = 1.5$ nM	29.2 16.8	1.60	[67]
	_SSS	–	$IC_{50} = 2.3$ nM	56.3 23.7	-1.46	[67]
	RWJ-37947 _RWJ	–	$IC_{50} = 21-130$ nM	18.2 14.4	1.90	[68]
	STB	1okm	–	33.9 26.5	0.52	[69]
	SAB555 SAB	1okn	–	34.3 26.0	0.24	[69]

Table 1. (Continued)








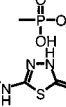


Compound	name	PDB	$K_i / K_d / IC_{50}$	Tanimoto Dorzolamide Acetazolamide	MlogP	lit.
	11e _TFU	–	$K_i = 1.24$ nM $IC_{50} = 5.6$ nM	35.0 20.8	-0.55	[70]
	_SFS	–	$IC_{50} = 3.0$ nM	23.4 18.2	0.10	[71]
	HAE555 _HAE	1am6	$IC_{50} = 47$ μ M	6.2 13.2	-0.81	[20]
	_HYF	–	$IC_{50} = 3.8$ μ M	6.8 11.6	-0.06	[20]
	_SCO	–	$IC_{50} = 7.0$ nM	23.8 20.9	0.33	[72]
	Al-7182 _AL-7182	1bnn	$K_d = 0.10$ nM	51.4 24.2	0.35	[19]
	AHR-16329 _AHR	–	$K_i = 7.0$ nM	19.3 16.6	0.37	[73]
	L-662,583 _TPSA	–	$IC_{50} = 0.7$ nM	53.0 26.8	-0.45	[74]
	_TFM	–	$K_i = 2.0$ nM	10.9 14.8	-1.05	[15]

preferred interaction sites in proteins by matching scatter plots or associated propensity maps onto exposed residues. The SuperStar maps show experimentally observed distributions around one central functional group in the spatial arrangement with other contacting groups. GRID is based on a force-field approach. The interaction potential is calculated for a variety of different probes at the intersections of a regularly spaced grid embedded in the binding pocket. Similar to GRID, DrugScore determines the putative contact preferences between ligand and protein atoms on a grid.⁴¹ Its potential function is based on the "inverse Boltzmann principle" evaluating the binding pocket according to frequencies for which a particular atom type of a

potential ligand is found in contact with a protein in known protein–ligand complexes.

The pocket of hCAII was systematically searched for favorable interactions with the following functional groups: C=O (H-bond acceptor), an amide NH group (H-bond donor), and a CH₃ group to describe hydrophobic interactions. In the case of GRID, a DRY probe has been selected for the analysis. In DrugScore, the following probes were used: O.3 and O.2 type oxygens (Sybyl atom type annotation²⁷) as acceptor, N.3 and N.ar type nitrogen as donor, and C.2, C.3, and C.ar for hydrophobic contacts. In addition, all lipophilic ring moieties observed in crystal structures of protein–ligand complexes with hCAII were superimposed and

Table 2. List of Anchor Groups Anticipated as Possible Zinc-Binding Groups, with SLN Code Given

Anchor	SLN-code
	C(=O)NH(OH)
	N(OH)C(=O)H
	C(=O)OH
	S(=O)(=O)NH2
	C(=O)NH2
	P(=O)NH2
	P(=O)(OH)OH
	N[1]NHC(SC=@1NHC)=S
	NH2NHC(N)=S
	S(N)(N)

extracted to highlight hydrophobic ligand portions. The different methods reveal qualitatively similar regions yet differ quantitatively because of differences in their functional form and relative weighting.

In Figure 2, the binding mode of dorzolamide is shown together with the contour maps of an H-bond acceptor (Figure 2a–c) and a donor (Figure 2d–f). The acceptor property shows a strong contour next to the zinc for all methods, where a H bond is formed from Thr199 NH to the inhibitor (Figure 1). In addition, an interaction with a ligand acceptor atom is predicted next to Gln92. Dorzolamide orients one of its sulfone oxygens into this area (Figure 1). The additional hot spot next to Asn62 and Asn67, revealed by GRID and SuperStar, corresponds to an area where two of the conserved water molecules (W325, W393) are found. Favorable interactions to a putative donor group are predicted next to zinc and Thr199 OH by all methods. For dorzolamide, the amino group of the sulfonamide coincides with this area. In addition, GRID and SuperStar highlight an area next to Asn67, where the conserved water molecules are localized.

In the case of hydrophobic interactions (Figure 2g–i), favorable ligand binding regions are predicted in a long extended area next to Leu198, Val121, Val143, Phe131, and Ile91. The bicyclic ring system of dorzolamide partially coincides with this area. The binding mode of other inhibitors (Figure 3) shows that the rendered area is accommodated. Furthermore, all methods indicate a hydrophobic pocket that is occupied in the case of dorzolamide by its ethylamino side chain. DrugScore differs from GRID and SuperStar in that only two compact areas are indicated as most favorable for hydrophobic interactions.

Translation of the indicated hot spots for H-bonding properties into a pharmacophore pattern is straightforward. Generic pharmacophore sites have been assigned to the geometric centers of the contoured "hot spots" (Figure 4). The localization of an area suitable for hydrophobic contacts is more problematic. We decided

to merge the information from the different contouring methods with that of the known binding modes (Figure 3). The ultimate pharmacophore model submitted to UNITY comprised five centers of deviating spherical tolerance (Figure 4). An acceptor property was placed next to Gln92, a second was adjacent to Thr199 NH, while a donor property was requested close to Thr199 OH. Two adjacent spheres were assigned to approximate the elliptical region favorable for hydrophobic interaction. Two of the centers were requested to be matched optionally. With this condition, all 35 inhibitors of the test set were retrieved successfully from the residual set of 5904 compounds.

3D Pharmacophore Query with Unity and Similarity Scoring with FlexS. Of the 5904 compounds selected in the first step, 3314 entries satisfied the pharmacophore query in UNITY through a flexible search. This remaining set of compounds was reranked by superimposing all entries with two potent hCAII inhibitors and computing their similarity scoring with respect to these reference molecules. FlexS⁴² flexibly superimposes a test ligand with a reference through optimizing the overlap of physicochemical properties described by associated Gaussian functions in space. The reference could either be a single molecule or the average over the associated properties of several molecules. Furthermore, we added a feature to FlexS to incorporate the spatial orientation of flanking binding-site residues in terms of repulsive centers (C. Lemmen, personal communication). The associated repulsive Gaussian functions restrict the adopted binding mode of particularly large inhibitors to the space actually accessible for ligands.⁴³ The merged reference was constructed from the crystallographically given binding modes of dorzolamide (1cil)²³ and E63 (1cny).⁴⁴ The latter inhibitor was selected for its extended shape, guiding the superpositioning of larger inhibitors toward the entrance of the binding pocket. FlexS computed a superposition for 2237 of the previously selected 3314 compounds. The thus-computed similarity score of the two reference molecules was subsequently used to order the obtained hits. To avoid arbitrary size dependencies in the similarity scores, we introduced a size normalization as described previously.⁴⁵ By adding the 35 test compounds to the analysis, we could show that nearly half of the test examples were ranked among the first 10% of the entire data sample. This finding convinced us that our strategy was efficient in retrieving and scoring the most favorable candidates from our previous UNITY pharmacophore search.

Final Docking with FlexX. In a final step, the 100 best-ranked hits from the FlexS filtering were docked into the binding pocket using FlexX.⁴⁶ Among these, half of the hits contained a terminal sulfonamide group. During docking, the previously "conserved" water molecules were considered as part of the protein to restrict the binding pocket sterically. Furthermore, the MapRef mode in FlexX was applied to achieve proper tetrahedral zinc coordination. The final docked binding modes in hCAII were ranked for their expected binding affinity either by the regression-based scoring function implemented in FlexX or by DrugScore. Visual inspection of the suggested binding modes of FlexX, together with the scoring values of FlexX, FlexS, and DrugScore, was used

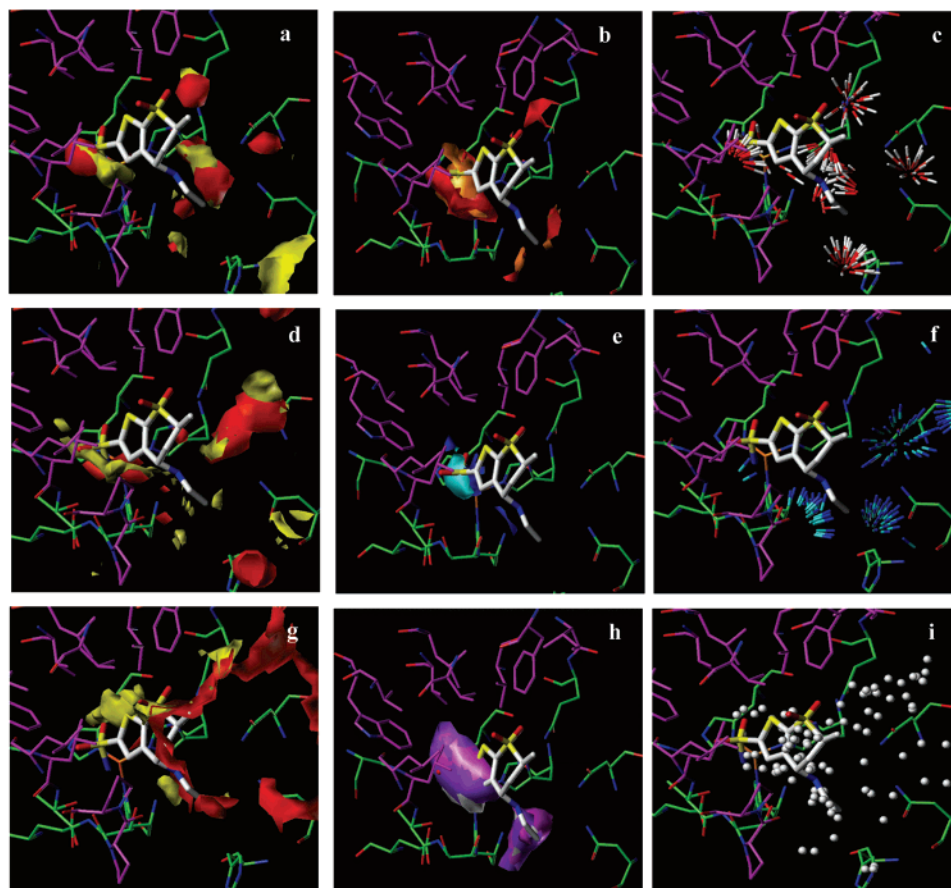


Figure 2. Mapping of putative binding “hot spots” in the active site of hCAII. Values 10% above the global minimum of the map have been contoured. For orientation, the binding geometry of the inhibitor dorzolamide, studied crystallographically, is also shown. Parts a–c highlight the properties of a putative hydrogen-bond acceptor group in a ligand: (a) carbonyl oxygen as probe in SuperStar (red) and in GRID (yellow); (b) DrugScore with O.2 (red) and O.3 (orange) as probes; (c) LUDI sites generated for a C=O group. Parts d–f highlight the properties of a putative hydrogen-bond donor group in a ligand: (d) NH amino group as a probe in (red) SuperStar and in (yellow) GRID; (e) DrugScore with N.3 (blue) and N.am (cyan) as probes; (f) LUDI sites generated for a N–H group. Parts g–i highlight the properties of hydrophobic groups in a putative ligand: (g) CH₃ group in SuperStar (red) and DRY probe in GRID (yellow); (h) DrugScore with C.2 (white), C.3 (violet), and C.ar (magenta) as probes; (i) LUDI sites generated for hydrophobic carbon atoms.

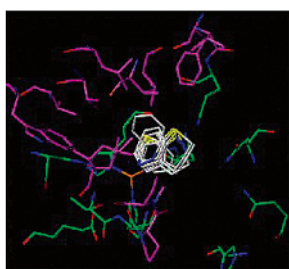


Figure 3. Superposition of the crystal structures of hCAII complexes (listed in Table 1 according to their PDB codes) containing ligands with aromatic/aliphatic five- and six-membered rings. For clarity, only the atoms of the hydrophobic ring systems are shown. Their spatial location has been used to place hydrophobic site points in a pharmacophore assignment.

to select a small set of compounds for ordering and subsequent biological testing. The following criteria were considered in the visual inspection to further select prospective hits from the list: (1) degree of occupancy of the amphiphilic binding pocket next to zinc, in particular with respect to the achieved surface complementarity between protein and ligand; (2) ligands with a minimal number of rotatable bonds to avoid entropically unfavorable binding due to pronounced conforma-

tional immobilization; (3) quality of the overall binding conformation generated by comparing with conformational preferences extracted from small-molecule crystal structures (Mimumba approach⁴⁵); (4) formation of hydrogen bonds to Glu92 or Thr200 to compensate for desolvation of these residues in the binding pocket. Finally, we focused on zinc anchoring groups that had already been observed to inhibit hCAII. In a recent comparative crystallographic study, we could show that the sulfonamide group (XSO₂NH) appears to be an ideal anchor for hCAII inhibitors⁴⁷ because of its stereochemical organization and its appropriate pK_a value. (Figure 1). Nevertheless, we also included two hydroxamates, since this class of compounds has been shown to bind at least with moderate affinity.²⁰ Our selection resulted in the 13 compounds listed in Table 3, of which 10 were from LeadQuest and the remaining three from Maybridge. They range in molecular weight from 201 to 385 Da, and the number of rotatable bonds varied from 3 to 9. The number of H-bond acceptors and donors varied from 2 to 6 and from 1 to 3, respectively. Topological similarities of the two established products dorzolamide and acetazolamide, determined using the Tanimoto index, ranged from 20% to 45% (dorzolamide) and from 10% to 30% (acetazolamide).

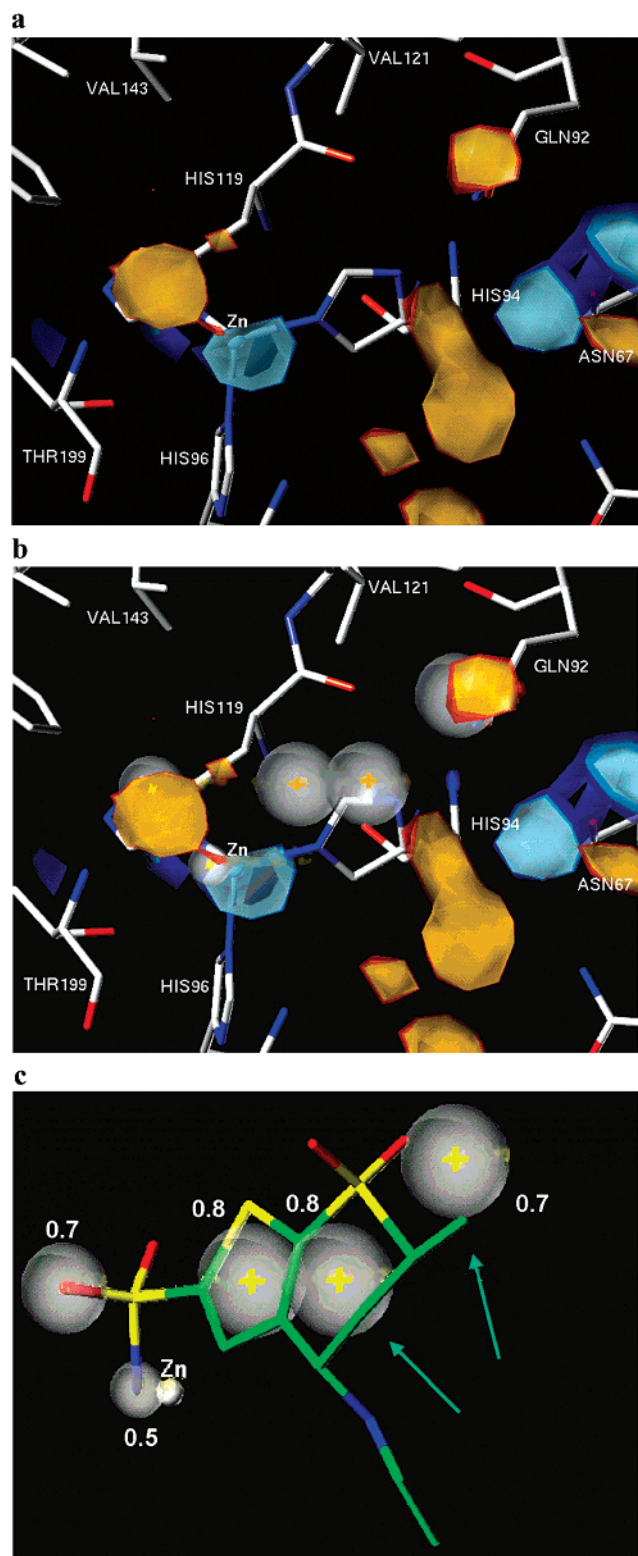


Figure 4. “Hot spots” indicated in the various maps in Figure 2 (top, H-bond acceptor (red/yellow) and donor (cyan/blue) properties) have been translated into a protein-based pharmacophore hypothesis (center). Spheres of adjustable size (gray, relative size given) have been assigned to the centers of the “hot spots”, and those indicated by arrows have been considered optionally in the subsequent UNITY search (bottom). To place the hydrophobic sites, data from Figure 3 have been included. In the UNITY search, putative ligands are considered as hits if their matching functional groups coincide with these pharmacophore spheres.

Biological Testing. hCAII for biological testing and subsequent crystallization was obtained from Sigma-Aldrich (C 6165). Photometric determination of IC_{50} values was performed as described by Pocker and Stone.⁴⁸ Acetazolamide was measured as a reference. An IC_{50} of 6.7 ± 0.9 nM was found in accordance with the reported value of 3.4 nM.⁴⁹ All values are given in Table 3.

Crystal Structure Determination. Suitable crystals of hCAII were grown under hanging drop conditions from a 1:1 mixture of 0.3 mM protein solution with 20 mM Tris-HCl and 20 mM NaCl at pH 8.0 and crystallization buffer (100 mM Tris-HCl, pH 8.5, 300 mM NaCl, 27 M $(NH_4)_2SO_4$, 5 mg of $CH_3HgO_2CCH_3$). Crystals were obtained at 6 °C over a period of 2–3 weeks. To soak the test compounds into hCAII crystals, saturated solutions of the test compounds in the crystallization buffer were made. After slow addition of 3 μ L of this solution to 14 μ L of the crystallization buffer containing hCAII crystals, the system was kept under hanging drop conditions over a reservoir of crystallization buffer for 3 days at 6 °C.

Suitable crystals of the ligand–protein complexes were mounted in glass capillaries and exposed to the X-ray beam on a Rigaku RU300 rotating anode generator using Cu K α radiation (50 kV/100 mA). Reflection intensities were registered on a RAXIS-IV image plate system (MSC) with a crystal-to-detector distance of 100 mm. Data were collected to a resolution of 2.4 or 2.3 Å and a completeness of 81.3% or 95.2%, respectively. Details of the data collection and refinement are given in Table 4. After data reduction with DENZO and SCALEPACK,⁵⁰ phases were calculated by using protein and water coordinates from 2cba.⁵¹ The protein and inhibitor structures of 2000-07780 and 0134-36 (Table 3) were refined, using X-PLOR,⁵² to an $R_{(F)}$ value of 0.192 for 8012 reflections with $F > 2\sigma$ (standard deviation of bond lengths, 0.007 Å; standard deviation of bond angles, 1.64°) for the first example or $R_{(F)} = 0.196$ for 10 124 reflections with $F > 2\sigma$ (0.008 Å, 1.44°) for the second complex. The atomic coordinates of the two crystal structures have been deposited with the Protein Data Bank (PDB) and can be accessed through the PDB file names 1kwq (complex with 2000-07790) and 1kwr (complex with 0134-36).

Complex Geometries of hCAII with Inhibitors 2000-07790 and 0134-36. The ligand 2000-07790 coordinates to zinc (Figure 5a) via its likely deprotonated terminal sulfonamide nitrogen (1.8 Å). In addition, this group donates a hydrogen bond to Thr199O γ (3.3 Å). The adjacent SO₂ group accepts a hydrogen bond from the backbone NH (3.1 Å). The nitro group at the central phenyl portion is oriented into a hydrophobic pocket formed by the side chains of Val121, Phe131, Val135, Leu141, and Val143. The terminal pyrrolidinone moiety adopts an orientation nearly perpendicular to the phenyl ring and forms a hydrogen bond to a water molecule (3.3 Å) accommodated in the binding pocket.

The binding geometry of 0134-36 with respect to zinc (Figure 5b) and Thr199 is very similar. The Zn–N distance of 1.8 Å is equivalent, whereas slightly shorter contacts to Thr199 O γ (2.9 Å) and Thr199 NH (2.9 Å) are detected. The thiazolinone moiety of the inhibitor

Table 3. List of the 13 Compounds from LeadQuest or Maybridge Discovered by Virtual Screening and Tested for hCAII Binding Together with the Experimentally Determined pIC₅₀ and the Computed Prediction by DrugScore

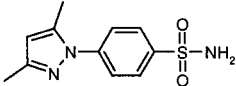
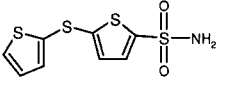
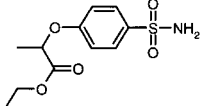
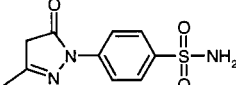
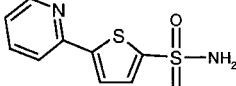
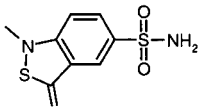
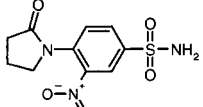
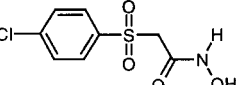
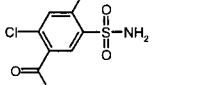
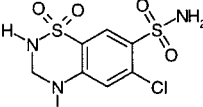
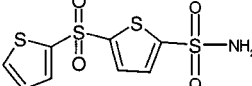
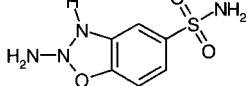
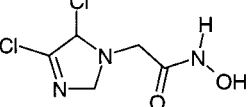
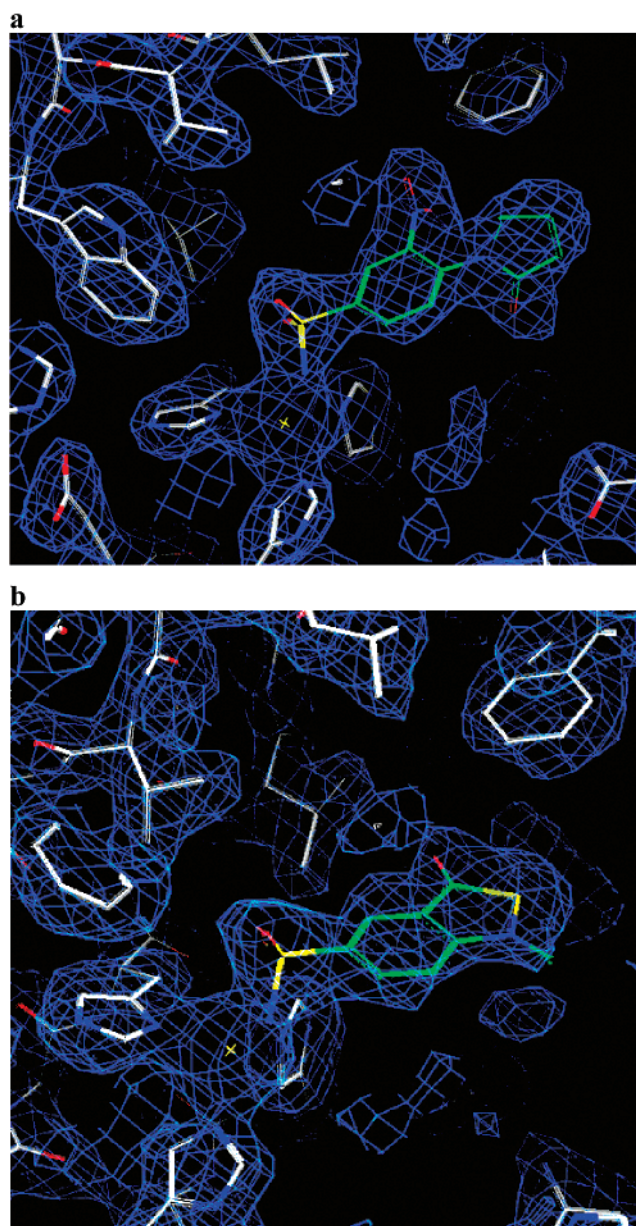
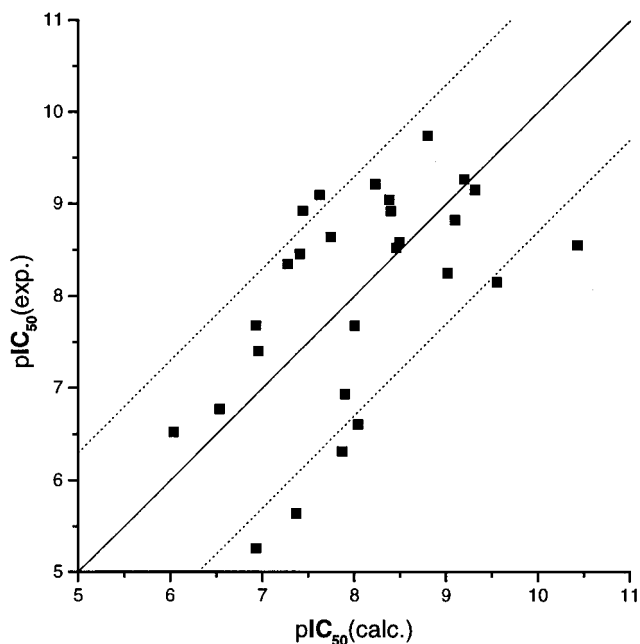
Compound	label	pIC ₅₀ (calc.)	pIC ₅₀ (exp.)	database
DRUGSCORE				
	0009-46	6.92	7.67	LeadQuest
	0049-01	8.38	9.05	LeadQuest
	0076-89	9.11	6.73	LeadQuest
	0114-01	7.87	6.31	LeadQuest
	0191-86	7.90	6.93	LeadQuest
	0134-36	8.04	6.61	LeadQuest
	2000-07790	8.23	9.21	LeadQuest
	0164-08	-	(>1mM)	LeadQuest
	2000-01951	6.54	6.77	LeadQuest
	JFD00715	6.93	5.26	Maybridge
	KM05745	7.62	9.09	Maybridge
	NRB01507	7.37	5.64	Maybridge
	KM01987	-	(>1mM)	Maybridge

Table 4. Crystallographic Data and Data Collection for the hCAII Complexes with 2000-07790 and 0134-36

crystal data	2000-07790	0134-36
space group	<i>P2</i> (1)	<i>P2</i> (1)
<i>a</i> (Å)	42.69	42.56
<i>b</i> (Å)	41.69	41.40
<i>c</i> (Å)	72.64	72.34
β (deg)	104.8	104.0
resolution (Å)	10–2.4	10–2.3
unique reflections	8012	10124
completeness (%)	81.3	95.2
$R_{(F)}$ free (%)	28.3	26.9
$R_{(F)}$ (%)	19.3	19.5
no. of waters	118	122
no. of Hg	1	0
angular rmsd (deg)	1.637	1.443
bond length rmsd (Å)	0.007	0.008

**Figure 5.** Crystallographically determined binding mode of 2000-07790 (top) and 0134-36 (bottom) in hCAII. The $2F_o - F_c$ density is shown on a 1σ contour level.

adopts an orientation that allows formation of a hydrogen bond to the terminal NH_2 group of Gln92 (2.8 Å).

**Figure 6.** Correlation of experimental and DrugScore predicted IC_{50} values for inhibitors taken from Tables 1 and 3.

Correlation of Measured and Predicted Binding Affinities. To assess the reliability of the predicted binding affinities, the calculated DrugScore ranks were correlated with the measured IC_{50} values for the 13 compounds evaluated in this study. The data were complemented by those entries from our original test for which IC_{50} values had been published. The regression analysis revealed $r^2 = 0.31$ with a standard deviation of 0.9 log units (Figure 6).

Discussion

In the present paper, we present a strategy for the computer screening of large compound libraries to obtain a limited set of prospective hits. Our approach successfully retrieved active compounds as revealed by IC_{50} values for 13 selected compounds and correctly predicted binding modes as shown by X-ray crystallography of two complexes. Of these 13, three could be shown to be subnanomolar and one nanomolar while a further seven were micromolar inhibitors. We considered entries from the LeadQuest and Maybridge database, both compiled from molecules that pass the conditions of the Lipinski rules. A larger quantity of hits was found in LeadQuest, highlighting this compound library as a relevant source for the discovery of novel hits as prospective leads. All selected compounds possessed a sulfonamide as zinc binding anchor; two with a hydroxamate have been considered. However, the latter exhibited only very reduced activity.

Until now, sulfonamides are the only functional group found to be suitable for high-affinity inhibition of hCAII. We have recently shown that a terminal sulfonamide group is ideally suited to coordinate the zinc while simultaneously donating an H bond to the neighboring Thr199 O_γ and accepting another from the backbone NH of this amino acid.⁴⁷ The electronic (charge and pK_a conditions) and geometric requirements for such binding will be difficult to match by other functional groups (Figure 1). In our virtual screening, we tried to expand

the scope to other zinc-binding anchors. However, already prior to our final docking step, half of the 100 candidate molecules contained a sulfonamide group. In addition, the final docking run confirmed this group as superior zinc anchor in hCAII. Although the sulfonamides were scored best in our VS, we decided to include two hydroxamides to our experimental testing. We believe the nearly exclusive retrieval of sulfonamides does not diminish the novelty of our discovered hits as potential new leads. First of all, the remaining skeleton besides the zinc-anchoring group still contributes significantly to the structure–activity relationship as demonstrated by a series of related sulfonamide inhibitors covering an affinity spread over several orders of magnitude. Selection of inappropriate side chains can destroy binding to hCAII. Furthermore, from a more pragmatic point of view, the actual discovery of new leads is best illustrated by the fact that a search in the patent literature showed that the subnanomolar compounds discovered comprise scaffolds that, to our knowledge, are not covered by existing patents.⁵³ Such scaffolds are generally accepted as novel leads in a drug discovery project.

We believe the success of the present VS arises from three aspects. First, we combined several approaches to map the structural prerequisites for high-affinity binding to hCAII. A consensus composite picture was compiled to retrieve “hot spots” of binding. Subsequently, these were translated into a pharmacophore query to be addressed in a UNITY search. Although in principle all retrieved candidate molecules could have been docked using FlexX into the binding site, a procedure often used in VS as high-throughput docking, this ignores the significant amount of structural knowledge about potential inhibitors already available for hCAII. The introduction of another filtering step provides a second advantage to our strategy. Superpositioning with FlexS in terms of molecular similarity with a composed reference, simultaneously regarding the approximate shape of the binding pocket by repulsive Gaussian functions, allowed the incorporation of information about the chemical structure of known leads. The concept of combining docking with simultaneous similarity considerations has recently been reported in other studies.^{37,54,55} Ideally in future developments, the knowledge about characterized ligands will be directly incorporated in docking by considering them in a tailored scoring function. Finally, we believe the application of DrugScore gave us a third advantage, not only for the determination of “hot spots” but also because DrugScore selects more reliably computed binding modes of FlexX, approximating a near-native geometry.³⁸

The detection of such ligand poses is essential for a reliable affinity scoring. The superior performance of DrugScore compared to the regression-based scoring function in FlexX is demonstrated by the crystal structures determined from two of the hits. The binding modes highlighted by DrugScore approximate the experimentally found structures much better than the original scoring function in FlexX. This function places these solutions only at rank 51 and 61, respectively. The best-scored solution for 2000-07790 (Table 1) from the regression-based ranking in FlexX deviates by rmsd =

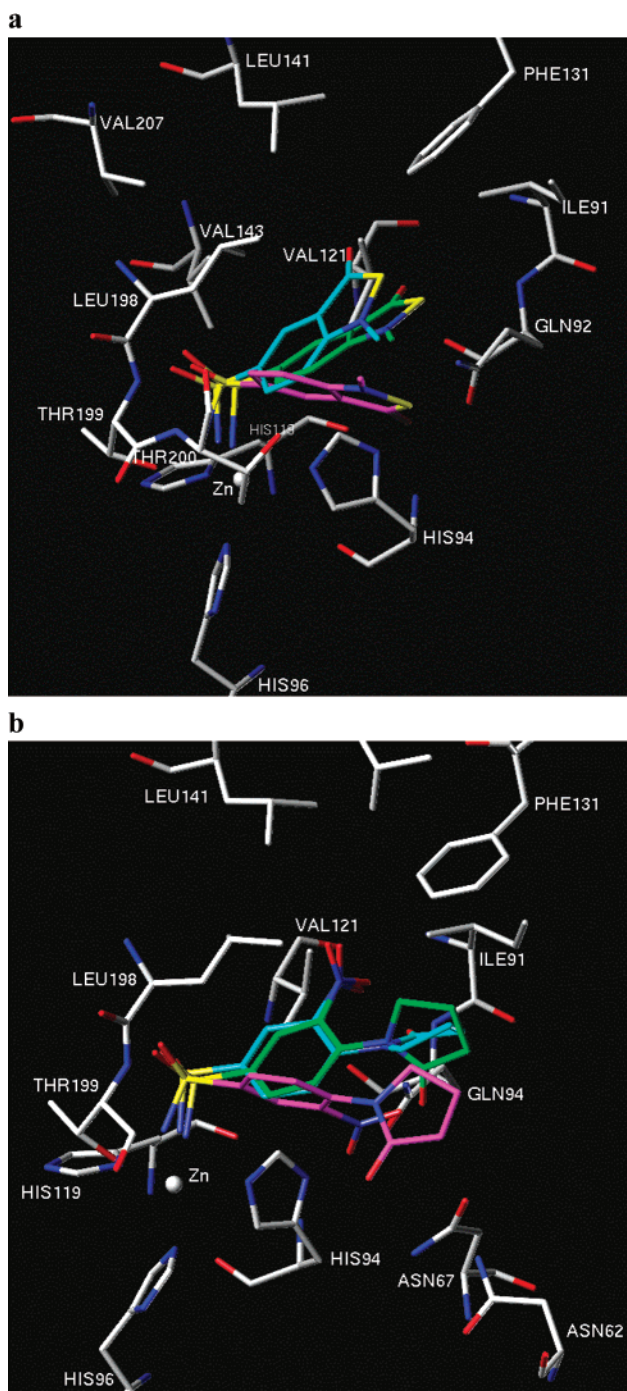


Figure 7. Binding modes of 2000-07790 (top) and 0134-36 (bottom) as found crystallographically (green) in hCAII. Furthermore, two docking solutions are shown: (magenta) solution ranked best by the original scoring function in FlexX (rmsd values to crystal structure are 2.7 Å (top) and 2.2 Å (bottom)); (cyan) solution ranked best by DrugScore (rmsd values to crystal structure are 1.4 Å (top) and 1.2 Å (bottom)).

2.7 Å from the crystal structure, whereas DrugScore suggested a solution that approximates experiment with rmsd = 1.4 Å. The residual deviation of the DrugScore solution shows a flip of the terminal pyrrolidinone ring of 150° compared to the crystal structure. Presumably, this difference results from the accommodation of an additional water molecule found in the crystal structure. This water was not considered in the docking run; however, it will clearly change the scoring of the complex.²² In the solution scored best by the regression-

based scoring function in FlexX, the entire ligand is tilted off from its position in the crystal structure. For the inhibitor 0134-36 (Table 1), the original FlexX scoring suggested a binding mode with 2.2 Å deviation, whereas DrugScore ranks best a solution with only 1.2 Å deviation. Again, the DrugScore solution approximates the experimental results better; however, the bicyclic thiazolinone moiety of the inhibitor is slightly tilted toward Leu198 and Leu204, apparently to optimize hydrophobic contacts to these residues. The short hydrogen bond found to the terminal amide of Gln92 (2.8 Å) is not realized in the docked solution.

Finally, we assessed the power of DrugScore to predict binding affinities. The correlation of measured and DrugScore-predicted affinities reveals, on the mean, an accuracy of about one logarithmic unit.⁴¹ This is probably too large to serve as the sole selection criterion. However, experience in the application of scoring functions shows that on the whole no better results can be achieved at present. To some degree, this results from the uncertainties in the experimental determination of binding data. However, we must also admit that the physical nature of ligand binding, in particular regarding the role of water in the process, is not yet fully understood.²² Undoubtedly, more profound insight into these phenomena is necessary for the development of an improved scoring function and the generation of better and more reliable binding modes, e.g., by considering the lacking water molecule next to the pyrrolidinone ring in the docked complex of 2000-07790 (Figure 7).

Acknowledgment. The authors are grateful to Tripos (Munich, Germany) for making their software and some of the test compounds available to us. The fruitful collaboration with Dr. B. Wendt (Tripos, Bude, U.K.) on the project is gratefully acknowledged. The present study was financially supported as part of the ReLiMo project by the German Minister of Education, Science, Research, and Technology under Grant No. 0311619. We thank Dr. M. Rarey (University of Hamburg, Germany) and Dr. C. Lemmen (BioSolvIt, Bonn, Germany) for some important program adaptations in FlexX and FlexS.

References

- Consortium, I. H. G. S. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921.
- Drews, J. Drug discovery: A historical perspective. *Science* **2000**, *287*, 1960–1964.
- Houston, J. G.; Banks, M. The chemical–biological interface: developments in automated and miniaturised screening technology. *Curr. Opin. Biotechnol.* **1997**, *8*, 734–740. Kenny, B. A.; Bushfield, M.; Parry-Smith, D. J.; Fogarty, S.; Treherne, J. M. The application of high-throughput screening to novel lead discovery. *Prog. Drug Res.* **1998**, *51*, 245–269. Hertzberg, R. P.; Pope, A. J. High-throughput screening: new technology for the 21st century. *Curr. Opin. Chem. Biol.* **2000**, *4*, 445–451.
- Lahana, R. How many leads from HTS? *Drug Discovery Today* **1999**, *4*, 447–448.
- Walters, W. P.; Stahl, M. T.; Murcko, M. A. Virtual screening—an overview. *Drug Discovery Today* **1998**, *3*, 160–178. Schneider, G.; Böhm, H. J. Virtual screening and fast automated docking methods. *Drug Discovery Today* **2002**, *7*, 64–70. Good, A. Structure-based virtual screening protocols. *Curr. Opin. Drug Discovery Dev.* **2001**, *4*, 301–307.
- Surgrue, M. F. Pharmacological and ocular hypotensive properties of topical carbonic anhydrase inhibitors. *Prog. Retinal Eye Res.* **2000**, *19*, 87–112.
- Hendlich, M. Databases for protein–ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 1178–1182.
- Mason, J. S.; Good, A. C.; Martin, E. J. 3-D pharmacophores in drug discovery. *Curr. Pharm. Des.* **2001**, *7*, 567–597. Klebe, G. *Virtual Screening: An Alternative or Complement to High Throughput Screening?*; Kluwer Academic: Dordrecht/Boston/London, 2000. Böhm, H. J.; Schneider, S. *Virtual Screening for Bioactive Molecules*; Wiley-VCH: Weinheim; 2000. Langer, T.; Hoffmann, R. D. Virtual screening: an effective tool for lead structure discovery? *Curr. Pharm. Des.* **2001**, *7*, 509–527. Lamb, M. L.; Burdick, K. W.; Toba, S.; Young, M. M.; Skillman, A. G.; Zou, X.; Arnold, J. R.; Kuntz, I. D. Design, docking, and evaluation of multiple libraries against multiple targets. *Proteins* **2001**, *42*, 296–318. Beroza, P.; Bradley, E. K.; Eksterowicz, J. E.; Feinstein, R.; Greene, J.; Grootenhuys, P. D.; Henne, R. M.; Mount, J.; Shirley, W. A.; Smellie, A.; Stanton, R. V.; Spellmeyer, C. Applications of random sampling to virtual screening of combinatorial libraries. *J. Mol. Graphics Modell.* **2000**, *18*, 335–342. Filikov, A. V.; Mohan, V.; Vickers, T. A.; Griffey, R. H.; Cook, P. D.; Abagyan, R. A.; James, T. L. Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 593–610. Perola, E.; Xu, K.; Kollmeyer, T. M.; Kaufmann, S. H.; Prendergast, F. G.; Pang, Y. P. Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads. *J. Med. Chem.* **2000**, *43*, 401–408. Chen, G. S.; Chang, C. S.; Kan, W. M.; Chang, C. L.; Wang, K. C.; Chern, J. W. Novel lead generation through hypothetical pharmacophore three-dimensional database searching: discovery of isoflavonoids as nonsteroidal inhibitors of rat 5 α -reductase. *J. Med. Chem.* **2001**, *44*, 3759–3763. Aronov, A. M.; Munagala, N. R.; Kuntz, I. D.; Wang, C. C.; Virtual screening of combinatorial libraries across a gene family: in search of inhibitors of *Giardia lamblia* guanine phosphoribosyltransferase. *Antimicrob. Agents Chemother.* **2000**, *45*, 2571–2576.
- Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035–1042.
- Charifson, P. S.; Corkerey, J. J.; Murcko, M. A.; Walters, W. P. Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking of Three-Dimensional Structures into Proteins. *J. Med. Chem.* **1999**, *42*, 5100–5109.
- Merz, K. M.; Banci, L. Binding Of Bicarbonate to Human Carbonic Anhydrase II: A Continuum of Binding States. *J. Am. Chem. Soc.* **1997**, *119*, 863–871. Silverman, D. N.; Lindskog, S. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc. Chem. Res.* **1988**, *21*, 30–36.
- Eriksson, A. E.; Jones, T. A.; Liljas, A. Refined Structure of Human Carbonic Anhydrase II at 2.0 Å Resolution. *Proteins* **1988**, *4*, 274–282.
- Supuran, C. T.; Scozzafava, A. Carbonic anhydrase inhibitors and their therapeutic potential. *Expert Opin. Ther. Pat.* **2000**, *10*, 575–600.
- Supuran, C. T.; Scozzafava, A. Carbonic anhydrase inhibitors and their therapeutic potential. *A. Expert Opin. Ther. Pat.* **2000**, *10*, 575–601. Supuran, C. T.; Scozzafava, A. Carbonic anhydrase inhibitors. *Curr. Med. Chem.: Immunol., Endoc. Metab. Agents* **2001**, *1*, 61–97. Christianson, D. W.; Cox, J. D. Catalysis by metal-activated hydroxide in zinc and manganese metalloenzymes. *Annu. Rev. Biochem.* **1999**, *68*, 33–49. Supuran, C. T.; Conroy, C. W.; Maren, T. H. Is cyanate a carbonic anhydrase substrate? *Proteins: Struct., Funct., Genet.* **1997**, *27*, 272–278. Briganti, F.; Mangani, S.; Scozzafava, A.; Vernaglion, G.; Supuran, C. T. Carbonic anhydrase catalyzes cyanamide hydration to urea: is it mimicking the physiological reaction? *J. Biol. Inorg. Chem.* **1999**, *4*, 528–536. Guerri, A.; Briganti, F.; Scozzafava, A.; Supuran, C. T.; Mangani, S. Mechanism of cyanamide hydration catalyzed by carbonic anhydrase II suggested by cryogenic X-ray diffraction. *Biochemistry* **2000**, *39*, 12391–12397. Liljas, A.; Hakansson, K.; Jonsson, B. H.; Xue, Y. Inhibition and catalysis of carbonic anhydrase. Recent crystallographic analyses. *Eur. J. Biochem.* **1994**, *219*, 1–10. Lindskog, S.; Silverman, D. W. In *The Carbonic Anhydrases—New Horizons*; Chegwidan, W. R., Edwards, Y., Carter, N., Eds.; Birkhäuser Verlag: Basel, Switzerland, 2000; pp 175–196.
- Maren, T. H.; Wiley, C. E. The vitro activity of sulfonamides against red cell carbonic anhydrase. Effect of ionic and substrate variation on the hydration reaction. *J. Med. Chem.* **1968**, *11*, 228–232.
- Maren, T. H.; Conroy, C. W. A new class of carbonic anhydrase inhibitor. *J. Biol. Chem.* **1993**, *268*, 26233–26239.
- Vidgren, J.; Liljas, A.; Walker, N. P. C. Refined Structure of the Acetazolamide Complex of Human Carbonic Anhydrase II at 1.9 Å. *Int. J. Biol. Macromol.* **1990**, *12*, 342–344.
- Smith, G. M.; Alexander, R. S.; Christianson, D. W.; McKeever, B. M.; Ponticello, G. S.; et al. Positions of His-64 and a bound water in human carbonic anhydrase II upon binding three structurally related inhibitors. *Protein Sci.* **1994**, *3*, 118–125.
- Boriack-Sjodin, P. A.; Zeitlin, S.; Chen, H. H.; Crenshaw, L.; Gross, S.; et al. Structural analysis of inhibitor binding to human carbonic anhydrase II. *Protein Sci.* **1998**, *7*, 2483–2489.

- (20) Scolnick, L. R.; Clements, A. M.; Liao, J.; Crenshaw, L.; Hellberg, M.; et al. Novel binding mode of hydroxamate inhibitors to human carbonic anhydrase II. *J. Am. Chem. Soc.* **1997**, *119*, 850–851.
- (21) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (22) Günther, J. Ph.D. Thesis, University of Marburg, Marburg, Germany, 2002.
- (23) Smith, G. M.; Alexander, R. S.; Christianson, D. W.; McKeever, B. M.; Ponticello, G. S.; Springer, J. P.; Randall, W. C.; Baldwin, J. J.; Habecker, C. N. Positions of His-64 and a bound water in human carbonic anhydrase II upon binding three structurally related inhibitors. *Protein Sci.* **1994**, *3*, 118–125.
- (24) Gerber, P. R.; Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 251–268.
- (25) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic Generation of 3D-Atomic Coordinates for Organic Molecules. *Tetrahedron Comput. Methodol.* **1990**, *3*, 537–547.
- (26) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity—A Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219–3228.
- (27) SYBYL Molecular Modelling Package, version 6.5; Tripos, Inc.: St. Louis, MO.
- (28) UNITY Chemical Information Software, version 4.1; Tripos, Inc.: St. Louis, MO.
- (29) Moriguchi, I.; Hirono, S.; Qian, L.; Nakagome, I.; Matsushita, Y. Simple Method of Calculating Octanol/Water Partition Coefficient. *Chem. Pharm. Bull.* **1992**, *40*, 127–130. Moriguchi, I.; Hirono, S.; Nakagome, I.; Hirano, H. Comparison of Reliability of log P Values for Drugs Calculated by Several Methods. *Chem. Pharm. Bull.* **1994**, *42*, 976–978.
- (30) Dean, P. M. *Molecular Similarity in Drug Design*; Blackie Academic & Professional: London, 1995.
- (31) Downs, G. M.; Willett, P. Similarity Searching in Databases of Chemical Structures. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 7, pp 1–66.
- (32) *Maybridge Database*, version August 99; Maybridge Chemical Co. Ltd.: U.K., 1999.
- (33) *LeadQuest Chemical Compound Libraries*; Tripos, Inc.: St. Louis, MO, 2000; Vols. 1–3.
- (34) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- (35) Böhm, H. J. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 593–606. Böhm, H. J. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 61–78.
- (36) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.
- (37) Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: a knowledge-based approach for identifying interaction sites in proteins. *J. Mol. Biol.* **1999**, *289*, 1093–1108.
- (38) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-Based Scoring Function To Predict Protein–Ligand Interactions. *J. Mol. Biol.* **2000**, *295*, 337–356.
- (39) Allen, F. H.; Davies, J. E.; Galloy, J. J.; Johnson, O.; Kennard, O.; et al. The development of version-3 and version-4 of the Cambridge Structural Database system. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187–204.
- (40) Klebe, G. The Use of Composite Crystal-Field Environments in Molecular Recognition and the “De-Novo” Design of Protein Ligands. *J. Mol. Biol.* **1994**, *237*, 212–235.
- (41) Gohlke, H.; Hendlich, M.; Klebe, G. Prediction of Binding Modes, Binding Affinities and “Hot Spots” for Protein–Ligand Complexes Using a Knowledge-Based Scoring Function. *Perspect. Drug Discovery Des.* **2000**, *20*, 115–144.
- (42) Lemmen, C.; Lengauer, T.; Klebe, G. FLEXS: a method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502–4520.
- (43) Schafferhans, A.; Klebe, G. Docking Ligands onto Binding Site Representations Derived from Proteins Built by Homology Modelling. *J. Mol. Biol.* **2001**, *307*, 407–427.
- (44) Boriack, P. A.; Christianson, D. W.; Kingery-Wood, J.; Whitesides, G. M. Secondary interactions significantly removed from the sulfonamide binding pocket of carbonic anhydrase II influence inhibitor binding constants. *J. Med. Chem.* **1995**, *38*, 2286–2291.
- (45) Klebe, G.; Mietzner, T.; Weber, F. Different approaches toward an automatic structural alignment of drug molecules: applications to sterol mimics, thrombin and thrombolysin inhibitors. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 751–778.
- (46) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (47) Abbate, F.; Supuran, C. T.; Scozzafava, A.; Orioli, P.; Stubbs, M. T.; Klebe, G. Nonaromatic sulfonamide group as an ideal anchor for potent human carbonic anhydrase inhibitors: role of hydrogen-bonding networks in ligand binding and drug design. *J. Med. Chem.* **2002**, *45*, 3583–3587.
- (48) Pocker, Y.; Stone, J. T. The Catalytic Versatility of Erythrocyte Carbonic Anhydrase. III. Kinetic Studies of the Enzym-Catalyzed Hydrolysis of *p*-Nitrophenyl Acetate. *Biochemistry* **1967**, *6*, 668–678.
- (49) Surgue, M. F.; Harris, A.; Adamsons, I. Dorzolamidehydrochloride: a topically active, carbonic anhydrase inhibitor the treatment of glaucoma. *Drugs Today* **1997**, *33*, 283–298.
- (50) Otwinowski, Z. *DENZO*; Yale University: New Haven, CT.
- (51) Hakansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. Structure of Native and Apo Carbonic Anhydrase II and Structure of Some of Its Anion-Ligand Complexes. *J. Mol. Biol.* **1992**, *227*, 1192–1204.
- (52) Brünger, A. T. *X-PLOR*, version 3.851; Yale University: New Haven, CT, 1996.
- (53) Grüneberg, S.; Wendt, B.; Klebe, G. Subnanomolar inhibitors from computer screening: A model study using human carbonic anhydrase II. *Angew. Chem.* **2001**, *113*, 404–408; *Angew. Chem., Int. Ed.* **2001**, *40*, 389–393.
- (54) Fradera, X.; Knegtel, R. M.; Mestres, J. Similarity-driven flexible ligand docking. *Proteins* **2000**, *40*, 623–636. Vieth, M.; Cummins, D. J. DoMCoSAR: Novel Approach for Establishing the Docking Mode That Is Consistent with the Structure–Activity Relationship. Application to HIV-1 Protease Inhibitors and VEGF Receptor Tyrosine Kinase Inhibitors. *J. Med. Chem.* **2000**, *43*, 3020–3032.
- (55) Gohlke, H. Ph.D. Thesis, University Marburg, Marburg, Germany, 2000.
- (56) Nair, S. K.; Elbaum, D.; Christianson, D. W. Unexpected binding mode of the sulfonamide fluorophore 5-(dimethylamino)-1-naphthalene sulfonamide to human carbonic anhydrase II. Implications for the development of a zinc biosensor. *J. Biol. Chem.* **1996**, *271*, 1003–1007.
- (57) Graham, S. L.; Shepard, K. L.; Anderson, P. S.; Baldwin, J. J.; Best, D. B.; et al. Topically active carbonic anhydrase inhibitors. 2. Benzo[*b*]thiophenesulfonamide derivatives with ocular hypotensive activity. *J. Med. Chem.* **1989**, *32*, 2548–2554.
- (58) Ponticello, G. S.; Freedman, M. B.; Habecker, C. N.; Lyle, P. A.; Schwam, H.; et al. Thienothiopyran-2-sulfonamides: a novel class of water-soluble carbonic anhydrase inhibitors. *J. Med. Chem.* **1987**, *30*, 591–597.
- (59) Hunt, C. A.; Mallorga, P. J.; Michelson, S. R.; Schwam, H.; Sondey, J. M.; et al. 3-Substituted thieno[2,3-*b*][1,4]thiazine-6-sulfonamides. A novel class of topically active carbonic anhydrase inhibitors. *J. Med. Chem.* **1994**, *37*, 240–247.
- (60) Baldwin, J. J.; Ponticello, G. S.; Anderson, P. S.; Christy, M. E.; Murcko, M. A.; et al. Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. *J. Med. Chem.* **1989**, *32*, 2510–2513.
- (61) Maren, T. H. Carbonic Anhydrase: General Perspectives and Advances in Glaucoma Research. *Drug Dev. Res.* **1987**, *10*, 255–276.
- (62) Stams, T.; Chen, Y.; Boriack-Sjodin, P. A.; Hurt, J. D.; Liao, J.; et al. Structures of murine carbonic anhydrase IV and human carbonic anhydrase II complexed with brinzolamide: molecular basis of isozyme-drug discrimination. *Protein Sci.* **1998**, *7*, 556–563.
- (63) Prugh, J. D.; Hartman, G. D.; Mallorga, P. J.; McKeever, B. M.; Michelson, S. R.; et al. New isomeric classes of topically active ocular hypotensive carbonic anhydrase inhibitors: 5-substituted thieno[2,3-*b*]thiophene-2-sulfonamides and 5-substituted thieno[3,2-*b*]thiophene-2-sulfonamides. *J. Med. Chem.* **1991**, *34*, 1805–1818.
- (64) Boriack, P. A.; Christianson, D. W.; Kingery-Wood, J.; Whitesides, G. M. Secondary interactions significantly removed from the sulfonamide binding pocket of carbonic anhydrase II influence inhibitor binding constants. *J. Med. Chem.* **1995**, *38*, 2286–2291.
- (65) Woltersdorf, O. W., Jr.; Schwam, H.; Bicking, J. B.; Brown, S. L.; deSolms, S. J.; et al. Topically active carbonic anhydrase inhibitors. 1. *O*-Acyl derivatives of 6-hydroxybenzothiazole-2-sulfonamide. *J. Med. Chem.* **1989**, *32*, 2486–2492.
- (66) Graham, S. L.; Hoffman, J. M.; Gautheron, P.; Michelson, S. R.; Scholz, T. H.; et al. Topically active carbonic anhydrase inhibitors. 3. Benzofuran- and indole-2-sulfonamides. *J. Med. Chem.* **1990**, *33*, 749–754.
- (67) Shepard, K. L.; Graham, S. L.; Hudcosky, R. J.; Michelson, S. R.; Scholz, T. H.; et al. Topically active carbonic anhydrase inhibitors. 4. *J. Med. Chem.* **1991**, *34*, 3098–3105.

- (68) Maryanoff, B. E.; Costanzo, M. J.; Nortey, S. O.; Greco, M. N.; Shank, R. P.; et al. Structure–activity studies on anticonvulsant sugar sulfamates related to topiramate. Enhanced potency with cyclic sulfate derivatives. *J. Med. Chem.* **1998**, *41*, 1315–1343.
- (69) Elbaum, D.; Nair, S. K.; Patchan, M. W.; Thompson, R. B.; Christianson, D. W. Structure-Based Design of a Sulfonamide Probe for Fluorescence Anisotropy Detection of Zinc with a Carbonic Anhydrase-Based Biosensor. *J. Am. Chem. Soc.* **1996**, *118*, 8381–8387.
- (70) Hartmann, M. The Important Role of Active Site Water in the Catalytic Mechanism of Human Carbonic Anhydrase—A Semiempirical MO Approach to the Hydration of CO₂. *J. Mol. Model.* **1998**, *4*, 355–365.
- (71) Scholz, T. H.; Sondey, J. M.; Randall, W. C.; Schwam, H.; Thompson, W. J.; et al. Sulfonylmethanesulfonamide inhibitors of carbonic anhydrase. *J. Med. Chem.* **1993**, *36*, 2134–2141.
- (72) Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: ureido and thioureido derivatives of aromatic sulfonamides possessing increased affinities for isozyme I. A novel route to 2,5-disubstituted-1,3,4-thiadiazoles via thioureas, and their interaction with isozymes I, II and IV. *J. Enzyme Inhib.* **1998**, *13*, 103–123.
- (73) Brechue, W. F.; Maren, T. H. Carbonic anhydrase inhibitory activity and ocular pharmacology of organic sulfamates. *J. Pharmacol. Exp. Ther.* **1993**, *264*, 670–675.
- (74) Sugrue, M. F.; Gautheron, P.; Mallorga, P.; Nolan, T. E.; Graham, S. L.; et al. L-662,583 is a topically effective ocular hypotensive carbonic anhydrase inhibitor in experimental animals. *Br. J. Pharmacol.* **1990**, *99*, 59–64.

JM011112J