3D QSAR Selectivity Analyses of Carbonic Anhydrase Inhibitors: Insights for the Design of Isozyme Selective Inhibitors

Alexander Weber,[†,||] Markus Böhm,[§] Claudiu T. Supuran,* Andrea Scozzafava,[‡] Christoph A. Sotriffer,[†] and Gerhard Klebe*,[†]

Department of Pharmaceutical Chemistry, University of Marburg, Marbacher Weg 6, D-35032 Marburg, Germany, Dipartimento di Chimica, Laboratorio di Chimica Bioinorganica, Università degli Studi di Firenze, Via della Lastruccia, 3, Room 188, I-50019 Sesto Fiorentino, Florence, Italy, Pfizer Global R&D, Eastern Point Road, Groton, Connecticut 06340, and Discovery IT, ALTANA Pharma AG, Konstanz, Germany

Received July 18, 2006

A 3D QSAR selectivity analysis of carbonic anhydrase (CA) inhibitors using a data set of 87 CA inhibitors is reported. After ligand minimization in the binding pockets of CA I, CA II, and CA IV isoforms, selectivity CoMFA and CoMSIA 3D QSAR models have been derived by taking the affinity differences (ΔpK_i) with respect to two CA isozymes as independent variables. Evaluation of the developed 3D QSAR selectivity models allows us to determine amino acids in the respective CA isozymes that possibly play a crucial role for selective inhibition of these isozymes. We further combined the ligand-based 3D QSAR models with the docking program AUTODOCK in order to screen for novel CA inhibitors. Correct binding modes are predicted for various CA inhibitors with respect to known crystal structures. Furthermore, in combination with the developed 3D QSAR models we could successfully estimate the affinity of CA inhibitors even in cases where the applied scoring function failed. This novel strategy to combine AUTODOCK poses with CoMFA/CoMSIA 3D QSAR models can be used as a guideline to assess the relevance of generated binding modes and to accurately predict the binding affinity of newly designed CA inhibitors that could play a crucial role in the treatment of pathologies such as tumors, obesity, or glaucoma.

INTRODUCTION

The enzyme carbonic anhydrase (CA, carbonate hydrolase, EC 4.2.1.1) is involved in a variety of physiological and physiopathological processes.[1−5] CA isozymes catalyze the hydration of carbon dioxide to hydrogen carbonate and a proton. Up to 16 CA isozymes were found in different tissues in the human body to deliver protons and hydrogen carbonate for physiological reactions. These different CA isozymes are either located in the cytosol (CA I, II, III, VII, VIII, X, XI, and XIII) or in mitochondria (CA VA and CA VB), secreted into saliva (CA VI), or are membrane-bound (CA IV, IX, XII, XIV, XV).[6] Selectivity issues in CA inhibition arise from the systemic application of nonselective CA inhibitors leading to severe side effects due to the widespread distribution of CA isozymes (e.g. eye, lungs, kidney, liver, stomach, pancreas, brain, and erythrocytes). Furthermore, nonselective CA inhibitors interfere not only with one isozyme in one particular tissue but also with multiple isozymes in different tissues.[4,7] Among a variety of CA inhibitor classes, sulfonamides are the most important clinical agents, e.g. used in the treatment and prevention of glaucoma, gastroduodenal ulcers, certain neurological disorders, motion and altitude sickness, tumors, and many other conditions.[8−17] One strategy to circumvent the insufficient selectivity issue was achieved for the treatment of glaucoma, by developing topical CA inhibitors.[18] They reduce side effects via direct application of the drug in eye drop formulation. However, due to the requirement of oral administration for other classes of CA inhibitors, the observed side effects remain to be considered seriously.

A large number of aromatic, heterocyclic, and more recently aliphatic sulfonamides have been synthesized and tested for their CA inhibitory activities. Some of them are widely used as clinical or investigational drugs, such as dorzolamide, brinzolamide, acetazolamide, methazolamide, ethoxzolamide, or dichlorophenamide.[5]

Early structure−activity relationships for this class of inhibitors demonstrate the importance of an unsubstitued sulfonamide group for CA inhibition.[1] Furthermore, crystal structures of sulfonamides in complex with CA II elucidate the main interactions of this compound class with the protein. In carbonic anhydrases, a zinc ion, critical for catalysis, is located in the binding pocket. It is coordinated by three histidine residues and one water molecule. Sulfonamides bind in their deprotonated form ($RSO2NH^-$) to the zinc ion by replacing the zinc-bound water molecule, thus perturbing the catalytic cycle.[1,4,5,10,19] A recent comprehensive review by Clare and Supuran summarizes QSAR studies on CA inhibitors.[20] Whereas a correlation between CA inhibition and empirical parameters, for example such as pK_a, lipophilicity, and electrostatic-potential-based charges on the atoms of the sulfonamide group were derived from these QSAR models, the extraction of selectivity discriminating features

* Corresponding author phone: +49 6421 28 21313; fax: +49 6421 28 28994; e-mail: klebe@mailer.uni-marburg.de (G.K.) and phone: +39-055-4573005; fax: +39-055-4573385; e-mail: claudiu.supuran@unifi.it (C.T.S.).

[†] University of Marburg.

[‡] Università degli Studi di Firenze.

[§] Pfizer Global R&D.

[||] Discovery IT, ALTANA Pharma AG.

10.1021/ci600298r CCC: $33.50 © 2006 American Chemical Society
Published on Web 09/20/2006

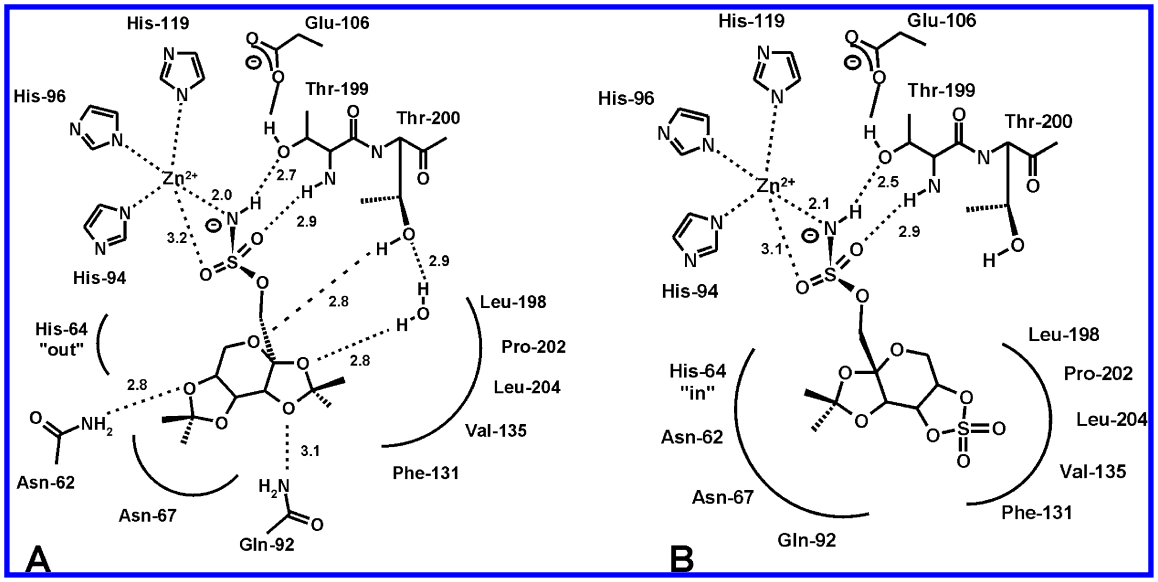


Figure 1. Binding mode of topiramate (**A**) and RWJ-37947 (**B**, PDB code: *1eou*) as observed in complex with CA II. The crystal structure of topiramate with CA II was kindly provided by Casini et al.[55] The sulfamate anchoring groups of both compounds form favorable hydrogen bonds interacting with the zinc ion and Thr199. In the complex structure of topiramate, additional hydrogen bond interactions are made with Asn62, Thr200, and a water molecule that forms an additional hydrogen bond to the side chain hydroxyl group of Thr200 (**A**). In the crystal structure of RWJ-37947 (**B**) the ring system is rotated about 180° and points with its cyclic sulfate group into the more hydrophobic pocket (Leu198, Pro202, Phe131);[57] compared to the topiramate complex no additional hydrogen bonds were formed. All amino acid side chains in both crystal structures showed similar side chain conformations, with the exception of His64, which points out of the binding pocket in case of the complex structure with topiramate (**A:** His64 "out"), whereas with the RWJ compound it is oriented toward the binding pocket (**B:** His64 "in").

as a guideline to synthesis of isozyme selective inhibitors still remains challenging.

In this paper, we present a 3D QSAR analysis on CA I, II, and IV, using a data set of 87 inhibitors. After minimizing the inhibitors individually within the binding pockets of CA I, II, and IV using the program MOLOC[21,22] CoMFA and CoMSIA models were developed (Tables 2 and 3). In addition, 3D QSAR selectivity models were obtained in a similar fashion as it was described for matrix metalloproteases and serine proteases.[23−25]

We further used the docking program AUTODOCK[26,27] to automatically place CA inhibitors in the binding pocket of CA II. Our interest in the automatic docking was stimulated by the observation of deviating binding modes of two structurally similar inhibitors, topiramate and RWJ-37947 in CA II (Figure 1).[28] The generation of correct poses for such ligands in CA II presents a challenging problem for docking algorithms and for drug design in general. With AUTODOCK, we were able to predict the correct binding mode of these two inhibitors. In addition to this correct binding mode prediction, a set of additional CA inhibitors taken from literature shows that in cases where the affinity prediction of the docked ligands based on the scoring function implemented in AUTODOCK fails, we could accurately predict their inhibitory potency using the 3D QSAR models developed in this contribution.

RESULTS AND DISCUSSION

To obtain insight into the selective inhibition of CA isozymes, the affinity differences (ΔpK_i) of inhibitors between two CA isozymes were regarded as independent variables for the establishment of QSAR models (selectivity models for CA I−CA II, CA I−CA IV, CA II−CA IV, respectively). Since every inhibitor was minimized in each individual CA binding pocket, we are able to draw conclusions with respect to the amino acid composition in the different binding pockets that might be of pivotal importance for selectivity (mainly CA I and CA II).

Data Analyses. A data set of 87 inhibitors (Table 1) with known inhibitory activities toward CA I, CA II, and CA IV was used for the 3D QSAR analyses described herein.[29−39] The compounds of the data set fall into six categories with different scaffolds: thiadiazole- (**A**), thienothiopyrane- (**B**), benzothiazole- (**C**), and benzenesulfonamides (**D**), hydroxamates (**E**), and hydroxysulfonamides (**F**) (Figure 2). The experimentally determined binding affinities of the CA inhibitors are given as pK_i ($-\log K_i$) values (Table 1). The pK_i values of the 87 compounds in the data set cover a range of more than four logarithmic units with respect to all three isozymes (CA I = 4.84 logarithmic units, CA II = 4.70, CA IV = 4.18; Table 1). This fulfills the prerequisite of a spread of at least 3 orders of magnitude[40] to ensure the development of statistically significant 3D QSAR models. The thiadiazolesulfonamide **A_11** shows the highest affinity toward all three isozymes in this series, with nanomolar affinity toward CA I ($pK_i = 8.70$) and CA IV ($pK_i = 8.52$) and even subnanomolar inhibition for CA II ($pK_i = 10.0$). In general, thiadiazole- (**A**, see Figure 2) and thienothiopyranesulfonamides (**B**) are medium to strong inhibitors of CA isozymes, whereas benzothiazole derivatives (**C**) belong to the strongest inhibitors in the data set. Benzenesulfonamides (**D**) are generally less active for the three isozymes compared to inhibitors with the previously mentioned scaffolds. While hydroxamates (**E**) inhibit all three CA isozymes in the nanomolar range, the hydroxysulfonamides (**F**) are equally potent for CA II and CA IV but less active for CA I.

Table 1. 87 CA Inhibitors Used for the Generation of the 3D QSAR Models

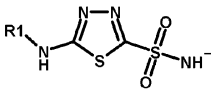
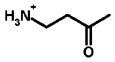
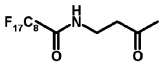
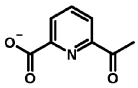
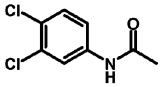
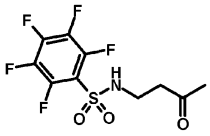
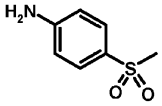
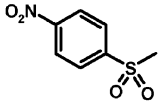
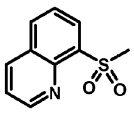
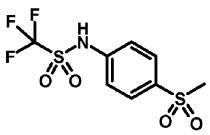
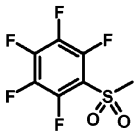
No.	Structure	Residue	Scaffold	pK_i CA I	pK_i CA II	pK_i CA IV	Charge
		R1 =					
1		-H	A_01	5.07	7.22	6.27	-1
2			A_02	6.34	8.52	6.90	0
3			A_03	6.52	8.30	7.89	-1
4			A_04	6.52	8.70	8.10	-1
5			A_05	6.26	7.68	7.25	-1
6			A_06	6.05	7.82	7.00	-1
7			A_07	7.44	8.22	7.68	-2
8			A_08	8.52	8.22	8.10	-1
9			A_09	6.70	9.40	8.05	-1
10			A_10	8.22	8.70	8.30	-1
11			A_11	8.70	10.00	8.52	-1
12			A_12	7.48	8.70	8.05	-1
13			A_13	6.00	8.52	7.60	-1
14			A_14	6.68	9.52	8.22	-1

Table 1. (Continued)

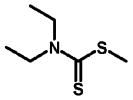
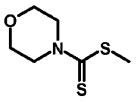
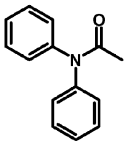
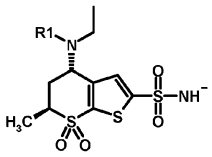
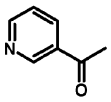
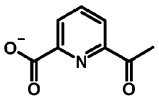
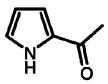
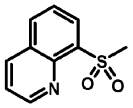
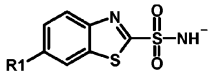
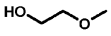
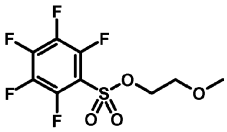
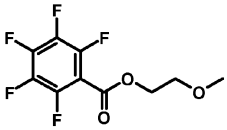
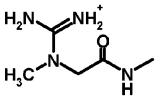
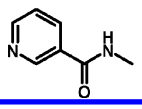
No.	Structure	Residue	Scaffold	pK_i CA I	pK_i CA II	pK_i CA IV	Charge
15			A_15	6.52	8.15	7.48	-1
16			A_16	6.62	8.22	7.60	-1
17			A_17	8.15	8.10	7.82	-1
18			A_18	6.05	7.92	6.66	-1
19			B_01	5.70	8.30	7.96	-1
20			B_02	6.40	8.30	7.77	-2
21			B_03	6.46	8.10	7.40	-1
22			B_04	6.77	8.30	8.00	-1
23		$-H_2$	B_05	4.30	8.05	7.37	0
24		-OH	C_01	7.26	8.10	7.77	-1
25			C_02	7.30	8.15	7.82	-1
26			C_03	7.74	9.30	8.15	-1
27			C_04	7.19	9.15	7.96	-1
28		$-NH_2$	C_05	7.15	8.05	7.72	-1
29			C_06	8.00	8.40	7.92	0
30			C_07	7.89	8.30	7.72	-1

Table 1. (Continued)

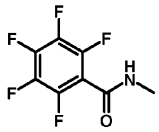
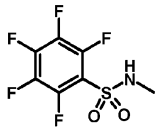
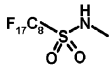
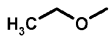
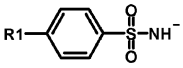
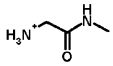
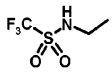
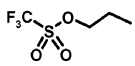
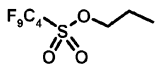
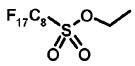
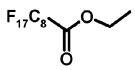
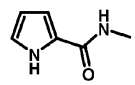
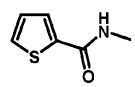
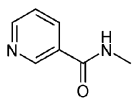
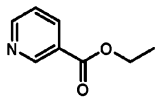
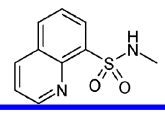
No.	Structure	Residue	Scaffold	pK_i CA I	pK_i CA II	pK_i CA IV	Charge
31			C_08	6.92	9.30	8.10	-1
32			C_09	7.39	9.70	8.30	-1
33			C_10	7.31	9.22	8.15	-1
34			C_11	7.60	8.10	7.89	-1
35		-NH_2	D_01	4.55	6.52	5.52	-1
36			D_02	4.11	6.49	5.49	-1
37			D_03	4.60	6.77	5.55	0
38			D_04	4.68	6.80	5.61	0
39			D_05	4.74	6.96	6.35	-1
40			D_06	4.82	6.91	6.76	0
41			D_07	3.86	5.30	4.61	-1
42			D_08	4.10	6.26	4.35	-1
43			D_09	4.68	6.41	5.04	-1
44			D_10	4.82	6.27	4.96	-1
45			D_11	4.82	6.26	4.96	-1
46			D_12	4.60	6.98	6.80	-1
47			D_13	4.94	7.21	6.99	-1
48			D_14	4.80	6.85	6.78	-1
49			D_15	5.67	7.10	6.85	-1
50			D_16	4.81	6.87	6.75	-1

Table 1. (Continued)

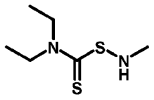
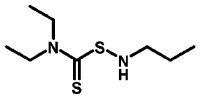
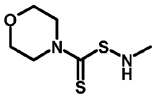
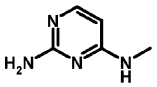
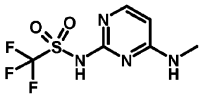
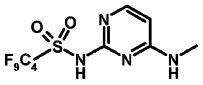
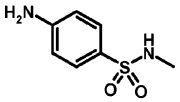
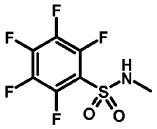
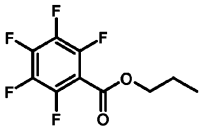
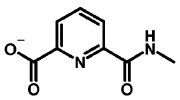
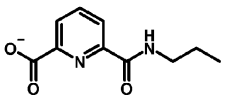
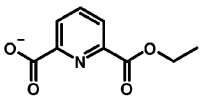
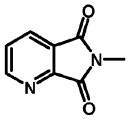
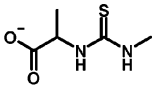
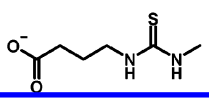
No.	Structure	Residue	Scaffold	pK_i CA I	pK_i CA II	pK_i CA IV	Charge
51			D_17	4.82	7.32	6.85	-1
52			D_18	6.20	7.96	7.52	-1
53			D_19	5.05	7.44	7.00	-1
54			D_20	6.16	7.92	6.81	-1
55			D_21	5.10	7.60	7.41	-1
56			D_22	5.22	7.70	7.52	-1
57			D_23	7.38	8.22	7.30	-1
58			D_24	6.05	8.00	6.39	-1
59			D_25	5.90	7.30	6.40	-1
60			D_26	4.82	6.92	6.82	-2
61			D_27	6.26	8.15	7.40	-2
62			D_28	5.69	7.70	6.92	-2
63			D_29	4.81	6.88	6.77	-1
64			D_30	7.27	7.92	7.44	-2
65			D_31	7.35	7.96	7.48	-2

Table 1. (Continued)

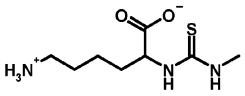
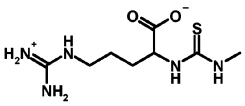
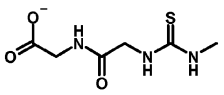
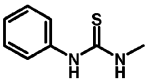
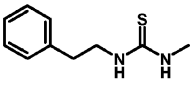
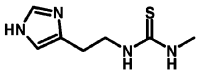
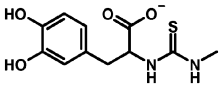
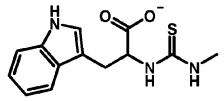
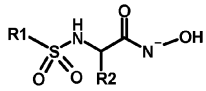
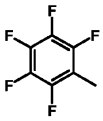
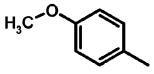
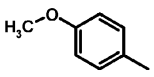
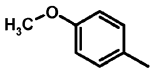
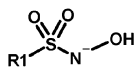
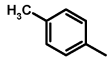
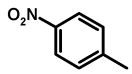
No.	Structure	Residue		Scaffold	pK_i CA I	pK_i CA II	pK_i CA IV	Charge
66				D_32	7.01	7.92	7.26	-1
67				D_33	6.94	7.82	7.21	-1
68				D_34	7.27	7.82	7.49	-2
69				D_35	7.52	8.22	7.89	-1
70				D_36	6.90	7.38	7.12	-1
71				D_37	7.04	7.48	7.25	-1
72				D_38	7.46	8.15	7.74	-2
73				D_39	7.33	7.96	7.62	-2
74		R_1 = 	R_2 = $-CH_3$	E_01	8.15	8.10	8.00	-1
75				E_02	8.10	7.96	7.89	-1
76				E_03	7.48	7.41	7.42	-1
77				E_04	7.30	8.30	7.41	-1
78		$-C_4F_9$	$-CH_3$	E_05	7.68	7.80	7.77	-1
79			$-CH_3$	E_06	7.49	7.46	7.52	-1
80		$-CCl_3$		F_01	4.64	6.82	6.68	-1
81		$-CF_3$		F_02	4.70	7.52	7.80	-1
82		$-C_4F_9$		F_03	5.74	7.82	7.57	-1

Table 1. (Continued)

No.	Structure	Residue	Scaffold	pK_i CA I	pK_i CA II	pK_i CA IV	Charge
83		$-C_8F_{17}$	F_04	5.80	8.05	7.68	-1
84			F_05	4.92	7.40	6.54	-1
85			F_06	4.29	7.15	6.90	-1
86			F_07	4.24	8.05	7.15	-1
87			F_08	4.74	7.59	7.13	-1
MEAN				6.21	7.81	7.19	
STDEV				1.27	0.84	0.88	
MAX				8.70	10.0	8.52	
MIN				3.86	5.30	4.35	
RANGE				4.84	4.70	4.18	

Since we are interested in generating selectivity models between pairs of CA isozymes, the correlation coefficients of the affinities for all 87 inhibitors in the data set with respect to two isozymes were calculated (Figure 3). The highest data correlation was observed for CA II and CA IV ($r = 0.86$), followed by CA I−CA IV ($r = 0.74$) and CA I−CA II ($r = 0.71$). As a general observation, the inhibitors in the data set are more active against CA II than CA I (Figure 3A) and CA IV (Figure 3C). Up to now, there are no inhibitors known that inhibit CA I or CA IV significantly higher than CA II. Furthermore, the inhibitors are generally less active for CA I compared to CA IV (Figure 3B). In summary, the inhibitors are most active against CA II, followed by CA IV and CA I inhibition (CA II > CA IV > CA I). The highly correlated data between CA II and CA IV ($r = 0.86$) also indicate that the interpretation of selectivity models between these two isozymes requires some caution. Nevertheless, the prerequisite for the generation of individual 3D QSAR models are fulfilled for each isozyme.[40]

Alignment. A superposition of over 150 CA protein crystal structures accommodating more than 40 different ligands deposited in the Protein Data Bank (PDB)[41−43] was performed using RELIBASE.[44,45] This analysis demonstrates that CAs are relatively rigid proteins. All amino acid side chains in the active site adopt a similar conformation, except His64, which is known to operate as "*proton shuttle*". In CA II (and other high activity isozymes, such as CA VII, CA IX, and CA IV) it is essential for catalysis, and the different conformations of His64 are due to its role in the proton transfer.[46−48] The superposition with RELIBASE shows that all inhibitors coordinate with an anchoring group (mostly a sulfonamide group) toward the catalytically active zinc ion and form additional favorable hydrogen bonds with Thr199 (Figure 4A). In CA II, further ligand contacts occur via *van der Waals* or polar interactions with amino acids in the nearby hydrophilic (Asn67, Glu69, Gln92) or hydrophobic (Phe131, Val135, Leu198, Pro202, Leu204; Figure 4A) pockets. The X-ray structure of CA II in complex with a hydroxamate ligand[49] (PDB code: *1am6*) reveals that it

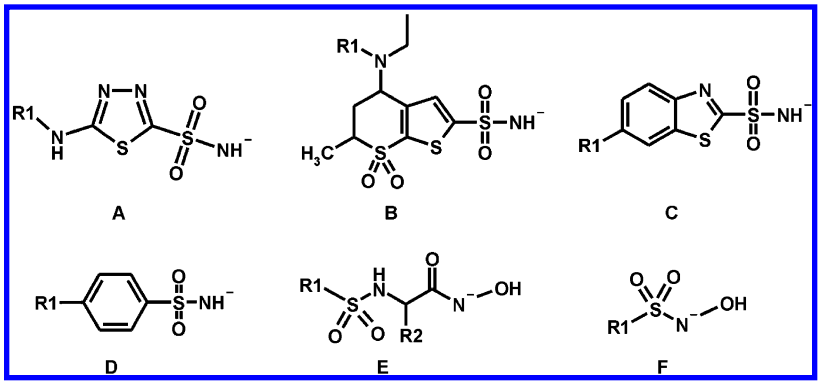


Figure 2. Core fragments of CA inhibitors that are used in the data set for 3D QSAR analyses (thiadiazol-(**A**), thienothiopyran-(**B**), benzothiazol-(**C**) and benzenesulfonamides (**D**), hydroxamates (**E**), and hydroxysulfonamides (**F**)). Sulfonamides coordinate with their deprotonated nitrogen atom to the catalytic zinc ion;[19] therefore, the atomic charge at the sulfonamide nitrogen atom was set to −1. To be consistent within the data set, the hydroxamate and hydroxysulfonamide nitrogen charge was also set to −1.

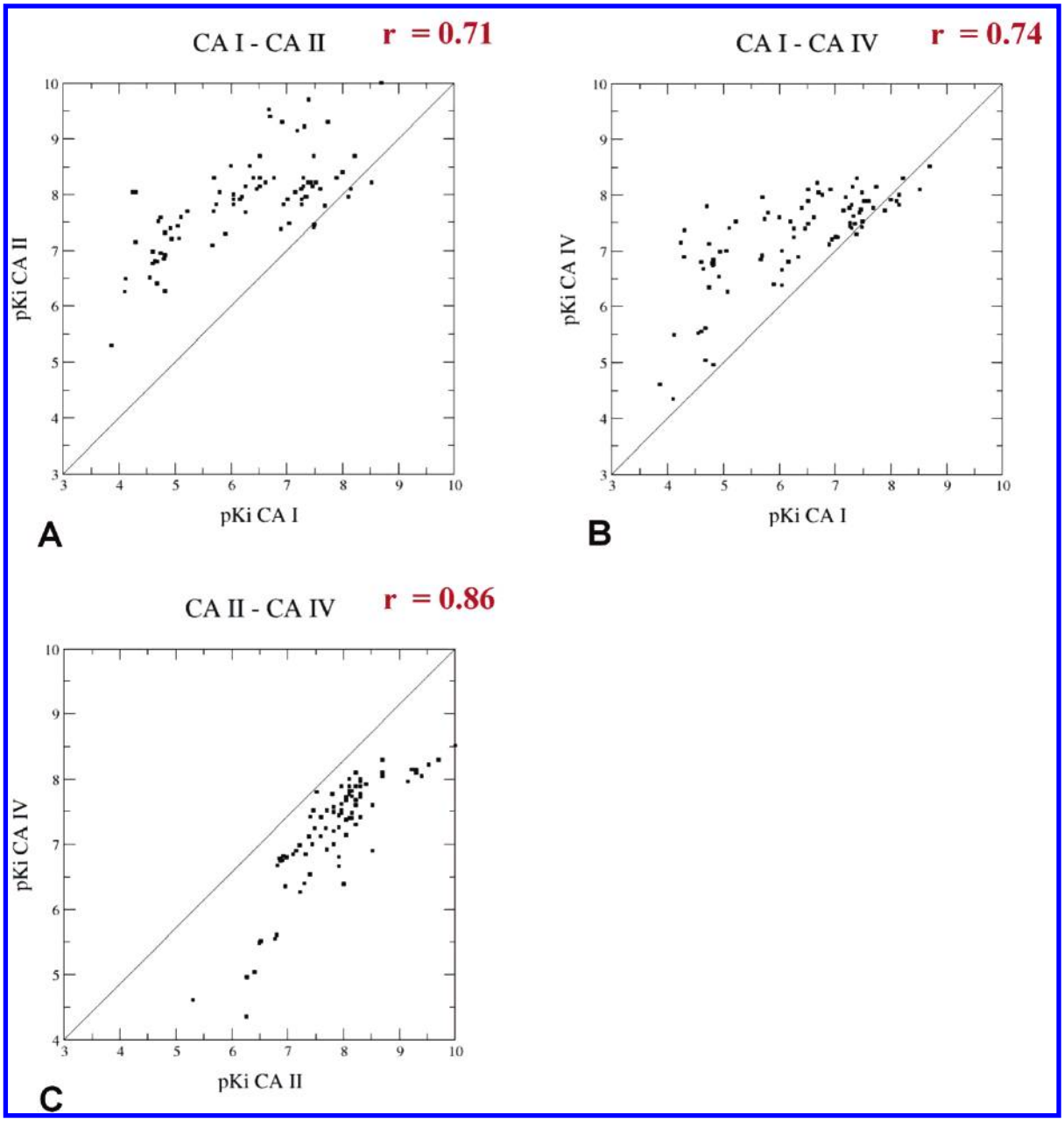


Figure 3. Affinity data of CA inhibitors for CA I, CA II, and CA IV are plotted against each other. Inhibitory activity values are plotted as pK_i. The inhibitors in the data set are more active against CA II than CA I (**A**) and CA IV (**C**) and less active for CA I compared to CA IV (**B**). The calculated correlation coefficient for the affinity data indicates the high correlation between these isozymes ($r = 0.71$ between CA I and CA II inhibition; $r = 0.74$ between CA I and CA IV inhibition), with the highest correlation between CA II and CA IV inhibition ($r = 0.86$) within this series.

coordinates with the deprotonated nitrogen toward the zinc ion and with the hydroxyl and carbonyl group toward Thr199 (Figure 4B). Due to this observation, the hydroxamate (and similarly also the hydroxysulfonamide) series in the data set was minimized according to the postulated and observed binding mode described by Scolnick,[49] neglecting the fact that different binding modes of these scaffolds might still occur.

To focus the alignment of the proteins predominantly on the binding pockets, 37 protein−ligand complexes (three complexes of CA I, 32 of CA II, two of CA IV) were selected, and 20 conserved amino acid residues in close proximity to the binding pocket were considered for a structural alignment within SYBYL (see the Experimental Section for more details). Analysis of the 32 protein−ligand complexes in CA II showed that all sulfonamides coordinate with the nitrogen atom of the sulfonamide portion toward the zinc ion (taken the atom assignment to the electron density for granted), whereas inhibitor side chains point toward the hydrophobic pocket, comprised of Phe131, Val135, Leu198, Pro202, and Leu204 (Figure 4A, Figure 5C). In CA I (three complexes, Figure 5A) and CA IV (two complexes, Figure 5E), only a few X-ray complex structures have been deposited in the PDB. Nevertheless, these structures served as templates and provided starting points for the subsequent minimization of ligand poses. The 87 inhibitors compiled from literature were minimized with MOLOC[21,22] in each of the three CA binding pockets, taking into account that essential interactions to the zinc ion and Thr199 are preserved during minimization. As a result, three

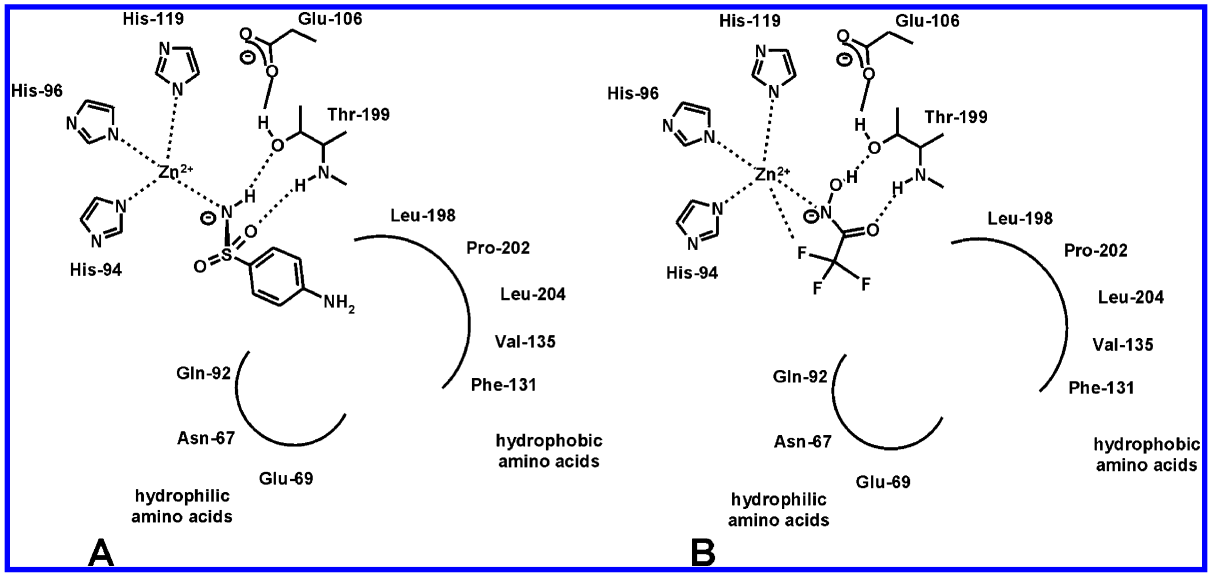


Figure 4. Schematic binding mode of sulfonamides (**A**) and hydroxamates (**B**) in CA II. Sulfonamides coordinate with their deprotonated sulfonamide group to the zinc ion and form additional hydrogen bonds with the side chain oxygen and the backbone nitrogen atom of Thr199 (**A**). In CA II one crystal structure was solved with a hydroxamate inhibitor (PDB code: *1am6*).[49] The hydroxamate group interacts with its hydroxyl and carbonyl group with the side chain oxygen and the backbone nitrogen of Thr199, whereas the nitrogen atom is coordinated toward the zinc ion (**B**).

distinct alignments, specifically adapted for CA I (Figure 5B), CA II (Figure 5D), and CA IV (Figure 5F) were generated, which were used for subsequent CoMFA and CoMSIA model generation.

Individual 3D QSAR Models. CoMFA[50] and CoMSIA[51] models were derived for each alignment in the corresponding binding pocket of CA I, CA II, and CA IV. Identical grid parameters using a grid spacing of 1 Å and 2 Å were applied for each alignment to ensure a straightforward comparison of the different models (see the Supporting Information). The statistical results of the cross-validated "leave-one-out" (LOO) analyses for individual CoMFA and CoMSIA models are shown in Tables 2 (2 Å grid spacing) and 3 (1 Å grid spacing). For each alignment in the three CA isozymes, CoMFA and CoMSIA models were calculated for CA I, CA II, and CA IV. With the exception of the CoMFA and CoMSIA models for CA II and CA IV based on the CA I alignment, all other 14 QSAR models revealed a high quality ($q^2 > 0.5$) using the 2 Å grid spacing. Also, the q^2 values greater than 0.46 for the derived CA II and CA IV models using the CA I alignment are close to the threshold generally accepted for good quality QSAR models. A comparison of the statistical parameters between the models derived with 1 Å and 2 Å grid spacing shows that only three CoMFA models show a significant difference. The q^2 value for the CA IV CoMFA model minimized in CA I and CA II as well as the CA I CoMFA model minimized in CA IV slightly increases for the 1 Å grid spacing (change in $q^2 > 0.07$). However, two of the models also show an increase in the number of components used for deriving the model, which could at least partially explain these enhancements. The subtle variations of the statistical parameters for the CoMFA procedure were described previously.[23] Comparing the different alignments in the corresponding binding pockets of CA I, CA II, and CA IV, one would expect that the best 3D QSAR model for each isozyme should be observed considering the alignment in its matching binding pocket. However, with respect to q^2 and S_{PRESS} values the best models were obtained with the alignment in CA II for both the 1 Å and 2 Å grid spacing (Tables 2 and 3). These observations could be explained with the larger number of template protein−ligand complexes in CA II (32 vs three for CA I and two for CA IV) that guided the manual alignment of the 87 inhibitors in the data set prior to minimization, thus enabling a more relevant alignment. In summary, the minimum prerequisite of individual high quality 3D QSAR models is satisfied in order to gain more insight into the selective inhibition of CA isozymes.

Validation of Individual 3D QSAR Models. The predictive power of each individual 3D QSAR model was further checked by "leave-5-out" (L5O) analyses, where five compounds were omitted from the data set during cross-validation. 3D QSAR models with the remaining compounds were calculated, and subsequently the affinities of the five omitted inhibitors were predicted. Several runs have been performed since the random selection of the five compounds produces slightly different results each time. The statistical results of 10 cross-validation runs using the CA II CoMFA and CoMSIA models confirmed the validity of the obtained QSAR models (data not shown).

A further test to validate the statistical significance of QSAR models is to scramble the affinity values of the data set inhibitors and subsequently calculate 3D QSAR models. The fact that no significant models ($q^2 < 0$; data not shown) were obtained using the CA inhibitor data set confirmed once more the statistical significance of the previously derived models.

In the next validation step, the data set was split into a training set of 60 compounds and a test set of 27 compounds. The SELECTOR module in SYBYL was used to define a subset of 60 diverse compounds (see the Experimental Section for more details). CoMFA and CoMSIA models were derived based on the 60 training set compounds only for each isozyme in their corresponding binding pocket (CA I model minimized in CA I, CA II model minimized in CA II, and CA IV model minimized in CA IV). The statistical

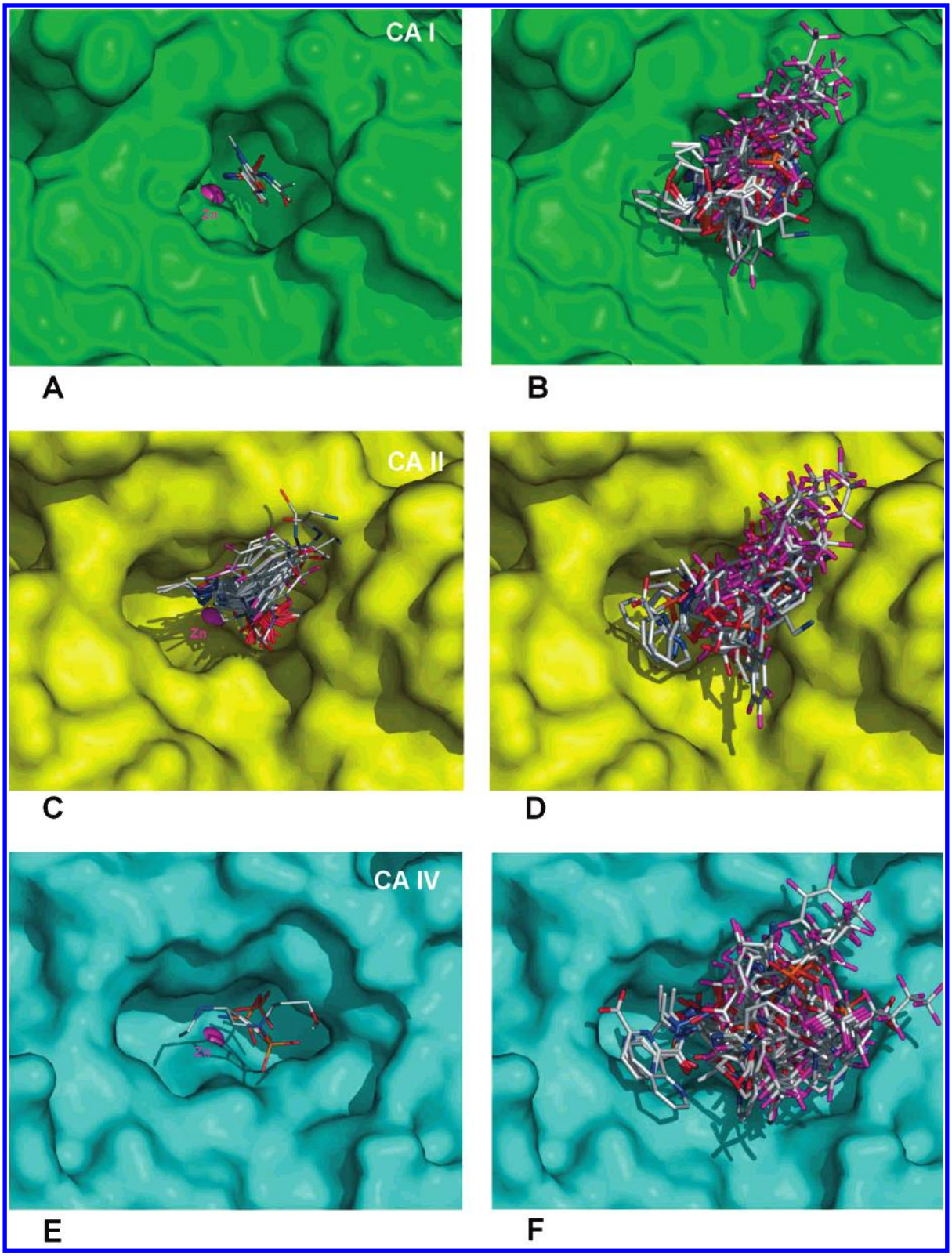


Figure 5. Binding mode of CA inhibitors in complex with CA I (**A**), CA II (**C**), and CA IV (**E**). The solvent accessible surface of CA I (green, 1azm), CA II (yellow, 1cil), and CA IV (cyan, 1znc) is shown together with three (CA I), 32 (CA II), and two (CA IV) ligand poses from crystal structures. The ligand complexes in these crystal structures guided the minimization of the inhibitors in the 3D QSAR data set. The 87 inhibitors of the 3D QSAR data set were minimized in the corresponding binding pocket (**B:** CA I; **D:** CA II; **F:** CA IV). Figures 5, 7, 8, and 10 were created by PyMol.[68]

results of the cross-validation procedure (LOO) for the CoMFA and CoMSIA models are summarized in Table 4. For each isozyme, statistically significant models were obtained ($q^2 > 0.3$). The predictive power of the various models was further confirmed by estimating the affinity of the remaining 27 inhibitors in the test set. The results are shown in Table 5 and Figure 6. The predictive r^2 value for each model indicates that the affinities of the test set

Table 2. Statistical Results of the 3D QSAR Analyses for the CoMFA and CoMSIA Method with a Grid Spacing of 2 Å

	CA I		CA II		CA IV	
	CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
Minimized in the Binding Pocket of CA I (Single Model)						
q^2	0.639	0.605	0.497	0.467	0.475	0.496
S_{PRESS}	0.784	0.820	0.607	0.622	0.655	0.641
r^2	0.880	0.896	0.856	0.818	0.828	0.857
S	0.451	0.422	0.325	0.364	0.374	0.342
F	150.7	175.8	121.6	124.1	98.9	122.4
components	4	4	4	3	4	4
Minimized in the Binding Pocket of CA II (Single Model)						
q^2	0.650	0.658	0.568	0.645	0.599	0.674
S_{PRESS}	0.776	0.762	0.563	0.513	0.582	0.522
r^2	0.929	0.887	0.884	0.911	0.963	0.938
S	0.350	0.438	0.292	0.257	0.177	0.228
F	211.6	161.4	156.2	166	293.6	201.7
components	5	4	4	5	7	6
Minimized in the Binding Pocket of CA IV (Single Model)						
q^2	0.589	0.635	0.513	0.514	0.504	0.562
S_{PRESS}	0.836	0.788	0.598	0.598	0.648	0.598
r^2	0.898	0.917	0.890	0.906	0.970	0.916
S	0.416	0.375	0.285	0.263	0.159	0.261
F	180.7	227.5	165.1	196.9	364.4	224.5
components	4	4	7	4	7	4
grid	2 Å					

Table 3. Statistical Results of the 3D QSAR Analyses for the CoMFA and CoMSIA Method with a Grid Spacing of 1 Å

	CA I		CA II		CA IV	
	CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
Minimized in the Binding Pocket of CA I (Single Model)						
q^2	0.649	0.600	0.500	0.484	0.591	0.512
S_{PRESS}	0.773	0.825	0.606	0.612	0.585	0.631
r^2	0.896	0.885	0.884	0.823	0.951	0.858
S	0.421	0.442	0.292	0.359	0.203	0.340
F	176.1	158.2	156.2	128.2	257.7	123.9
components	4	4	4	3	6	4
Minimized in the Binding Pocket of CA II (Single Model)						
q^2	0.684	0.669	0.609	0.665	0.664	0.686
S_{PRESS}	0.737	0.781	0.536	0.499	0.530	0.513
r^2	0.944	0.888	0.906	0.912	0.965	0.939
S	0.310	0.437	0.263	0.256	0.172	0.226
F	274.5	162.2	197.7	167.8	364.9	204.1
components	5	4	4	5	6	6
Minimized in the Binding Pocket of CA IV (Single Model)						
q^2	0.671	0.638	0.532	0.512	0.506	0.560
S_{PRESS}	0.757	0.785	0.586	0.599	0.647	0.599
r^2	0.974	0.918	0.899	0.907	0.973	0.917
S	0.215	0.373	0.273	0.261	0.153	0.261
F	491.4	229.6	181.5	200.5	399.7	225.3
components	6	4	4	4	7	4
grid	1 Å					

compounds were accurately predicted for all models (predictive $r^2 > 0.5$; Figure 6). In the case of CA I, the largest deviation in affinity prediction by more than one logarithmic unit was observed for the thiadiazole derivative **A_13** and the benzenesulfonamide **D_16** (CoMFA and CoMSIA). The affinity for **A_13** was predicted too high ($pK_{i,actual} = 6.0$; CoMFA: $pK_{i,}$predict $= 7.47$; CoMSIA: $pK_{i,predict} = 7.73$), due to the fact that structurally similar compounds in the training set possess higher affinity values (**A_10:** $pK_{i,actual} = 8.22$; **A_11:** $pK_{i,actual} = 8.70$; **A_12:** $pK_{i,actual} = 7.48$; Table 1). Also, the overprediction of compound **D_16** can be explained by the fact that compounds with similar side chain scaffolds (**A_12, B_04**) have higher affinity values for CA I (**D_16**: $pK_{i,actual} = 4.81$; CoMFA: $pK_{i,predict} = 6.58$; CoMSIA: $pK_{i,predict} = 6.25$; **A_12:** $pK_{i,actual} = 7.48$; **B_04:** $pK_{i,actual} = 6.77$). For the CA II models, only CoMFA predicts the affinity of **D_16** too high, whereas for all CA IV models the deviations of the predicted affinity values of the test compounds are all within an acceptable range of one logarithmic unit.

Table 4. Statistical Results of the PLS Analysis of the 60 Training Data Set Compounds[a]

	CA I minimized in CA I		CA II minimized in CA II		CA IV minimized in CA IV	
	CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
q^2	0.573	0.522	0.413	0.572	0.326	0.425
S_{PRESS}	0.868	0.91	0.687	0.598	0.767	0.721
r^2	0.903	0.882	0.857	0.952	0.749	0.933
S	0.413	0.452	0.339	0.201	0.468	0.246
F	128.7	139.7	111.7	212.9	85	191.3
components	4	3	3	5	3	4
grid	2 Å					

[a] Results are plotted for CA I, CA II, and CA IV minimized in the corresponding binding pocket.

3D QSAR Selectivity Models. For obtaining 3D QSAR selectivity models, the affinity differences (ΔpK_i) for each compound between two particular CA isozymes were calculated. Analogous to the model generation of the individual models, the affinity difference values should spread at least over three logarithmic units in order to allow for the establishment of statistically significant 3D QSAR selectivity models. The affinities and affinity differences of all 87 inhibitors for each of the three selectivity models are summarized in Table 6. Affinity differences were calculated between CA I and CA II (CA I−CA II selectivity model), CA I and CA IV (CA I−CA IV selectivity model), and CA II and CA IV (CA II−CA IV selectivity model). For example, for the CA I−CA II selectivity model the CA II pK_i values of the inhibitors were subtracted from the CA I pK_i values. Negative numbers indicate that these compounds are stronger inhibitors for CA II, whereas CA I selective compounds exhibit positive numbers. The same calculations were performed for the CA I−CA IV and CA II−CA IV selectivity models. As highlighted in Table 6, the values for the affinity differences spread over three logarithmic units for the CA I−CA II (4.10 log units) and CA I−CA IV (3.52 log units) models, satisfying the minimum criteria mentioned above. Unfortunately, the range of affinity differences between CA II and CA IV is only 2.19 log units. Obviously, there are no compounds present in the data set that inhibits either CA II or CA IV by at least 1000-fold better than the other isozymes. Due to this fact as well as considering the high correlation between CA II and CA IV data ($r = 0.86$; Figure 3C), the chances to obtain statistically significant CA II−CA IV selectivity 3D QSAR models appear to be rather limited. Another important aspect to note is the fact that there are no compounds in the data set inhibiting CA I or CA IV significantly stronger than CA II (maximum ΔpK_i of 0.30 in favor of CA I and 0.28 for CA IV, respectively). Nevertheless, for each alignment in CA I, CA II, and CA IV CoMFA and CoMSIA selectivity models were derived. The statistical results using the 2 Å grid spacing are summarized in Table 7. Similar results were obtained with the 1 Å grid spacing (data not shown). Significant selectivity models were obtained for **CA I−CA II** evaluating all alignments; however, the best statistical parameters resulted

Table 5. Predicted Affinities of the 27 Compounds in the Test Data Set

scaffold	CA I $pK_{i,act}$	CA I CoMFA	CA I CoMSIA	CA II $pK_{i,act}$	CA II CoMFA	CA II CoMSIA	CA IV $pK_{i,act}$	CA IV CoMFA	CA IV CoMSIA
A_02	6.34	−0.23	−0.46	8.52	−0.03	0.06	6.90	−0.40	−0.66
A_04	6.52	−0.36	−0.50	8.70	0.02	0.19	8.10	0.09	0.24
A_10	8.22	0.87	0.44	8.70	−1.05	−0.78	8.30	0.55	0.36
A_13	6.00	−1.47	−1.73	8.52	−0.55	−0.57	7.60	−0.53	−1.11
A_16	6.62	0.32	−0.18	8.22	0.03	0.01	7.60	0.09	0.26
A_17	8.15	1.34	0.54	8.10	−0.18	0.18	7.82	−0.21	0.24
B_01	5.70	−0.07	−0.35	8.30	−0.09	0.06	7.96	−0.38	0.20
B_03	6.46	0.56	0.36	8.10	−0.30	−0.04	7.40	−1.05	−0.46
C_04	7.19	0.33	0.41	9.15	0.02	0.71	7.96	0.26	0.77
C_05	7.15	0.54	0.22	8.05	0.01	−0.28	7.72	−0.21	0.02
C_08	6.92	0.04	−0.03	9.30	0.90	0.92	8.10	0.07	0.22
C_11	7.60	0.94	0.83	8.10	−0.34	0.01	7.89	−0.23	0.34
D_02	4.11	−1.02	−0.66	6.49	−0.04	−0.37	5.49	−0.39	−0.57
D_04	4.68	0.51	−0.28	6.80	0.13	−0.31	5.61	−0.55	−0.94
D_09	4.68	0.30	0.15	6.41	−0.52	0.08	5.04	−1.18	−0.67
D_12	4.60	−0.20	−0.78	6.98	0.23	0.42	6.80	0.18	0.03
D_16	4.81	−1.77	−1.44	6.87	−1.40	−0.62	6.75	−0.39	−0.22
D_19	5.05	−0.36	−1.29	7.44	0.04	−0.18	7.00	0.23	0.24
D_21	5.10	−1.28	−0.73	7.60	−0.22	−0.81	7.41	0.97	0.15
D_29	4.81	−0.11	−0.68	6.88	−0.08	−0.11	6.77	0.14	0.57
D_30	7.27	0.76	1.03	7.92	0.67	0.51	7.44	0.38	0.40
D_31	7.35	0.85	0.58	7.96	0.13	0.29	7.48	0.41	0.37
D_33	6.94	0.24	−0.44	7.82	−0.15	−0.20	7.21	−0.10	−0.39
E_02	8.10	1.43	0.98	7.96	0.49	0.38	7.89	0.51	0.34
E_03	7.48	0.58	−0.10	7.41	−0.42	−0.44	7.42	−0.45	−0.50
E_04	7.30	0.01	−0.19	8.30	0.74	0.55	7.41	−0.39	−0.45
F_08	4.74	0.18	0.22	7.59	0.03	0.39	7.13	0.32	0.51
MEAN	6.29	0.11	−0.15	7.86	−0.07	0.00	7.27	−0.08	−0.03
STDEV	1.25	0.78	0.69	0.75	0.48	0.44	0.79	0.48	0.49
MAX	8.22	1.43	1.03	9.30	0.90	0.92	8.30	0.97	0.77
MIN	4.11	−1.77	−1.73	6.41	−1.40	−0.81	5.04	−1.18	−1.11
RANGE	4.11	3.20	2.76	2.89	2.30	1.73	3.26	2.15	1.88

using the alignment in CA I ($q^2 = 0.555$; CoMFA). As expected from the previous data analysis and due to the high correlation between **CA II and CA IV**, selectivity models for CA II−CA IV are below the commonly accepted threshold for statistically significant models of $q^2 > 0.3$ for each alignment. The best models were derived for the alignment in CA II (CoMFA: $q^2 = 0.281$; CoMSIA: $q^2 = 0.296$). Comparing the different alignments in each CA isozyme, the q^2 values observed for all models aligned in CA II show the smallest deviation between all selectivity models (CA II alignment: $q^2_{min} = 0.281$, $q^2_{max} = 0.460$; CA I: $q^2_{min} = 0.062$, $q^2_{max} = 0.555$; CA IV: $q^2_{min} = 0.011$, $q^2_{max} = 0.470$), indicating that the alignment in CA II seems to be the most robust one. Similar to the analysis of the individual 3D QSAR models, this observation can be explained by the fact that most protein−ligand complexes used for the alignment have been experimentally determined with CA II. Obviously, the more protein−ligand template structures are available the more reliable the positioning of modeled inhibitor side chains can be performed during the alignment process.

Graphical Interpretation of the Selective 3D QSAR Models. CoMFA Steric Fields. Although steric, electrostatic, hydrophobic, hydrogen bond donor, and acceptor fields were considered during the model generation process, only the steric and hydrogen bond acceptor properties will be discussed in the following, since they allow interpretation of the field contributions in terms of the protein binding pocket environment, whereas the other fields do not suggest clear-cut contours in this case. The isocontours of the field contributions *(*"stdev*coeff") for the **steric** properties as found by CoMFA and the **hydrogen** bond acceptor properties suggested by CoMSIA are illustrated together with exemplary ligands. The best 3D QSAR CA I−CA II selectivity model was derived with the alignment in CA I (CoMFA: $q^2 = 0.555$). The steric properties derived from this selectivity model are shown in Figure 7. Areas indicated by green contours correspond to regions where steric occupancy with bulky ligand groups increases selectivity toward CA I. In contrast, areas highlighted in yellow correspond to regions where steric occupancy with bulky inhibitor side chains will increase selectivity toward CA II. CA I-selective inhibitors lack bulky side chains in close proximity to yellow contours, whereas in CA II-selective inhibitors bulky groups are absent adjacent to regions highlighted in green. In addition, the amino acid residues in the active site of CA I and CA II are displayed that are in close proximity to the contoured isopleths (**CA I** (green): Phe91, Leu131, His200, Tyr204; **CA II** (yellow): Ile91, Phe131, Thr200, Leu204). Figure 7A shows one of the most active CA I inhibitors (**A_17**) with a pK_i of 8.15 for CA I and 8.10 for CA II. In Figure 7B, the CA II selective compound **A_09** is depicted (pK_i CA I = 6.70; pK_i CA II = 9.40). For clarity, in Figure 7A only contours (green) indicating selectivity toward CA I and in Figure 7B only contours (yellow) indicating CA II selectivity are plotted. One of the most CA II-selective compounds is dorzolamide (DZM; **B_05**, Table 1) with a pK_i of 4.30 for CA I and 8.05 for CA II. The weak inhibition of CA I by DZM can be related to amino acid differences between CA I and CA II (see the Discussion section below). Close to the sterically selective areas of CA I (green) bulky amino acid side chains of CA II point toward these CA I selectivity enhancing regions. Noticeable, Phe131 in CA II penetrates into the CA I-selectivity enhancing contour,

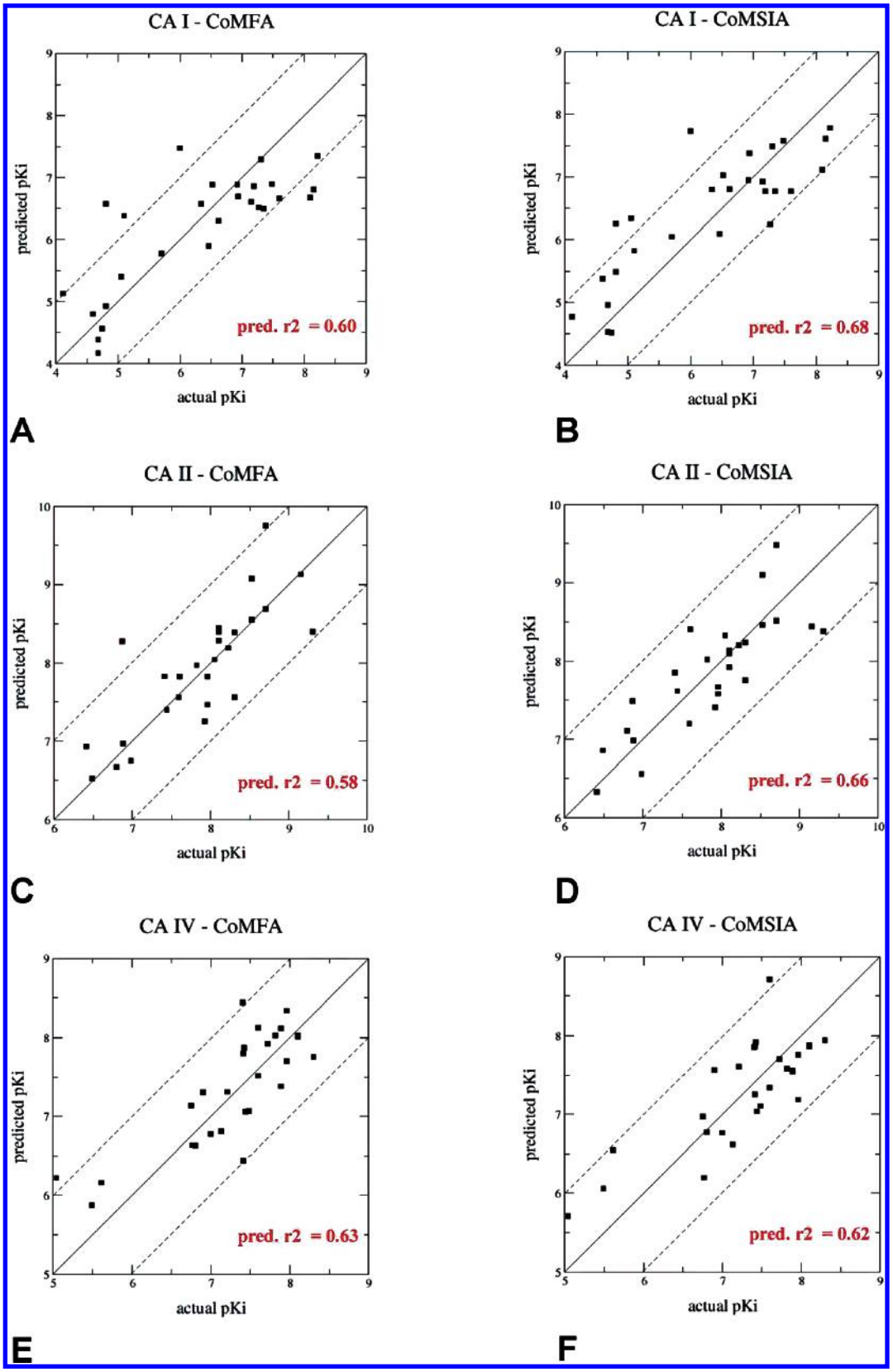


Figure 6. Predicted affinities for the 27 test data set compounds for CA I (**A, B**), CA II (**C, D**), and CA IV (**E, F**) obtained with CoMFA and CoMSIA are plotted against the actual values. The good predictive r^2 values (all models > 0.5) confirm the usefulness of the 3D QSAR models for predicting the affinity for CA I, CA II, and CA IV inhibition.

whereas in CA I a sterically less demanding Leu131 is found in this position (Figure 7A). On the other hand, close to the contour enhancing CA II selectivity due to steric reasons (yellow), in CA I bulky amino acid side chains point toward this area (His200, Tyr204, and Phe91). In contrast, sterically less demanding amino acids are exposed to this region in CA II (Thr200, Leu204, Ile91; Figure 7B). The binding modes of **A_17** (Figure 7A) and **A_09** (Figure 7B) based

Table 6. Affinity of CA Inhibitors toward CA I, CA II, and CA IV, and Their Affinity Differences

no.	scaffold	CA I pK_i	CA II pK_i	CA IV pK_i	CA I−CA II ΔpK_i	CA I−CA IV ΔpK_i	CA II−CA IV ΔpK_i	no.	scaffold	CA I pK_i	CA II pK_i	CA IV pK_i	CA I−CA II ΔpK_i	CA I−CA IV ΔpK_i	CA II−CA IV ΔpK_i
1	A_01	5.07	7.22	6.27	−2.15	−1.20	0.95	47	D_13	4.94	7.21	6.99	−2.27	−2.05	0.22
2	A_02	6.34	8.52	6.90	−2.18	−0.56	1.62	48	D_14	4.80	6.85	6.78	−2.05	−1.98	0.07
3	A_03	6.52	8.30	7.89	−1.78	−1.37	0.41	49	D_15	5.67	7.10	6.85	−1.43	−1.18	0.25
4	A_04	6.52	8.70	8.10	−2.18	−1.58	0.60	50	D_16	4.81	6.87	6.75	−2.06	−1.94	0.12
5	A_05	6.26	7.68	7.25	−1.42	−0.99	0.43	51	D_17	4.82	7.32	6.85	−2.50	−2.03	0.47
6	A_06	6.05	7.82	7.00	−1.77	−0.95	0.82	52	D_18	6.20	7.96	7.52	−1.76	−1.32	0.44
7	A_07	7.44	8.22	7.68	−0.78	−0.24	0.54	53	D_19	5.05	7.44	7.00	−2.39	−1.95	0.44
8	A_08	8.52	8.22	8.10	0.30	0.42	0.12	54	D_20	6.16	7.92	6.81	−1.76	−0.65	1.11
9	A_09	6.70	9.40	8.05	−2.70	−1.35	1.35	55	D_21	5.10	7.60	7.41	−2.50	−2.31	0.19
10	A_10	8.22	8.70	8.30	−0.48	−0.08	0.40	56	D_22	5.22	7.70	7.52	−2.48	−2.30	0.18
11	A_11	8.70	10.00	8.52	−1.30	0.18	1.48	57	D_23	7.38	8.22	7.30	−0.84	0.08	0.92
12	A_12	7.48	8.70	8.05	−1.22	−0.57	0.65	58	D_24	6.05	8.00	6.39	−1.95	−0.34	1.61
13	A_13	6.00	8.52	7.60	−2.52	−1.60	0.92	59	D_25	5.90	7.30	6.40	−1.40	−0.50	0.90
14	A_14	6.68	9.52	8.22	−2.84	−1.54	1.30	60	D_26	4.82	6.92	6.82	−2.10	−2.00	0.10
15	A_15	6.52	8.15	7.48	−1.63	−0.96	0.67	61	D_27	6.26	8.15	7.40	−1.89	−1.14	0.75
16	A_16	6.62	8.22	7.60	−1.60	−0.98	0.62	62	D_28	5.69	7.70	6.92	−2.01	−1.23	0.78
17	A_17	8.15	8.10	7.82	0.05	0.33	0.28	63	D_29	4.81	6.88	6.77	−2.07	−1.96	0.11
18	A_18	6.05	7.92	6.66	−1.87	−0.61	1.26	64	D_30	7.27	7.92	7.44	−0.65	−0.17	0.48
19	B_01	5.70	8.30	7.96	−2.60	−2.26	0.34	65	D_31	7.35	7.96	7.48	−0.61	−0.13	0.48
20	B_02	6.40	8.30	7.77	−1.90	−1.37	0.53	66	D_32	7.01	7.92	7.26	−0.91	−0.25	0.66
21	B_03	6.46	8.10	7.40	−1.64	−0.94	0.70	67	D_33	6.94	7.82	7.21	−0.88	−0.27	0.61
22	B_04	6.77	8.30	8.00	−1.53	−1.23	0.30	68	D_34	7.27	7.82	7.49	−0.55	−0.22	0.33
23	B_05	4.30	8.05	7.37	−3.75	−3.07	0.68	69	D_35	7.52	8.22	7.89	−0.70	−0.37	0.33
24	C_01	7.26	8.10	7.77	−0.84	−0.51	0.33	70	D_36	6.90	7.38	7.12	−0.48	−0.22	0.26
25	C_02	7.30	8.15	7.82	−0.85	−0.52	0.33	71	D_37	7.04	7.48	7.25	−0.44	−0.21	0.23
26	C_03	7.74	9.30	8.15	−1.56	−0.41	1.15	72	D_38	7.46	8.15	7.74	−0.69	−0.28	0.41
27	C_04	7.19	9.15	7.96	−1.96	−0.77	1.19	73	D_39	7.33	7.96	7.62	−0.63	−0.29	0.34
28	C_05	7.15	8.05	7.72	−0.90	−0.57	0.33	74	E_01	8.15	8.10	8.00	0.05	0.15	0.10
29	C_06	8.00	8.40	7.92	−0.40	0.08	0.48	75	E_02	8.10	7.96	7.89	0.14	0.21	0.07
30	C_07	7.89	8.30	7.72	−0.41	0.17	0.58	76	E_03	7.48	7.41	7.42	0.07	0.06	−0.01
31	C_08	6.92	9.30	8.10	−2.38	−1.18	1.20	77	E_04	7.30	8.30	7.41	−1.00	−0.11	0.89
32	C_09	7.39	9.70	8.30	−2.31	−0.91	1.40	78	E_05	7.68	7.80	7.77	−0.12	−0.09	0.03
33	C_10	7.31	9.22	8.15	−1.91	−0.84	1.07	79	E_06	7.49	7.46	7.52	0.03	−0.03	−0.06
34	C_11	7.60	8.10	7.89	−0.50	−0.29	0.21	80	F_01	4.64	6.82	6.68	−2.18	−2.04	0.14
35	D_01	4.55	6.52	5.52	−1.97	−0.97	1.00	81	F_02	4.70	7.52	7.80	−2.82	−3.10	−0.28
36	D_02	4.11	6.49	5.49	−2.38	−1.38	1.00	82	F_03	5.74	7.82	7.57	−2.08	−1.83	0.25
37	D_03	4.60	6.77	5.55	−2.17	−0.95	1.22	83	F_04	5.80	8.05	7.68	−2.25	−1.88	0.37
38	D_04	4.68	6.80	5.61	−2.12	−0.93	1.19	84	F_05	4.92	7.40	6.54	−2.48	−1.62	0.86
39	D_05	4.74	6.96	6.35	−2.22	−1.61	0.61	85	F_06	4.29	7.15	6.90	−2.86	−2.61	0.25
40	D_06	4.82	6.91	6.76	−2.09	−1.94	0.15	86	F_07	4.24	8.05	7.15	−3.81	−2.91	0.90
41	D_07	3.86	5.30	4.61	−1.44	−0.75	0.69	87	F_08	4.74	7.59	7.13	−2.85	−2.39	0.46
42	D_08	4.10	6.26	4.35	−2.16	−0.25	1.91	MEAN		6.21	7.81	7.19	−1.60	−0.98	0.62
43	D_09	4.68	6.41	5.04	−1.73	−0.36	1.37	STDEV		1.27	0.84	0.88	0.89	0.86	0.46
44	D_10	4.82	6.27	4.96	−1.45	−0.14	1.31	MAX		8.70	10.0	8.52	0.30	0.43	1.91
45	D_11	4.82	6.26	4.96	−1.44	−0.14	1.30	MIN		3.86	5.30	4.35	−3.80	−3.10	−0.28
46	D_12	4.60	6.98	6.80	−2.38	−2.20	0.18	RANGE		4.84	4.70	4.18	4.10	3.52	2.19

on the alignment in CA I and CA II display two examples which may explain the importance of these amino acids for inhibitor binding. In CA II, the binding pocket is split by **Phe131** in two portions. As indicated by most of the modeled protein−ligand complexes, aliphatic or aromatic inhibitor side chains interact with hydrophobic amino acids in the hydrophobic part of the pocket (Figures 4A and 7B). The observation that increasingly hydrophobic inhibitor side chains form favorable *van der Waals* interactions with the hydrophobic pocket correlates with increased affinity toward CA II (especially for F-substituted aromatic side chains; A_09, A_14, C_03, C_04, C_08). On the other hand, inhibitors with less flexible side chains, such as **A_17**, are not able to fully accommodate the hydrophobic pocket, thus no optimal interactions are formed with Phe131 in CA II. In CA I, **A_17** fits perfectly into the more open CA I pocket (Leu131 instead of Phe131), which explains the higher affinity for CA I for this type of inhibitor compared to other benzenesulfonamides (Figure 7A). Furthermore, the amino acid residues at position 200 likely influences the selectivity inhibition profile of CA I versus CA II. **His200** in CA I restricts the space for more bulky side chains, whereas **Thr200** in CA II is sterically less demanding and provides more space for inhibitor functional groups.

The binding mode of dorzolamide (**B_05**) in CA II has been described in detail,[52] also taking the rearrangement of the flexible "*proton-shuttle*" His64 into account. Here, we collect computational evidence that more bulky side chains in this region not only explain the more effective CA II binding but also suggest weaker binding toward CA I. Assuming the same binding mode as observed in CA II, the ethylamino side chain of **B_05** forms unfavorable interactions with His200 in CA I. In consequence, dorzolamide derivatives lacking this ethylamino side chain are expected to be more potent inhibitors of CA I than dorzolamide itself. The analysis of the CA I−CA II selectivity model based on the alignment in CA II (CoMFA: $q^2 = 0.460$) results in similar contour plots as described above (data not shown).

CoMSIA Acceptor Fields. In addition to CoMFA's steric and electrostatic fields, hydrophobic, hydrogen bond donor

Table 7. Statistical Results of the 3D QSAR Selectivity Analyses for the CoMFA and CoMSIA Method with a Grid Spacing of 2 Å

	CA I−CA II		CA I−CA IV		CA II−CA IV	
	CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
Minimized in the Binding Pocket of CA I (Selectivity Model)						
q^2	0.555	0.441	0.406	0.253	0.062	0.102
S_{PRESS}	0.607	0.683	0.682	0.756	0.447	0.438
r^2	0.832	0.842	0.849	0.683	0.318	0.340
S	0.373	0.364	0.344	0.492	0.381	0.375
F	136.5	109.1	115.2	90.5	39.6	43.7
components	3	4	4	2	1	1
Minimized in the Binding Pocket of CA II (Selectivity Model)						
q^2	0.460	0.392	0.391	0.341	0.281	0.296
S_{PRESS}	0.668	0.705	0.691	0.71	0.401	0.395
r^2	0.804	0.639	0.828	0.662	0.821	0.787
S	0.403	0.543	0.367	0.508	0.2	0.217
F	113.3	74.3	98.4	82.3	74.3	75.9
components	3	2	4	2	5	4
Minimized in the Binding Pocket of CA IV (Selectivity Model)						
q^2	0.312	0.470	0.250	0.427	0.263	0.011
S_{PRESS}	0.754	0.666	0.762	0.682	0.403	0.459
r^2	0.754	0.869	0.751	0.954	0.826	0.308
S	0.451	0.331	0.439	0.193	0.196	0.384
F	84.8	135.7	83.6	235.9	97.3	87.9
components	3	4	3	7	4	1
grid	2 Å					

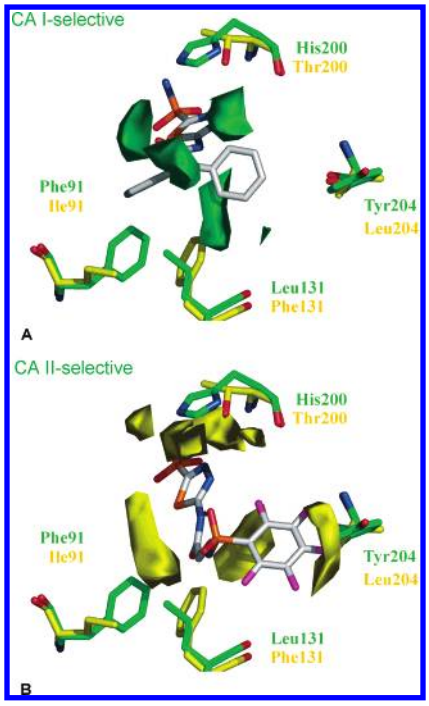


Figure 7. Steric selective contour plots obtained with the CoMFA model (CA I alignment; q^2 value = 0.555) for CA I (green, contour level: 0.026) and CA II (yellow, contour level: −0.014). Amino acids in **CA I** (green, clockwise from the top: His200, Tyr204, Leu131, Phe91) and **CA II** (yellow, Thr200, Leu204, Phe131, Ile91) that are in close proximity to these selective steric regions are displayed. Areas indicated by green contours correspond to regions where steric occupancy with bulky ligand groups increases selectivity toward CA I (**A**). By contrast, areas highlighted in yellow correspond to regions where steric occupancy with bulky inhibitor side chains will increase selectivity toward CA II (**B**). One of the most active CA I inhibitors (**A: A_17**, pK_i CA I = 8.15, pK_i CA II = 8.10) and a selective CA II inhibitor (**B: A_09**, pK_i CA I = 6.70, pK_i CA II = 9.40) are shown.

and acceptor fields were calculated with CoMSIA.[51] Analysis of the hydrogen bond acceptor contours based on the CoMSIA selectivity model suggest correlations of the amino acid composition with respect to selectivity discriminating features between CA I and CA II. Figure 8 illustrates the acceptor contours that discriminate between CA I and CA II. Areas highlighted in red indicate that acceptor ligand atoms in close proximity increase selectivity toward CA II, whereas acceptor atoms nearby blue colored regions increase selectivity toward CA I (Figures 8A and 8B). Viewed from the protein, donor atoms in the amino acid residues falling next to red contours should increase the selectivity toward CA II, whereas donor atoms in residues nearby regions highlighted in blue should increase affinity toward CA I. Amino acids in CA I are displayed in green (Figure 8A), and those in CA II are highlighted in yellow (Figure 8B). In both figures, the binding mode of DZM (**B_05**) is shown as observed in the crystal structure of CA II (PDB code: *1cil*).

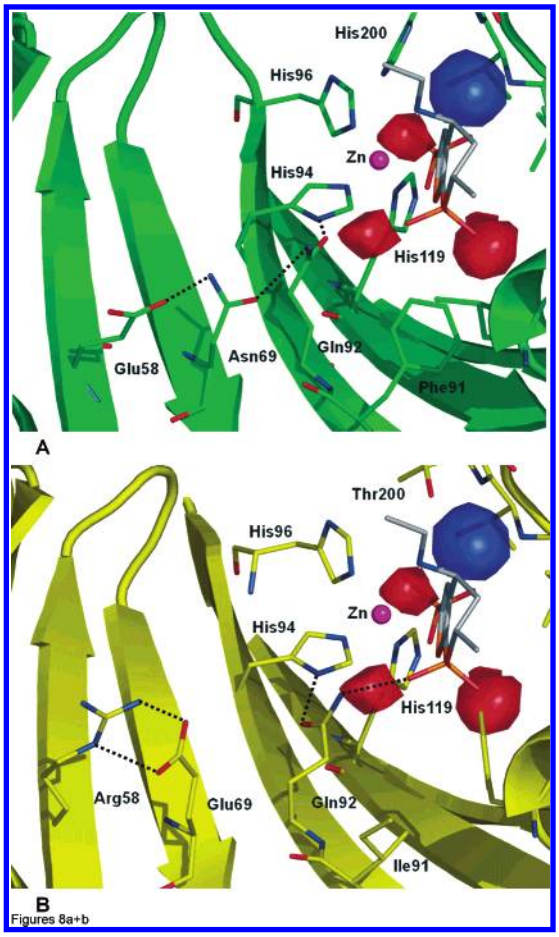


Figure 8. Acceptor contour plots obtained with the CoMSIA model for CA I (blue, contour level: 0.027) and for CA II (red, contour level: −0.0018) that discriminates between both isozymes. Selected amino acids in CA I (**A**, green) and CA II (**B**, yellow) are shown. Contour plots highlighted in red indicate that acceptor ligand atoms in close proximity increase selectivity toward CA II, whereas acceptor atoms nearby blue colored regions increase selectivity toward CA I. A red CA II selective region close to Gln92 suggests that nearby ligand acceptor atoms increase the selectivity toward CA II. Gln92 in CA II forms a single hydrogen bond with one nitrogen of the imidazole ring of His94 (dotted lines). The nitrogen atom of the Gln92 side chain forms no further hydrogen bonds with the protein and therefore could interact with ligand acceptor atoms (**B**). In CA I, the amide nitrogen of the Gln92 side chain is involved in a hydrogen-bonding network with the neighboring Asn69 side chain oxygen (**A**). Due to the different hydrogen-bonding networks in CA I and CA II, Gln92 adopts different conformations that can influence ligand binding. The binding mode of dorzolamide in complex with CA II is shown (PDB code: *1cil*).

CA II acceptor contours (red) are in close proximity of **Gln92**. In both isozymes, the carbonyl oxygen of the terminal amide group of this residue forms a hydrogen bond to one imidazole nitrogen of His94. However, while the terminal amide nitrogen of Gln92 forms in CA I a hydrogen bond to the side chain carbonyl oxygen of **Asn69**, in CA II this amide nitrogen forms no further hydrogen bonds with the protein. Instead, the neighboring **Glu69** in CA II interacts via twinned hydrogen bonds with **Arg58**; therefore, no hydrogen bond can be formed between Glu69 and Gln92 in CA II (Figure 8B). In CA I, **Glu58** interacts with the amino group of **Asn69** (Figure 8A). Due to these different hydrogen bonding networks in CA I and CA II, the side chain of Gln92 adopts different conformations. In the case of CA II, it points via its amino group to the binding pocket. Close to this amino group a selectivity enhancing acceptor field for CA II is indicated by CoMSIA, suggesting that acceptor ligand atoms could form favorable interactions with the amino group of Gln92. No comparable favorable interactions are observed for acceptor ligand atoms in CA I, since here the amino group of Gln92 is involved in a protein-internal hydrogen-bonding network (Figure 8A). This pattern seems highly conserved, and only Arg58 in CA II exhibits a different side chain conformation in some structures, exposing this residue toward the solvent but still interacting with Glu69, at least via a single hydrogen bond. Distances between amino acids in CA I and CA II that are involved in the hydrogen bond network are summarized and provided in the Supporting Information.

The statistical analysis of the **CA II−CA IV** selectivity models indicate that no significant correlations were obtained for the alignments in CA I and CA IV (Table 7). These unsatisfactory results could evolve due to pronounced data correlation of CA II and CA IV and the modest spread of the affinity differences. Based on the alignment in CA II only, selectivity models were derived that are slightly below the significance threshold of $q^2 = 0.3$ (CoMFA: $q^2 = 0.281$, CoMSIA: $q^2 = 0.296$). Evaluation of the CoMSIA hydrophobic fields indicates that hydrophobic ligand portions placed next to the hydrophobic pocket in CA II (Leu198, Leu204, Cys206, Val135, and Phe131) possibly improve the selectivity toward CA II (data not shown). A superposition of CA IV (PDB code: *1znc*) with CA II shows that in CA IV predominantly hydrophilic amino acids (Asp204, Lys206, Asp136, Glu140, and Glu123) are next to this region. However, a more thorough analysis would require improved statistical significance of the selectivity models between CA II and CA IV.

Binding Mode of Thienothiopyrane Derivatives. Considering the inhibition profiles of dorzolamide (**B_05**) toward CA I and CA II proves this ligand as one of the most potent CA II selective inhibitors ($\Delta p K_i = -3.75$). Modification of the thienothiopyrane scaffold leads to an increase in CA I inhibition (B_01, B_02, B_03, B_04; pK_i CA I between 5.70 and 6.77), whereas with respect to CA II these compounds do not differ significantly (pK_i between 8.05 and 8.30, Table 1). Figure 9 illustrates the predicted binding mode of B_05 and B_03 after minimization in CA I (Figures 9A and 9C) and CA II (Figures 9B and 9D). The selectivity profile between these isozymes reveals B_05 to be a more CA II selective compound ($\Delta p K_i = -3.75$) compared to B_03 ($\Delta p K_i = -1.64$; Table 6). Up to now, there are no crystal structures described with B_03 in complex with CA I or CA II; for dorzolamide (B_05) only the crystal structure with CA II is known. Nevertheless, after minimization of B_03 and dorzolamide in CA I and CA II, different protein−ligand interactions are obtained with the respective isozymes which could give some indications for the selectivity enhancement of B_05 toward B_03.

The binding modes of dorzolamide (B_05) in CA I (Figure 9A) and CA II (Figure 9B) show that the sulfonamide group coordinates in both cases to the zinc ion and forms important hydrogen bonds to Thr199. CA I and CA II differ in position 200 (**CA I: His200; CA II: Thr200**). In CA II, the nitrogen of the ethylamino substituent of B_05 interacts with Thr200, and an additional hydrogen bond is formed between one oxygen of the cyclic sulfone and the amide nitrogen of Gln92 (Figure 9B). However, after minimizing dorzolamide in CA I no hydrogen bonds are indicated to His200. In addition, bulky side chains are more favorable at that position for CA II inhibition (Figure 7B, yellow contour plots). Accordingly, CA I selective inhibitors should avoid sterically demanding substituents next to His200. Furthermore, ligands with acceptor atoms in close proximity to Gln92 enhance CA II inhibition due to the possible involvement in hydrogen bond interactions to Gln92 (Figure 8B), whereas in CA I, Gln92 is engaged in hydrogen bonds to Asn69 (Figure 8A).

In the proposed binding mode of B_03 in CA I the ligand forms a hydrogen bond via its pyrrol ring to His67 in CA I (Figure 9C). In CA II, a similar interaction to Asn67 (Figure 9D) was observed, but in consequence no further direct hydrogen bond to Thr200 could be accomplished. Thus, B_03 can form one additional hydrogen bond in CA I compared to CA II.

Docking CA Inhibitors into CA II. The established 3D QSAR models were further evaluated to assess their predictive power to estimate binding affinity of docked CA inhibitors. Docking generally faces two main problems: the accuracy of the docked binding poses (the "docking problem"), and the correct prediction of the binding affinities (the "scoring problem").[53] Topiramate[54,55] and RWJ-37947[56,57] (Figure 1) show a similar overall topology, but they obey slightly different substitution patterns. One of the diisopropylidene moieties of topiramate is replaced by a cyclic sulfate group in RWJ-37947. Solely based on bonding topology, it is tempting to assume similar binding properties of both molecules, resulting in comparable binding modes. However, experiment shows the opposite with two clearly alternative orientations. In addition, acetazolamide (AZM, **A_18**), dorzolamide (DZM, **B_05**), brinzolamide (BZM), and a benzenesulfonamide (**D_01**) were selected for this validation study.

The program AUTODOCK was used for docking the ligands into CA II. This program performs docking on a precalculated grid storing potential values from any sort of interaction field, e.g. Lennard-Jones and Coulomb potentials. It accomplishes multiple stochastic searches on the potential hypersurface; accordingly, the frequency of occurrence of certain docking solutions can be used as an additional figure-of-merit to assess the relevance of the generated solutions. With the exception of His64, CA II is virtually a rigid protein. Two conformations of the "*proton shuttle*" His64 are observed: "*in*" conformation with His64 pointing toward the active site and "*out*" with His64 orienting to the solvent. For docking, the crystal structure of CA II with His64 in

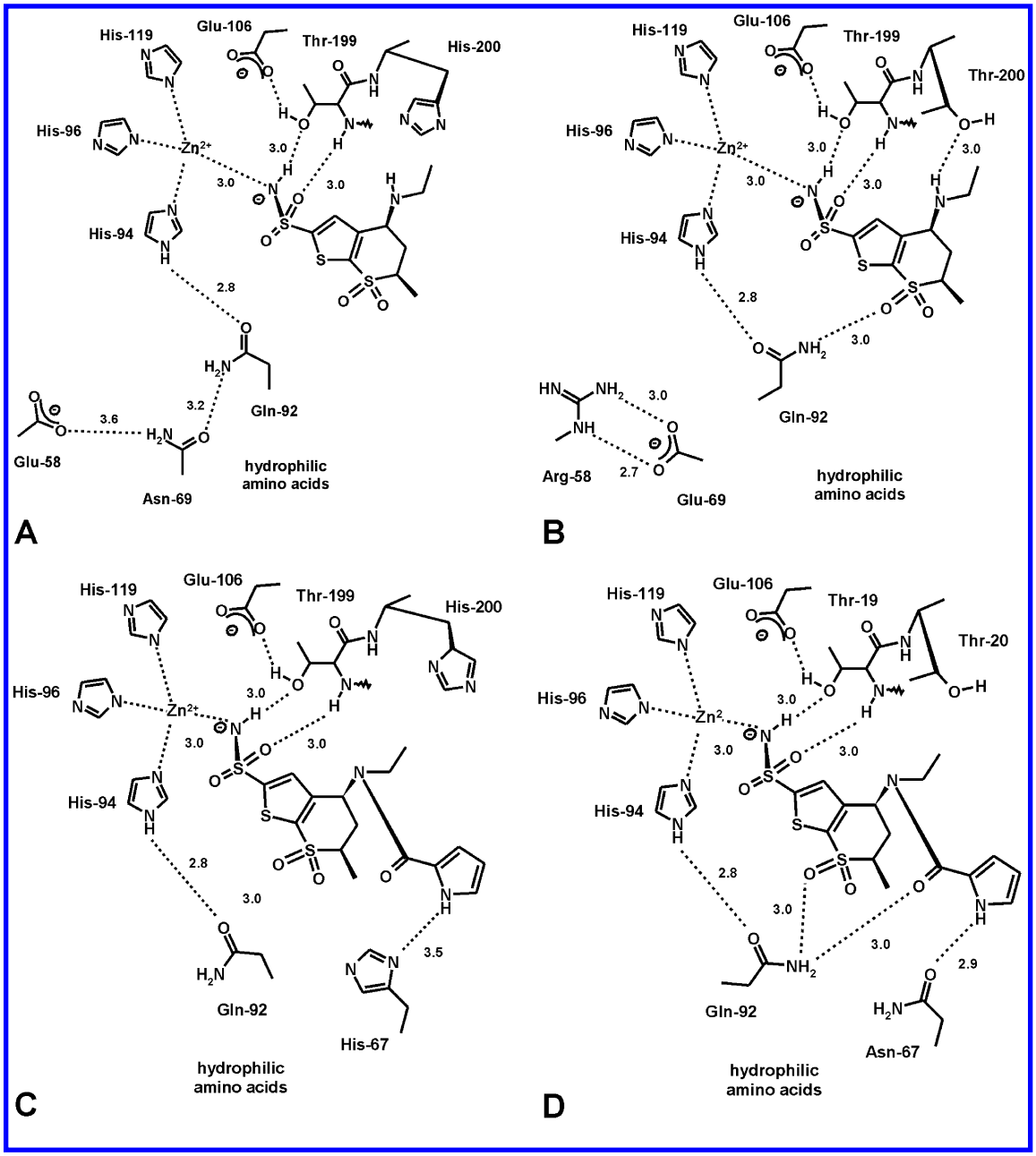


Figure 9. Predicted binding mode of thienothiopyrane derivatives (**B_05** (top), **B_03** (bottom)) in **CA I** (**A, C**) and CA II (**B, D**) after minimization in both isozymes (CA I: *1azm*; CA II: *1cil*). Possible hydrogen bonds and the coordination to the zinc ion are indicated in dotted lines (**B_05**: pK_i CA I = 4.30, pK_i CA II = 8.05; **B_03**: pK_i CA I = 6.46, pK_i CA II = 8.10).

the "out" conformation was used (*1cil*) to allow for enough space during ligand placement. Up to 100 docked poses were stored for each ligand. Results differing by less than 1 Å rmsd (root-mean-square deviation) were merged together into one cluster. We then analyzed the top ranked ligand pose of all 100 docked solutions as well as the first ligand pose of the largest cluster. The rmsd values of the docked solutions with respect to the crystal structure and the corresponding energy values are summarized in Table 8. With the exception of topiramate, the best-ranked cluster shows unsatisfactory rmsd values of more than 3.5 Å; the corresponding cluster size, however, is rather small (less than 11 cluster members). When the first rank in the largest cluster is considered, the rmsd values are all below 1.4 Å, implying that the resulting binding modes are close to the crystal structure (Figure 10). Comparing the scoring values for the first rank of the top and largest cluster, respectively, the differences are small (<0.5 kcal/mol) and fall into the degeneracy range previously identified for the same docking methodology.[58] The binding modes of CA inhibitors observed in the crystal structures (green) and the first rank of the largest cluster obtained with AUTODOCK (atom type color-coded) are shown in Figure 10. For all six inhibitors the predicted binding mode is close to the pose observed in the crystal structure (**A**: acetazolamide, rmsd = 1.23 Å; **B**: dorzolamide, 0.54; **C**: brinzolamide, 1.34; **D**: benzenesulfonamide D_01, 1.48; **E**: topiramate, 1.34; **F**: RWJ-37947, 0.54; Figure 10 and Table 8). In the case of acetazolamide (**A**) the docked solution is compared to the binding mode of acetazolamide in CA I (1azm), assuming that the binding mode of this compound in CA I is transferable to CA II. Surprisingly, also the deviating binding modes of the related topiramate and RWJ-

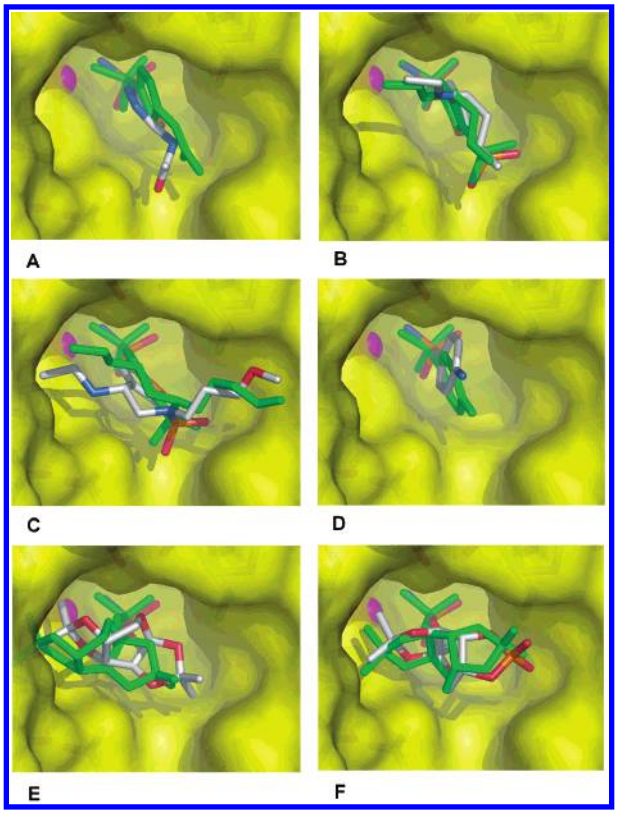


Figure 10. Binding mode prediction of selected CA inhibitors with AUTODOCK. Acetazolamide (**A**), dorzolamide (**B**), brinzolamide (**C**), a benzenesulfonamide (**D**), topiramate (**E**), and RWJ-37947 (**F**) were docked into the binding pocket of CA II (PDB code: *1cil*). For each compound, 100 docked solutions were stored, but only the top rank pose of the largest cluster is shown together with the binding mode observed from the crystal structure (green). The rmsd value between the top rank pose of the largest cluster and the crystal structure is 1.23 Å for acetazolamide (**A**), 0.54 Å for dorzolamide (**B**), 1.34 Å for brinzolamide (**C**), 1.48 Å for the benzenesulfonamide D_01 (**D**), 1.34 Å for topiramate (**E**), and 0.54 Å for RWJ-37947 (**F**) (see also Table 8).

Table 8. CA II Docking Results of Six Selected Protein−Ligand Complexes Using AUTODOCK[a]

inhibitor	cluster	cluster scoring rank[b]	cluster size	best energy in cluster [kcal/mol]	RMSD [Å] X-ray structure
AZM	top	1	11	−7.30	4.38
	largest	7	50	−6.90	1.23
DZM	top	1	6	−8.61	3.55
	largest	4	60	−8.45	0.54
BZM	top	1	1	−9.86	3.87
	largest	2	32	−9.71	1.34
TPM	top	1	71	−8.46	1.34
	largest	1	71	−8.46	1.34
RWJ-37947	top	1	1	−8.40	4.20
	largest	5	81	−7.95	0.54
D_01	top	1	4	−6.57	4.74
	largest	2	78	−6.52	1.48

[a] Rmsd and the corresponding energy values of the docked poses with respect to the observed binding mode in crystal structures for the top ranked ligand pose and for the first ligand pose of the largest cluster are shown. [b] The total number of clusters obtained in each case is 13 (AZM), 5 (DZM), 14 (BZM), 11 (TPM), 9 (RWJ), and 5 (D_01).

37947 were correctly generated by AUTODOCK considering the first rank of the largest cluster. The latter results indicate that AUTODOCK successfully generates the correct binding mode for thiadiazole-, thienothiopyrane-, and benzenesulfonamides following the descibed protocol.

Affinity Prediction with Individual CoMFA and CoMSIA Models. To assess whether the above-described 3D QSAR models can be used to quantitatively predict binding affinities of automatically docked ligands (e.g. as a result from a virtual screening run) a data set of 12 known CA inhibitors was compiled from the literature (Table 9, Figure 11)[17,59] and docked into the binding pocket of CA II using AUTODOCK following the described protocol. For affinity prediction, we used the scoring function implemented in AUTODOCK 3.0 as well as our derived CoMFA and CoMSIA models for CA II. Since acetazolamide (A_18), dorzolamide (B_05), and the benzenesulfonamide D_01 are included in the original QSAR data set of 87 inhibitors, they were omitted, and novel CA II CoMFA and CoMSIA models were derived (84 compounds; CoMFA: LOO, $q^2 = 0.557$, 4 components; CoMSIA: LOO, $q^2 = 0.633$, 5 components). These models were used for affinity prediction of the docked poses based on the first pose of the largest cluster. In Table 9, the actual and predicted inhibitory activities of the investigated compounds are summarized. In Figure 12, the predicted affinities are plotted against the experimental values of the examined compounds. Obviously, the affinity prediction using the AUTODOCK scoring function is less accurate (negative predictive r^2 value), whereas the CoMFA model reveals an acceptable estimate (predictive $r^2 = 0.263$). Even better results are obtained based on the CoMSIA model (predictive $r^2 = 0.497$). In summary, these results show accurately predicted binding modes of e.g. Celecoxib (rmsd = 1.40 Å) and binding affinities (also for the cyclooxygenase inhibitors) using our 3D QSAR models described in this contribution.

CONCLUSIONS

In summary, we successfully generated 3D QSAR CoMFA and CoMSIA models ($q^2 > 0.5$) for CA I, II, and IV, that were used as a starting point for the generation of selectivity 3D QSAR models for CA isozymes. Taking the affinity difference (ΔpK_i) for each compound with respect to two particular CA isozymes as independent variable, steric and hydrogen bond acceptor properties allowed the interpretation of field contributions in terms of the protein binding pocket environment (particularly between CA I and CA II). Even though no protein information was included in deriving the selectivity 3D QSAR models, we were able to spatially locate amino acids in CA I and CA II that take discriminative influence on the selective binding with respect to either one or the other isozyme. Whereas the amino acid residues 91, 131, 200, and 204 determine selectivity toward CA I and CA II based on their steric properties, Gln92 is expected to make a significant impact via its hydrogen bond acceptor properties on bound ligands. Computationally the observation could be rationalized that His200 in CA I restricts the space for more bulky side chains, whereas Thr200, present at this position in CA II, is sterically less demanding, thus offering more space for inhibitor functional groups. This explains the binding mode of several inhibitors (Figure 7) as well as those of dorzolamide and related derivatives (Figure 9) in both isozymes. The importance of Gln92 for selective inhibitor binding can be explained by differences

Table 9. Predicted Inhibitory Activities of Known and Novel CA Inhibitors[a]

inhibitor/no.	CA II $pK_{i,act}$	AutoDock energy [kcal/mol]	AutoDock $pK_{i,pred}$	CoMFA $pK_{i,pred}$	CoMSIA $pK_{i,pred}$
AZM	7.92	−6.90	4.29 (3.63)	7.69 (0.23)	7.01 (0.91)
DZM	8.05	−8.45	5.55 (2.50)	7.17 (0.88)	7.78 (0.27)
BRZ	8.49	−9.71	5.22 (3.27)	7.90 (0.59)	8.09 (0.40)
TPM	8.30	−8.46	5.00 (3.30)	7.91 (0.39)	8.56 (−0.26)
RWJ-37947	7.44	−7.95	4.93 (2.51)	7.53 (−0.09)	7.72 (−0.28)
D_01	6.52	−6.52	4.50 (2.02)	7.26 (−0.74)	6.81 (−0.29)
88	7.82	−11.12	6.63 (1.19)	6.83 (0.99)	6.77 (1.05)
89	8.00	−10.11	6.79 (1.21)	7.28 (0.72)	7.52 (0.48)
90	7.89	−10.11	6.68 (1.21)	7.12 (0.77)	7.23 (0.66)
91	7.68	−7.59	5.10 (2.58)	7.71 (−0.03)	7.16 (0.52)
92 (Valdecoxib)	7.37	−9.68	6.17 (1.20)	7.20 (0.17)	7.11 (0.26)
93 (Celecoxib)	7.68	−8.88	5.59 (2.09)	8.08 (−0.40)	7.85 (−0.17)

[a] The inhibitors were automatically docked into CA II with AUTODOCK. The top rank pose of the largest cluster was used for affinity prediction with the scoring function implemented in AUTODOCK 3.0 and our previously generated CoMFA and CoMSIA models (deviations from experimental values in parentheses).

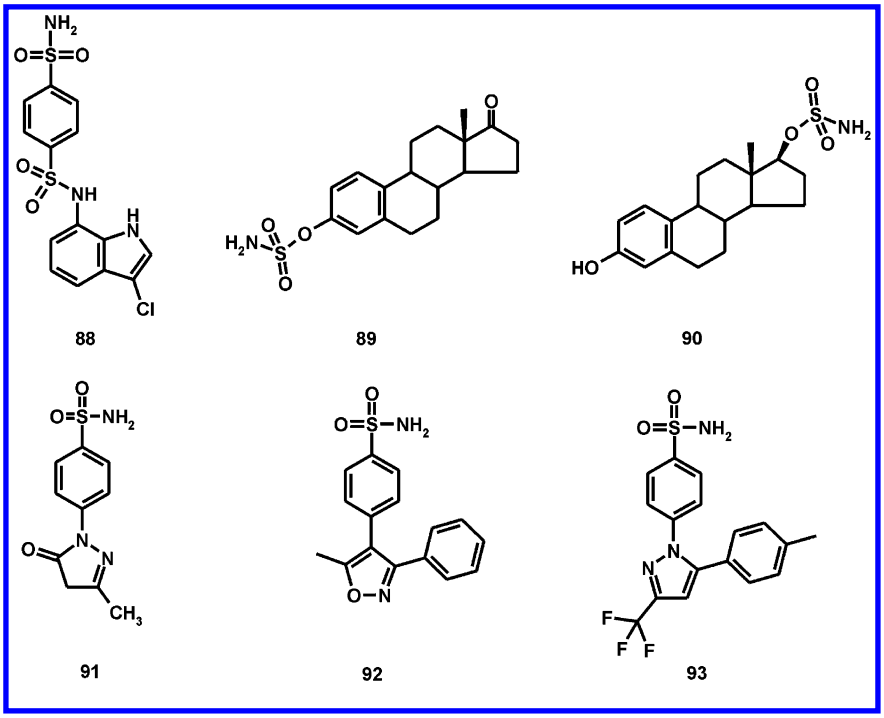


Figure 11. E7070 (**88**), two steroid sulfatase inhibitors (**89, 90**), a compound found with a virtual screening approach (**91**), and the cyclooxygenase inhibitors valdecoxib (**92**) and celecoxib (**93**) are shown.

in the hydrogen-bond network formed across the amino acids in CA I and CA II exposed toward the binding sites (Figure 8). Considering that the ligand alignment for CA II is supposedly more solid as many more reference protein−ligand complex structures are available, it serves as a basis in combination with affinity data for other CA isozymes (CA V, CA IX) to obtain further insight into selective inhibition of other CA isozymes that are, for example, involved in obesity or cancer.

We further established a docking protocol regarding the best ranked solution of the largest cluster obtained from 100 AUTODOCK-generated ligand poses to correctly predict the binding mode of six selected CA inhibitors. This conclusion is based on a comparison with crystal structures, and it shows that even the deviating binding modes of the two topologically rather similar compounds topiramate and RWJ-37947 were generated correctly (Figure 10).

For virtual screening purposes not only prediction of the correct binding mode is a challenging task but also the correct affinity prediction of the docked ligands must be achieved. In our analysis, we combined the AUTODOCK-generated poses of 12 selected ligands (including cyclooxygenase inhibitors) with the previously derived 3D QSAR models for CA II. This strategy allows successful prediction of the binding affinity based on the docked poses, even for those cases where the scoring function, implemented in AUTODOCK, failed.

Since structural water molecules, indicated by crystal structure analysis as relevant, were neglected in our approach, future work will incorporate these waters which are most likely conserved for the binding mode and affinity prediction, especially in the handling of ligands with flexible side chains. In conclusion, the described docking protocol in combination with the 3D QSAR models established in parallel could be

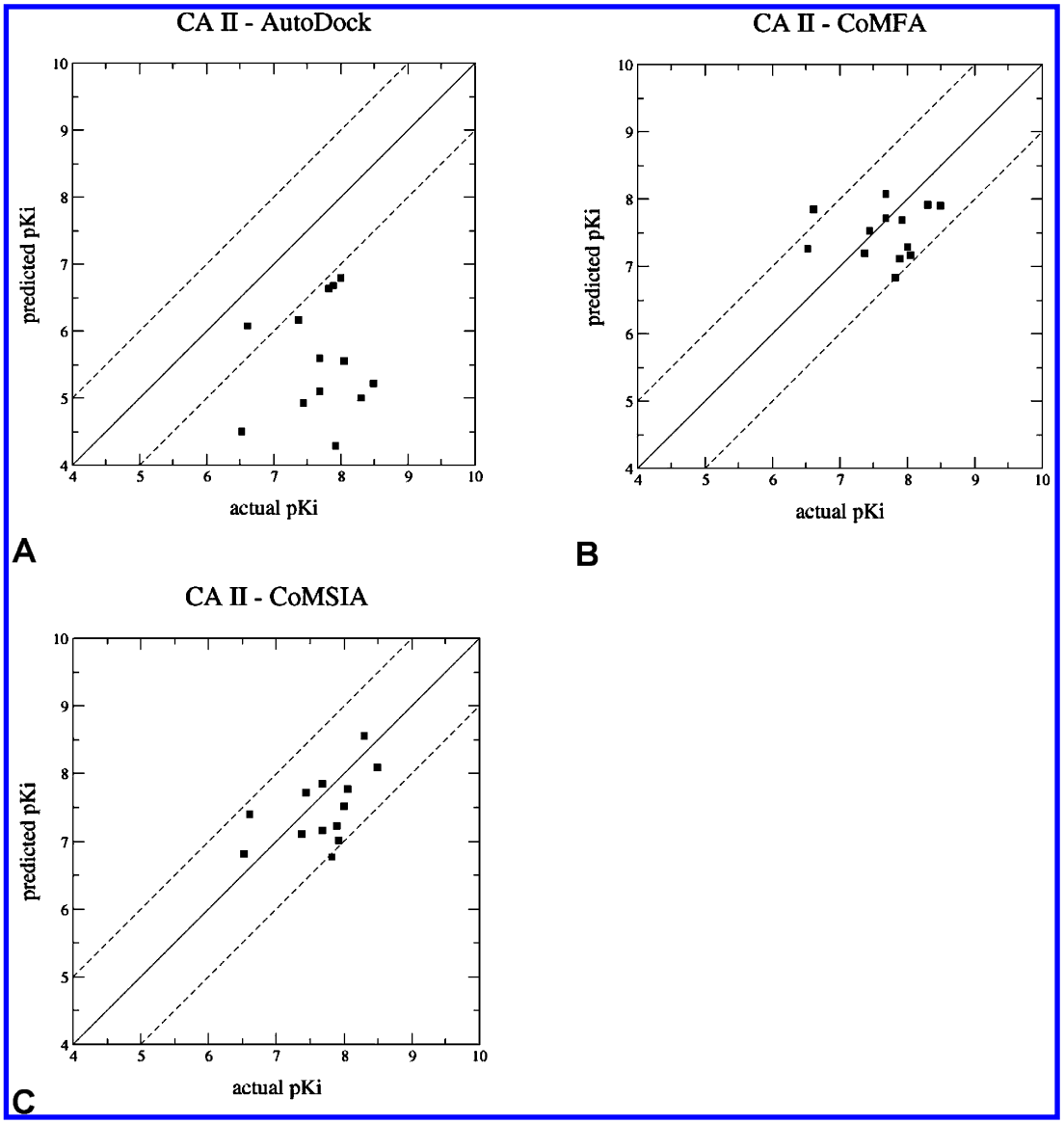


Figure 12. Affinity prediction of CA inhibitors after docking into the CA II binding pocket with AUTODOCK. Predicted inhibitory activities obtained with AUTODOCK (**A**), CoMFA (**B**), and CoMSIA (**C**) are plotted against actual values.

used to enhance the predictive power of further virtual screening campaigns or to shape up selectivity profiles, including other CA isozymes that are involved in pathological situations such as tumors, obesity, or glaucoma.

EXPERIMENTAL SECTION

Protein alignment and comparative molecular field evaluations were performed using SYBYL[60] version 6.8 running on a Silicon Graphics workstation. The manual minimization of the inhibitor data set in CA isozymes was performed using MOLOC[21,22] version 2003. CA inhibitors were automatically docked into CA II using AUTODOCK 3.0.[26,27]

Data Set and Alignment. A data set of 87 inhibitors (Table 1) with six different scaffolds (Figure 2) was used for CoMFA and CoMSIA analyses.[29−39] The data set is composed of 18 thiadiazole- (**A**), 5 thienothiopyrane- (**B**), 11 benzothiazole- (**C**), and 39 benzenesulfonamides (**D**), 6 hydroxamates (**E**), and 8 hydroxysulfonamides (**F**). NMR spectroscopy shows that sulfonamides predominantly coordinate with the monoprotonated and negatively charged nitrogen of the sulfonamide group to the zinc ion.[19] Therefore, the formal charge of the sulfonamide nitrogen atom was set to −1. The experimentally determined affinities of the CA inhibitors are given as pK_i (−logK_i; Table 1).

For the protein alignment 37 protein−ligand complexes of CA I, CA II, and CA IV were selected from the PDB: (**CA I:** *1azm*, *1bzm*, *1czm*; **CA II:** *1a42*, *1am6*, *1bcd*, *1bn1*, *1bn3*, *1bn4*, *1bnm*, *1bnn*, *1bnq*, *1bnt*, *1bnu*, *1bnv*, *1bnw*, *1cil*, *1cim*, *1cin*, *1cnw*, *1cnx*, *1cny*, *1g1d*, *1g45*, *1g46*, *1g48*, *1g4j*, *1g4o*, *1g52*, *1g53*, *1g54*, *1okl*, *1okm*, *1okn*, *2h4n*; **CA IV:** *1znc*, *3znc*). To focus the alignment to the binding pocket, 20 conserved amino acids in CA I, II, and IV in close proximity to the active site were chosen for alignment *(Ser29, Pro30, Asn61, Gly63, Gln92, His94, His96, Trp97, Glu106, Glu117, His119, His122, Ala142, Val143, Ser197, Thr199, Pro201, Glu205, Trp209, Asn244*). The crystal structure of CA II in complex with dorzolamide (*1cil*, 1.58 Å resolution) was used as a reference protein. The coordinates were centered, and all remaining 36 protein−ligand complexes were aligned with SYBYL based on an all non-hydrogen

atom fits of the previously mentioned 20 conserved amino acids. The conformations of the ligand scaffolds as seen in complex with the three isozymes were used as templates for the 87 inhibitors in the 3D QSAR data set, assuming that the conformations from crystal structures should resemble the bioactive conformation. The 87 inhibitors in the data set were set up by adding or exchanging side chains of the crystal structure template molecules. The inhibitors were initially minimized with the Tripos force field[61] with 100 steps of "*steepest decent*" minimization. The protonation of functional groups (aliphatic amine protonated using atom type *N.4*, carbonic acid deprotonated using atom type *O.co2*) was selected as expected under physiological conditions (pH 7.4).

The initial alignment was optimized by minimizing the 87 inhibitors in the corresponding protein environment with MOLOC.[21,22] For each CA isozyme a reference protein was used (CA I: *1azm*, CA II: *1cil*, CA IV: *1znc*), water molecules and bound ligand were deleted, and all 87 inhibitors of the data set were minimized in the binding pockets of the three isozymes. The protein was kept rigid and the ligand atoms flexible. The inhibitors were minimized in the binding pocket with the MAB force field within MOLOC taking into account that essential hydrogen bond interactions with the zinc ion and Thr199 are preserved during minimization (<3.0 Å). The side chains of the 87 inhibitors were oriented as observed in the template molecules from the known X-ray structures and afterward minimized in each binding pocket. The initial minimization within MOLOC was performed with CA II, since most template molecules are available in complex with this isozyme (CA I: 3 crystal structures, CA II: 32, CA IV: 2; Figure 5A,C,E). The minimized conformations in CA II were used as a starting point for the minimization in CA I and CA IV, taking into account that in both isozymes possible different interactions with different amino acids guide the alignment in the corresponding binding pocket. As a result, three distinct alignments were generated, one alignment for CAI, CA II, and CA IV, respectively (Figure 5B,D,F).

CoMFA and CoMSIA Analysis. Steric and electrostatic CoMFA fields were calculated as implemented in SYBYL using Lennard-Jones- and Coulomb-potentials.[50] Partial atomic charges were determined using the AM1 Hamiltonian[62] with the semiempirical package MOPAC.[63] Standard parameters were used within SYBYL for the calculation of steric and electrostatic fields (sp^3-hybridized carbon atom with a charge of +1, energetic "cutoff" = 30 kcal/mol). For CoMSIA steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor properties were calculated using a probe atom with 1 Å radius, charge +1, and hydrophobicity and hydrogen bond property values of +1. The attenuation factor α was set to 0.3. To allow for a comparison of all models equal box coordinates were used for all calculations (see the Supporting Information). To test the influence of grid spacing, lattices with 1 Å and 2 Å distances were used. The box sizes were chosen such that all molecules of the three different alignments were embedded in the grid with a margin of at least 4 Å. For each alignment in CA I, CA II, and CA IV, three models were generated (CA I, CA II, CA IV) with CoMFA and CoMSIA. In total, 36 models were obtained using two different grid spacings (Tables 2 and 3). The statistical significance of the models was checked by performing the "leave-one-out"-(LOO) analysis using the SAMPLS[64] method implemented in SYBYL. The optimal number of components was chosen such that each additional component had to increase the q^2 values (cross-validated r^2 value) by at least 5%. In almost every case, the number of components resembles the minimal S_{PRESS} value. The optimal number of components was used to derive the final QSAR model. To enhance the calculation time, in case of CoMFA analyses the SYBYL option "*column filtering*" was set to 2.0 kcal/mol ensuring that at least 10% of the variables were included in the calculation. Variations of this value lead to quite similar statistical results so that the less time-consuming 2.0 kcal/mol filtering option was used. The statistical parameters (q^2, S_{PRESS}, r^2, S, F) are summarized in Tables 2 and 3 for the individual 3D QSAR models. Statistical results for the training data sets are summarized in Table 4. 3D QSAR selectivity models were calculated using the affinity differences between the isozymes as independent variables; the statistical results of these models are summarized in Table 7.

Docking. AUTODOCK[26,27,65] was used for automatic placement of CA inhibitors in the binding pocket of CA II (PDB code: *1cil*). The correct prediction of the binding mode was validated based on six inhibitors (Figure 10). For five inhibitors, the crystal structure in complex with CA II is known (*1a42*, *1cil*, *1eou*, *1cim*; the structure of CA II in complex with topiramate was kindly provided by Casini et al.). The complex of CA I with acetazolamide (AZM, *1azm*) was used to predict the binding mode of this inhibitor in CA II (assuming that the binding mode of AZM in CA I is transferable to CA II). All six protein−ligand complexes were superimposed as described above (Data Set and Alignment section). Comparison of the amino acids in the binding pocket of the superimposed structures indicates conformational flexibility only for His64. This fact is further confirmed by superimposing the complex structures of CA II with RELIBASE,[44,45] which reveals that CA II has a fairly rigid binding site with the exception of His64.[66] To allow as much space as possible for ligand placement, the complex of CA II with dorzolamide (*1cil*) was used for ligand placements. The structure was set up for docking by removing all water and ligand atoms, adding polar hydrogens, and assigning AMBER[67] atomic charges and solvation parameters as required by the AUTODOCK program.[27] For preparing the ligand structures, Gasteiger−Marsili charges were assigned to the ligand atoms, and rotatable bonds were explicitly defined. The largest ring system of the inhibitor was chosen as root fragment, and all rotatable bonds were kept flexible. Docking was then carried out using an empirical free energy function and the Lamarckian genetic algorithm.[27] One hundred independent docking runs were performed for each ligand, applying a standard AUTODOCK protocol, with a grid spacing of 1 Å, an initial population of 50 randomly placed individuals, a maximum number of 1.5×10^6 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1. Results differing by less than 1 Å rmsd were merged together into one cluster.

ACKNOWLEDGMENT

We would like to thank Dr. Jochen Antel (Solvay Pharmaceutical, Hannover, Germany) for his current interest and support of this project.

Supporting Information Available: Parameters of the grid boxes used for CoMFA and CoMSIA analyses (Table S01) and distances between atoms that are involved in a hydrogen bond network in CA I and CA II (Table S02). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Maren, T. H. Relations between structure and biological activity of sulfonamides. *Annu. Rev. Pharmacol. Toxicol.* **1976**, *16*, 309−327.

(2) Supuran, C. T. Carbonic anhydrase inhibitors and their therapeutical potential. *Exp. Opin. Ther. Patents* **2000**, *10*, 575−600.

(3) Supuran, C. T.; Scozzafava, A. Carbonic Anhydrase Inhibitors. *Curr. Med. Chem.: Immun., Endocr. Metab. Agents* **2001**, *1*, 61−97.

(4) Supuran, C. T.; Scozzafava, A. Applications of carbonic anhydrase inhibitors and activators in therapy. *Expert Opin. Ther. Pat.* **2002**, *12*, 217−242.

(5) Supuran, C. T.; Scozzafava, A.; Casini, A. Carbonic anhydrase inhibitors. *Med. Res. Rev.* **2003**, *23*, 146−189.

(6) Chegwidden, W. R.; Carter, N. D. Introduction to the carbonic anhydrases. *The Carbonic Anhydrases − New Horizons*; Birkhäuser Verlag: Basel, 2000; pp 13−28.

(7) Maren, T. H. Carbonic anhydrase: chemistry, physiology, and inhibition. *Physiol. Rev.* **1967**, *47*, 595−781.

(8) Evans, W. O.; Robinson, S. M.; Houstman, D. H.; Jackson, R. E.; Weinkopf, R. B. Amelioration of the symptoms of acute mountain sickness by staging and acetazolamide. *Aviat., Space Environ. Med.* **1976**, *47*, 512−516.

(9) Forster, P. Methazolamide in acute mountain sickness. *Lancet* **1982**, *1*, 1254.

(10) Lindskog, S.; Wistrand, P. J. Inhibitors of carbonic anhydrase. *Design of Enzyme Inhibitors as Drugs*; Oxford University Press: 1989; pp 698−723.

(11) Owa, T.; Yoshino, H.; Okauchi, T.; Yoshimatsu, K.; Ozawa, Y. et al. Discovery of novel antitumor sulfonamides targeting G1 phase of the cell cycle. *J. Med. Chem.* **1999**, *42*, 3789−3799.

(12) Owa, T.; Nagasu, T. Novel sulfonamide derivatives for the treatment of cancer. *Expert Opin. Ther. Pat.* **2000**, *10*, 1725−1740.

(13) Abbate, F.; Supuran, C. T.; Scozzafava, A.; Orioli, P.; Stubbs, M. T. et al. 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.

(14) Casini, A.; Scozzafava, A.; Mastrolorenzo, A.; Supuran, L. T. Sulfonamides and sulfonylated derivatives as anticancer agents. *Curr. Cancer Drug Targets* **2002**, *2*, 55−75.

(15) Scozzafava, A.; Owa, T.; Mastrolorenzo, A.; Supuran, C. T. Anticancer and antiviral sulfonamides. *Curr. Med. Chem.* **2003**, *10*, 925−953.

(16) Supuran, C. T. Indisulam: an anticancer sulfonamide in clinical development. *Expert Opin. Invest. Drugs* **2003**, *12*, 283−287.

(17) Abbate, F.; Casini, A.; Owa, T.; Scozzafava, A.; Supuran, C. T. 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.* **2004**, *14*, 217−223.

(18) Maren, T. H. The Development of Topical Carbonic Anhydrase Inhibitors. *J. Glaucoma* **1995**, *4*, 49−62.

(19) Lindskog, S. Structure and mechanism of carbonic anhydrase. *Pharmacol. Ther.* **1997**, *74*, 1−20.

(20) Clare, B. W.; Supuran, C. T. QSAR studies of sulfonamide carbonic anhydrase inhibitors. *Carbonic Anhydrase − Its inhibitors and activators*; Taylor & Francis: London & New York, 2004; pp 149−182.

(21) 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.

(22) Gerber, P. R. Charge distribution from a simple molecular orbital type calculation and nonbonding interaction terms in the force field MAB. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 37−51.

(23) Böhm, M.; Stürzebecher, J.; Klebe, G. Three-dimensional quantitative structure−activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* **1999**, *42*, 458−477.

(24) Matter, H.; Schwab, W.; Barbier, D.; Billen, G.; Haase, B. et al. Quantitative structure−activity relationship of human neutrophil collagenase (MMP-8) inhibitors using comparative molecular field analysis and X-ray structure analysis. *J. Med. Chem.* **1999**, *42*, 1908−1920.

(25) Matter, H.; Schwab, W. Affinity and selectivity of matrix metalloproteinase inhibitors: a chemometrical study from the perspective of ligands and proteins. *J. Med. Chem.* **1999**, *42*, 4506−4523.

(26) Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 293−304.

(27) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E. et al. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(28) Antel, J.; Weber, A.; Sotriffer, C. A.; Klebe, G. Multiple binding modes observed in X-ray structures of carbonic anhydrase inhibitor complexes and other systems: consequenses for structure-based drug design. *Carbonic Anhydrase − Its inhibitors and activators*; Taylor & Francis: London & New York, 2004; pp 45−65.

(29) Borras, J.; Scozzafava, A.; Menabuoni, L.; Mincione, F.; Briganti, F. et al. Carbonic anhydrase inhibitors: synthesis of water-soluble, topically effective intraocular pressure lowering aromatic/heterocyclic sulfonamides containing 8-quinoline-sulfonyl moieties: is the tail more important than the ring? *Bioorg. Med. Chem.* **1999**, *7*, 2397−2406.

(30) Casini, A.; Scozzafava, A.; Mincione, F.; Menabuoni, L.; Ilies, M. A. et al. Carbonic anhydrase inhibitors: water-soluble 4-sulfamoylphenylthioureas as topical intraocular pressure-lowering agents with long-lasting effects. *J. Med. Chem.* **2000**, *43*, 4884−4892.

(31) Ilies, M.; Supuran, C. T.; Scozzafava, A.; Casini, A.; Mincione, F. et al. Carbonic anhydrase inhibitors: sulfonamides incorporating furan-, thiophene- and pyrrole-carboxamido groups possess strong topical intraocular pressure lowering properties as aqueous suspensions. *Bioorg. Med. Chem.* **2000**, *8*, 2145−2155.

(32) Scozzafava, A.; Briganti, F.; Mincione, G.; Menabuoni, L.; Mincione, F. et al. Carbonic anhydrase inhibitors: synthesis of water-soluble, aminoacyl/dipeptidyl sulfonamides possessing long-lasting intraocular pressure-lowering properties via the topical route. *J. Med. Chem.* **1999**, *42*, 3690−3700.

(33) Scozzafava, A.; Menabuoni, L.; Mincione, F.; Briganti, F.; Mincione, G. et al. Carbonic anhydrase inhibitors. Synthesis of water-soluble, topically effective, intraocular pressure-lowering aromatic/heterocyclic sulfonamides containing cationic or anionic moieties: is the tail more important than the ring? *J. Med. Chem.* **1999**, *42*, 2641−2650.

(34) Scozzafava, A.; Supuran, C. T. Carbonic anhydrase and matrix metalloproteinase inhibitors: sulfonylated amino acid hydroxamates with MMP inhibitory properties act as efficient inhibitors of CA isozymes I, II, and IV, and N− hydroxysulfonamides inhibit both these zinc enzymes. *J. Med. Chem.* **2000**, *43*, 3677−3687.

(35) Scozzafava, A. Carbonic Anhydrase Inhibitors: Synthsis of N-Morpholyl-thiocarbonylsulfenylamino Aromatic/Heterocyclic Sulfonamides and their Interaction with Isozymes I, II and IV. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1117−1120.

(36) Scozzafava, A.; Menabuoni, L.; Mincione, F.; Briganti, F.; Mincione, G. et al. Carbonic anhydrase inhibitors: perfluoroalkyl/aryl-substituted derivatives of aromatic/heterocyclic sulfonamides as topical intraocular pressure-lowering agents with prolonged duration of action. *J. Med. Chem.* **2000**, *43*, 4542−4551.

(37) Supuran, C. T. Carbonic anhydrase inhibitors − part 70. Synthesis and ocular pharmacology of a new class of water-soluble, topically effective intraocular pressure lowering agents derived from nicotinic acid and aromatic/heterocyclic sulfonamides. *Eur. J. Med. Chem.* **1999**, *34*, 799−808.

(38) Supuran, C. T.; Scozzafava, A. Carbonic anhydrase inhibitors- -Part 94. 1,3,4-thiadiazole-2- sulfonamidederivatives as antitumor agents? *Eur. J. Med. Chem.* **2000**, *35*, 867−874.

(39) Supuran, C. T.; Briganti, F.; Tilli, S.; Chegwidden, W. R.; Scozzafava, A. Carbonic anhydrase inhibitors: sulfonamides as antitumor agents? *Bioorg. Med. Chem.* **2001**, *9*, 703−714.

(40) Thibaut, U.; Folkers, G.; Klebe, G.; Kubinyi, H.; Merz, A. et al. Recommendations for CoMFA Studies and 3D QSAR Publications. *3D QSAR in Drug Design: Theory Methods and Applications*; ESCOM: Leiden, 1993; pp 711−716.

(41) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E. et al. The Protein Data Bank. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58*, 899−907.

(42) Berman, H. M.; Bourne, P. E.; Westbrook, J. The Protein Data Bank: A Case Study in Management of Community Data. *Curr. Proteomics* **2004**, *1*, 49−57.

(43) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N. et al. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(44) Hendlich, M.; Bergner, A.; Günther, J.; Klebe, G. Relibase: Design and Development of a Database for Comprehensive Analysis of Protein−Ligand Interactions dagger. *J. Mol. Biol.* **2003**, *326*, 607−620.

(45) Günther, J.; Bergner, A.; Hendlich, M.; Klebe, G. Utilising structural knowledge in drug design strategies: applications using Relibase. *J. Mol. Biol.* **2003**, *326*, 621−636.

(46) Duda, D.; Tu, C.; Qian, M.; Laipis, P.; Agbandje-McKenna, M. et al. Structural and kinetic analysis of the chemical rescue of the proton-transfer function of carbonic anhydrase II. *Biochemistry* **2001**, *40*, 1741−1748.

(47) Duda, D.; Govindasamy, L.; Agbandje-McKenna, M.; Tu, C.; Silverman, D. N. et al. The refined atomic structure of carbonic anhydrase II at 1.05 Å resolution: implications of chemical rescue of proton transfer. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 93−104.

(48) Christianson, D. W.; Cox, J. D. Catalysis by metal-activated hydroxide in zinc and manganese metalloenzymes. *Annu. Rev. Biochem.* **1999**, *68*, 33−57.

(49) Scolnick, L. R. Novel Binding Mode of Hydroxamate Inhibitors to Human Carbonic Anhydrase II. *J. Am. Chem. Soc.* **1997**, *119*, 850−851.

(50) Cramer, R. D., III.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(51) Klebe, G. Comparative Molecular Similarity Indices Analysis: CoMSIA. *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 87−104.

(52) 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.

(53) Sotriffer, C. A.; Stahl, M.; Klebe, G. The Docking Problem. *Handbook of Chemoinformatics*; Wiley-VCH: Weinheim, 2003; pp 1732−1766.

(54) Masereel, B.; Rolin, S.; Abbate, F.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: anticonvulsant sulfonamides incorporating valproyl and other lipophilic moieties. *J. Med. Chem.* **2002**, *45*, 312−320.

(55) Casini, A.; Antel, J.; Abbate, F.; Scozzafava, A.; David, S. et al. Carbonic anhydrase inhibitors: SAR and X-ray crystallographic study for the interaction of sugar sulfamates/sulfamides with isozymes I, II and IV. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 841−845.

(56) 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.

(57) Recacha, R.; Costanzo, M. J.; Maryanoff, B. E.; Chattopadhyay, D. Crystal structure of human carbonic anhydrase II complexed with an anti- convulsant sugar sulphamate. *Biochem. J.* **2002**, *361*, 437−441.

(58) Sotriffer, C. A.; Gohlke, H.; Klebe, G. Docking into knowledge-based potential fields: a comparative evaluation of DrugScore. *J. Med. Chem.* **2002**, *45*, 1967−1970.

(59) Winum, J. Y.; Vullo, D.; Casini, A.; Montero, J. L.; Scozzafava, A. et al. Carbonic anhydrase inhibitors. Inhibition of cytosolic isozymes I and II and transmembrane, tumor-associated isozyme IX with sulfamates including EMATE also acting as steroid sulfatase inhibitors. *J. Med. Chem.* **2003**, *46*, 2197−2204.

(60) *SYBYL molecular modeling package*; version 6.8 ed.; Tripos Inc.: 1699 South Hanley Road, Suite 303, St. Louis, MO 63144.

(61) Clark, M.; Cramer, R. D.; van Opdenbosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982−1012.

(62) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(63) Stewart, J. J. MOPAC: a semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−105.

(64) Bush, B. L.; Nachbar, R. B., Jr. Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587−619.

(65) Goodsell, D. S.; Olson, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins* **1990**, *8*, 195−202.

(66) Klebe, G. From structure to recognition principles: mining in crystal data as a prerequisite for drug design. *Ernst Schering Res. Found Workshop* **2003**, *42*, 103−126.

(67) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C. et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.

(68) DeLano, W. L. *The PyMOL Molecular Graphics System;* DeLano Scientific: San Carlos, CA, 2002. http://www.pymol.org.

CI600298R