RESEARCH ARTICLE

Chromatographic evaluation and QSAR optimization for benzoic acid analogues against carbonic anhydrase III

Muhammed Alzweiri[1], Qosay Al-Balas[2] and Yusuf Al-Hiari[1]

[1]Department of Pharmaceutical Sciences, Faculty of Pharmacy, The University of Jordan, Amman, Jordan and [2]Department of Medicinal Chemistry and Pharmacognosy, Faculty of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan

Abstract

An HPLC-size exclusion method was developed as an assay method to evaluate the binding of tested compounds with carbonic anhydrase III (CAIII) enzyme. Inhibition of CAIII by a group of benzoic acid analogues was characterized by vacancy (negative) peak intensity representing the fraction of the compounds bound with CAIII enzyme. Interestingly, p-hydroxyl benzoic acid and aspirin were found potent inhibitors against CAIII with affinity constants of 9954 and 9013 M^{-1} respectively. Affinity values of twenty training compounds were modeled against thirty-five descriptors derived from their structures. Strong correlation was obtained between the affinity values and the formal charge of the molecules. Docking studies on training set compounds generated consensus scores having a strong agreement with affinity factors obtained from the chromatographic analysis.

Introduction

Carbonic anhydrases (CA) catalyze the inter-conversion between carbon dioxide and bicarbonate across the body organs[1–3]. They contain zinc atom in their prosthetic groups[1]. Three classes of CA were isolated from animal and plant kingdoms namely; α, β and γ[4]. More than fourteen of α-CA isoenzymes were isolated from higher vertebrates, seven of them were found in human[1,4–6]. Isoenzymes such as I, II, III and VII are cytoplasmic, others like IV, IX, XII and XIV are bound to the cell membrane, whereas V was found in mitochondria and VI in saliva[1]. In addition to their activity in carbon dioxide hydration, they can hydrate cyanate to carbamate and cyanamide to urea[1]. Inhibition of CAs is a target in treatment of many diseases including, hypertension, glaucoma, epilepsy and edema[7–9].

CAIII governs a quarter of the total protein content in human adipocytes and lacks a distinctive known function[10–12]. Although CAIII structure resembles the most active isoform (CAII; 1). It has a low carbon dioxide hydratase activity, equals 0.3% of the CAII activity[1]. Instead of one of histidine groups around the zinc atom of CAII, there is a lysine moiety in CAIII which reduces its aptitude in proton shuttling and consequently the activity in carbon dioxide hydration[1]. Also its cone-shape active site is partially blocked by phenylalanine residue which makes it insensitive to sulfonamides such as acetazolamide[1]. However, small phenols were found moderately potent inhibitors against isoenzymes I, II, IV, IX, XII and XIV and particularly against CAIII[13–16].

Our research group had previously found that vanillic acid has strong inhibition potency against CAIII[17]. The binding assessment was conducted by measuring the intensity of negative peaks generated from vacancy chromatography using size exclusion separation. This method has advantage over currently available colorimetric methods in its stability against false results generated from pH fluctuation due to intrinsic acidity or basicity of analyte compounds[17–19]. Additionally, colorimetric methods are not considered sensitive in evaluating the CAIII inhibition due to the limited catalytic activity of this isoform in carbon dioxide hydration[1].

Affinity values of binding for a set of benzoic acid analogues against CAIII, obtained from chromatographic analysis, were modeled against a set of structure descriptors by multiple regression analysis to generate quantitative structural activity relationship (QSAR). Subsequently, a docking study was conducted to optimize the results of QSAR model.

Experimental

Instrumentation

Thermo scientific HPLC system (Tewksbury, MA) equipped with Spectra system UV1000 detector, Spectra system P1000 pump, and SN 4000 controller. Chromatographic data were acquired and processed by ChromQuest 3.1.6 software. BioSep-SEC Phenomenex column of s2000, 300 × 7.8 mm (MA) was used for size exclusion separation. The mobile phase flow was maintained at 1 ml/min during the run time. Wavelengths used for analyte groups were maximum lambda of each analyte.

Chemicals

Carbonic anhydrase III, from Biovar Ltd (Yerevan, Armenia), was stored and handled according to the producer instructions. 4-Hydroxybenzaldehyde was purchased from Riedel-deHaen (Seelze, Germany) whereas 3-aminophenol was purchased from

Address for correspondence: Muhammed Alzweiri, Department of Pharmaceutical Sciences, Faculty of Pharmacy, The University of Jordan, Amman, Jordan. E-mail: m.alzweiri@ju.edu.jo.

Merck-schuchardt (Darmstadt, Germany). Gallic acid, caffeic acid and ferulic acid were purchased from Molekula (Dorset, UK) and also salicylic acid was purchased from Daejung (Gyeonggi-do, Korea). Other chemicals, used in testing and training sets, were purchased from Sigma-Aldrich (St. Louis, MO) including: phenol, benzoic acid, catechol, 5-hydroxy-1,4-naphthoquinone, 4-aminobenzoic acid, 4-nitrophenol, 4-hydroxybenzoate, N-(4-hydroxyphenyl) acetamide, 4-aminophenol, 4-chlorocresol, 5-hydroxyanthranilic acid, aspirin, 3,4-dihydroxybenzaldehyde, 4-aminosalicylic acid, vanillic acid and 1,3,5-trihydroxybenzene. HPLC grade methanol and acetonitrile, from Fisher scientific (Loughborough, UK), were used without further purification. Sodium dihydrogen phosphate monohydrate from Gainland (Flintshire, UK), Hydrochloric acid from Carlo Erba (Milano, Italy), sodium hydroxide from Lonver (England), and deionized water were used in mobile phase preparation. 4-Hydroxy-3-methoxy-5-sulfamoylbenzoic acid was synthesized and fully characterized by the research group using vanillic acid as a starting material.

Mobile phase and sample preparation

Each compound from training and testing sets was dissolved in a mixture of acetonitrile: phosphate buffer, pH 6.0 (1:9 v/v) to obtain concentrations of 0.24, 0.20 and 0.18 mM. Subsequently, mobile phases were filtered and degassed by a sonic bath prior the use. An excess concentration of CAIII in water (6.3 mM) was injected in each run.

QSAR model

Training set which embraces twenty molecules was treated by using specialized protocol in Discovery Studio 3.5 software (CA) to extract the values of molecular properties. These values were studied in details and the most promising descriptors were used effectively in building up the QSAR model, thirty five descriptors were carefully selected and measured according to their potential against CAIII. SIMCA-P software (version 12.0, CA) was used to refine the descriptors according to their correlation with the K_f values as dependent variables by applying Partial Least Square Regression analysis (PLSR) and also it was employed to accomplish clustering analysis according to the underlined properties of the molecules. The strongest descriptors were then treated by multiple linear regression and ANOVA through Minitab software (version 16, Coventry, UK). The results of the training set were validated by permutations plot and jack-knifing internal validation using SIMCA-P. Furthermore, external set of compounds were used to launch external validation for the QSAR model.

Docking study

The crystal structure of carbonic anhydrase III (PDB code: 3UYQ)[20] which has a very good resolution of 1.7 Å has been downloaded from the Protein Data Bank and was prepared for the following steps using Prepare Protein protocol implemented in Discovery Studio 3.5 software (DS 3.5) from Accelrys® (CA). The preparation steps include standardization of atom names, insertion of missing atoms in residues and removal of alternate conformations, insertion of missing loop regions, calculation of pKa values and protonation of the structure functionalities. Once the protein is prepared it becomes ready for docking steps.

Within the docking procedure of metalloenzymes, it is essential to determine the boundaries of active site around the metallic ion. A virtual sphere of 7 Å was suggested which is enough to accommodate the natural substrate and the suggested inhibitors of both the training and the test sets.

Consensus scoring strategy[21,22] has been implemented in this study by employing different docking programs namely; CDOCKER, LibDock, LigandFit and GOLD. Several docking programs were used because each one operates different pose fitting methods of the ligand inside the active site and also each one uses its own energy calculation methodology. Thus, using all these techniques should minimize any bias that could arise from a single docking program.

CDOCKER uses a CHARMm-based molecular dynamics (MD) scheme to dock ligands into a receptor binding site; then random conformation will be produced using high temperature molecular dynamics. Once these conformations are translated to the active site, candidate poses are then created using random rigid-body rotations followed by simulated annealing[23]. Libdock high-throughput algorithm where the ligand will be aligned in the active site based on polar and apolar interaction sites termed as ''Hotspots''. For more accurate docking, a CHARMm minimization step was activated[24]. Additionally, LigandFit docks ligands into an active site using a shape filter and Monte-Carlo ligand conformation. The docked poses will be minimized with CHARMm and evaluated with a set of scoring functions such as PLP1, PLP2, Jain, PMF and PMF04[25,26]. GOLD parameters were set to give maximum flexibility to the active site and the ligands in addition to using four scoring functions namely, Goldscore, Chemscore, ASP and CHEMPLP to reach optimal accuracy for candidate compound selection. GOLD recognizes zinc atom by allowing the coordination geometry number to be 4,5,6 which correlate to tetrahedral, trigonal bipyramidal and octahedral geometries respectively[27]. The consensus scoring step is finally performed by summing up all the values of the twelve scoring functions used in this study (four for GOLD, six for LigandFit, one for Libdock and one for CDOCKER).

Results and discussion

HPLC method of analysis

Recently, we have reported vanillic acid as a novel potent inhibitor for CAIII[17]. It was found more potent than the reported phenolic inhibitors[14–16]. Owing to the susceptibility of the reported colorimetric assays of CA to the analytes' acidity and alkalinity and also the low activity of CAIII in carbon dioxide hydration, a Hummel-dryer chromatographic method was developed and used by our research group previously in evaluating vanillic acid affinity to CAIII[17]. The method was further optimized in this work to evaluate the affinity of benzoic acid analogues against CAIII since the chromatography provides a relatively enough space for the protein-analyte interaction and consequently low possibility of protein self-aggregation. Additionally the overwhelming effect of mobile phase pH and its ionic strength dominate any possibility of misinterpretation due to the intrinsic properties of analytes. Moreover, chromatographic analysis relied on vacancy peak formation occurred due to the subtraction of the CAIII-bound portion of analyte from the mobile phase. This will be independent on the enzyme catalytic activity and consequently provides a sensitive mean for inhibition activity in comparison with the currently available colorimetric methods. HPLC analysis of CAIII binding with candidate inhibitors generates chromatograms such as those appeared in Figure 1. When weak inhibitors were injected in the chromatographic system, they created weak vacancy (negative) peaks independent of the CAIII concentration (Figure 1A). In contrary, binding of potent inhibitor with CAIII leads to intense negative peak at the expected retention time of the analyte standard (Figure 1B). The intensity of the peak is proportional with relatively low concentrations of CAIII solutions. Due to the
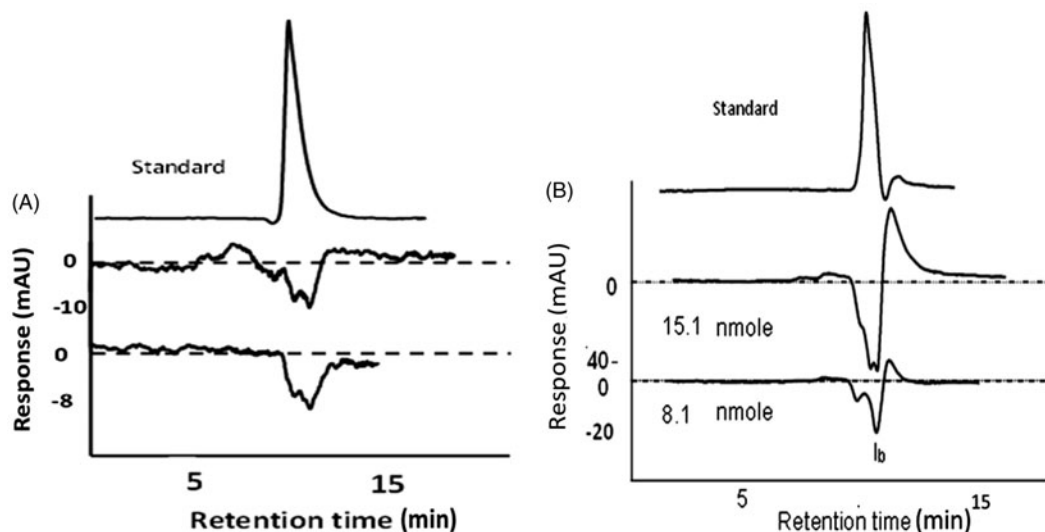
Figure 1. Examples of chromatograms generated from the vacancy chromatography for phenol as a weak inhibitor (A) and ferulic acid as a strong one (B). Strong inhibitors were found as CAIII concentration dependent.

mechanism of size exclusion of the column free and complex forms of CAIII co-elute prior to the ligand peak. CAIII forms retained low absorptivity at lambda maxima of the chromophoric ligands used in training and testing sets here.

pH of the mobile phase was chosen neutral to mimic the physiological value with better chance of *in vivo* correlation. Additionally, the lowest percentage of organic modifier (10% acetonitrile) was used to elute the analytes from the hydrophobic column. It was also used to avoid adhesiveness of the macromolecules with the column. Temperature was kept constant at ambient temperature because the temperature fluctuation affects the degree of binding.

Affinity bindings of benzoic acid analogues with CAIII

Injection of a group of CAIII solutions with different concentrations in HPLC containing constant concentration of a certain analyte generated vacancy peaks. In case of potent inhibitors, intensities of vacancy peaks were directly proportion to low amounts of CAIII and then steady state was occurred at higher CAIII amounts. The binding ratios with CAIII were found 1:1 for all tested analytes. The binding ratios were deduced from scatchard plot, Figure 2 is an example. Where $\overline{r}$ stands for the average molar ratio of the inhibitor bound to CAIII and [E] represents the concentration of injected enzyme. The inflection points were found at around $\overline{r}$ is one for the tested analytes.

Thus, affinity constant (k_f) of the inhibitors were calculated at equimolar ratios. A modified form of an equation introduced by Flood et al was adopted to calculate affinity constants of tested compounds[28].

$$k_f \frac{I_B}{[I](E-I_B)} \qquad (1)$$

Where I_B represents the bound amount of inhibitor, E is the amount of enzyme injected whereas [I] is the concentration of inhibitor in mobile phase.

QSAR model

Group of structural descriptors containing thirty-five properties were selected according to their eligibility in binding with metalloenzyme active site. They were *in silico* calculated for twenty compounds (training set) as shown in Table 1. The values, displayed in Table 1, were correlated with affinity constants of
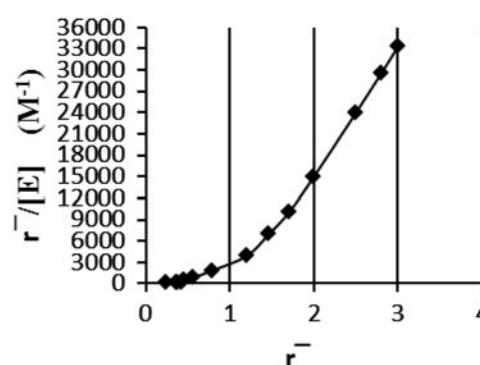

Figure 2. Scatchard plot indicates the inflection value at equimolar ratio of inhibitor to CAIII. Where $\overline{r}$ stands for the average values of bind ratio and [E] stands for concentration of injected CAIII into the HPLC system.

tested molecules extracted from the chromatographic analysis (actual K_f values, Table 2). The correlation process was carried out by implementing partial least square regression (PLSR) through SIMCA-P software to avoid any possible bias might occur by missed values in data matrices. PLSR is suitable to deal with noisy character of chemical data represented by sinusoidal behavior of results (very-high values appeared simultaneously with very-low values). The PLSR correlation was able to express the majority of variation among the descriptors. This was proved by showing accumulative values of fit goodness (R^2) above 0.75 for the first two principal components (eigen vectors) and also the cross validation coefficients (Q^2) were above the third quartile[29,30]. Correlation coefficients were shown in a histogram of Figure 3. The data, which have absolute mean values of correlation above 0.05 as a threshold, were gathered and introduced into Minitab software to apply multiple linear regression (MLR) and analysis of variance (ANOVA) for the correlation of each independent variable with affinity constants' set. It was found that ''formal charge'' and ''molecular polar surface area'' overwhelm the other correlations with affinity constants as appeared in Figure 4 and according to the following QSAR equation
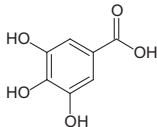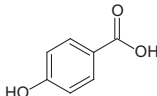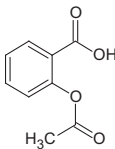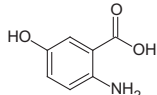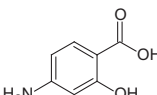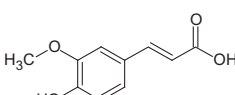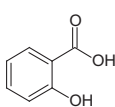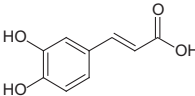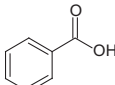
$$K_f = 46.3\,M - 6189\,F - 1699 \qquad (2)$$

where ''M'' stands for ''molecular polar surface area'' and ''F'' stands for ''formal charge''.

Table 1. Structural descriptors of training set compounds calculated *in silico* by Discovery Studio 3.5 software. (Abbreviations of molecules' names are listed in Table 2.)

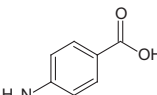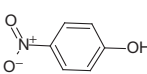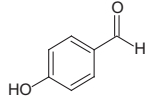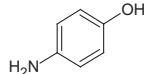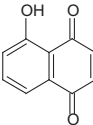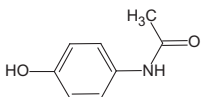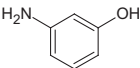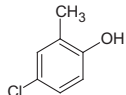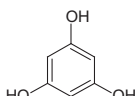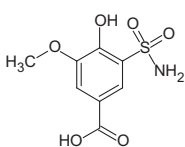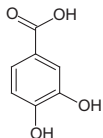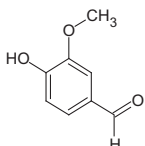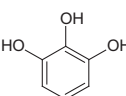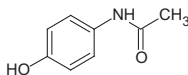| Name	P	B	CA	4HB	HN	AB	NP	HB	HPA	4AP	3AP	CC	HA	S	A	AS	G	C	F	THB
Initial RMS Gradient	0	0	0	0	0	49	53	54	42	45	45	42	49	49	42	48	55	48	56	52
CHARMm Energy	-2	-3	-12	0	-5	-12	9	-15	-17	-13	0	-4	-4	11	-9	12	-39	-25	-15	13
Initial Potential Energy	-2	-3	-12	0	-5	18	36	15	13	10	23	14	31	39	34	45	16	23	60	43
Improper Energy	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Angle Energy	0	0	1	0	0	0	2	0	1	0	0	1	1	1	2	1	2	2	3	1
Dihedral Energy	0	0	0	1	0	0	1	1	2	1	1	1	3	1	2	1	0	1	1	0
Electrostatic Energy	-4	-7	-16	-3	-9	-15	2	-20	-21	-15	-2	-8	-9	6	-13	7	-42	-30	-21	9
Potential Energy	-2	-3	-12	0	-5	-12	9	-15	-17	-13	0	-4	-4	11	-9	12	-39	-25	-15	13
Van der Waals Energy	2	3	1	1	3	3	3	3	1	1	1	1	1	2	0	2	1	2	2	2
RMS Gradient	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bond Energy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALogP	2	0	1	1	2	-1	1	0	1	1	1	3	-1	0	0	-1	-1	0	0	1
ALogP98	2	0	1	1	2	-1	1	0	1	1	1	3	-1	0	0	-1	-1	0	0	1
FormalCharge	0	-1	0	0	0	-1	0	-1	0	0	0	0	-1	-1	-1	0	0	-1	-1	0
Molecular_Weight	94	121	110	122	174	136	139	137	151	109	109	143	152	137	179	152	169	179	193	126
Molecular_Mass	94	121	110	122	174	136	139	137	151	109	109	142	152	137	179	152	169	179	193	126
Molecular_Solubility	-1	-1	-1	-1	-2	-1	-2	-1	-1	-1	-1	-2	-1	-1	-2	-1	0	-2	-2	0
HBA_Count	0	0	0	1	2	0	0	0	1	0	0	0	0	0	2	0	0	0	1	0
HBD_Count	1	0	2	1	1	1	1	1	2	2	2	1	2	1	0	2	3	2	1	3
Num_Bonds	7	9	8	9	14	10	10	10	11	8	8	9	11	10	13	11	12	13	14	9
Num_RotatableBonds	0	1	0	1	0	0	1	1	1	0	0	0	1	1	3	1	1	2	3	0
Num_Rings	1	1	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Num_H_Acceptors	1	2	2	2	3	3	3	3	2	2	2	1	4	3	4	4	5	4	4	3
Num_H_Donors	1	0	2	1	1	1	1	3	2	2	2	1	2	1	0	2	3	2	1	3
Num_H_Acceptors_Lipinski	1	2	2	2	3	3	4	3	3	2	2	1	4	3	4	4	5	4	4	3
Num_H_Donors_Lipinski	1	0	2	1	1	2	1	1	2	3	3	1	3	1	0	3	3	2	1	3
Molecular_Volume	59	71	69	75	106	80	79	80	99	70	69	95	86	80	110	92	97	108	123	78
Molecular_PolarSurfaceArea	20	40	40	37	54	66	66	60	49	46	46	20	86	60	66	86	101	81	70	61
Molecular_FractionalPolarSurfaceArea	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0
Molecular_PolarSASA	52	79	88	79	106	132	125	114	94	105	105	52	167	114	112	167	185	150	121	123
Molecular_FractionalPolarSASA	1	7	2	1	1	11	4	11	4	1	1	0	9	10	9	13	15	20	17	2
Dipole_mag	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0
Molecular_3D_SAVol	214	240	229	248	289	256	250	253	286	233	235	272	264	248	305	267	274	314	337	245
Energy	2	1	4	2	3	7	4	3	9	8	8	2	15	3	41	10	7	6	5	6
PMI_mag	59	111	77	132	202	161	164	164	228	91	85	159	168	128	210	178	221	400	447	114
Molecular_3D_SASA	242	267	258	277	321	286	279	281	324	265	266	297	295	277	343	299	304	348	377	276
Molecular_3D_PolarSASA	53	79	93	95	111	139	136	128	99	115	114	44	174	115	105	181	215	179	147	157

Table 2. Affinity and inhibition constants of tested compounds against CAIII obtained from chromatographic analysis (actual K_f and K_i values). And those which were resulted from the refined form of QSAR model (calculated K_f and K_i values). The values of the most important two descriptors of QSAR model were presented (F and M).

Name	Abbreviation	Structure	Actual $K_f \pm 100$ (M^{-1})	F	M	Calculated K_f (M^{-1})	Actual K_i (μM)	Calculated K_i (μM)
Gallic acid	G		12547	−1	100.820	9158	80	109
4-Hydroxybenzoic acid	HB		9954	−1	60.360	7285	100	137
Aspirin	A		9013	−1	66.430	7566	111	132
5-Hydroxyanthranilic acid	HA		8987	−1	86.380	8489	111	118
4-Aminosalicylic acid	AS		7419	−1	86.380	8489	135	118
Ferulic acid	F		7351	−1	69.590	7712	136	130
Salicylic acid	S		7017	−1	60.360	7285	143	137
Caffeic acid	C		5729	−1	80.590	8221	175	122
Benzoic acid	B		5684	−1	40.120	6348	176	158
4-Aminobenzoic acid	AB		4431	−1	66.140	7552	226	132
4-Nitrophenol	NP		1469	0	66.050	1359	681	736
4-Hydroxybenzaldehyde	4HB		1283	0	37.290	28	779	>10 000
4-Aminophenol	4AP		701	0	46.250	442	1427	2262

(*continued*)

Table 2. Continued

Name	Abbreviation	Structure	Actual $K_f \pm 100$ (M^{-1})	F	M	Calculated K_f (M^{-1})	Actual K_i (μM)	Calculated K_i (μM)
Phenol	P		<100	0	20.230	–762	>10000	>10 000
Catechol	CA		<100	0	40.460	174	>10 000	5747
5-Hydroxy-1,4-naphthoquinone	HN		<100	0	54.370	818	>10 000	1222
N-(4-hydroxyphenyl) acetamide	HPA		<100	0	49.330	585	>10 000	1709
3-Aminophenol	3AP		<100	0	46.250	442	>10 000	2262
4-Chlorocresol	CC		<100	0	20.230	–762	>10 000	>10 000
1,3,5-Trihydroxybenzene	THB		<100	0	60.690	1111	>10 000	900
4-Hydroxy-3-methoxy-5-sulfamoylbenzoic acid*	–		5538	–1	138	10 885	181	92
3,4-Dihydroxybenzoic acid*	–		3877	–1	80.59	8221	258	122
4-Hydroxy-3-methoxybenzaldehyde*	–		863	0	66.76	1392	1159	718
1,2,3-Trihydroxybenzene*	–		120	0	60.690	1111	8333	900
4-Acetamidophenol*	–		110	0	49.33	585	9091	1709
1-Naphthol*	–		<100	0	20.23	–762	>10 000	>10 000

F: is formula charge; M: is Molecular Polar Surface Area; *: is a molecule used in testing set.
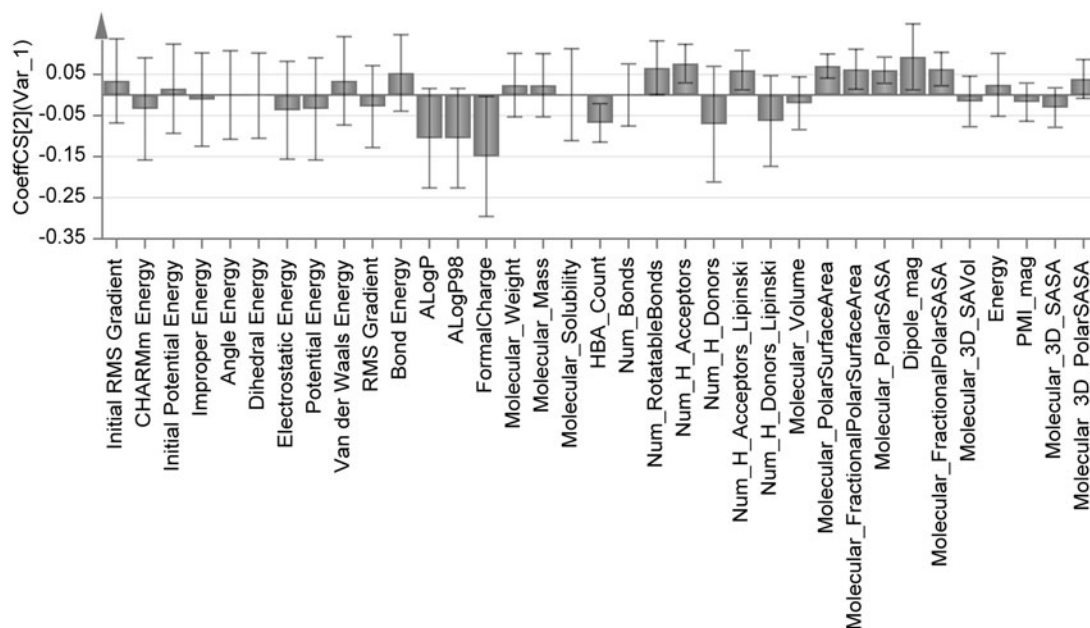
Figure 3. Histogram displayed the correlation coefficient for each structural descriptor with actual values of affinity constants extracted from chromatographic analysis.

Figure 4. Histogram of correlation coefficients with affinity constants of the most two important descriptors obtained from the finally refined QSAR model. It depicts the importance of formula charge in molecular binding with CAIII enzyme.
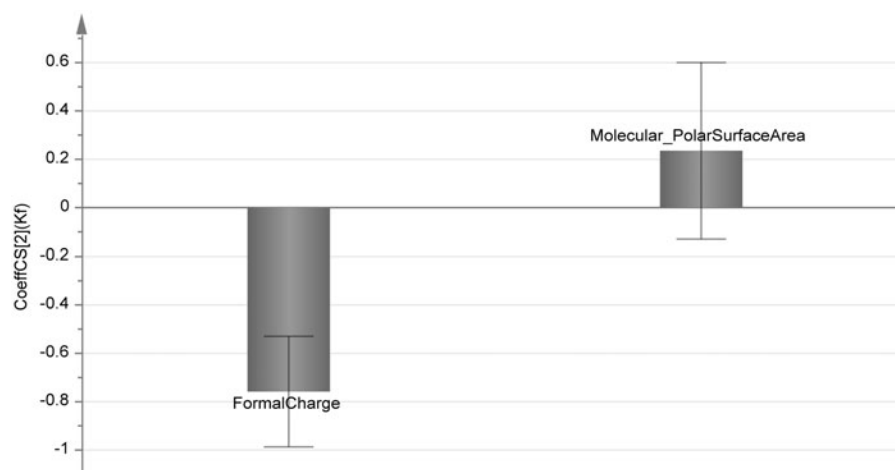
The both descriptors were related to the polarity of molecules and to the distribution of charges on them. It can be deduced from the above how the negative charge of carboxylate group is important in the binding. This drove the attention to the carboxylate ligand formation with zinc atom at CAIII active site. Substitution of carboxylic group with aldehydic, phenolic and acetamido groups lessen the activity (Table 2). Moreover, the presence of other substitution on the benzoic acid and its analogues enhances the activity, particularly with hydroxyl groups such as gallic acid and 4-hydroxybenzoic acid.

ANOVA of regression, exhibited p values below 0.05, indicated the capability of the model in expressing the variation from the descriptors[31]. The validity of the model was tested by internal and external validation strategies. Internal validation was fulfilled by recalculating the values of affinity constants according to the obtained QSAR equation against the actual values by implementing Jack-knifing method which generates accumulative results by leaving-one of descriptor out in each permutation[32]. Plotting the actual affinity values against the calculated ones produced two distinctive clusters without any overlaps as shown in Figure 5. The absence of overlapping is a proof of the model validity obtained from the training set values[33]. Additionally,
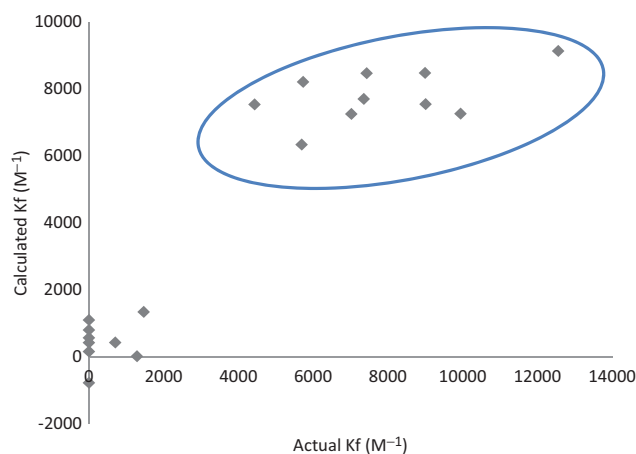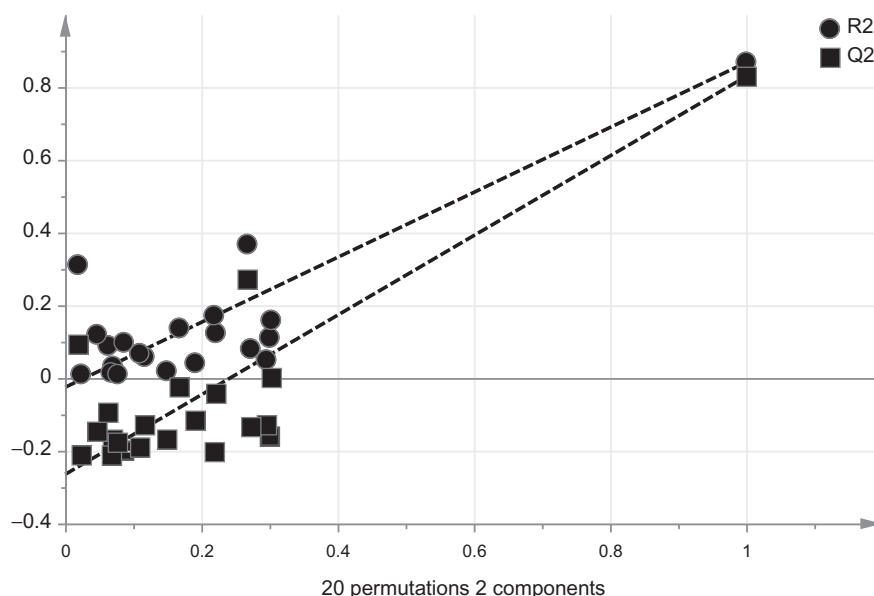

Figure 5. Internal validation chart resulted from plotting actual affinity values against the calculated ones by leaving-one-out strategy. The absence of overlapping between the clusters is a requirement for validity.

Table 3. Results of docking studies using different programs and the consensus scores for training set and relevant compounds. Ferulic acid, caffeic acid and interestingly aspirin depicted the best score values.

Compound Name	CDOCKER	Libdock	PLP1	PLP2	Jain	PMF	PMF04	Dcock score	Gold Score	ASP	Chem PLP	Chemscore	Consensus Score
Ferulic_acid	56.2	77.9	37.5	48.3	4.9	82.9	74.4	91.0	53.4	26.2	53.9	26.6	633.0
Caffeic_acid	52.4	80.0	23.8	45.5	4.3	78.7	67.0	89.4	50.7	22.8	55.7	27.1	597.4
Aspirin	43.1	70.2	47.5	44.8	3.3	70.6	61.7	108.1	50.9	24.4	44.0	20.3	588.8
gallic_acid	45.7	66.9	32.5	42.5	4.6	64.5	57.6	106.2	53.9	24.2	39.7	17.7	556.0
Vanillic_acid	45.7	70.4	29.0	37.0	4.9	60.3	56.9	105.6	55.0	22.7	40.9	19.5	547.8
5-Hydroxyanthranilic_acid	40.3	67.5	37.0	37.8	3.6	59.4	47.8	99.4	41.0	23.6	39.1	21.4	518.0
4-Aminosalicylic_acid	38.4	71.5	29.8	35.7	2.9	53.2	52.2	99.0	37.7	27.8	41.4	23.9	513.6
4-Hydroxybenzoic acid	43.6	67.4	22.7	29.4	3.5	51.8	44.6	103.3	51.4	22.1	39.3	19.3	498.3
4-Hydroxycumarin	25.1	71.0	50.2	57.4	2.5	75.6	63.8	30.9	37.8	22.7	37.4	22.7	497.0
Salicylic_acid	37.0	60.5	33.7	38.0	2.1	51.3	40.9	99.6	45.0	20.1	41.9	23.0	493.1
Para-amino_benzoic_acid	41.4	66.3	25.3	30.0	3.3	50.7	44.1	96.8	37.3	22.3	39.0	22.0	478.5
N-(4-hydroxyphenyl)acetamide	24.6	71.0	43.2	41.3	2.5	78.2	65.2	18.8	43.1	22.7	40.7	19.5	470.8
Benzoic acid	40.1	59.8	24.8	28.3	3.5	48.3	43.9	97.5	44.5	18.5	33.8	20.6	463.6
4-Hydroxy-3-methoxybenzaldehyde	25.5	71.5	35.5	35.3	1.8	70.1	57.8	36.9	42.8	21.2	41.1	19.8	459.2
5-Hydroxy-1_4-naphthoquinone	23.5	67.0	40.7	43.7	2.8	74.1	52.6	31.7	37.8	21.9	34.7	21.1	451.7
7-Hydroxycumarin	25.2	73.8	19.5	42.4	2.5	79.7	61.2	22.2	41.5	21.4	39.1	20.0	448.5
2-Chloro-4-nitrophenol	23.0	68.2	33.4	31.2	0.3	48.3	49.9	36.4	48.1	20.3	33.9	18.5	411.5
4-Hydroxy_benzaldehyde	25.4	63.3	28.9	31.1	1.3	54.1	43.8	31.6	40.0	19.8	38.3	19.9	397.5
4-Nitrophenol	22.8	67.6	26.0	22.9	0.4	48.9	46.9	35.6	39.6	20.1	37.1	20.9	388.7
1_3_5-Trihydroxybenzene	22.3	53.2	29.9	35.1	1.1	54.2	45.9	26.4	32.5	20.9	35.2	20.9	377.6
3-Aminophenol	17.3	58.6	35.1	35.4	1.9	54.8	47.1	22.5	32.4	19.8	33.8	18.8	377.4
4-Chlorocresol	18.5	58.3	38.2	39.1	1.5	49.2	41.1	24.4	33.5	18.1	35.0	20.2	377.0
Catechol	15.5	54.6	25.6	30.9	0.8	54.8	31.7	27.4	33.6	18.9	36.2	24.9	343.8
Resorcinol	20.0	57.0	26.2	32.0	1.2	42.2	35.2	24.7	31.7	18.7	34.9	17.4	341.2
Phenol	10.8	52.9	23.4	31.4	0.9	47.4	39.5	22.3	31.7	16.1	32.8	22.6	331.7
4-Aminophenol	18.9	57.6	23.7	22.8	0.2	35.8	35.5	23.4	31.0	18.6	30.9	18.2	316.7

Figure 6. Permutation plot obtained by calculation of random order of affinity coefficients against the original set of descriptors. The model is not valid unless the goodness of fit values (R^2 and Q^2) at left are below the two reference points on the right and also the regression line of fir (Q^2) should intersect y-axis at or below zero point.

permutation plot, which is calculated by changing the order of coefficient values randomly against constant matrix of descriptors, was used to test the model predictability. Goodness of fit values (R^2 and Q^2) for the permuted y-values at the left side of Figure 6 were below the reference values at the right[34,35]. This indicates the model reliability and also the y-intercept of regression line of Q^2 should intersect y-axis at or below zero point for good predictability[34,35]. As shown in Figure 6, the intercept was found equal to −0.26.

External validation of the model was tested by analyzing six molecules through the chromatographic method and compared with the corresponding predicted values from the equation of QSAR model (Equation 1). Those which acquire strong actual affinity values were also predicted as strong inhibitors

(Figure 7) and also the weak inhibitors were satisfactorily predicted.

Docking analysis

Twenty-six compounds were used in the training and the testing sets. All these compounds have been docked using molecular modeling software and their results were in correlation to the experimental results performed in the enzyme. For example; ferulic acid has the highest consensus docking score of 633 (Table 3), while its assay results were considered of the top molecules of an actual K_f value of 7351 and calculated K_f value of 7712 (Table 2). In addition, it is noticeable that the top ranked molecules of both molecular docking results and the

RIGHTSLINK()

Figure 7. Comparison for the calculated and actual values of affinity constant of molecules of testing set used to accomplish external validation for the QSAR model.
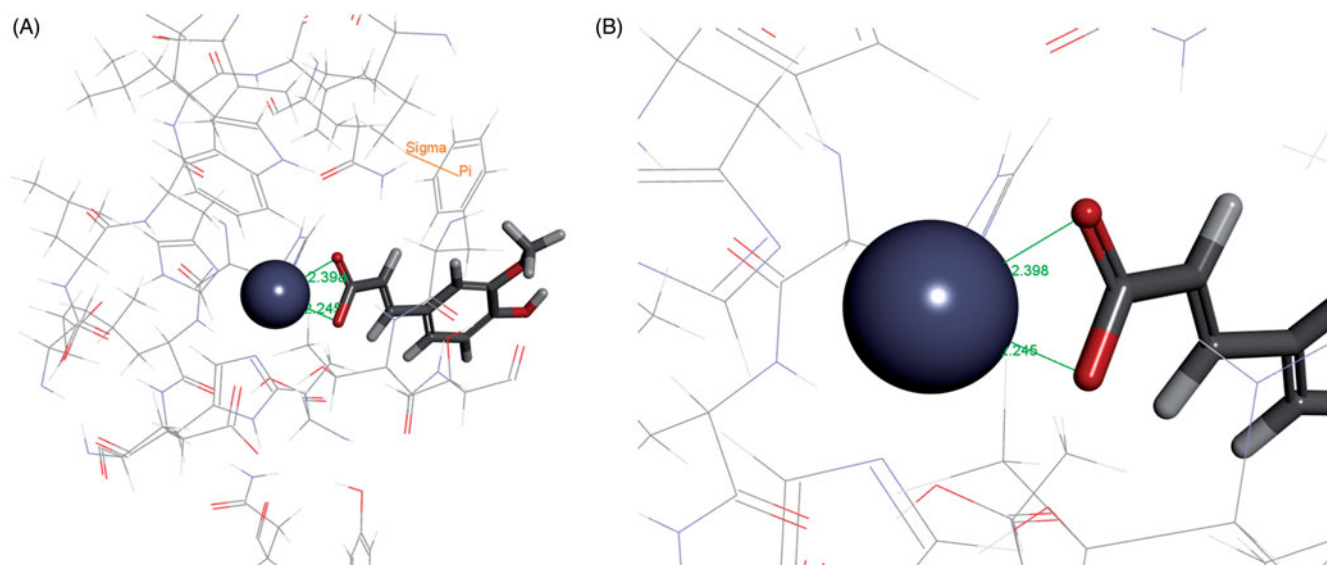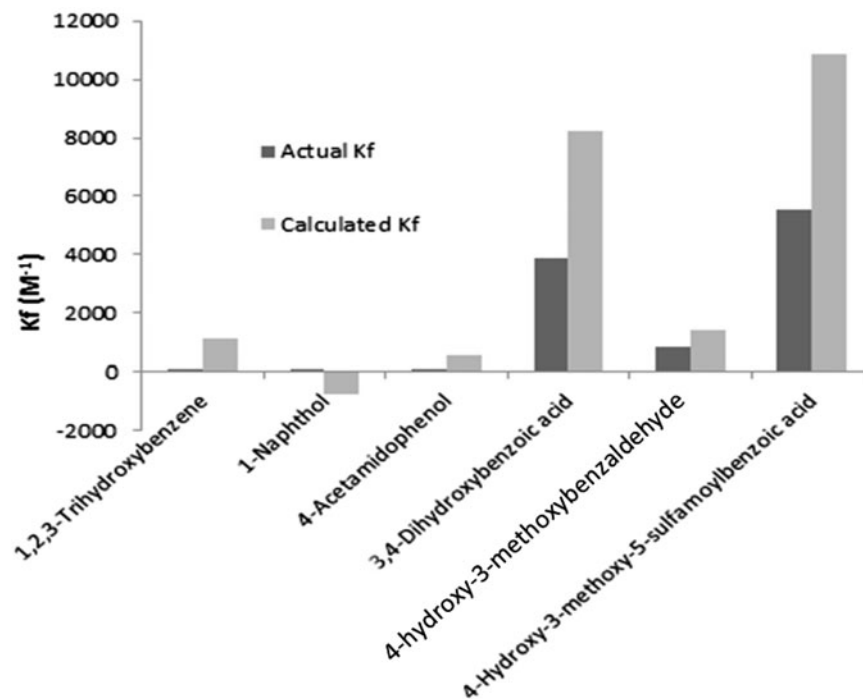

Figure 8. Ferulic acid docked inside the active site of CA-III showing the COO– group in optimum orientation with zinc atom with Oxygens-Zn distances of 2.398 and 2.246 respectively. [B is a closer look to zinc atom for clarification purposes].

chromatographic assay were these compounds that possess carboxylic acid groups indicating their essentiality. Moreover, addition of an extra phenolic group to the carboxylic acid containing molecules will potentiate the activities of these molecules. The carboxylic group is thought to perform coordinate interactions between its negatively charged oxygens of the vacant orbitals of the zinc atom in the active site (Figure 8). This finding is correlated to a publication performed by Nagata in 2012 where he discovered that the most common functional group that coordinated zinc atom inside the active site was COOH[36]. Additionally, the docking study revealed that hydroxyl group of ferulic acid is able to bind with the guanido group of Arg67.

Presence of only phenolic or aniline moiety on the tested compound will be weakly correlated to the activity which is proved by the consensus scoring results and the experimental data (Table 3).

Conclusion

It can be concluded that the vacancy chromatography is suitable for enzymatic inhibition studies against CAIII. The method was working independently of the enzyme catalytic activity. Screening of small molecules accessible to the narrow active site of CAIII enzyme displayed the necessity of carboxylate ion for reaching strong binding with the zinc atom in the enzyme. This work reveals novel inhibitors of CAIII enzyme. The carboxylic acid group of benzoic acid analogues potentiated with phenolic group at p-position represented the maximum activity. The role of the negative charge of the candidate inhibitors is emphasized by a significant correlation coefficient of molecular charge with affinity constants resulted from the valid QSAR model. Docking study also exhibited the importance of the negatively charged carboxylates. Co-crystallization of the enzyme with the tested compounds will be further studied in the future by single crystal

X-ray to gather more information of binding enough to design potential synthetic drugs possessing stronger activity against CAIII.

Declaration of interest

References

1. Supuran CT, Scozzafava A, Casini A. Carbonic anhydrase inhibitors. Med Res Rev 2003;23:146–89.
2. Migliardini F, De Luca V, Carginale V, et al. Biomimetic CO_2 capture using a highly thermostable bacterial α-carbonic anhydrase immobilized on a polyurethane foam. J Enzyme Inhib Med Chem 2014;29:146–50.
3. Supuran CT. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. Nat Rev Drug Discov 2008;7:168–81.
4. Neri D, Supuran CT. Interfering with pH regulation in tumours as a therapeutic strategy. Nat Rev Drug Discov 2011;10:767–77.
5. Supuran CT, Scozzafava A. Carbonic anhydrases as targets for medicinal chemistry. Bioorg Med Chem 2007;15:4336–50.
6. Gilmour KM. Perspectives on carbonic anhydrase. Comp Biochem Physiol A Mol Integr Physiol 2010;157:193–7.
7. Imtaiyaz Hassan M, Shajee B, Waheed A, et al. Structure, function and applications of carbonic anhydrase isozymes. Bioorg Med Chem 2012;21:1570–82.
8. Thiry A, Rolin S, Vullo D, et al. Indanesulfonamides as carbonic anhydrase inhibitors and anticonvulsant agents: structure-activity relationship and pharmacological evaluation. Eur J Med Chem 2008;43:2853–60.
9. Supuran CT, Scozzafava A, Menabuoni L, et al. Carbonic anhydrase inhibitors. Part 71: Synthesis and ocular pharmacology of a new class of water-soluble, topically effective intraocular pressure lowering sulfonamides incorporating picolinoyl moieties. Eur J Pharm Sci 1999;8:317–28.
10. Vuotikka P, Uusimaa P, Niemelä M, et al. Serum myoglobin/carbonic anhydrase III ratio as a marker of reperfusion after myocardial infarction. Int J Cardiol 2003;91:137–44.
11. Nishita T, Igarashi S-I, Asari M. Determination of carbonic anhydrase-III by enzyme-immunoassay in liver, muscle and serum of male rats with streptozotocin-induced diabetes mellitus. Int J Biochem Cell Biol 1995;27:359–64.
12. Mitterberger MC, Kim G, Rostek U, et al. Carbonic anhydrase III regulates peroxisome proliferator-activated receptor-γ^2. Exp Cell Res 2012;318:877–86.
13. Innocenti A, Vullo D, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: interactions of phenols with the 12 catalytically active mammalian isoforms (CA I-XIV). Bioorg Med Chem Lett 2008;18:1583–7.
14. Innocenti A, Vullo D, Scozzafava A, Supuran CT. Carbonic anhydrase inhibitors: inhibition of mammalian isoforms I-XIV with a series of substituted phenols including paracetamol and salicylic acid. Bioorg Med Chem Lett 2008;16:7424–8.
15. Carta F, Vullo D, Maresca A, Scozzafava A, Supuran CT. Mono-/dihydroxybenzoic acid esters and phenol pyridinium derivatives as inhibitors of the mammalian carbonic anhydrase isoforms I, II, VII, IX, XII and XIV. Bioorg Med Chem 2012;21:1564–9.
16. Davis RA, Innocenti A, Poulsen S-A, Supuran CT. Carbonic anhydrase inhibitors. Identification of selective inhibitors of the human mitochondrial isozymes VA and VB over the cytosolic isozymes I and II from a natural product-based phenolic library. Bioorg Med Chem 2010;18:14–18.
17. Alzweiri M, Al-Hiari Y. Evaluation of vanillic acid as inhibitor of carbonic anhydrase isozyme III by using a modified Hummel-Dreyer method: approach for drug discovery. Biomed Chromatogr 2013;27:1157–61.
18. Rengel Z. Carbonic anhydrase activity in leaves of wheat genotypes differing in Zn efficiency. J Plant Physiol 1995;147:251–6.
19. Conroy CW, Buck RH, Maren TH. The microchemical detection of carbonic anhydrase in corneal epithelia. Exp Eye Res 1992;55:637–40.
20. Elder I, Fisher Z, Laipis PJ, et al. Structural and kinetic analysis of proton shuttle residues in the active site of human carbonic anhydrase III. Proteins: Struct Funct Bioinform 2007;68:337–43.
21. Al-Balas Q, Hassan M, Al-Oudat B, et al. Generation of the first structure-based pharmacophore model containing a selective ''zinc binding group'' feature to identify potential glyoxalase-1 inhibitors. Molecules 2012;17:13740–58.
22. Clark RD, Strizhev A, Leonard JM, et al. Consensus scoring for ligand/protein interactions. J Mol Graph Model 2002;20:281–95.
23. Wu G, Robertson DH, Brooks CL, Vieth M. Detailed analysis of grid-based molecular docking: a case study of CDOCKER-A CHARMm-based MD docking algorithm. J Comput Chem 2003;24:1549–62.
24. Rao SN, Head MS, Kulkarni A, LaLonde JM. Validation studies of the site-directed docking program LibDock. J Chem Inf Model 2007;47:2159–71.
25. Muegge I. PMF scoring revisited. J Med Chem 2006;49:5895–902.
26. Böhm H-J. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. J Comput Aided Mol Des 1998;12:309.
27. Verdonk ML, Cole JC, Hartshorn MJ, et al. Improved protein-ligand docking using GOLD. Proteins: Struct Funct Bioinform 2003;52:609–23.
28. Flood KG, Reynolds ER, Snow NH. Determination of apparent association constants of steroid-cyclodextrin inclusion complexes using a modification of the Hummel-Dreyer method. J Chromatogr A 2001;913:261–8.
29. Alzweiri M, Watson DG, Parkinson JA. Metabonomics as a clinical tool of analysis: LC-MS approaches. J Liquid Chromatogr Relat Technol 2012;36:94–115.
30. Alzweiri M, Sills GJ, Leach JP, et al. Response to drug treatment in newly diagnosed epilepsy: A pilot study of 1H NMR-and MS-based metabonomic analysis. Epilepsy Res 2010;88:189–95.
31. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biometrical J 2008;50:346–63.
32. Martens H, Høy M, Westad F, et al. Analysis of designed experiments by stabilised PLS regression and jack-knifing. Chemometr Intell Lab 2001;58:151–70.
33. Shiyab S, Shatnawi M, Shibli R, et al. Influence of developmental stage on yield and composition of *Origanum syriacum* L. oil by multivariate analysis. J Med Plants Res 2012;6:2985–94.
34. Li X, Lu X, Tian J, et al. Application of fuzzy c-means clustering in data analysis of metabolomics. Anal Chem 2009;81:4468–75.
35. Gilbody S, Bower P, Whitty P. Costs and consequences of enhanced primary care for depression: systematic review of randomised economic evaluations. Br J Psychiatry 2006;189:297–308.
36. Kawai K, Nagata N. Metal-ligand interactions: an analysis of zinc binding groups using the Protein Data Bank. Eur J Med Chem 2012;51:271–6.