European Journal of Medicinal Chemistry 46 (2011) 1593–1603

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Pharmacophore-based virtual screening and Bayesian model for the identification of potential human leukotriene A4 hydrolase inhibitors

Sundarapandian Thangapandian, Shalini John, Sugunadevi Sakkiah, Keun Woo Lee*

Division of Applied Life Science (BK21 Program), Environmental Biotechnology National Core Research Center (EB-NCRC), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University (GNU), 900 Gazwa-dong, Jinju 660-701, Republic of Korea

ARTICLE INFO

Article history:
Received 20 August 2010
Received in revised form
31 January 2011
Accepted 3 February 2011
Available online 5 March 2011

Keywords:
Leukotriene A4 hydrolase
Inflammation
Pharmacophore
Database searching
Molecular docking
Bayesian model

ABSTRACT

Leukotriene A4 hydrolase (LTA4H), an enzyme involved in the conversion of LTA4 to LTB4, is an emerging and important anti-inflammatory target. This study demonstrates the development of quantitative pharmacophore hypothesis and Bayesian model and their applications in identification of potential human LTA4H (hLTA4H) inhibitors. The best hypothesis with a high correlation coefficient value of 0.951 was validated using different methods including a test set containing 136 compounds. It was further used as a three-dimensional query in database searching to retrieve virtual leads for hLTA4H inhibition. Molecular docking study was employed to identify the compounds that bind the active site with high affinity. Developed Bayesian model suggested molecular features favoring and not favoring the inhibition of hLTA4H.

1. Introduction

A ubiquitously present 64 kDa metal (Zn^{2+}) containing cytosolic human leukotriene A4 hydrolase (hLTA4H) is a bi-functional enzyme with epoxide hydrolase and aminopeptidase activities utilizing the same Zn present active site [1]. The development and regulation of inflammation are maintained by a complex network of variety of cellular and soluble factors. These factors majorly contain eicosanoids (structurally similar paracrine hormones produced along the arachidonic acid (AA) pathway) which includes the prostaglandins, the leukotrienes (LT), and the lipoxins [2]. LT are a group of lipid mediators associated with acute and chronic inflammatory diseases, particularly asthma, rhinitis, and atherosclerosis [3–5]. Biosynthesis of LT promotes the phosphorylation and membrane translocation of cytosolic phospholipase A2 (cPLA2) and 5-lipoxygenase (5-LO) which are the major enzymes in AA pathway. cPLA2 releases the AA from membrane lipids followed by the action of 5-LO enzyme assisted by five lipoxygenase activating protein (FLAP) to form the unstable epoxide LTA4. This key intermediate is subsequently converted in to LTB4 and LTC4 by the hydrolase activity of LTA4H and by glutathione transferase activity of LTC4 synthase (LTC4S) enzymes, respectively [6]. The very little known aminopeptidase activity of LTA4H has recently speculated that the enzyme may process peptides related to inflammation and host defense [7,8]. LTB4 is a potent pro-inflammatory activator of inflammatory responses mediated through G-protein-coupled receptors, namely, BLT1 and BLT2. LTB4 plays an important role in amplification of many inflammatory disease states such as asthma [9], inflammatory bowel disease [10], chronic obstructive pulmonary disease [11,12], arthritis [13,14], psoriasis [15], and atherosclerosis [16]. It is also recently reported that increased production of LTB4 is associated with the increased risk for myocardial infarction and stroke [17]. Therefore, a therapeutic agent that inhibits the response of cells to LTB4 or the biosynthesis of LTB4 may be useful for the treatment of various inflammatory conditions. Inhibition of hLTA4H as therapeutic strategy is exemplified by the development of multiple inhibitors from different chemotypes [17–20]. There is handful of LT biosynthesis inhibitors and their receptor antagonists are in clinical trials showing promising

Abbreviations: LTA4H, leukotriene A4 hydrolase; hLTA4H, human leukotriene A4 hydrolase; AA, arachidonic acid; LT, leukotrienes; cPLA2, cytosolic phospholipase A2; 5-LO, 5-lipoxygenase; FLAP, five lipoxygenase activating protein; LTC4S, LTC4 synthase; SAR, structure–activity relationship; DS, discovery studio; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; HY, hydrophobic; RA, ring aromatic; PI, positive ionizable; RMSD, root mean square deviation; PDB, protein data bank; ADMET, absorption, distribution, metabolism, excretion and toxicity; GOLD, genetic optimization for ligand docking; E, enrichment factor; GH, goodness of hit; NCI, national cancer institute; FCFP_6, function class fingerprints of maximum diameter 6.

* Corresponding author. Tel.: +82 55 751 6276; fax: +82 55 752 7062.
E-mail address: kwlee@gnu.ac.kr (K.W. Lee).

0223-5234/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2011.02.007

efficacy and some of them were launched as therapeutic products such as zileuton [21], zafirlukast [22] and montelukast [23]. There, still, is a need for novel and potent LTA4H inhibitors to treat various inflammatory diseases. Numbers of crystal structures are available for hLTA4H with different kinds of inhibitors bound at the active site [17,24–28]. The catalytic domain of LTA4H shows high structural similarity to the bacterial metalloprotease thermolysin including conserved zinc binding active site motif. The metal ion is in co-ordination with H295, H299 and a carboxyl atom of E318 [24]. The active site of LTA4H is 17 Å long and L-shaped cavity capable of binding the LTA4 substrate (Fig. 1) [25].

The main objective of this study is to develop a pharmacophore model based on the structure–activity relationship (SAR) from the experimentally known hLTA4H inhibitors. The developed pharmacophore hypotheses were validated for the predictive ability of biological activities of compounds diverse to the compounds used in hypothesis generation. Using this validated model, chemical databases were searched to retrieve structures that can be used as new potentially active candidates. The retrieved candidates were subsequently docked in to the active site of hLTA4H to ensure their binding affinity. Finally two compounds were listed as potent virtual leads for novel hLTA4H inhibitor design as effective anti-inflammatory drugs. Bayesian model developed using 2D descriptors has listed molecular features that are good and bad in terms of enzyme inhibition. This information can be used in further optimization of existing and new inhibitors.

2. Methods and materials

2.1. Collection of dataset

A set of more than 500 chemical compounds experimentally known for the inhibition of LTA4H was retrieved from various literature resources. Variety of assay procedures was reported in literature to observe the LTA4H inhibitory activity of small molecules [18,19,29]. A set of 162 compounds along with the experimental activity expressed in IC_{50} values (concentration of a compound required to inhibit 50% of the hydrolase activity of LTA4H) was selected as a dataset to perform pharmacophore modeling calculations. The inhibitory activities of these compounds were reported employing the same assay procedure [29–35]. Of these 162 compounds, 26 diverse compounds were taken as training set (Fig. 2) by considering structural diversity and a wide range of experimental activity. A test set was prepared with rest of the 136 compounds.

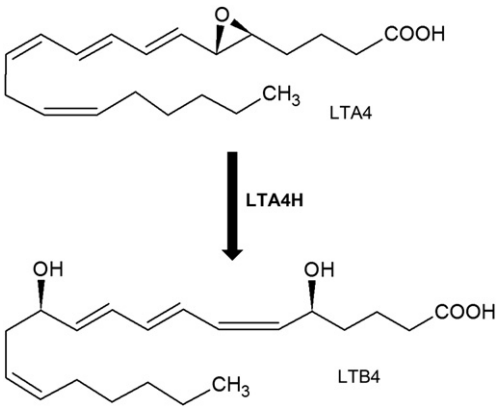

Fig. 1. The rate limiting step catalyzed by LTA4 hydrolase.

2.2. Molecular modeling and conformational generation

The two-dimensional chemical structures of the selected inhibitors were sketched using *ChemSketch* program (Advanced Chemistry Development Inc., Toronto, Canada) and saved in MDL-mol format. Subsequently, they were imported to *Accelrys Discovery Studio* 2.5 (DS) (Accelrys Inc., San Diego, USA) and converted into corresponding standard 3D structures. These compounds in the training and test set were then energy minimized to the closest local minimum using molecular mechanics force field and *Smart Minimizer* tool that performs 1000 steps of *Steepest Descent* followed by *Conjugate Gradient* minimization available in DS. Diverse conformational models were generated for every training set compound as it is needed to produce a good representation of different areas of the conformation space accessible to a molecule within a given energy range. *Diverse Conformation Generation* protocol implemented in DS adopting the *Best flexible conformation generation* option was employed to generate conformational models for each compound. Default parameters were employed in the conformation generation procedure, i.e., conformational models were generated with an energy threshold of 20 kcal/mol from the locally minimized structure and a maximum limit of 255 conformers per molecule [36]. This step probably identifies the best spatial arrangement of chemical groups explaining the activity variations among the training set [37,38].

2.3. Generation of pharmacophore hypotheses

A pharmacophore hypothesis is a representation of generalized molecular features including 3D (hydrophobic groups, charged/ionizable groups, and hydrogen bond donors/acceptors), 2D (substructures), and 1D (physical or biological properties) aspects that are considered to be responsible for a desired biological activity [39]. All 26 compounds in the training set associated with the diverse conformational models were utilized in hypotheses generation using HypoGen protocol available in DS. Two different methodologies (HipHop and HypoGen) are employed in automated generation of pharmacophore hypotheses. HypoGen is based on the SAR in a set of compounds spanning activities of 4–5 orders of magnitude. HypoGen tries to find hypotheses that are common among the active compounds of the training set but do not reflect the inactive ones. It generates and ranks the pharmacophores that correlate best the three-dimensional arrangement of features in a given set of training compounds with the corresponding pharmacological activities (IC_{50} or K_i). This generation and ranking by HypoGen is executed in three steps: the constructive phase, the subtractive phase and the optimization phase [40]. In the constructive phase, hypotheses that are common to the most active set of compounds are identified. In the subtractive phase, all pharmacophore configurations that are also present in the least active set of molecules are removed. All compounds whose activity is by default 3.5 orders of magnitude less than that of the most active compound are considered to represent the least active molecules. During the optimization phase, the hypothesis score is improved. The optimization involves a variation of features and/or locations to optimize activity prediction via a simulated annealing approach. Various cost parameters were also calculated for each new hypothesis. When the optimization process no longer improves the score, HypoGen stops and reports the top scoring 10 unique pharmacophores. HipHop method generates pharmacophore hypotheses based on the common features present in most active compounds. The activity of the compounds used is not considered during model generation using this methodology [41]. Both the methodologies use diverse set of conformational models of every training set compound as input

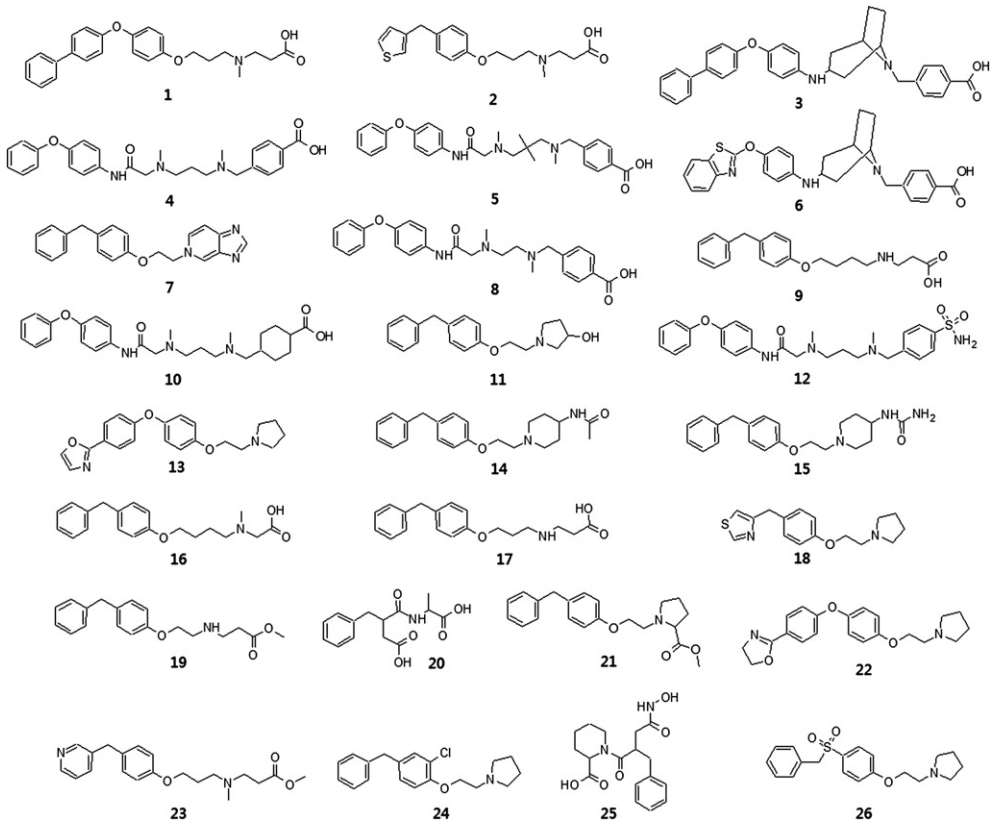

Fig. 2. Training set compounds used in pharmacophore hypotheses generation.

to generate pharmacophore models. The retrieved pharmacophore models are expected to discriminate between active and inactive compounds. The next step in pharmacophore modeling after selecting the training set is choosing possible features to be considered during pharmacophore hypotheses generation. In this study, common feature pharmacophore hypotheses were generated using only most active compounds in the training set to identify the basic common features that are present in most active compounds and selected the identified common features during HypoGen pharmacophore generation procedure. Features such as hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), hydrophobic (HY), positive ionizable (PI) and ring aromatic (RA) were selected based on the common feature pharmacophore generation results. The minimum and maximum number of every chemical feature used in pharmacophore model generation was set to 0 and 5, respectively. The number of out coming models was set to a minimum and maximum value of 4 and 5. The default uncertainty value 3 was changed to 2 as the activity range in the training set barely spans the minimum requirement of four orders of magnitude [42]. All other parameters were set at their default values. Ten pharmacophore hypotheses with significant parameters were generated by HypoGen. The activity of each training set compound is estimated using regression and geometric fit over the generated pharmacophore hypothesis. Error values are also calculated to show the deviation between the experimental and estimated activity. A positive error value is obtained where the estimated activity is higher than the experimental activity and vice versa. The great advantage of using HypoGen is the estimation of the pharmacological activity of a hit compound from the database or newly designed compound for a particular target. The activity estimation is completely depends on the geometric fit of the particular compound.

2.4. Pharmacophore validation

The generated pharmacophore hypotheses were validated using four different methods based on cost analysis, test set prediction, Fischer randomization test and decoy set. The main purpose of validation is to determine the ability of the pharmacophore hypothesis to identify active compounds and predict the activity accurately. The validated pharmacophore hypothesis can be used in database screening to identify novel compounds for designing new LTA4H inhibitors. The first validation was based on the cost functions calculated during pharmacophore generation. HypoGen attempts to minimize three cost functions during pharmacophore generation namely weight cost, error cost and configuration cost. Weight cost is a value that increases as the feature weight in a model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values add to the error cost [40,42,43]. The third function, configuration cost penalizes the complexity of the hypothesis. This is a fixed cost, which is equal to the entropy of the hypothesis space. The more the numbers of features (a maximum of five) in a generated hypothesis, the higher is the entropy with subsequent increase in this cost. The overall cost called total cost is calculated for every generated pharmacophore by summing over these three cost factors. However, error cost is the main contributor to total cost. HypoGen also calculates two other cost value, one for the null hypothesis and the second one for the ideal hypothesis. Null hypothesis which is also known as no correlation cost is generated assuming that there is no relationship in the data and that experimental activities are normally distributed about their mean. Ideal hypothesis is the most likely hypothesis to correlate the data well. The greater the difference from the null hypothesis the more likely the hypothesis does not reflect a chance correlation. In

second validation, a test set containing 136 compounds that are structurally diverse to that of training set and with a wide range of activity values was utilized and the inhibitory activity values were estimated for every test set compound. Test set compounds were classified on the basis of the activity values as active ($\leq$7 nM, ++++), moderately active (>7 $\leq$ 70 nM, +++), less active (>70 $\leq$ 700 nM, ++) and inactive (>700 nM, +). *Ligand Pharmacophore Mapping* protocol of DS with *BEST flexible* searching option was employed to map the test set compound upon the generated pharmacophore hypothesis. Third validation procedure involves Fischer randomization methodology with a goal to check whether there is a strong correlation between the chemical structures and the biological activity in the training set. During this validation number of pharmacophore hypotheses was generated by randomizing the activity data of the training set compounds using the same parameters used to generate the original hypothesis. In this study, 19 random pharmacophore generation runs were generated to achieve 95% confidence level. None of the randomly generated hypothesis should be generated with better cost functions, root mean square deviation (RMSD) and correlation than the original hypothesis to prove that it has not been generated by chance. As a fourth validation procedure, *E* value and other statistical parameters were calculated using a small database containing set of experimentally known active and inactive compounds (decoy set). Compounds with the IC_{50} value less than 7 nM and greater than 700 nM were selected as active and inactive compounds, respectively.

2.5. Database searching

Virtual screening of chemical databases can serve the purpose of finding novel, potential virtual leads suitable for further development. Database searching methodology provides the advantage that the retrieved compounds can be obtained easily for biological testing when compared to any *de novo* design methods [44]. Pharmacophore hypothesis can be used as structural query to search 3D databases to retrieve structures that fit the hypothesis or forecast the activities of novel compounds. A molecule must fit on all the features of the pharmacophore model that is used as 3D query in database searching to be retained as hit. Two database searching options such as *Fast*/*Flexible* and *Best*/*Flexible* search are available in DS. Better results can be achieved using *Best*/*Flexible* search option during database screening. In our study, we performed all database searching experiments using *Best*/*Flexible* search option. The validated pharmacophore model was used as a 3D structural query in database searching. Three publically available databases, namely, NCI2000, Maybridge and Chembridge containing huge set of diverse chemical compounds were deployed in database searching. The retrieved compounds from the databases were filtered based on their estimated activity values and drug-like properties. Estimated activity values were calculated for every hit compound screened from the database based on the geometric fit of the compound upon the pharmacophore hypothesis. Lipinski's rule of five was calculated for every hit compound from the database to reject the nondrug-like database hits. The compounds scored better estimated activity values and not violated any of the Lipinski rules were carried further for molecular docking.

2.6. Molecular docking

Molecular docking experiments were performed using GOLD (Genetic Optimization for Ligand Docking) program version 4.1. GOLD is an automated docking program that uses genetic algorithm to explore the ligand conformational flexibility with partial flexibility of the active site [45]. The algorithm was tested on a dataset of over 300 complexes extracted from the Protein Data Bank (PDB). GOLD succeeded in more than 70% cases in reproducing the experimental bound conformation of the ligand [46,47]. A high resolution (1.58 Å) crystal structure of hLTA4H bound with an inhibitory molecule (PDB code 3FUN) was selected as protein molecule for molecular docking calculation. All the water molecules and other hetero atoms other than the bound inhibitor and the catalytic metal ion present in the protein were removed and hydrogen atoms were added to the protein structure using CHARMM force field. A binding site was defined with the radius of 10 Å around the inhibitor present in the crystal structure. All hit compounds along with training set compounds were docked in to the created binding site. Early termination option was set to 5 from the default value of 3 to stop the program or move on to the next compound if any of the 5 docked conformations of a particular compound predicted within the RMSD value of 1.5 Å and the *Maximum save conformations* was set to 10. All other options were kept default during docking experiments. At the end of each run GOLD separates and ranks all the generated docked conformations based on the fitness score. The fitness score function that is implemented in GOLD consists of H-bonding, complex energy and the ligand internal energy terms. Molecular interactions were observed using DS and Molegro virtual docker [48] programs.

2.7. Bayesian model development using 2D descriptors

A Bayesian model with high classifying capability on active and inactive compounds for hLTA4H inhibitors was developed using DS [49–51]. The same training set containing 26 compounds used in HypoGen pharmacophore generation was used in generating Bayesian model. The following 2D descriptors were used as independent variables during the model generation: molecular function class fingerprints of maximum diameter 6 (FCFP_6), AlogP, molecular weight, number of rotatable bonds, number of rings, number of aromatic rings, number of HBDs, number of HBAs and molecular fractional polar surface area. The 'Calculate Molecular Properties' protocol was used in descriptors calculation and 'Create Bayesian Model' was used for model generation. A leave-one-out cross-validation receiver operating curve (ROC) was obtained. The bad and good molecular features of training set compounds that favored inhibition and did not favor inhibition were identified.

3. Results and discussion

3.1. Pharmacophore generation

The main objective of the present work is to develop a pharmacophore hypothesis to identify novel virtual leads to design potent hLTA4H inhibitors from chemical databases. Common feature

Table 1
Statistical results of the generated pharmacophore models.

Hypo	Total cost	Cost difference[a]	RMSD	Error cost	Correlation	Features
1	110.649	93.299	1.172	93.326	0.951	HBA HY RA RA
2	113.427	90.521	1.276	96.658	0.914	HBA HY RA RA
3	113.496	90.452	1.269	96.414	0.915	HBA HY RA RA
4	118.152	85.796	1.413	101.419	0.893	HBA HY RA RA
5	123.225	80.723	1.544	106.471	0.871	HBA HY RA RA
6	123.61	80.338	1.555	106.910	0.869	HBA HY RA RA
7	123.641	80.307	1.555	106.907	0.869	HBA HY RA RA
8	123.644	80.304	1.553	106.831	0.869	HBA HY RA RA
9	123.916	80.032	1.563	107.230	0.868	HBA HY RA RA
10	126.794	77.154	1.627	109.899	0.856	HBA HY RA RA

Null cost = 203.948; Fixed cost = 92.165; Configuration cost = 15.561.
[a] Cost difference = Null cost − Total cost.

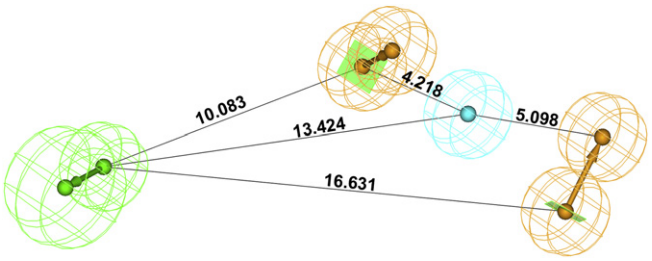

Fig. 3. The best pharmacophore, Hypo1, is shown with its inter-feature distance constraints in angstrom (Å). HBA, green; RA, orange and HYP, cyan color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pharmacophore (HipHop) hypotheses were developed prior to the SAR based (HypoGen) pharmacophore hypotheses using only the most active compounds with the IC_{50} values less than 10 nM. All the generated common feature pharmacophore hypotheses comprised the same chemical features, namely, HBA, HY, and RA indicating that these features are repeatedly present in all the most active hLTA4H inhibitors. However, we cannot use these models to predict the biological activities of the compounds retrieved from the database screening. Therefore, HypoGen pharmacophore hypotheses were generated using the chemical features identified in most active compounds. Ten pharmacophore models confirming the importance of HBA, HY, and RA were generated using 26 structurally diverse training set compounds with the IC_{50} values for hLTA4H inhibition spanning 0.5 nM–4000 nM. Each compound in the training set ought to provide new structural information to develop a good model of statistical significance and predictive power. All ten

Table 2
Comparison of experimental and estimated activity values of the training set compounds based on the best pharmacophore model Hypo1.

Name	Fit value	IC_{50} nM		Error[a]	Activity scale[b]	
		Experimental	Estimated		Experimental	Estimated
1	8.06	0.5	0.29	−1.7	++++	++++
2	6.74	2	1.43	−1.4	++++	++++
3	6.74	5	6.1	1.2	++++	++++
4	6.54	6	9.5	1.6	++++	+++
5	6.48	7	11	1.6	++++	+++
6	6.13	9	25	2.7	+++	+++
7	6.19	11	21	1.9	+++	+++
8	6.19	15	22	1.4	+++	+++
9	6.27	18	18	1	+++	+++
10	6.07	25	28	1.1	+++	+++
11	5.89	30	43	1.4	+++	+++
12	6.44	33	12	−2.7	+++	+++
13	6.46	43	12	−3.7	+++	+++
14	5.38	50	140	2.8	+++	++
15	5.42	87	130	1.5	++	++
16	5.65	110	74	−1.5	++	++
17	5.7	140	66	−2.1	++	++
18	4.68	190	690	3.7	++	++
19	5.43	220	120	−1.8	++	++
20	4.63	300	790	2.6	++	++
21	4.68	450	700	1.5	++	++
22	4.68	590	700	1.2	++	++
23	5.13	820	250	−3.3	+	+
24	4.68	1200	700	−1.8	+	++
25	4.67	1800	1710	−1.1	+	+
26	4.67	4000	2710	1.5	+	+

[a] Positive value indicates that the estimated activity is higher than experimental activity and negative value indicates that the estimated activity is lower than experimental activity.

[b] LTA4H enzyme inhibitory activity: active, ++++ ($IC_{50} \leq 7$ nM); moderately active, +++ ($IC_{50} > 7 \leq 70$ nM); less active, ++ ($IC_{50} > 70 \leq 700$ nM); inactive, + ($IC_{50} > 700$ nM).

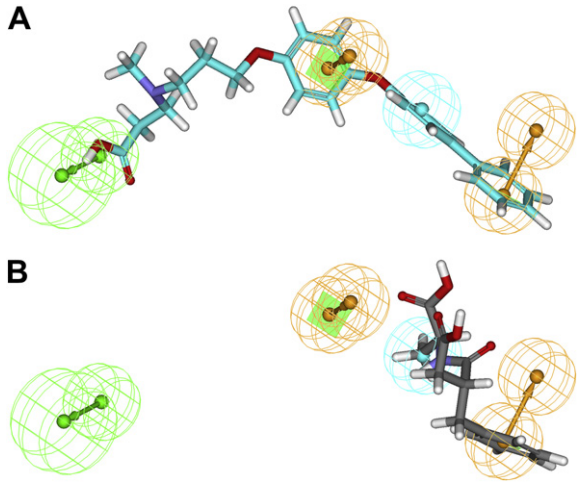

Fig. 4. Pharmacophore overlay of most (compound 1) and least active (compound 26) in the training set. HBA, green; RA, orange and HYP, cyan color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

generated hypotheses contain the same features: one HBA, one HY and two RA features and represent their importance.

3.2. Pharmacophore validation

Four different methods were used to validate the generated pharmacophore hypotheses. First validation is based on the various statistical parameters including cost values calculated during the pharmacophore generation. A significant pharmacophore hypothesis should have a great difference between its total cost and the null cost values. A difference of 40–60 bits between the total cost and the null cost hypotheses shows a 75–90% chance of representing a true correlation in the data [52]. In this study, the total cost values for all the generated pharmacophore hypotheses were between 110.649 and 126.794. The calculated fixed and null cost values were 92.165 and 203.948, respectively. The cost difference between the total and null cost values for the generated hypotheses ranged from 93.299 to 77.154 and represented that they all can correlate the data more than 90%. The first pharmacophore hypothesis (Hypo1) was developed with a high cost difference value of 93.299 than other hypotheses. Further evaluation of generated pharmacophore hypotheses was based on the correlation coefficient. Correlation values of these ten pharmacophore

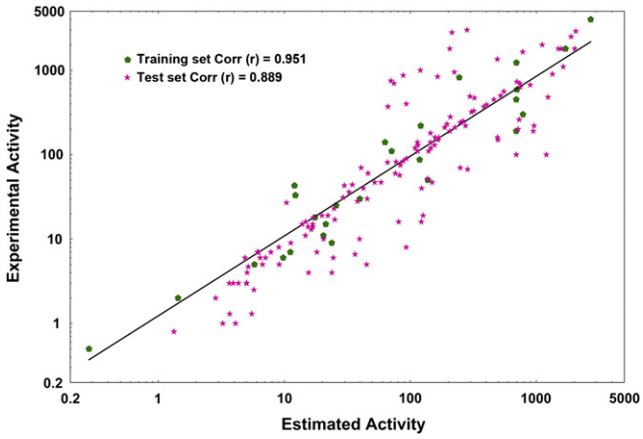

Fig. 5. Correlation plot derived between experimental and estimated activities of training and test set compounds based on Hypo1.

Table 3
Test set prediction based on the best pharmacophore model Hypo1.

Name	IC_{50} nM		Error[c]	Activity scale[d]		Name	IC_{50} nM		Error	Activity scale	
	Experimental[a]	Estimated[b]		Experimental	Estimated		Experimental	Estimated		Experimental	Estimated
27	0.8	1.3	1.7	++++	++++	95	75	83.5	1.1	++	++
28	1	3.2	3.2	++++	++++	96	81	65.9	−1.2	++	++
29	1	4.1	4.1	++++	++++	97	82	77.1	−1.1	++	++
30	1.3	3.7	2.8	++++	++++	98	84	87.8	1.0	++	++
31	1.3	5.5	4.2	++++	++++	99	90	93.4	1.0	++	++
32	2	2.8	1.4	++++	++++	100	100	693.1	6.9	++	++
33	2.5	5.7	2.3	++++	++++	101	100	1202.0	12.0	++	+
34	3	3.7	1.2	++++	++++	102	110	139.7	1.3	++	++
35	3	3.9	1.3	++++	++++	103	110	114.6	1.0	++	++
36	3	5.0	1.7	++++	++++	104	120	109.0	−1.1	++	++
37	3	4.3	1.4	++++	++++	105	120	145.7	1.2	++	++
38	3	5.0	1.7	++++	++++	106	130	115.0	−1.1	++	++
39	4	5.1	1.3	++++	++++	107	130	155.7	1.2	++	++
40	4	15.6	3.9	++++	+++	108	140	144.5	1.0	++	++
41	4	23.8	6.0	++++	+++	109	140	114.0	−1.2	++	++
42	4.7	5.2	1.1	++++	++++	110	150	166.3	1.1	++	++
43	5	9.1	1.8	++++	+++	111	150	496.0	3.3	++	++
44	5	45.0	9.0	++++	+++	112	160	492.6	3.1	++	++
45	5	6.7	1.3	++++	++++	113	160	157.3	−1.0	++	++
46	6	24.5	4.1	++++	+++	114	160	168.1	1.1	++	++
47	6	7.1	1.2	++++	+++	115	180	144.6	−1.2	++	++
48	6	6.4	1.1	++++	++++	116	190	946.1	5.0	++	+
49	6	4.9	−1.2	++++	++++	117	190	206.4	1.1	++	++
50	6.6	36.4	5.5	++++	+++	118	200	717.2	3.6	++	+
51	7	6.1	−1.1	++++	++++	119	210	226.0	1.1	++	++
52	7	6.3	−1.1	++++	++++	120	210	187.5	−1.1	++	++
53	7	7.8	1.1	++++	+++	121	220	960.0	4.4	++	+
54	7	18.1	2.6	++++	+++	122	220	274.2	1.2	++	++
55	8	9.0	1.1	+++	+++	123	230	195.4	−1.2	++	++
56	8	92.8	11.6	+++	++	124	240	248.8	1.0	++	++
57	9	11.3	1.3	+++	+++	125	250	261.6	1.0	++	++
58	10	20.5	2.0	+++	+++	126	260	731.8	2.8	++	+
59	10	39.5	4.0	+++	+++	127	280	207.7	−1.3	++	++
60	11	14.7	1.3	+++	+++	128	320	305.4	−1.0	++	++
61	13	16.4	1.3	+++	+++	129	330	318.9	−1.0	++	++
62	14	15.4	1.1	+++	+++	130	370	383.3	1.0	++	++
63	14	16.8	1.2	+++	+++	131	370	66.4	−5.6	++	++
64	15	13.9	−1.1	+++	+++	132	390	403.5	1.0	++	++
65	15	16.8	1.1	+++	+++	133	400	92.9	−4.3	++	++
66	16	122.4	7.6	+++	++	134	450	460.7	1.0	++	++
67	16	14.8	−1.1	+++	+++	135	470	322.5	−1.5	++	++
68	16	80.8	5.0	+++	++	136	480	1240.1	2.6	++	+
69	17	25.2	1.5	+++	+++	137	490	297.3	−1.6	++	++
70	18	17.0	−1.1	+++	+++	138	500	520.4	1.0	++	++
71	19	22.2	1.2	+++	+++	139	560	553.2	−1.0	++	++
72	19	19.9	1.0	+++	+++	140	630	745.5	1.2	++	+
73	19	126.3	6.6	+++	++	141	670	894.0	1.3	++	+
74	23	24.8	1.1	+++	+++	142	700	74.1	−9.4	++	++
75	27	10.4	−2.6	+++	+++	143	700	740.2	1.1	++	+
76	28	38.1	1.4	+++	+++	144	730	701.6	−1.0	+	+
77	30	45.5	1.5	+++	+++	145	750	70.1	−10.7	+	++
78	31	29.2	−1.1	+++	+++	146	840	164.5	−5.1	+	++
79	31	29.2	−1.1	+++	+++	147	870	87.9	−9.9	+	++
80	36	32.7	−1.1	+++	+++	148	900	1344.6	1.5	+	+
81	40	42.4	1.1	+++	+++	149	950	223.7	−4.2	+	++
82	40	122.8	3.1	+++	++	150	1000	120.1	−8.3	+	++
83	43	30.0	−1.4	+++	+++	151	1100	1635.9	1.5	+	+
84	44	34.7	−1.3	+++	+++	152	1350	490.9	−2.8	+	++
85	47	52.1	1.1	+++	+++	153	1650	780.4	−2.1	+	+
86	47	148.9	3.2	+++	++	154	1800	205.5	−8.8	+	++
87	47	58.6	1.2	+++	+++	155	1800	2020.1	1.1	+	+
88	52	137.3	2.6	+++	++	156	1800	1498.3	−1.2	+	+
89	57	82.4	1.4	+++	+++	157	1800	1581.5	−1.1	+	+
90	60	46.1	−1.3	+++	+++	158	2000	1114.2	−1.8	+	+
91	60	76.1	1.3	+++	+++	159	2500	1888.9	−1.3	+	+
92	67	284.9	4.3	+++	++	160	2800	214.0	−13.1	+	++
93	70	249.1	3.6	+++	++	161	2900	2059.6	−1.4	+	+
94	70	40.7	−1.7	+++	+++	162	3000	282.1	−10.6	+	++

[a] Experimental activity;
[b] Estimated activity.
[c] Positive value indicates that the estimated activity is higher than experimental activity and negative value indicates that the estimated activity is lower than experimental activity.
[d] LTA4H enzyme inhibitory activity: active, ++++ ($IC_{50} \leq 7$ nM); moderately active, +++ ($IC_{50} > 7 \leq 70$ nM); less active, ++ ($IC_{50} > 70 \leq 700$ nM); inactive, + ($IC_{50} > 700$ nM).

models were greater than 0.856 and the first three pharmacophore models were generated with the correlation values >0.9. This value proves the capability of the pharmacophore model to predict the experimental activity of training set compounds. Hypo1 showed the highest correlation coefficient value of 0.951 and thereby showing the high predictive ability of Hypo1. In addition, RMSD values for the top four pharmacophore models were less than 1.5 Å, which further support the predictive ability of the top pharmacophore models. Among the ten pharmacophore models, Hypo1 was developed with better statistical values including higher correlation, greater cost difference, lower RMSD (1.172) and lesser error (93.326) and configuration cost (15.561) values (Table 1). Based on this validation results, Hypo1 was considered as the best pharmacophore hypothesis to be carried for further analyses. Hypo1 consists of one HBA, one HY and two RA features with appropriate inter-feature distance constraints (Fig. 3). Hypo1 has also predicted the training set compounds accurately with the error values less than 10 representing not more than one order of magnitude difference between experimental and estimated activity values (Table 2). An analysis regarding the molecular overlay of the training set compounds upon the pharmacophore hypothesis, Hypo1, was performed. This analysis revealed that the most active compound in the training set has mapped all the pharmacophoric features of Hypo1 whereas the least active compound mapped only on two of the four features (Fig. 4). The only carboxyl moiety and the three phenyl rings present in the most active compound overlaid on HBA and rest (one HY and two RA) of the features, respectively. In terms of the least active compound, the only phenyl ring overlaid on one of the RA features and the central alkyl portion overlaid on HY feature.

Second validation procedure was based on the prediction of the biological activities of a set of test set compounds for which the biological activities are already known via the same assay procedure that was used to determine the activity values of the training set compounds. A test set containing 136 compounds that are structurally diverse to the training set compounds was utilized in this validation. Hypo1 has predicted the biological activities of most of the test set compounds within the same activity scale with the correlation coefficient value of 0.889 (Fig. 5). The biological activities of only 5 out of 136 compounds were predicted with the error values greater than 10 (Table 3). Distinctly, 73.53% (100 of 136 compounds) of the test set compounds were predicted within their activity scales. In third validation procedure, Fischer randomization calculations were carried out to achieve 95% confidence level. The results of the pharmacophore hypotheses generated from the 19 random generation runs were compared with the original pharmacophore hypothesis, Hypo1 and found that none of the randomly generated hypotheses has scored better statistical results than Hypo1 (Table 4). This result has clearly shown that Hypo1 has not been generated by any chance correlation. As a final validation procedure, a small database (D) containing 197 compounds including 45 active and 152 inactive compounds for hLTA4H inhibition was screened using Hypo1 as a 3D structural query. Various statistical parameters including *E*value and goodness of hit (GH) score were calculated to analyze the prediction power of Hypo1. Out of 48 screened compounds (Ht), 42 were active compounds (Ha). Therefore, the calculated *E* value from the screening data was 3.83 representing Hypo1 around four times more probable to pick an active compound from the database than an inactive one. The calculated GH value was 0.854 for which a value greater than 0.5 is significant for any pharmacophore hypothesis (Table 5). Based on these validation results, Hypo1 was considered best and selected to be used in database searching for novel virtual leads that can be used in the designing of potent hLTA4H inhibitors.

Table 4
Validation results for Hypo1 using Fischer randomization test.

Validation No.	Total cost	Fixed cost	RMSD	Correlation	Configuration cost
Hypo1	110.649	92.165	1.172	0.930	15.561
Randomization results					
Trial 1	145.204	89.594	2.059	0.756	12.990
Trial 2	156.718	93.347	2.196	0.716	16.743
Trial 3	167.982	93.347	2.383	0.653	16.743
Trial 4	138.697	89.798	1.887	0.802	13.194
Trial 5	121.246	92.165	1.485	0.882	15.561
Trial 6	153.423	89.793	2.150	0.733	13.188
Trial 7	149.945	93.347	1.985	0.781	16.743
Trial 8	166.963	89.657	2.423	0.638	13.053
Trial 9	170.822	94.360	2.419	0.639	17.756
Trial 10	144.933	88.922	2.019	0.769	12.318
Trial 11	144.971	93.682	1.877	0.807	17.078
Trial 12	115.006	89.793	1.365	0.901	13.188
Trial 13	150.522	88.922	2.128	0.739	12.318
Trial 14	162.221	89.205	2.356	0.662	12.600
Trial 15	142.638	90.390	2.002	0.771	13.786
Trial 16	143.574	91.998	1.949	0.786	15.394
Trial 17	153.842	93.682	2.145	0.731	17.078
Trial 18	168.566	89.594	2.456	0.625	12.990
Trial 19	134.211	90.390	1.832	0.813	13.786

Table 5
Enrichment factor and goodness of hit score validation for Hypo1.

Parameters	Values
Total molecules in database (D)	197
Total Number of actives in database (A)	45
Total hits (Ht)	48
Active hits (Ha)	42
% Yield of actives [(Ha/Ht) × 100]	87.50
% Ratio of actives [(Ha/A) × 100]	93.33
Enrichment factor (E) [(Ha × D)/(Ht × A)]	3.83
False negatives [A − Ha]	3
False positives [Ht − Ha]	6
Goodness of hit score (GH)[a]	0.854

[a] [(Ha/4HtA) (3A + Ht)) × (1 − ((Ht − Ha)/(D − A))]; GH score of >0.7 indicates a very good model.

3.3. Database searching and drug-likeness prediction

The best pharmacophore hypothesis, Hypo1, was used as a 3D query to search three chemical databases, namely, NCI (260,071 compounds), Maybridge (59,632) and Chembridge (50,000) containing totally 369,703 compounds. *Search 3D Database* protocol with the *Best Search* option as available in DS was employed to search these databases. Inhibitory activity values were estimated for the compounds obtained from the database screening. As

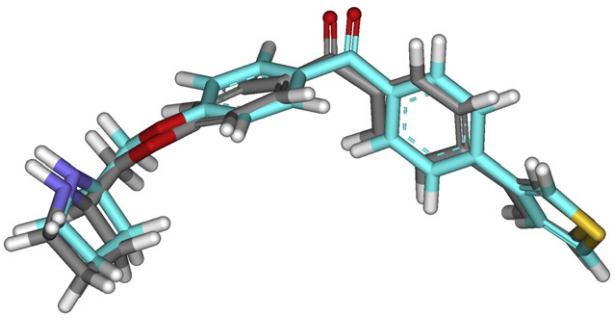

Fig. 6. Overlay of the inhibitor compound with its crystal structure (3FUN) bound (gray) and GOLD predicted (cyan) conformations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

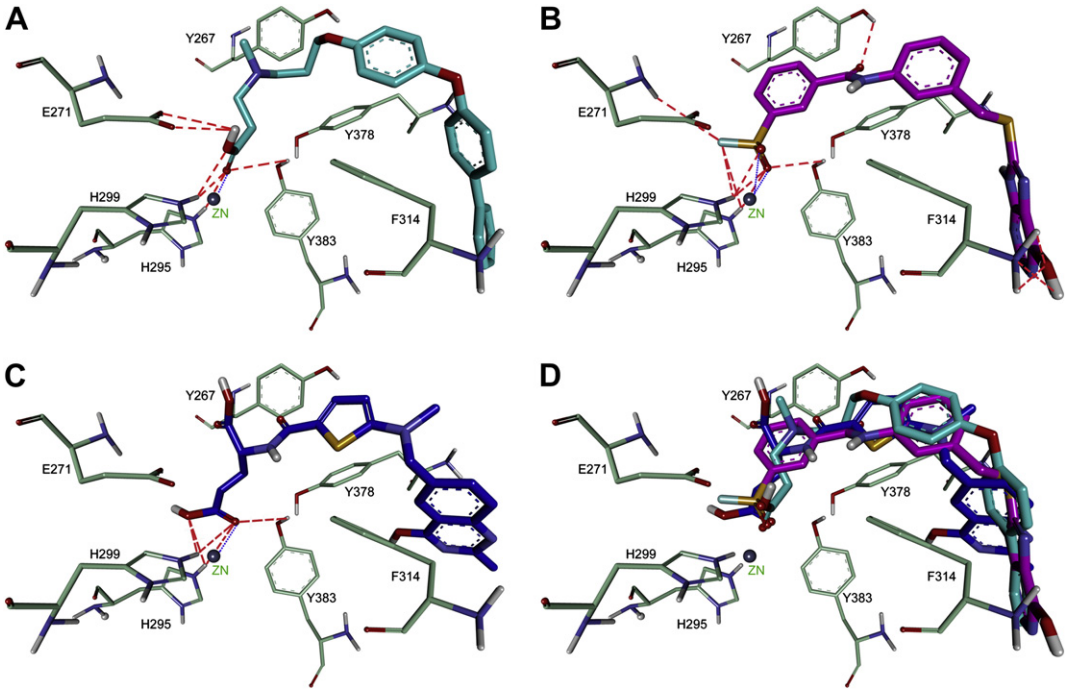

Fig. 7. Molecular docking results. (A) Compound 1 from the training set (B) NCI0211238 (C) NCI0639186 and (D) molecular overlay of docked poses of compound 1, NCI0211238 and NCI0639186. Hydrogen bond interactions are shown in red dashed lines whereas metal ion co-ordinations are shown in blue dotted lines. Only polar hydrogens are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a result, 5858, 909 and 621 (a total of 7388) compounds from NCI, Maybridge and Chembridge, respectively, were mapped upon all the pharmacophoric features present in Hypo1. Out of these, 1526 compounds have scored the *HypoGen* estimated activity value less than 2 nM (better than second most active compound in the training set) and thus considered best for further studies. In addition, violations of Lipinski's rule of five were calculated for the database hit compounds and 124 compounds that obeyed Lipinski's rule of five were subjected to molecular docking studies.

3.4. Molecular docking

Molecular docking program GOLD has been validated on a huge set of protein complexes and proved better than any other docking program for most of the proteins [46,47]. However, it also performed poor for many protein ligand complexes. Thus, we have performed a validation to check whether or not GOLD can perform better for hLTA4H by docking an inhibitor molecule for which the binding conformation is known crystallographically. Crystal structure with the PDB code 3FUN bound with an inhibitor molecule was chosen as receptor and the active site was defined around the bound inhibitor with a radius of 10 Å. The docked conformations were compared with the conformation of the crystal structure-bound conformation. The top conformation predicted by GOLD program was very close to the crystal structure-bound conformation with an all atom RMSD value of 0.62 Å and proved that GOLD can perform better for hLTA4H (Fig. 6). After this validation, all the drug-like final hits identified by the database searching procedure along with 26 training set compounds were docked into the active site of hLTA4H using GOLD 4.1 program. The most important active site functionalities are divalent metal ion (Zn^{2+}), H295, H299, Y378 and Y383 including other residues [24]. The most active compound in the training set has scored a best GOLD fitness score value of

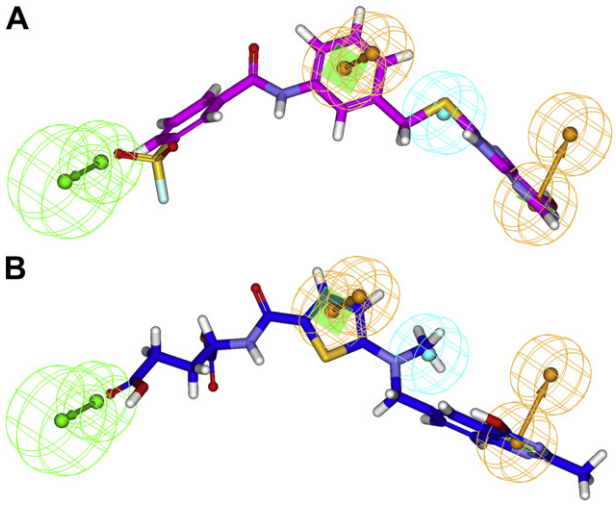

Fig. 8. Pharmacophore mapping of identified hits upon the selected pharmacophore hypothesis, Hypo1. (A) NCI0211238 (B) NCI0639186.

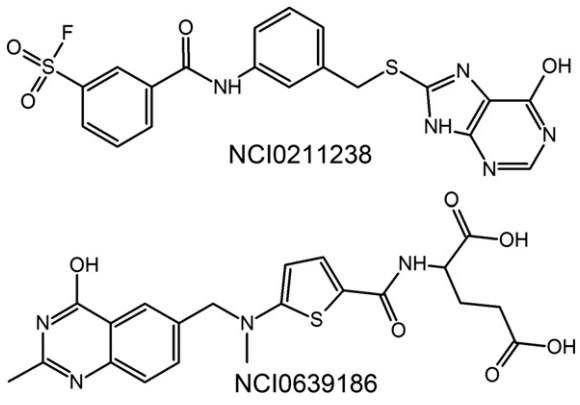

Fig. 9. Two-dimensional representations of the final hit compounds.

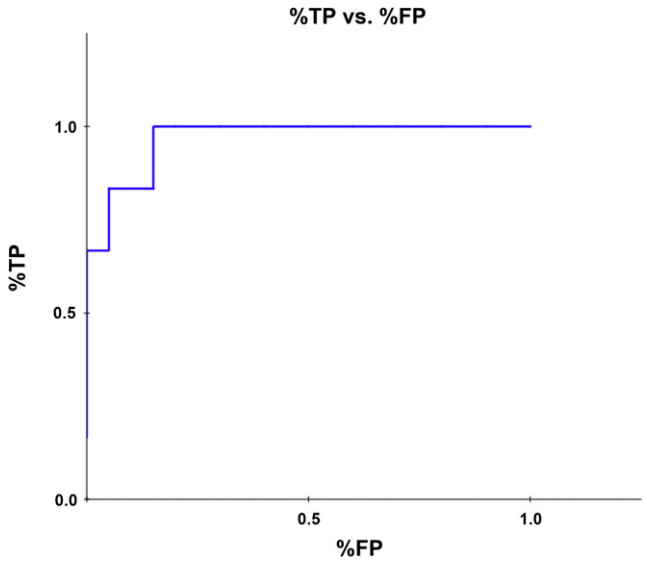

Fig. 10. Cross-validation ROC curve generated by leave-one-out method.

Table 6
Summary table for the Bayesian model.

XV ROC AUC	Best split	TP/FN FP/TN	# In category
0.842	4.180	5/1 3/17	6

90.99 and co-ordinated with the metal ion. It has also formed a hydrogen bond network with E271, H295, H299 and Y383 along with the hydrophobic interactions with Y267, Y378, S379, F314 and W315 (Fig. 7A). Hit compounds that scored a GOLD fitness score >90 were selected and checked for the valid binding modes and interaction with the catalytic functionalities. HBA feature was generated in Hypo1 corresponding to the expected metal binding chemical moieties in the training set compounds [53]. Thus, the parts of the hit compounds that overlaid on HBA feature of Hypo1 during database searching should be located within the interacting distance to metal ion (Zn^{2+}) at the active site of hLTA4H. Therefore, the binding modes not similar to their pharmacophore overlay were discarded during binding mode analysis. The hit compounds whose pharmacophore overlaid conformations substantially close to their binding modes were considered as best candidates for further studies. Hit compounds with similar scaffolds were also rejected based on their GOLD score and active site interactions. Finally, two compounds, namely, NCI0211238 and NCI0639186 scoring GOLD fitness score values of 95.397 and 90.69, respectively, were selected as best hits based on their binding modes and molecular interactions at the active site. The docked conformations of these hits were quite similar to their pharmacophore overlay (Figs. 7 and 8). The first hit compound NCI0211238 from NCI database has shown hydrogen bond interactions with E271, Y267, H295, H299, F314, W315 and Y383. It also interacted with the metal ion through co-ordinate interaction (Fig. 7B) and formed a π−σ interaction with F314. In addition to this the phenyl ring connected to sulfonyl fluoride, the central phenyl ring and the fused ring are positioned to form π−π interaction with Y267, Y378, and P382 (not shown in figure). The second hit compound which was also retrieved from NCI database has shown hydrogen bond interactions with Q134, H295, H299, S379, Y383 and formed co-ordinate interaction with the metal ion (Fig. 7C). Adding to that the central thiophene ring was positioned close to F314 and the quinazoline ring was located close to Y378 and P382 residues to form π−π interactions (not shown in figure). Molecular overlay of these hits with the most active compound of the training set revealed their similar binding orientations at the active site (Fig. 7D). Chemical structures of the final hits confirm that they are of diverse scaffolds. NCI0211238 is a benzamide derivative with a bulky purine substitution whereas NCI0639186 is a thiophene derivative with a bulky isoquinoline substitution (Fig. 9). The novelty search using *SciFinder Scholar* [54] and *PubChem structure* search [55] confirmed that they were not predicted for hLTA4H inhibition earlier.

3.5. Bayesian model with 2D descriptors

Nine descriptors including FCFP_6 and eight other interpretable descriptors were used in developing a Bayesian model with the same 26 training set compounds. Out of 26 compounds, 6 most active compounds (<10 nM) were designated as 'active' with a value '1' and rest of the compounds were given a value '0' to designate them 'not active'. Once the model was built using 26 compounds, it was validated using a leave-one-out cross-validation. Each compound was

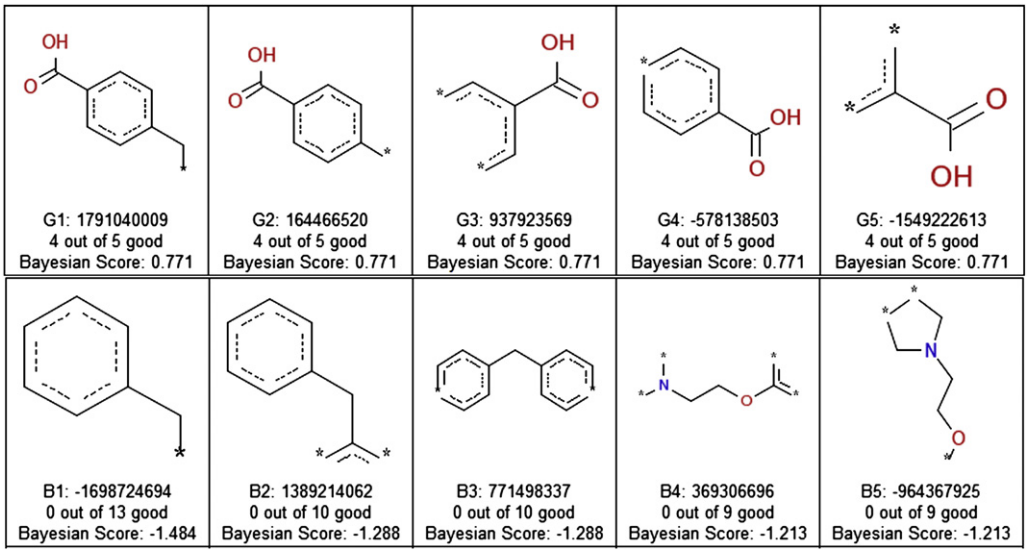

Fig. 11. Top 5 good (G1-G5) and bad (B1-B5) molecular fingerprints identified by FCFP_6 fingerprint descriptor.

left out one at a time, and a model built using the results of the compounds, and that model used to predict the left-out compound. Once all the compounds had predictions, a ROC plot was generated, and the area under the curve calculated. The model with a leave-one-out cross-validation ROC statistic of 0.842 and enrichments that suggested that most active hLTA4H compounds from less and inactive compounds (Fig. 10). The 'Best Split' value was calculated by picking the split that minimized the sum of the percent misclassified for inhibitors and for non-inhibitors, using the cross-validated score for each compound. Using this split, a contingency table is constructed, containing the number of true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN) (Table 6). Using FCFP_6 descriptors in the model generation helped in identifying the molecular features that favor the inhibition and features that not favoring the inhibition were identified. This has shown top 20 fingerprints that positively contributed to the model as well as the top 20 fingerprints that negatively contributed to the model (Figs. S1 and S2). Top 5 fingerprints of both categories are displayed in Fig. 11. The carboxyl moiety substituted in aromatic ring or alkyl chains greatly contributed for the inhibitory activity where as at the other hand aromatic rings with no substitution negatively contributed. Thus, carboxylic group present in most of the training set compounds, which was responsible in metal binding, is very important for the inhibition of the hLTA4H. This information can be considered in further optimization of lead compounds and *de novo* designing of potent hLTA4H inhibitors.

4. Conclusion

In this work we have developed a quantitative pharmacophore model for an important class of hLTA4H inhibitors as anti-inflammatory drugs. The best pharmacophore hypothesis was validated using different methods to evaluate its predicting power over the diverse test set compounds. It has predicted the biological activities of >70% of all test set compounds within the error value of one order of magnitude. This highly predictive hypothesis was further used in database searching for new hLTA4H inhibitors. Three diverse chemical databases were used in database searching. The hits from the database searching were filtered based on the estimated activity values (<2 nM) and Lipinski's rule of five. The resulted drug-like compounds were docked into the active site of hLTA4H using GOLD molecular docking program. Compounds showing better binding compared to the active compounds in the training set were selected as positives from the molecular docking study. Novelty of these hits was confirmed with *Scifinder Scholar* and *PubChem structure* search. Combining all these results, two new compounds were presented as possible lead candidates to be used in novel and potent hLTA4H inhibitors. Further biological testing of hits would be necessary to absolutely determine the success rate of this work and optimize the hits subsequently. Molecular fingerprints obtained from generated Bayesian model revealed the portions that favor and don't favor the inhibition of hLTA4H and be used in future designing of inhibitors.

Acknowledgments

This research was supported by Basic Science Research Program (2009-0073267), Pioneer Research Center Program (2009-0081539), and Environmental Biotechnology National Core Research Center program (20090091489) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST). And all students were recipients of fellowship from the BK21 Program of MEST.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in online version at doi:10.1016/j.ejmech.2011.02.007.

References

[1] C.A. Grice, K.L. Tays, B.M. Savall, J. Wei, C.R. Butler, F.U. Axe, S.D. Bembenek, A.M. Fourie, P.J. Dunford, K. Lundeen, F. Coles, X. Xue, J.P. Riley, K.N. Williams, L. Karlsson, J.P. Edwards, J. Med. Chem. 51 (2008) 4150–4169.
[2] C.D. Funk, Science 294 (2001) 1871–1875.
[3] M. Peters-Golden, W.R. Henderson Jr., N. Engl. J. Med. 357 (2007) 1841–1854.
[4] J.M. Drazen, E. Israel, P.M. O'Byrne, N. Engl. J. Med. 340 (1999) 197–206.
[5] C.Q. Meng, Curr. Top. Med. Chem. 6 (2006) 93–102.
[6] A. Rinaldo-Matthis, J.Z. Haeggstrom, Biochimie 92 (2010) 676–681.
[7] F. Tholander, F. Kull, E. Ohlson, J. Shafqat, M.M. Thunnissen, J.Z. Haeggstrom, J. Biol. Chem. 280 (2005) 33477–33486.
[8] J.Z. Haeggstrom, J. Biol. Chem. 279 (2004) 50639–50642.
[9] J.M. Drazen, Pharmacotherapy 17 (1997) S22–S30.
[10] P. Sharon, W.F. Stenson, Gastroenterology 86 (1984) 453–460.
[11] P.J. Barnes, Nat. Rev. Drug Discov. 1 (2002) 437–446.
[12] S. Gompertz, R.A. Stockley, Chest 122 (2002) 289–294.
[13] F. Tsuji, K. Oki, K. Fujisawa, A. Okahara, M. Horiuchi, S. Mita, Life Sci. 64 (1999) PL51–PL56.
[14] R.J. Griffiths, E.R. Pettipher, K. Koch, C.A. Farrell, R. Breslow, M.J. Conklyn, M.A. Smith, B.C. Hackman, D.J. Wimberly, A.J. Milici, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 517–521.
[15] L. Iversen, K. Kragballe, V.A. Ziboh, Skin Pharmacol. 10 (1997) 169–177.
[16] M. Back, D.X. Bu, R. Branstrom, Y. Sheikine, Z.Q. Yan, G.K. Hansson, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 17501–17506.
[17] V. Sandanayaka, B. Mamat, R.K. Mishra, J. Winger, M. Krohn, L.M. Zhou, M. Keyvan, L. Enache, D. Sullins, E. Onua, J. Zhang, G. Halldorsdottir, H. Sigthorsdottir, A. Thorlaksdottir, G. Sigthorsson, M. Thorsteinnsdottir, D.R. Davies, L.J. Stewart, D.E. Zembower, T. Andresson, A.S. Kiselyov, J. Singh, M.E. Gurney, J. Med. Chem. 53 (2010) 573–585.
[18] X. Jiang, L. Zhou, D. Wei, H. Meng, Y. Liu, L. Lai, Bioorg. Med. Chem. Lett. 18 (2008) 6549–6552.
[19] J.H. Hogg, I.R. Ollmann, J.Z. Haeggstrom, A. Wetterholm, B. Samuelsson, C.H. Wong, Bioorg. Med. Chem. 3 (1995) 1405–1415.
[20] H. Enomoto, Y. Morikawa, Y. Miyake, F. Tsuji, M. Mizuchi, H. Suhara, K. Fujimura, M. Horiuchi, M. Ban, Bioorg. Med. Chem. Lett. 19 (2009) 442–446.
[21] C.C. Zouboulis, Dermato-Endocrinol. 1 (2009) 188–192.
[22] F. Wunder, H. Tinel, R. Kast, A. Geerts, E.M. Becker, P. Kolkhof, J. Hutter, J. Erguden, M. Harter, Br. J. Pharmacol. 160 (2010) 399–409.
[23] N. Gueli, W. Verrusio, A. Linguanti, W. De Santis, N. Canitano, F. Ippoliti, V. Marigliano, M. Cacciafesta, Arch. Gerontol. Geriatr. 52 (2010) e36–e39.
[24] D.R. Davies, B. Mamat, O.T. Magnusson, J. Christensen, M.H. Haraldsson, R. Mishra, B. Pease, E. Hansen, J. Singh, D. Zembower, H. Kim, A.S. Kiselyov, A.B. Burgin, M.E. Gurney, L.J. Stewart, J. Med. Chem. 52 (2009) 4694–4715.
[25] M.M. Thunnissen, P. Nordlund, J.Z. Haeggström, Nat. Struct. Biol. 8 (2001) 131–135.
[26] F. Tholander, A. Muroya, B.P. Roques, M.C. Fournie-Zaluski, M.M. Thunnissen, J.Z. Haeggstrom, Chem. Biol. 15 (2008) 920–929.
[27] M.M. Thunnissen, B. Andersson, B. Samuelsson, C.H. Wong, J.Z. Haeggstrom, FASEB J. 16 (2002) 1648–1650.
[28] T.A. Kirkland, M. Adler, J.G. Bauman, M. Chen, J.Z. Haeggstrom, B. King, M.J. Kochanny, A.M. Liang, L. Mendoza, G.B. Phillips, M.M. Thunnissen, L. Trinh, M. Whitlow, B. Ye, H. Ye, J. Parkinson, W.J. Guilford, Bioorg. Med. Chem. 16 (2008) 4963–4983.
[29] S.K. Khim, J. Bauman, J. Evans, B. Freeman, B. King, T. Kirkland, M. Kochanny, D. Lentz, A. Liang, L. Mendoza, G. Phillips, J.L. Tseng, R.G. Wei, H. Ye, L. Yu, J. Parkinson, W.J. Guilford, Bioorg. Med. Chem. Lett. 18 (2008) 3895–3898.
[30] T.D. Penning, N.S. Chandrakumar, B.N. Desai, S.W. Djuric, A.F. Gasiecki, C.D. Liang, J.M. Miyashiro, M.A. Russell, L.J. Askonas, J.K. Gierse, E.I. Harding, M.K. Highkin, J.F. Kachur, S.H. Kim, D. Villani-Price, E.Y. Pyla, N.S. Ghoreishi-Haack, W.G. Smith, Bioorg. Med. Chem. Lett. 12 (2002) 3383–3386.
[31] T.D. Penning, N.S. Chandrakumar, B.B. Chen, H.Y. Chen, B.N. Desai, S.W. Djuric, S.H. Docter, A.F. Gasiecki, R.A. Haack, J.M. Miyashiro, M.A. Russell, S.S. Yu, D.G. Corley, R.C. Durley, B.F. Kilpatrick, B.L. Parnas, L.J. Askonas, J.K. Gierse, E.I. Harding, M.K. Highkin, J.F. Kachur, S.H. Kim, G.G. Krivi, D. Villani-Price, E.Y. Pyla, W.G. Smith, N.S. Ghoreishi-Haack, J. Med. Chem. 43 (2000) 721–735.
[32] T.D. Penning, N.S. Chandrakumar, B.N. Desai, S.W. Djuric, A.F. Gasiecki, J.W. Malecha, J.M. Miyashiro, M.A. Russell, L.J. Askonas, J.K. Gierse, E.I. Harding, M.K. Highkin, J.F. Kachur, S.H. Kim, D. Villani-Price, E.Y. Pyla, N.S. Ghoreishi-Haack, W.G. Smith, Bioorg. Med. Chem. Lett. 13 (2003) 1137–1139.
[33] B. Ye, J. Bauman, M. Chen, D. Davey, S.K. Khim, B. King, T. Kirkland, M. Kochanny, A. Liang, D. Lentz, K. May, L. Mendoza, G. Phillips, V. Selchau, S. Schlyer, J.L. Tseng, R.G. Wei, H. Ye, J. Parkinson, W.J. Guilford, Bioorg. Med. Chem. Lett. 18 (2008) 3891–3894.
[34] T.D. Penning, M.A. Russell, B.B. Chen, H.Y. Chen, C.D. Liang, M.W. Mahoney, J.W. Malecha, J.M. Miyashiro, S.S. Yu, L.J. Askonas, J.K. Gierse, E.I. Harding, M.K. Highkin, J.F. Kachur, S.H. Kim, D. Villani-Price, E.Y. Pyla, N.S. Ghoreishi-Haack, W.G. Smith, J. Med. Chem. 45 (2002) 3482–3490.

[35] T.D. Penning, L.J. Askonas, S.W. Djuric, R.A. Haack, S.S. Yu, M.L. Michener, G.G. Krivi, E.Y. Pyla, Bioorg. Med. Chem. Lett. 5 (1995) 2517–2522.
[36] M.A.C. Neves, T.C.P. Dinis, G. Colombo, M.L.S. Melo, Eur. J. Med. Chem. 44 (2009) 4121–4127.
[37] P.W. Sprague, R. Hoffman, Catalyst pharmacophore models and their utility as queries for searching 3D databases. in: H. Waterbeemd, B. Testa, G. Folkers (Eds.), Computer-assisted Lead Finding and Optimization-Current Tools for Medicinal Chemistry. VHCA, Basel, Switzerland, 1990, pp. 230–240.
[38] D. Barnum, J. Greene, A. Smellie, P. Sprague, J. Chem. Inf. Comput. Sci. 36 (1996) 563–571.
[39] E.M. Krovat, T. Langer, J. Med. Chem. 46 (2003) 716–726.
[40] Y. Kurogi, O.F. Guner, Curr. Med. Chem. 8 (2001) 1035–1055.
[41] S. Thangapandian, S. John, S. Sakkiah, K.W. Lee, Eur. J. Med. Chem. 45 (2010) 4409–4417.
[42] C. Ya-dong, J. Yong-Jun, Z. Jian-Wei, Y. Qing-Sen, Y. Qi-Dong, J. Mol. Graph. Model. 26 (2008) 1160–1168.
[43] K. Poptodorov, T. Luu, R. Hoffmann, in: T. Langer, R.D. Hoffmann (Eds.), Methods and Principles in Medicinal Chemistry, Pharmacophores and Pharmacophores Searches, Wiley-VCH, Weinheim 2, 2006, pp. 17–47.
[44] S. Theodora, L.J. Thierry, J. Chem. Inf. Comput. Sci. 44 (2004) 1849–1856.
[45] J. Gareth, W. Peter, C.G. Robert, R.L. Andrew, T. Robin, J. Mol. Biol. 267 (1997) 727–748.
[46] J.M.N. Willem, M. Chris, H. Mike, L.V. Marcel, C.C. Jason, T. Robin, Proteins 49 (2002) 457–471.
[47] M.J. Hartshorn, M.L. Verdonk, G. Chessari, S.C. Brewerton, W.T.M. Mooij, P.N. Mortenson, C.W. Murray, J. Med. Chem. 50 (2007) 726–741.
[48] R. Thomsen, M.H. Christensen, J. Med. Chem. 49 (2006) 3315–3321.
[49] A.E. Klon, J.F. Lowrie, D.J. Diller, J. Chem. Inf. Model. 46 (2006) 1945–1956.
[50] P. Prathipati, N.L. Ma, T.H. Keller, J. Chem. Inf. Model. 48 (2008) 2362–2370.
[51] D. Rogers, R.D. Brown, M. Hahn, J. Biomol. Screen. 10 (2005) 682–686.
[52] S. John, S. Thangapandian, S. Sakkiah, K.W. Lee, Eur. J. Med. Chem. 45 (2010) 4004–4012.
[53] S. Thangapandian, S. John, S. Sakkiah, K.W. Lee, J. Mol. Graph. Model. 29 (2010) 382–395.
[54] A.B. Wagner, J. Chem. Inf. Model. 46 (2006) 767–774.
[55] Y. Wang, E. Bolton, S. Dracheva, K. Karapetyan, B.A. Shoemaker, T.O. Suzek, J. Wang, J. Xiao, J. Zhang, S.H. Bryant, Nucleic Acids Res. 38 (2010) D255–D266.