Designing of new multi-targeted inhibitors of spleen tyrosine kinase (Syk) and zeta-associated protein of 70 kDa (ZAP-70) using hierarchical virtual screening protocol

Maninder Kaur, Archna Kumari, Malkeet Singh Bahia, Om Silakari *

Molecular Modeling Lab (MML), Department of Pharmaceutical Sciences and Drug Research, Punjabi University, Patiala, Punjab 147002, India

ARTICLE INFO

ABSTRACT

In the present study, diverse inhibitor molecules of two protein tyrosine kinases *i.e.* Syk and ZAP-70 were considered for the pharmacophore and docking analyses to design new multi-targeted agents for these enzymes. These enzymes are non-receptor protein tyrosine kinases and both are expressed mainly in B and T-lymphocytes where they play a crucial role in immune signaling. The role of these two enzymes in inflammatory and autoimmune diseases makes them potential therapeutic targets for the designing of new multi-targeted agents to combat disease conditions associated with them. The pharmacophore models were developed for Syk and ZAP-70 inhibitors using PHASE module of Schrödinger software. The generated pharmacophore models for both enzymes were clustered and top five models for each target were selected on the basis of *survival minus inactive* score that were subsequently used for the 3D-QSAR analysis. The best model for Syk (ADHR.45-5) and ZAP-70 (AADRR.265-3) were selected corresponding to highest value of Q^2. Both models were employed for the screening of a PHASE database of approximately 1.5 million compounds, subsequently the retrieved hits were screened employing docking simulations with Syk and ZAP-70 proteins. Finally, the screened compounds having structural features of both pharmacophore models and displaying essential interactions with both proteins were investigated for ADME properties. Thus, the new leads obtained in this way would show inhibitory activity against Syk and ZAP-70, and may serve as novel therapeutic agents for the treatment of inflammatory disorders.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Protein tyrosine kinases (PTKs) are important enzymes that play a crucial role in the regulation of cellular activation, proliferation, differentiation and T-cell signaling [1–3]. PTKs are of two types, *i.e.* receptor type and non-receptor type. Receptor type PTKs contains an extracellular ligand binding domain whereas non-receptor type PTKs contains an intracellular ligand binding domain. The members of receptor type PTKs include platelet derived growth factor receptor (PDGFR), epidermal growth factor receptor (EGFR), fibroblast growth factor (FGR) while non-receptor type PTKs include c-Src, Lck, Syk and ZAP-70. Among non-receptor type PTKs, Syk and ZAP-70 are the members of Syk kinase family [4,5]. Syk is mainly expressed in platelets, B-lymphocytes, mast cells, basophils, neutrophils, dendritic cells, macrophages and monocytes whereas ZAP-70 in T-lymphocytes and natural killer cells. As observed from the cellular location, both Syk and ZAP-70 play an important role in the immune response signaling [6–8]. Therefore, the molecules that can simultaneously inhibit Syk and ZAP-70 would serve as more efficacious therapeutic agents for the treatment of various inflammatory and auto-immune disease conditions.

In the present study, two different pharmacophore models were developed for inhibitors of Syk and ZAP-70 enzymes, subsequently the models were employed for the screening of a PHASE molecular database. The integration of pharmacophore based virtual screening employing developed models and structure based virtual screening employing 3D structures of Syk and ZAP-70 may result in the identification of new multi-targeted agents for Syk and ZAP-70 [9]. Therefore, this systematic hierarchical VS protocol including pharmacophore models and docking simulations could be used to dig out the valuable information for the designing of new dual inhibitors for Syk and ZAP-70.

2. Material and methods

2.1. Molecular dataset: selection and processing

For the present study, two datasets consisted of 116 Syk (Table S1, Supplementary information) and 63 ZAP-70 (Table S2, Supplementary information) inhibitors were selected

* Corresponding author. Tel.: +91 9501542696.
  E-mail address: omsilakari@rediffmail.com (O. Silakari).

from the literature [10–16]. Both datasets spanned over a wide range of biological activity, *i.e.* 0.004–4.102 μM and 0.011–13.2 μM for Syk and ZAP-70 inhibitors respectively. The biological activities of the datasets are reported as IC_{50} values where IC_{50} represents the dose in mole required to produce 50% inhibition of the corresponding enzyme. The reported IC_{50} value of each molecule was converted into pIC_{50} value by taking negative logarithm of IC_{50} value ($pIC_{50} = -\log IC_{50}$). All study molecules were sketched and cleaned using '*builder tools*' option of *Maestro 9.2* molecular modeling environment [17]. The sketched molecules were optimized in '*ligprep*' module using OPLS_2005 force field and the ionization states were generated at pH value of 7.0 ± 0.1 and 7.6 ± 0.2 for Syk and ZAP-70 inhibitors respectively [18].

2.2. Pharmacophore generation

For the generation of pharmacophore models, the optimized molecules were imported into PHASE 3.3 and the conformational space of each molecule was sampled using Monte Carlo Minimization Method (MCMM)/Low Mode Docking (LMOD) algorithm [19,20]. The conformational space of each molecule was utilized for the generation of pharmacophore models that are ranked on the basis of scoring parameters, *i.e. survival* and *survival minus inactive* (S-I) score. The generated hypotheses were clustered and the representative hypothesis of each cluster was selected on the basis of highest S-I score. The selected hypotheses were utilized to align the non-model molecules (*i.e.* not considered for the pharmacophore generation) and subsequently to develop 3D-QSAR models for the prediction of molecules.

2.3. 3D-QSAR analysis

3D-QSAR analysis is a computational method that correlates the dependent variables (biological activity) with independent variables (ligand descriptors) using chemometric technique (PLS) [21]. For the generation of 3D-QSAR models, all training set molecules aligned over the common pharmacophoric sites were placed into a regular grid of cubes (1 Å). Each cube was allocated 0 or 1 "*bits*" to account for the different types of atomic features in training set molecules that occupy the cube. Each occupied cube give rise to one or more *volume bits*, where a separate bit is allocated for each different categories of atom that occupy the cubes. The total number of volume bits assigned to a given cube is based on occupation from training set molecules. Hence, a single molecule may be represented by a string of *zeros* or *ones*, according to cubes it occupy. A large pool of binary values (0 and 1) was formed for the data set molecules that was treated as independent variable for the generation of QSAR models [19,20].

Partial least square (PLS) analysis, a combination of principal component and multiple regression analysis, was employed to generate 3D QSAR models [21]. To perform PLS analysis '*t-value*' (eliminate variable) was set to less than 2 in order to improve signal to noise ratio. PHASE creates a series of regression models incorporating progressively more PLS factors with maximum number of factors being not larger than 1/5th the number of molecules in training set. The best QSAR model was selected on basis of the highest values of test (coefficient of determination for test set molecules). The best model should also has high values of R^2_{train} (coefficient of determination for training set molecules), *F*-value (fisher test value) and low SD (standard deviation). The selected QSAR model was validated for its external predictive ability by calculating Pearson-*r*. Pearson coefficient of correlation determines the predictive reliability of the generated model for external data set that has not been considered for the development of model, *i.e.* test set molecules. Additionally, the external predictive ability of the generated model was also validated by calculating a set of parameters *i.e.* R^2_o or R'^2_o close to R^2,

$[(R^2 - R^2_o)/R^2] < 0.1$ or $[R^2 - R'^2_o/R^2] < 0.1$, and the corresponding $0.85 \le k \le 1.15$ or $0.85 \le k' \le 1.15$ [22].

To determine the external predictibility of generated models the applicability domain (APD) was also calculated. It defines the similarity threshold between training and test set compounds. Similarity measurements were calculated on the basis of Euclidean distances between all pairs of training set molecules. If the similarity of test set molecules does not lie within the threshold (APD), model is considered to be unreliable for the prediction of new compounds [23].

$$APD = \langle d \rangle + \sigma Z$$

To calculate APD, initially average of all Euclidean distances between the training set compounds was calculated [24]. Then, new average $\langle d \rangle$ and standard deviation σ of training set molecules with distances lower than previously calculated average value were computed. Z is an empirical cutoff with default value of 0.5.

Additionally, the robustness of generated model was analyzed by performing *Y*-randomization test [25]. In this, the activity data of training set molecules was scrambled randomly to generate different training sets from the original training set. Then, R^2_{train} value was determined for random sets and average value was reported as R^2_{scramble}. This value signifies the degree to which the generated models can fit meaningless data (scrambled) and must has value low as compared to R^2_{train} for original training set of the selected model.

2.4. Virtual screening (VS)

Virtual screening (VS) of the multi-conformational molecular database is a major technique in the drug discovery for new hit identification. VS can broadly be of two types *i.e.* structure based virtual screening (SBVS) and ligand based virtual screening (LBVS) [9]. Since, a precise scoring function which can accurately predict the binding conformations of a molecule for the protein is not defined yet; the discovery of high quality leads is hampered during SBVS. The database screening using pharmacophore model, a method of LBVS, is superior over SBVS due to its ability to screen millions of multi-conformational compounds in comparatively lesser time and retrieve structurally diverse hits. LBVS lacks the ability to predict precise activity of hits and identify false hits. These drawbacks can be overcome employing SBVS and LBVS as an integrated approach in drug discovery protocol.

A PHASE database of about 1.5 million molecules was utilized for the screening with developed pharmacophore models for Syk and ZAP-70 inhibitors to identify new hits. Hits are the compounds which possess chemical groups that spatially overlap (map) with corresponding structural features of the query pharmacophore model. The hits obtained from pharmacophore screening were further subjected to docking simulations with Syk and ZAP-70 proteins to remove false positive and negative molecules. Finally, the hits displaying essential interactions with both proteins were processed in '*QikProp*' program to calculate their ADME properties and observe pharmacokinetic profile.

3. Results and discussion

3.1. Pharmacophore development

For the generation of pharmacophore models, the optimized molecules were imported into PHASE and divided into actives and inactives. Syk inhibitors with pIC_{50} value more than 6.860 and less than 4.887 were considered as actives and inactives respectively. In this, total 10 molecules were obtained as actives and 6 as inactives whereas remaining molecules were considered as moderate

Table 1
Statistical results of PHASE generated pharmacophore hypotheses for Syk and ZAP-70 inhibitors.

Hypotheis ID	Survival score	S-I score[a]	Site	Vector	Volume	# Matches
For Syk inhibitors						
ADHR.43	75.004	73.485	0.690	0.886	0.542	10
ADHR.44	75.072	73.464	0.740	0.899	0.549	10
ADHR.45	75.056	73.413	0.730	0.857	0.587	10
ADHR.12	75.049	73.339	0.660	0.860	0.645	10
ADHR.130	75.010	73.325	0.660	0.868	0.596	10
For ZAP-70 inhibitors						
AADHR.112	69.271	68.164	0.950	0.980	0.777	9
AAAAR.1374	68.885	68.011	0.810	0.826	0.691	9
AADHR.98	69.241	67.988	0.950	0.972	0.756	9
AADRR.265	68.997	67.980	0.780	0.901	0.758	9
AADHR.88	69.198	67.949	0.940	0.976	0.721	9

[a] S-I, survival minus inactive score corresponds to *score inactives* and high value of S-I represents the hypothesis that is more likely to pick active molecules than inactives. A, hydrogen bond acceptor; D, hydrogen bond donor; H, hydrophobic; R, ring aromatic.

active. In case of ZAP-70 inhibitors, the molecules with pIC_{50} value greater than 6.560 were considered as actives and those with less than 5.000 as inactives; thereby total 9 molecules were obtained as actives and 5 as inactives. The conformational space of each molecule was then sampled using MCMM/LMOD algorithm. A maximum of 1000 conformers were generated for each molecule with relative energy difference of 20 kcal/mol. RMSD value of 1 Å was set to remove the redundant conformers. The pharmacophoric sites were created for all molecules that resulted in total six kind of features *i.e.* acceptor, donor, hydrophobic, positive ionizable, negative ionizable and ring aromatic. For the generation of pharmacophore models, the software was restricted to explore minimum of 4 and maximum of 5 pharmacophoric sites for both Syk and ZAP-70 models. The models for Syk and ZAP-70 inhibitors were restricted to match 10 out of 10 and 9 out of 9 active molecules respectively. Total 16 hypotheses were obtained for Syk and 195 for ZAP-70 inhibitors that were ranked accordingly on the basis of their statistical parameters *i.e. survival* score and *survival minus inactive* (S-I) score. Survival score corresponds to *score active*, S-I to *score inactive*. The pharmacophore model corresponding to the highest value of *score active* has ability to identify preferentially active molecules whereas the model with highest value of *score inactive* has ability to discriminate active and inactive molecules. The generated several hypotheses look much alike and have very similar score. To avoid the selection of similar hypotheses for analysis, all generated hypotheses were clustered and the representative hypothesis of each cluster was selected on the basis of highest S-I score. In this manner, total five hypotheses were selected for each Syk and ZAP-70 that belong to different clusters and display high values of S-I score. The statistical results of the selected best hypotheses for Syk and ZAP-70 inhibitors are displayed in Table 1.

3.2. 3D-QSAR analysis

For the generation of 3D-QSAR models, the molecules of both datasets were divided randomly into training and test set considering uniform variation of structural diversity and biological activity of the molecules (Table 2). The training set molecules aligned over the selected pharmacophore hypotheses were utilized for the generation of 3D-QSAR models employing 7 and 5 as PLS factors for Syk and ZAP-70 inhibitors respectively. For the generation of both models grid spacing was set to 1 Å.

The best QSAR model for Syk inhibitors was selected corresponding to PLS factor 5 and pharmacophore hypothesis ADHR.45 on the basis of highest value of Q^2 *i.e.* 0.506. The best model showed R^2_{train} 0.945, SD 0.227 and F-value 216.800. The stereoview of ADHR.45 and its mapping over the highest active molecule Syk58 is displayed in Fig. 1A and B. For ZAP-70 inhibitors best QSAR model was obtained corresponding to PLS factor 3 and pharmacophore

hypothesis AADRR.265 that displayed Q^2 0.695, R^2_{train} 0.901, SD 0.279 and F-value 103.100. The stereoview of AADRR.265 and its mapping over the highest active molecule ZAP35 are displayed in Fig. 1C and D respectively. The correlation graphs between the predicted activity of training set molecules from the best model and experimental activity are displayed in Fig. 2A (Syk) and C (ZAP-70). The statistical results of generated QSAR models for Syk and ZAP-70 inhibitors on the basis of selected pharmacophore models are mentioned in Table 3. The selected best models were further validated for their external predictive reliability by calculating Pearson coefficient of correlation (Pearson-r). The best model for Syk (ADHR.45-5) and ZAP-70 inhibitors (AADRR.265-3) showed Pearson-r value of 0.736 and 0.837 respectively. The correlation graphs between predicted activity of test set molecules from the best model and experimental activity are displayed in Fig. 2B (Syk) and D (ZAP-70). Additionally, ADHR.45-5 and AADRR.265-3 displayed satisfactory values of k, 0.997; k', 0.993; R^2_o, 0.998; R'^2_o, 0.997; $[(R^2 - R^2_o)/R^2]$, −0.056 or $[R^2 - R'_o/R^2]$, −0.111 and k, 0.995; k', 1.001; R^2_o, 0.997; R'^2_o, 0.999; $[(R^2 - R^2_o)/R^2]$, −0.107 or $[R^2 - R'_o/R^2]$, −0.098 respectively. These all parameters strengthen the external prediction reliability of the selected models.

Applicability domain calculated for all test set molecules of selected models for Syk and ZAP-70 was within limit of calculated APD, hence the selected two models were considered to be reliable for external prediction (Tables 4 and 5). In Y-randomization test for checking the robustness of selected models, the average value of $R^2_{scramble}$ for selected two models of Syk and ZAP-70 was calculated to be 0.814 and 0.573 respectively. The calculated lower values of $R^2_{scramble}$ for both selected models showed that the correlation of statistical data is real not by chance.

3.3. Validation of pharmacophore models by docking analysis

For the validation of generated pharmacophore models, the structural features of developed models were compared with the results of docking analysis. For this purpose, the highest active molecules Syk58 and ZAP35 were docked with the corresponding proteins and docking interactions were compared with the chemical features of the pharmacophore models (Fig. 3A and B). The mapping of ADHR.45-5 over the highest active molecule Syk58 showed that the HBA and HBD features are mapped to carbonyl 'O' and NH_2 of amide group; and these two features are complementing the docking interactions with Ala451 (O···N, 2.061 Å), Arg498 (O···N, 2.972 Å) and Glu449 (N···O, 1.941 Å). The hydrophobic feature of the pharmacophore model is mapped to the CH_3 of methoxy group which show interactions with Leu377 in docking analysis. The mapping of the highest active molecule ZAP35 over AADRR.265-3 displays that the HBD feature is mapped over the linker NH connecting pyrimidine and benzene ring, and the

Table 2
Details of dataset for 3D-QSAR of Syk and ZAP-70 inhibitors.

Name of target	Total number of molecules	Training set	Test set
Syk	116	69	47
ZAP-70	68	38	25
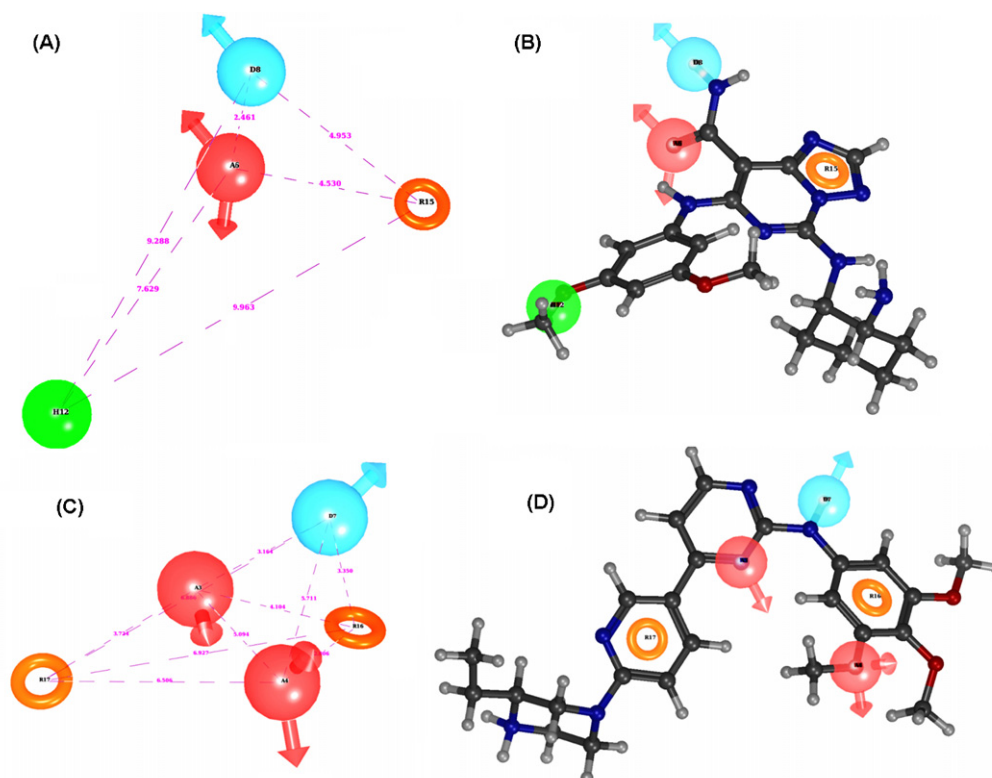
Fig. 1. Stereoviews of the best pharmacophore model ADHR.45 (A), stereoview of AADRR.265 (C), mapping of ADHR.45 over the highest active molecule Syk58 (B) and mapping of AADRR.265 over the highest active molecule ZAP35 (D).

same feature is also displaying an H-bonding interaction with Ala417 (N···O, 3.150 Å) amino acid residue in the docking analysis. The chemical features of both pharmacophore models exert complementary fit with the docking results, thus indicates that the developed models are of high quality.

3.4. Virtual screening and ADME properties calculation

Initially, ADHR.45-5 pharmacophore model generated for Syk inhibitors was challenged to screen PHASE molecular database

of approximately 1.5 million compounds using '*find matches*' option that retrieved 1000 diverse molecules as hits. These 1000 molecules were further screened with second pharmacophore model AADRR.265-3 generated for ZAP-70 inhibitors that retrieved 261 molecules having structural features of two query pharmacophore models. These molecules may consist of false positive and negative hits also, thus needed to be filtered. For this purpose, the docking simulations were conducted for 261 molecules with both proteins and the molecules displaying essential interactions with both proteins were scrutinized. Prior to docking analysis,

Table 3
Statistical results of the generated 3D-QSAR models for Syk and ZAP-70 inhibitors.

Model ID	#PLS factors	SD	R^2_{train}	*F*-value	Stability	RMSE	Q^2_{test}	Pearson-*r*
For Syk								
ADHR.43	5	0.244	0.936	184.600	0.497	0.813	0.2376	0.493
ADHR.44	7	0.182	0.966	244.100	0.253	0.830	0.2052	0.467
ADHR.45	5	0.227	0.945	216.800	0.477	0.654	0.5064	0.736
ADHR.12	1	0.699	0.444	53.500	0.849	0.803	0.2569	0.526
ADHR.130	1	0.779	0.309	29.900	0.905	0.821	0.2233	0.498
For ZAP-70								
AADHR.112	4	0.142	0.975	322.900	0.677	0.512	0.545	0.753
AAAAR.1374	1	0.556	0.582	50.200	0.927	0.533	0.507	0.719
AADHR.98	5	0.137	0.977	277.300	0.574	0.491	0.580	0.770
AADRR.265	3	0.279	0.901	103.100	0.790	0.419	0.695	0.837
AADHR.88	1	0.579	0.547	43.500	0.902	0.509	0.550	0.760

SD, standard deviation; R^2_{train}, coefficient of determination for training set molecules; *F*-value, Fisher test; RMSE, root means squared error; Q^2_{test}, coefficient of determination for test set molecules; Pearson-*r*, Pearson coefficient of correlation for test set molecules.
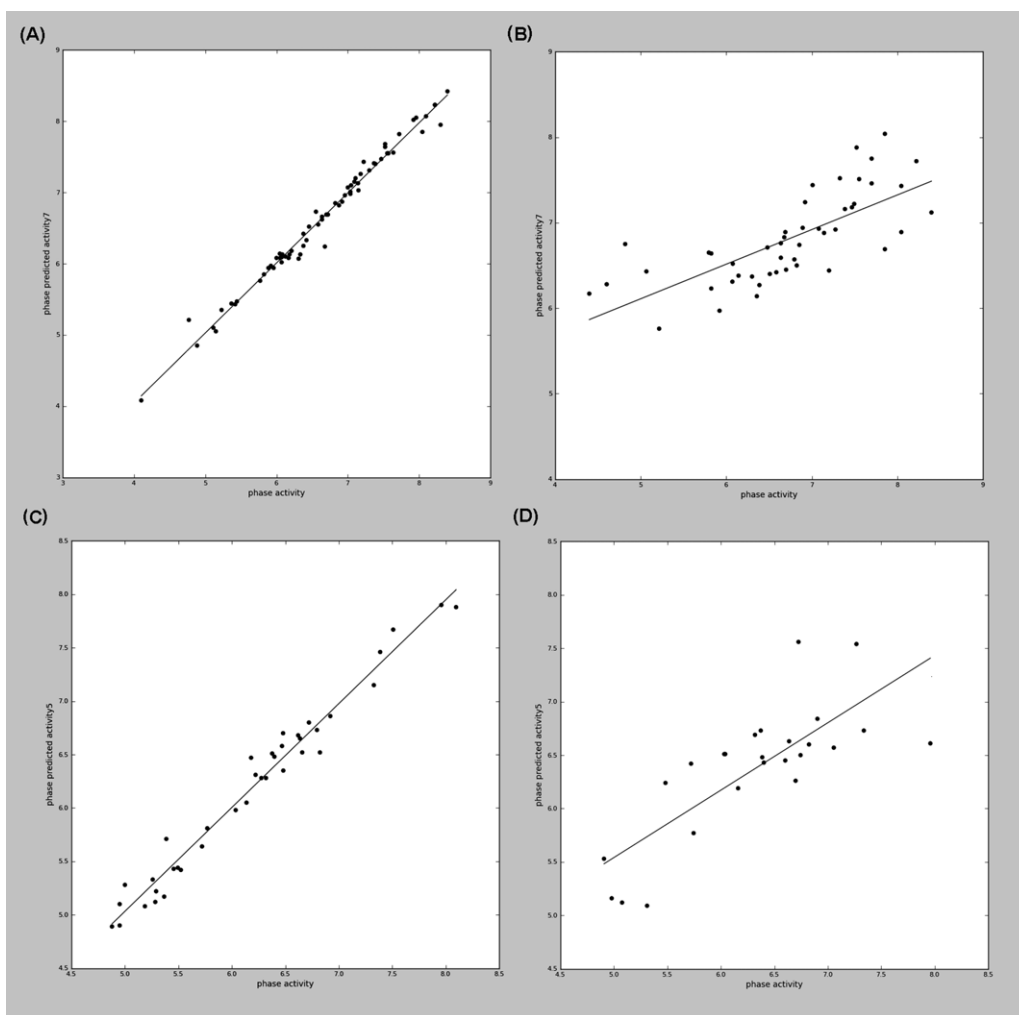
Fig. 2. Correlation graph between experimental and predicted activities of: training set molecules of Syk (A), test set molecules of Syk (B), training set molecules of ZAP-70 (C) and test set molecules of ZAP-70 (D).

all available crystal structures of Syk and ZAP-70 were downloaded from protein databank (PDB) and investigated for essential amino acids required for activity. The crystal structures 1XBB, 1XBC, 3EMG, 3FQE, 3FQH and 3FQS for Syk and 1U59 for ZAP-70 were investigated to check important amino acids for inhibitory activity [26–29]. The investigation of crystal structures revealed

that the H-bonding interactions with Ala451 and Asp512 are crucial for Syk inhibitory activity whereas the two interactions with Ala417 and Arg465 amino acid residues are important for ZAP-70 inhibitory activity. After the investigation of essential amino acid residues of both proteins, all 261 molecules were docked with protein Syk (PDB ID 1XBB) using Glide XP program and those
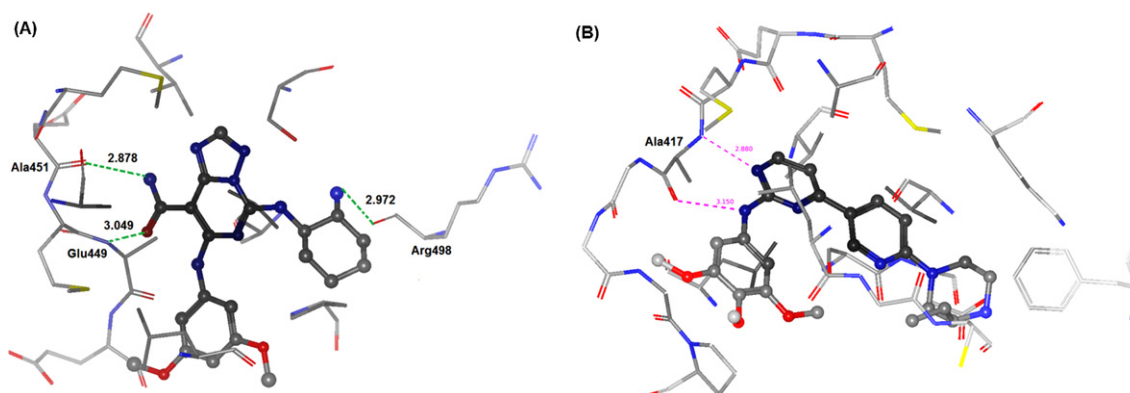
Fig. 3. Docking interactions of the highest active molecule Syk58 with crystal structure 1XBB of protein Syk (A) and of ZAP35 with crystal structure 1U59 of protein ZAP-70 (B).

Table 4
Applicability domain for the test set of Syk.

Compound	Distance (APD = 4.648)
Syk 1	2.449
Syk 3	3.464
Syk 5	2.646
Syk 6	3.162
Syk 10	2
Syk 11	3
Syk 12	2.236
Syk 13	2.828
Syk 16	2.828
Syk 19	1.414
Syk 20	3
Syk 23	2
Syk 27	2.449
Syk 28	1
Syk 29	2.646
Syk 35	3.464
Syk 36	3.606
Syk 42	2.236
Syk 44	2.646
Syk 47	2.646
Syk 52	2.236
Syk 54	2.828
Syk 56	2
Syk 57	3.873
Syk 61	2.828
Syk 66	4
Syk 68	1.732
Syk 69	2.236
Syk 95	2.336
Syk 74	2.336
Syk 82	2.336
Syk 84	2.336
Syk 85	2.449
Syk 88	2.646
Syk 90	4.472
Syk 91	4.358
Syk 92	3.464
Syk 93	2.236
Syk 94	2.646
Syk 96	3.742
Syk 97	2.645
Syk 98	3.162
Syk 100	2.236
Syk 104	2.646
Syk 106	1.732
Syk 109	1.732
Syk 112	1.732
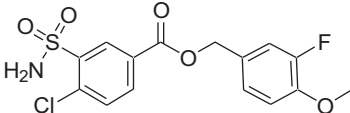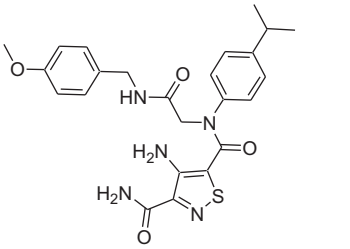
Table 5
Applicability domain for the test set of Zap-70.

Compound	Distance (APD = 4.078)
ZAP-70 4	2.236
ZAP-70 5	2.828
ZAP-70 6	2.828
ZAP-70 18	2.449
ZAP-70 22	2.236
ZAP-70 24	2
ZAP-70 25	3
ZAP-70 26	3.162
ZAP-70 27	1.732
ZAP-70 28	2.236
ZAP-70 31	2.449
ZAP-70 33	2.236
ZAP-70 34	2.236
ZAP-70 36	2.236
ZAP-70 42	2.646
ZAP-70 43	2.646
ZAP-70 45	1.414
ZAP-70 48	2
ZAP-70 50	2.236
ZAP-70 52	2.828
ZAP-70 53	2.449
ZAP-70 56	2.236
ZAP-70 58	2
ZAP-70 61	2.646
ZAP-70 62	3

molecules showing important interactions with Syk protein were further docked with protein ZAP-70 (PDB ID 1U59) [30,31]. Finally, 44 molecules having structural features of both pharmacophore models and displaying important interactions with Syk and ZAP-70 protein were obtained. The chemical structures, predicted Syk and ZAP-70 activities on the basis of corresponding best pharmacophore models and essential docking interactions of the two best molecules with Syk and ZAP-70 proteins are mentioned in Table 6. The information about remaining 42 molecules is given in Table S3 (Supplementary information).

Determination of ADME properties of the molecules is an essential guide for lead generation that can avoid the failure of candidates in later stages due to poor pharmacokinetic profile. The '*drug likeness*' of newly designed multi-targeted inhibitors was analyzed by calculating various ADME parameters using '*QikProp*' program of maestro. The calculated parameters then determine

Table 6
Predicted activities and docking interactions of best molecules after VS.

S. no.	Phase database ID	Structure	Predicted activity[a]	Predicted activity[b]	Docking interactions for Syk Ala451	Asp512	ZAP-70 Ala417	Arg465
1	CACDB_2007a_0466556		6.087	5.621	√	√	√	√
2	CACDB_2007a_1411951		6.818	6.076	√	√	√	√

[a] Activity predicted by ADHR.45-5.
[b] Activity predicted by AADRR.265-3.

Table 7
ADME properties of 44 newly designed multikinase inhibitor molecules using Qikprop module of Maestro 9.2.

S. no.	ID	QPlogP$_{o/w}$[a]	QPlogS[b]	QPlogHERG[c]	QPPCaco[d]	QPPMDCK[e]	Percent human oral absorbtion[f]
1	0466556	1.934	−4.143	−5.238	143.813	212.691	76.891
2	1411951	1.255	−3.503	−5.040	118.632	115.449	71.047
3	0468908	1.649	−4.266	−5.723	111.664	89.703	73.254
4	0493741	1.992	−4.286	−5.220	59.749	59.832	70.404
5	0547657	1.327	−3.693	−5.689	182.764	80.885	75.198
6	0592835	1.879	−4.670	−5.042	143.411	77.212	76.546
7	1407819	2.993	−3.627	−5.904	85.102	502.132	79.012
8	0450522	1.876	−3.951	−5.423	293.965	275.910	82.106
9	0016678	4.273	−5.768	−5.218	3355.600	4512.8	100.00
10	0047774	2.277	−5.122	−6.186	165.970	71.006	80.013
11	0162185	5.206	−7.347	−6.779	1010.800	6563.2	85.292
12	0211884	2.867	−4.947	−6.099	205.811	183.561	59.225
13	0295593	1.587	−3.873	−5.607	148.657	237.844	75.115
14	0480818	1.790	−4.326	−5.752	230.315	101.180	79.709
15	1066534	3.305	−4.315	−5.527	628.525	299.475	96.384
16	0609941	3.113	−5.613	−6.238	101.688	82.918	68.141
17	0623863	2.091	−4.202	−5.089	132.816	271.426	77.190
18	0701419	1.719	−3.951	−5.597	166.266	341.859	76.757
19	0883785	3.368	−4.380	−5.801	782.034	379.262	100.00
20	0883826	4.585	−5.758	−6.877	1304.000	659.141	100.00
21	0917587	3.653	−4.975	−5.058	512.830	1058.400	96.838
22	1231773	4.836	−5.415	−5.325	2907.500	4697.700	100.000
23	1343390	2.651	−4.823	−5.455	311.090	140.030	87.084
24	0151588	3.035	−6.484	−7.497	64.910	114.865	77.151
25	0383267	2.173	−4.557	−5.687	132.543	243.224	77.655
26	0535149	3.121	−3.868	−5.457	621.025	295.614	95.209
27	1066534	3.599	−4.518	−5.593	996.390	492.779	100.00

[a] Predicted octanol/water partition co-efficient log P (acceptable range: −2.0 to 6.5).
[b] Predicted aqueous solubility: S in mol/L (acceptable range: −6.5 to 0.5).
[c] Predicted IC$_{50}$ value for blockage of HERG K$^+$ channels (concern below −5.0).
[d] Predicted Caco-2 cell permeability in nm/s (acceptable range: <25 is poor and >500 is great).
[e] Predicted apparent MDCK cell permeability in nm/s.
[f] Percentage of human oral absorption (<25% is poor and >80% is high).

the absorption, distribution, metabolism and excretion (ADME) of molecules were partition coefficient (QPlogP$_{o/w}$), water solubility (QPlogS), cell permeability (QPPCaco) and percentage human oral absorption. The calculated ADME parameters of 27 out of 44 molecules were within the acceptance range *i.e.* QPlogP$_{o/w}$, 1.255–5.206; QPlogS, −3.503 to −7.347; QPlogHERG, −5.040 to −6.877; QPPCaco, 59.749–3355.6, QPPMDCK, 59.832–6563.200, percentage human oral absorption, 68.141 to 100%; thus these 27 molecules can theoretically be considered as orally active for human use. The calculated parameters of all 27 molecules are mentioned in Table 7.

3.5. Validation of virtual screening hits by pharmacophore and docking analysis

Two best hits CACDB_2007a_0466556 and CACDB_2007a_1411951 were obtained using the hierarchical virtual screening protocol. In molecule CACDB_2007a_0466556, the HBA and HBD features of ADHR.45-5 and AADRR.265-3 are mapped over the 'O' of sulfone and NH of ligand respectively (Fig. 4A and B). These two features complement the H-bonding interactions of this molecule with Syk (Fig. 5A) *i.e.* NH is displaying a H-bond with carbonyl 'O' of Ala451 (N···O, 2.108 Å) and 'O' of sulfone group with NH of Asp512 (O···N, 2.060 Å). This molecule is also showing three H-bonds with the active site amino acids of ZAP-70 (Fig. 5B) *i.e.* two between NH of ligand and carbonyl 'O' of Agr465 (N···O, 3.232 Å) and Asn466 (N···O, 2.879 Å), third bond between 'O' of ligand and NH of Ala417 (O···N, 3.153 Å) that are complementing the structural features of the pharmacophore model.

The mapping of second molecule CACDB_2007a_1411951 over ADHR.45-5 (Fig. 4C) exhibits that the HBA and HBD are mapped to the linker amide group and this HBD feature complements the docking analysis (Fig. 5C) *i.e.* H-bond is formed between NH of ligand and carbonyl 'O' of Ala451 (N···O, 2.141 Å). The molecule is also displaying two more interactions: one between NH of ligand and carbonyl 'O' Glu420 (N···O, 1.982 Å) and second between carbonyl 'O' of ligand and NH of Asp512 (O···N, 2.005 Å). The mapping of same molecule over AADRR.265-3 shows that two HBA groups are mapped over two carbonyl groups and the HBD group is mapped over the NH group of the ligand (Fig. 4D). These pharmacophoric features are not complemented to the docking results of the same molecule but the molecule is displaying all essential interactions with active site amino acid residues of ZAP-70 (Fig. 5D) *i.e.* one interaction between carbonyl 'O' of ligand and NH of Arg465 (O···N, 1.704 Å), second between NH of ligand and carbonyl 'O' of the Asp479 (N···O, 1.799 Å), third between NH of ligand and carbonyl 'O' of the Leu344 (N···O, 1.980 Å) and fourth between 'O' of methoxy group and NH of Ala417 (O···N, 2.336 Å).

3.6. Contour analysis

Additional insight into the inhibitory activity can be gained by visualizing the QSAR model in the context of most active and least active compounds. The contribution maps obtained from the hypotheses shows how 3D-QSAR methods can identify features important for the interaction between ligands and their target protein. Such maps allow identification of those positions that require a particular physicochemical property to enhance the bioactivity of a ligand. A pictorial representation of the contours
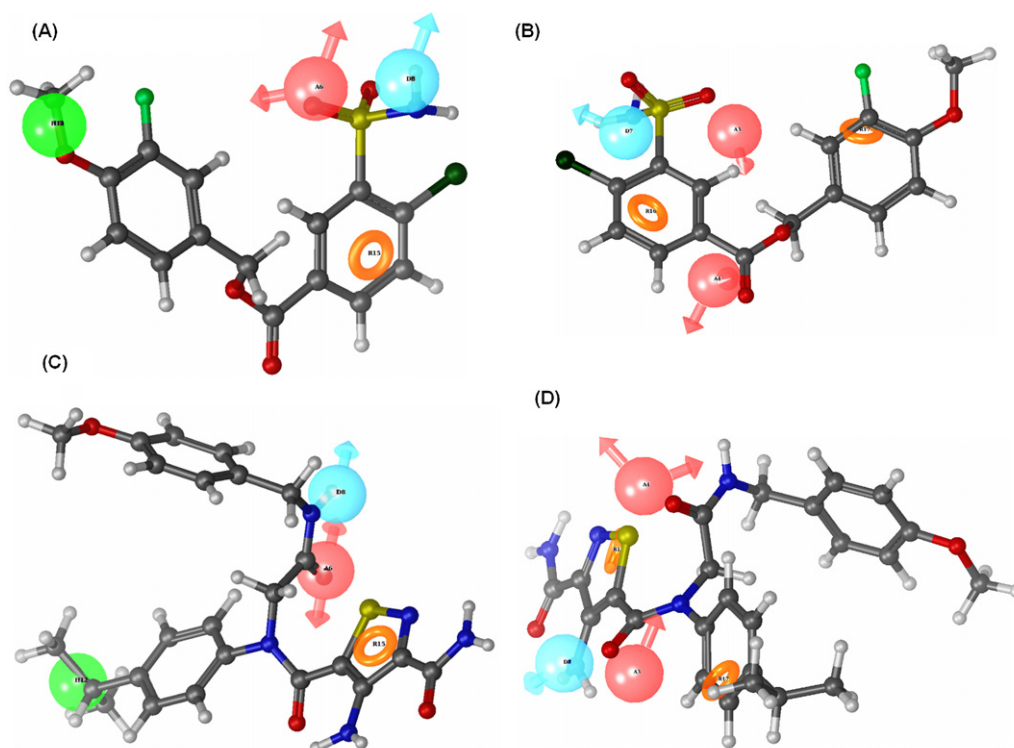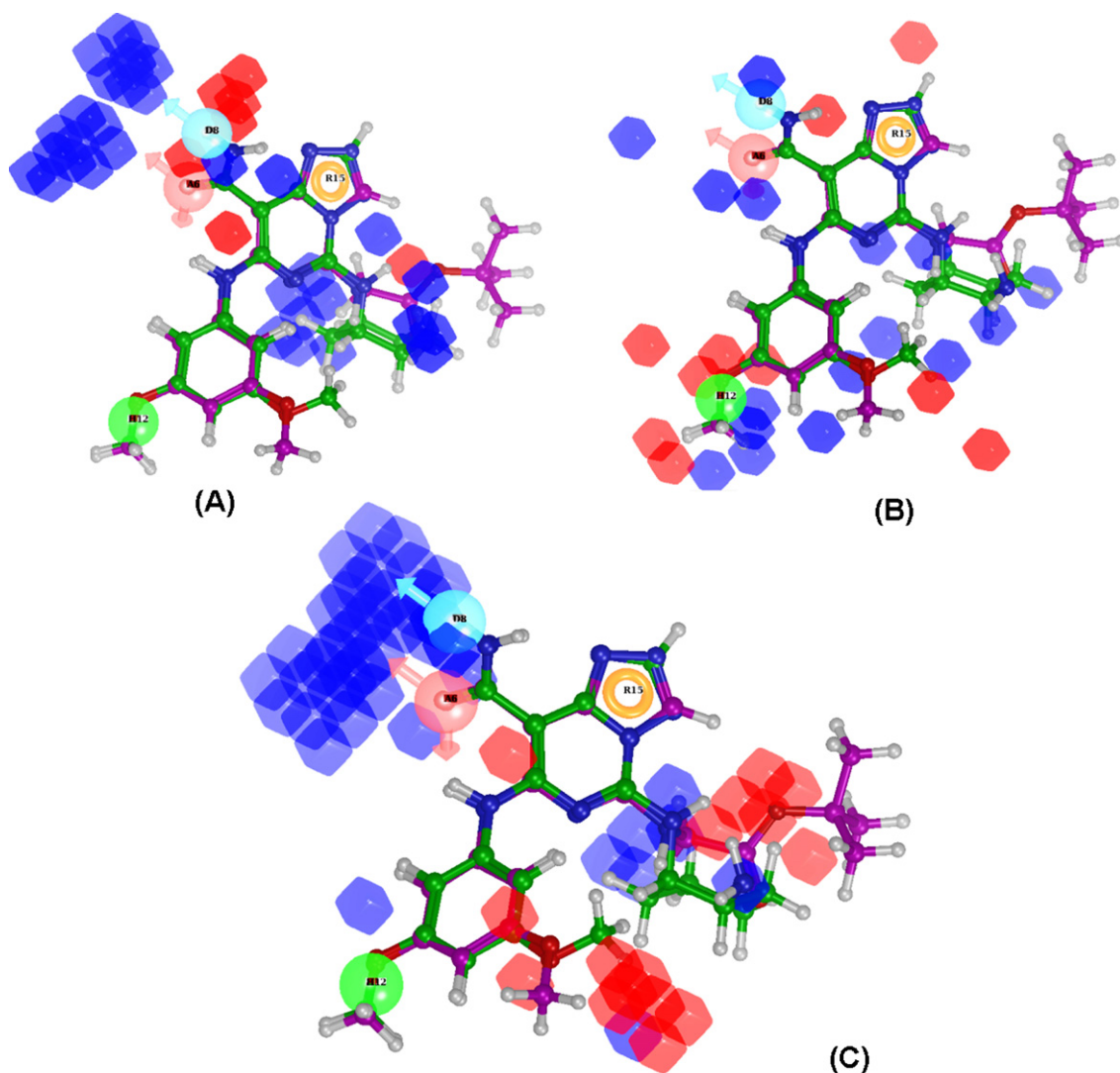
Fig. 4. Pharmacophore mapping of virtual screening retrieved molecules: CACDB_2007a_0466556 over ADHR.45 (A) and AADRR.265 (B); mapping of CACDB_2007a_1411951 over ADHR.45 (C) and AADRR.265 (D).
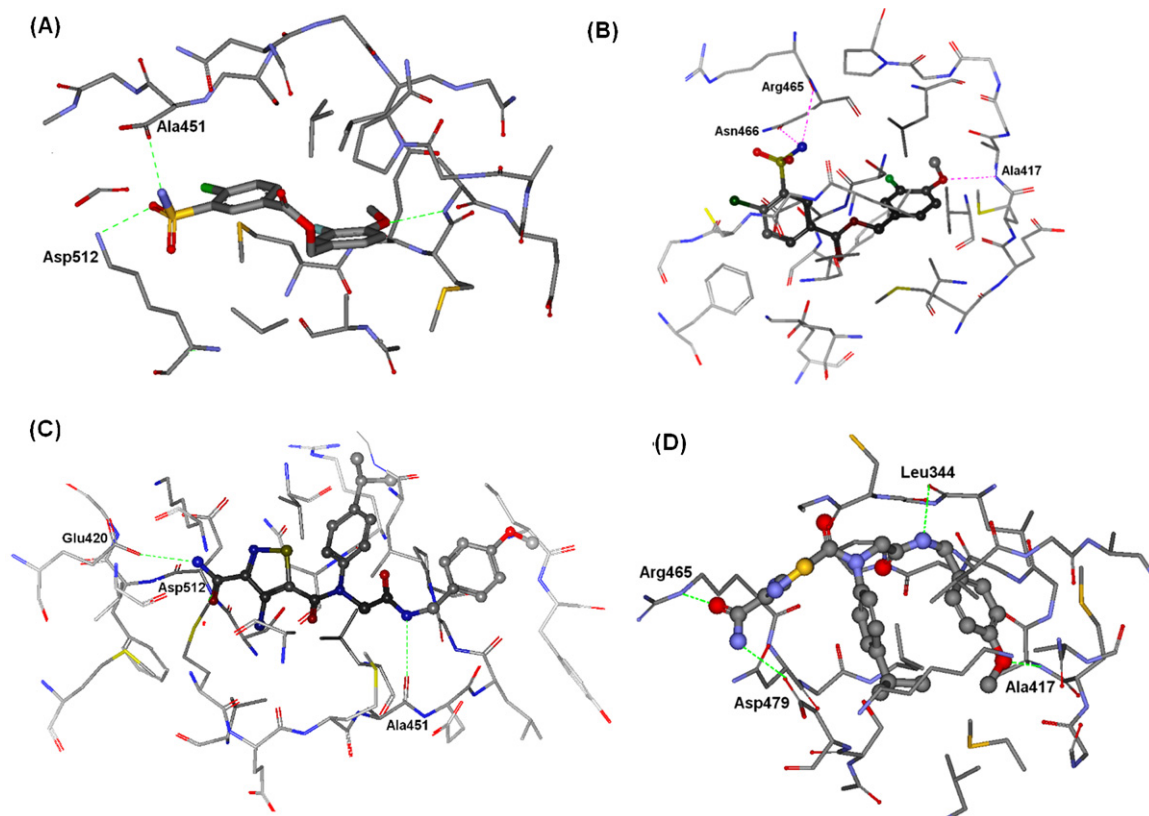
Fig. 5. Docking interactions of virtual screening retrieved molecules: CACDB_2007a_0466556 with Syk (A) and ZAP-70 (B); docking of CACDB_2007a_1411951 with Syk (C) and ZAP-70 (D).

Fig. 6. Contour maps for 3D-QSAR models generated against Syk: hydrogen bond donor (A), hydrophobic (B) and electron withdrawing property (C). The highest active molecule Syk58 (green) and the least active molecule Syk86 (purple) are displayed as reference. Blue contours correspond to the regions where the substitution of corresponding property group enhances the activity and *vice versa* for the red contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

generated for Syk and ZAP-70 is shown in Fig. 6A–C and 7A–D respectively. In these representations, the highest active and least active molecules for Syk and ZAP-70 *i.e.* Syk58, Syk86, ZAP35 and ZAP19 are shown as reference and the blue cubes indicate favorable regions while red cubes indicate unfavorable regions for activity.

3.6.1. Contours for Syk model

In Fig. 6A for HBD contour, the most active molecule displays blue contours near NH_2 group substituted at cyclohexyl ring whereas in the least active molecule this cyclohexyl ring was absent. Fig. 6B displays contour map for hydrophobic property and the blue contours are present near cyclohexyl ring in the highest active molecule whereas in the least active molecule this ring is absent. This clearly indicates the importance of this cyclohexyl ring for the activity. Fig. 6C displaying contour map for electron withdrawing property describes that

the favorable blue contours are present near to the carbonyl group present in both active and inactive molecules indicating the importance of this carbonyl group for kinase activity of all molecule.

3.6.2. Contours for ZAP-70 model

In Fig. 7A for HBD property, the blue contour is present near NH of piperidine ring whereas the ring is missing in the least active molecule. Fig. 7B displays the hydrophobic property contour map and the blue contours are present near to piperidine that is missing in the least active molecule. The positive ionizable property contour map displayed in Fig. 7C exhibits the blue contour near NH of piperidine ring. In Fig. 7D for electron withdrawing property contours, the favorable blue contours are displayed near to N atom of the pyrimidine ring whereas the ring is absent in the least active molecule.
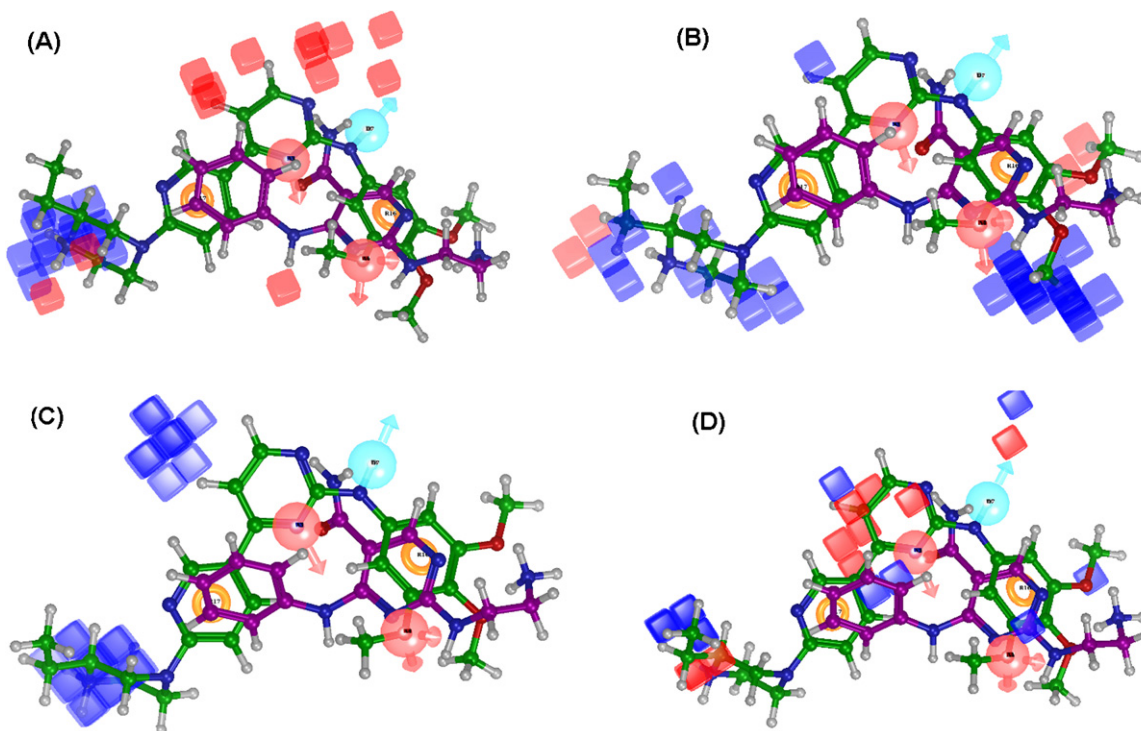
Fig. 7. Contour maps for 3D-QSAR models generated against ZAP-70: hydrogen bond donor (A), hydrophobic (B), positive ionizable (C) and electron withdrawing property (D). The highest active molecule ZAP35 (green) and the least active molecule ZAP19 (purple) are displayed as reference. Blue contours correspond to the regions where the substitution of corresponding property group enhances the activity and *vice versa* for the red contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

4. Conclusions

Two statistically stable and reliable pharmacophore models were generated for diverse Syk and ZAP-70 inhibitors, subsequently the 3D-QSAR analysis was performed for both datasets that helped to identify the important structural features of molecules required for corresponding biological activity. The generated pharmacophore models coupled with docking simulations were subjected to sequential multi-step systematic virtual screening and total 27 dual inhibitors were obtained as hits. These compounds were then analyzed by Qikprop to calculate their different ADME parameters.

The success of combined pharmacophore, hierarchical virtual screening and QSAR strategy to identify 27 dual inhibitors lends firm support for the identification of new lead in the multi-targeted drug discovery program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2012.11.011.

References

[1] J. Schlessinger, A. Ullrich, Growth factor signaling by receptor tyrosine kinases, Neuron 9 (1992) 383–391.

[2] P. Van der Geer, T. Hunter, R.A. Lindberg, Receptor protein-tyrosine kinases and their signal transduction pathways, Annual Review of Cell Biology 10 (1994) 251–337.

[3] A. Levitzki, A. Gazit, Tyrosine kinase inhibition: an approach to drug development, Science 267 (1995) 1782–1788.

[4] T. Taniguchi, T. Kobayashi, J. Kondo, K. Takahashi, H. Nakamura, J. Suzuki, K. Nagai, T. Yamada, S. Nakamura, H. Yamamura, Molecular cloning of a porcine gene syk that encodes a 72-kDa protein-tyrosine kinase showing high susceptibility to proteolysis, Journal of Biological Chemistry 266 (1991) 15790–15796.

[5] A.C. Chan, M. Iwashima, C.W. Turck, A. Weiss, ZAP-70: a 70 kDa protein-tyrosine kinase that associates with the TCR zeta chain, Cell 71 (1992) 649–662.

[6] B.R. Wong, E.B. Grossbard, D.G. Payan, E.S. Masuda, Targeting Syk as a treatment for allergic and autoimmune disorders, Expert Opinion on Investigational Drugs 13 (2004) 743–762.

[7] R. Singh, E.S. Masuda, Spleen tyrosine kinase (Syk) biology, inhibitors and therapeutic applications, Annual Reports in Medicinal Chemistry 42 (2007) 379–391.

[8] S. Faruki, R.L. Geahlen, D.J. Asai, Syk-dependent phosphorylation of microtubules in activated B-lymphocytes, Journal of Cell Science 113 (2000) 2557–2565.

[9] A.K. Saxena, K.K. Roy, Hierarchical virtual screening: identification of potential high affinity and selective β3-adrenergic receptor agonists, SAR and QSAR in Environmental Research (2012), http://dx.doi.org/10.1080/1062936X.2012.664824.

[10] C.L. Cywin, B.P. Zhao, D.W. McNeil, M. Hrapchak, A.S. Prokopowicz, D.R. Goldberg, T.M. Morwick, A. Gao, S. Jakes, M. Kashem, R.L. Magolda, R.M. Soll, M.R. Player, M.A. Bobko, J. Rinker, R.L. DesJarlaisb, M.P. Winters, Discovery and SAR of novel [1,6]naphthyridines as potent inhibitors of spleen tyrosine kinase (SYK), Bioorganic and Medicinal Chemistry Letters 13 (2003) 1415–1418.

[11] A. Hirabayashi, H. Mukaiyama, H. Kobayashi, H. Shiohara, S. Nakayama, M. Ozawa, E. Tsuji, K. Miyazawa, K. Misawa, H. Ohnota, M. Isaji, Structure activity relationship studies of imidazo[1,2-c]pyrimidine derivatives as potent and orally effective Syk family kinases inhibitors, Bioorganic and Medicinal Chemistry 16 (2008) 9247–9260.

[12] H. Hisamichi, R. Naito, A. Toyoshima, N. Kawano, A. Ichikawa, A. Orita, M. Orita, N. Hamada, M. Takeuchi, M. Ohtaa, S. Tsukamotoa, Synthetic studies on novel Syk inhibitors. Part 1: synthesis and structure activity relationships of pyrimidine-5-carboxamide derivatives, Bioorganic and Medicinal Chemistry 13 (2005) 4936–4951.

[13] A. Hirabayashi, H. Mukaiyama, H. Kobayashi, H. Shiohara, S. Nakayama, M. Ozawa, K. Miyazawa, K. Misawa, H. Ohnota, M. Isaji, A novel Syk

family kinase inhibitor: design, synthesis, and structure activity relationship of 1,2,4-triazolo[4,3-c]pyrimidine and 1,2,4-triazolo[1,5-c]pyrimidine derivatives, Bioorganic and Medicinal Chemistry 16 (2008) 7347–7357.

[14] J.Y.Q. Lai, P.J. Cox, R. Patel, S. Sadiq, D.J. Aldous, S. Thurairatnam, K. Smith, D. Wheeler, S. Jagpal, S. Parveen, G. Fenton, T.K.P. Harrison, C. McCarthy, P. Bamborough, Potent small molecule inhibitors of spleen tyrosine kinase (Syk), Bioorganic and Medicinal Chemistry Letters 13 (2003) 3111–3114.

[15] D. Moffat, P. Davis, M. Hutchings, J. Davis, M. Berg, M. Batchelor, J. Johnson, J. O'Connell, R. Martin, T. Crabbe, J. Delgado, M. Perry, 4-Pyridin-5-yl-2-(3,4,5-trimethoxyphenylamino)pyrimidines: potent and selective inhibitors of ZAP-70, Bioorganic and Medicinal Chemistry Letters 9 (1999) 3351–3356.

[16] A. Hirabayashi, H. Mukaiyama, H. Kobayashi, H. Shiohara, S. Nakayama, M. Ozawa, K. Miyazawa, K. Misawa, H. Ohnota, M. Isaji, Structure activity relationship studies of 5-benzylaminoimidazo[1, 2-c] pyrimidine-8-carboxamide derivatives as potent, highly selective ZAP-70 kinase inhibitors, Bioorganic and Medicinal Chemistry 17 (2009) 284–294.

[17] Maestro, Version 9.2, User Manual, Schrödinger, LLC, New York, 2011.

[18] Ligprep, Version 2.5, User Manual, Schrödinger, LLC, New York, 2011.

[19] PHASE, Version 3.3, Schrödinger, LLC, New York, NY, 2011.

[20] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception. 3D QSAR model development, and 3D database screening. 1. Methodology and preliminary results, Journal of Computer-Aided Molecular Design 20 (2006) 647–671.

[21] J. Verma, V.M. Khedkar, E.C. Coutinho, 3D-QSAR in drug design: a review, Current Topics in Medicinal Chemistry 10 (2010) 95–115.

[22] A. Golbraikh, A. Tropsha, Beware of q2!, Journal of Molecular Graphics and Modelling 20 (2002) 269–276.

[23] V.D. Mouchlis, G. Melagraki, T. Mavromoustakos, G. Kollias, A. Afantitis, Molecular modeling on pyrimidine-urea inhibitors of TNF-α production: an integrated approach using a combination of molecular docking, classification techniques, and 3D-QSAR CoMSIA, Journal of Chemical Information and Modeling 52 (2012) 711–723.

[24] S. Zhang, A. Golbraikh, S. Oloff, H. Kohn, A. Tropsha, A novel automated lazy learning QSAR (ALL-QSAR) approach: method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models, Journal of Chemical Information and Modeling 46 (2006) 1984–1995.

[25] R.D. Cramer, J.D. Bunce, D.E. Patterson, I.E. Frank, Crossvalidation, bootstrapping, and partial least squares compared with multiple regression in conventional QSAR studies, Quantitative Structure–Activity Relationship 7 (1988) 18–25.

[26] S. Atwell, J.M. Adams, J. Badger, M.D. Buchanan, I.K. Feil, K.J. Froning, X. Gao, J. Hendle, K. Keegan, B.C. Leon, H.J. Muller-Deickmann, V.L. Nienaber, B.W. Noland, K. Post, K.R. Rajashankar, A. Ramos, M. Russell, S.K. Burley, S.G. Buchanan, A novel mode of Gleevec binding is revealed by the structure of spleen tyrosine kinase, Journal of Biological Chemistry 279 (2004) 55827–55832.

[27] L.J. Farmer, G. Bemis, S.D. Britt, J. Cochran, M. Connors, E.M. Harrington, T. Hoock, W. Markland, S. Nanthakumar, P. Taslimi, E.T. Haar, J. Wang, D. Zhaveri, F.G. Salituro, Discovery and SAR of novel 4-thiazolyl-2-phenylaminopyrimidines as potent inhibitors of spleen tyrosine kinase (SYK), Bioorganic and Medicinal Chemistry Letters 18 (2008) 6231–6235.

[28] A.G. Villasenor, R. Kondru, H. Ho, S. Wang, E. Papp, D. Shaw, J.W. Barnett, M.F. Browner, A. Kuglstatter, Structural insights for design of potent spleen tyrosine kinase inhibitors from crystallographic analysis of three inhibitor complexes, Chemical Biology and Drug Design 73 (2009) 466–470.

[29] L. Jin, S. Pluskey, E.C. Petrella, S.M. Cantin, J.C. Gorga, M.J. Rynkiewicz, P. Pandey, J.E. Strickler, R.E. Babine, D.T. Weaver, K.J. Seidl, The three-dimensional structure of the ZAP-70 kinase domain in complex with staurosporine: implications for the design of selective inhibitors, Journal of Biological Chemistry 279 (2004) 42818–42825.

[30] Glide, Version 5.6, User Manual, Schrödinger, LLC, New York, NY, 2010.

[31] R.A. Friesner, J. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, D.E. Shaw, M. Shelley, J.K. Perry, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, Journal of Medicinal Chemistry 47 (2004) 1739–1749.