Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Molecular Graphics and Modelling 23 (2004) 129–138

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Modeling of p38 mitogen-activated protein kinase inhibitors using the Catalyst^{TM} HypoGen and *k*-nearest neighbor QSAR methods

Zhiyan Xiao [a,1,2], Shikha Varma [b,2], Yun-De Xiao [a,3], Alexander Tropsha [a,*]

[a] *The Laboratory for Molecular Modeling, School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA*
[b] *Accelrys Inc., 9685 Scranton Rd, San Diego, CA 92121-3752, USA*

Received 20 October 2003; received in revised form 1 May 2004; accepted 1 May 2004

Available online 2 June 2004

Abstract

We have employed in parallel the Catalyst HypoGen pharmacophore modeling approach and the variable selection *k*-nearest neighbor quantitative structure–activity relationship (*k*NN QSAR) method to model a diverse data set of p38 mitogen-activated protein (MAP) kinase inhibitors. The HypoGen pharmacophore model, developed from a novel automated training set selection protocol, identified chemical functional features that were characteristic of the active compounds and differentiated the active from the inactive inhibitors. The *k*NN QSAR modeling employed topological descriptors and afforded predictive QSAR models with consistently high values of both leave-one-out cross-validated R^2 for the training set and predictive R^2 for the test set. The results of both modeling approaches were sensitive to the selection of the training and test sets used for model development and validation. The resulting Catalyst pharmacophore and *k*NN QSAR models can be used concurrently for rapid virtual screening of chemical databases to identify novel p38 MAP kinase inhibitors.

Keywords: p38 MAP kinase inhibitors; Quantitative structure–activity relationships; HypoGen; *k*NN QSAR; Rational selection of training and test sets

1. Introduction

The p38 mitogen-activated protein (MAP) kinase is a member of the MAP kinase family. In response to a variety of stress stimuli (heat, UV light, lipopolysaccharide, high osmolarity), p38 MAP kinase is activated via dual phosphorylation of the TGY motif in the activation loop of the enzyme [1]. Upon activation, p38 MAP kinase phosphorylates a number of downstream substrates, thereby regulating the synthesis of several important pro-inflammatory cytokines [2] such as tumor necrosis factor-α (TNF-α) and interleukin-1 (IL-1). Therefore, the inhibition of p38 MAP kinase would potentially prevent the underlying pathophysiology in the inflammatory diseases, which makes p38 MAP kinase an attractive target for drug discovery [3,4].

The pyridinylimidazole compounds, exemplified by SB203580 [5] (Fig. 1), were originally developed as inflammatory cytokine synthesis inhibitors [6], and were later identified as selective inhibitors of p38 MAP kinase [7]. Thereafter, structurally diverse p38 MAP kinase inhibitors have been synthesized and tested extensively to seek potential therapy for inflammatory diseases resulting from excess cytokine production [8–11].As the number of synthetic inhibitors of p38 MAP kinase increases, it becomes important to elucidate the structure–activity relationships (SAR) of these diverse compounds. A pharmacophore model derived from several series of tri- and tetrasubstituted imidazole p38 MAP kinase inhibitors was previously described [7]. Briefly, the central imidazole acted as a scaffold to optimally present the substitutions at positions 1, 2, 4, and 5 (Fig. 1). The following SAR trends for this class of compounds have been established: (1) bulky groups can be well tolerated at N_1, and lipophilic substitutions at N_1 enhance p38 MAP kinase binding; (2) polar groups at the para position of the C_2 phenyl ring lead to improved p38 MAP kinase binding; (3) the nitrogen with lone pair at 3 position is essential for p38 MAP kinase binding; (4) aromatic ring at C_4 is required, which prefers lipophilic substituents and tolerates bulky groups better at meta than para position;

[*] Corresponding author. Tel.: +1 919 966 2955.
E-mail address: alex_tropsha@unc.edu (A. Tropsha).
[1] Present address: Department of Synthetic Medicinal Chemistry, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China.
[2] These authors have contributed equally to the work presented in this paper.
[3] Present address: Targacept Inc., 200 East First Street, Suite 300, Winston-Salem, NC 27101-4165.

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.05.001

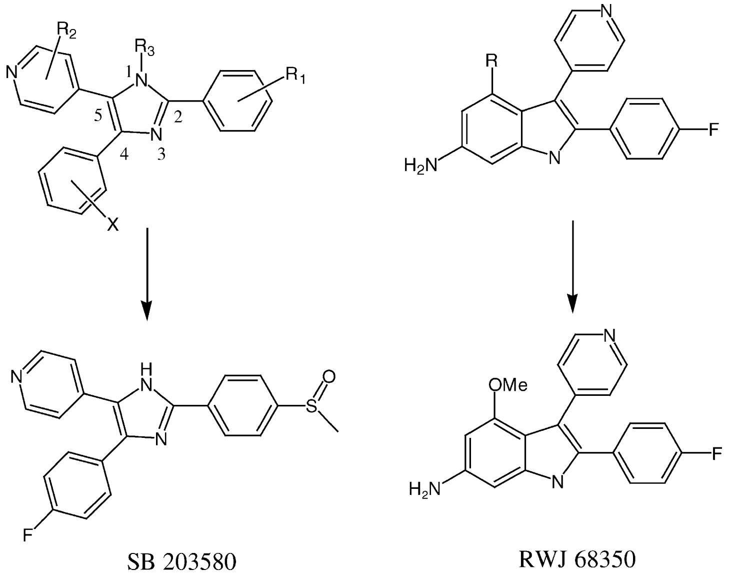


Fig. 1. Two sub-structural classes and corresponding examples of p38 MAP kinase inhibitors.

(5) the pyridine ring at C_5 is crucial for p38 MAP kinase binding, and the 4-pyridinyl nitrogen is required.

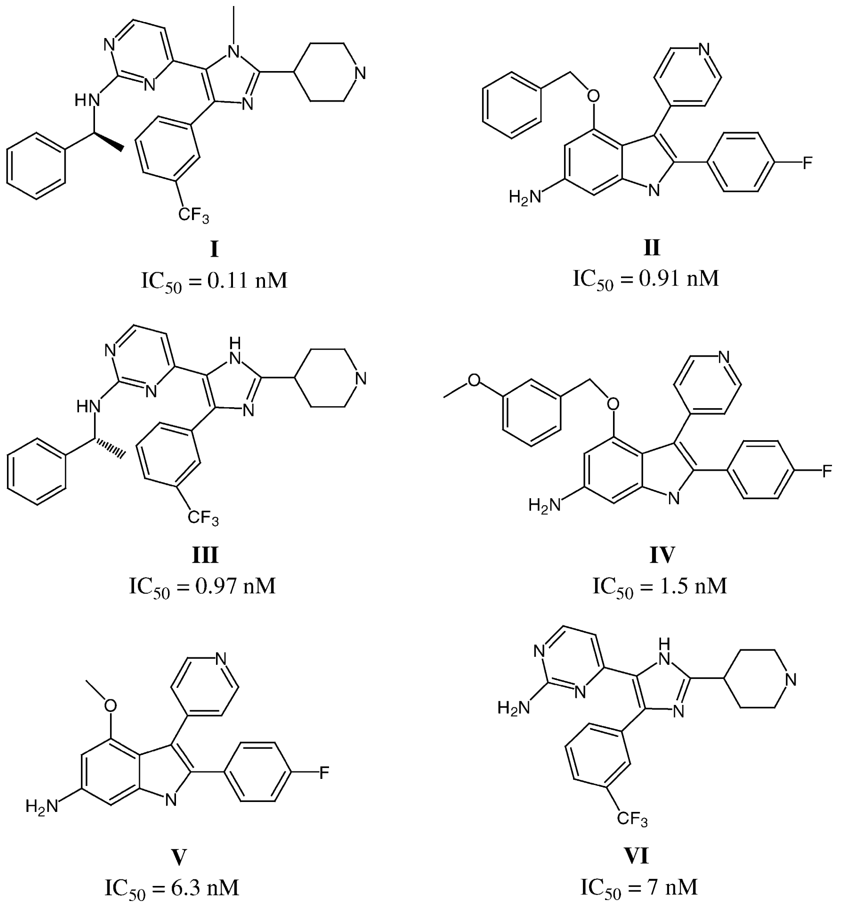
The SAR trends discussed above for the pyridinylimidazole compounds are consistent with observations derived from the analysis of several X-ray characterized structures of pyridinylimidazole inhibitor-p38 MAP kinase complexes [12–14]. Three important features are observed in all these structures: (1) a hydrophobic aryl binding pocket behind and orthogonal to the binding site normally occupied by the adenine ring of ATP, which accommodates the C_4 aromatic ring, (2) a hydrogen bond between the 4-pyridinyl nitrogen and the amide N-H of Met^{109}, and (3) a hydrogen bond between Lys^{53} and the imidazole N_1.

The Catalyst HypoGen approach [15] was recently applied in combination with a novel training set selection method to generate a pharmacophore model for 131 diverse p38 MAP kinase inhibitors belonging to the pyridinylimidazole or 3-(4-pyridyl)-1*H*-pyrrolo[2,3-*b*]pyridine (e.g. RWJ68350) classes (Fig. 1) [16]. The robustness of the model was highly dependent on the selection of the training set molecules used for the model generation. Rational selection of training and test sets resulted in pharmacophore models that were validated by both internal (for training set molecules) and external (for test set molecules) correlation coefficients with the highest values of 0.76 and 0.74, respectively [16]. The hypotheses were also readily mapped with both the active compounds and the X-ray crystal structure of p38 MAP kinase in a chemically meaningful way compatible with the known SAR.

The Catalyst HypoGen pharmacophore models are helpful in obtaining the molecular alignments of active conformers, which is a required first step for most three-dimensional (3D) QSAR approaches. In addition, these models serve as generalized queries for identifying novel, potentially active compounds through rapid virtual screening of chemical databases. In this paper, we have extended the earlier study on p38 MAP kinase inhibitors [16] to further refine and validate the previous HypoGen pharmacophore models and facilitate future design of novel p38 MAP kinase inhibitors or focused libraries based on the pyridinylimidazole or 3-(4-pyridyl)-1*H*-pyrrolo[2,3-*b*]pyridine skeletons. Concurrently, we have applied the *k*-nearest neighbor quantitative structure–activity relationship (*k*NN QSAR) [17] method to the same data set to formulate validated and predictive QSAR models using alignment free molecular operational environment (MOE) descriptors [18] as well as molecular topological indices [19]. Different training and test set selection techniques were exploited to evaluate the merits and limitations of the two QSAR approaches. Both Catalyst HypoGen and *k*NN QSAR methods afforded statistically significant training set models. However, our results underscore the significance of the rational selection of training and test sets in order to obtain QSAR models with high external predictive power.

2. The dataset

The chemical structures and their biological activities were taken from published data [8–11,20,21]. The activities of 131 available p38 MAP kinase inhibitors were expressed as IC_{50} values for the inhibition of p38 MAP kinase ranging from 0.11 to 114,000 nM. These compounds can be catalogued into two sub-structural classes: pyridinylimidazoles (e.g. SB203580, Fig. 1) and 3-(4-pyridyl)-1*H*-pyrrolo[2,3-*b*]pyridines (e.g. RWJ 68354, Fig. 1). The structures of the six most active compounds are shown in Fig. 2.



I
IC_{50} = 0.11 nM

II
IC_{50} = 0.91 nM

III
IC_{50} = 0.97 nM

IV
IC_{50} = 1.5 nM

V
IC_{50} = 6.3 nM

VI
IC_{50} = 7 nM

Fig. 2. Most active p38 MAP kinase inhibitors.

3. Computational methods

3.1. Catalyst HypoGen modeling

Structures and conformers were generated and Catalyst HypoGen calculations were performed with the Catalystv 4.6 software [15]. The default Catalyst settings were used except as otherwise noted. All calculations were performed on a Silicon Graphics Octane workstation.

The algorithm underlying the HypoGen option in Catalyst was described previously [22]. Briefly, HypoGen attempts to construct the simplest pharmacophore hypothesis that best explains the activities. The goal is to select a subset of the pharmacophoric elements that are present in the highly active compounds, but missing from the least active ones in the training set.

The predictive models are generated in three steps: constructive, subtractive, and optimization phases. The constructive phase identifies hypotheses that are common to the compounds in the most active set. In contrast, the subtractive phase recognizes the pharmacophoric features shared by both the most and the least active sets of molecules and removes these features from the putative pharmacophoric configurations developed in the constructive phase. In the optimization phase, small perturbations are applied to those hypotheses that survived the subtractive phase and the hypotheses are scored based on errors in the activity prediction from the regression and the complexity of the hypothesis. Upon completion of this phase, HypoGen reports 10 unique pharmacophores with the best scores.

The Catalyst program then fits each compound to a hypothesis and reports back a series of ‘Fit’ scores. The ‘Fit’ scores are calculated with Eq. (1):

$$\text{Fit} = \sum_{\text{mapped hypofunctions}} w \left[1 - \sum_{\text{spheres}} \left(\frac{\text{Disp}}{\text{Tol}} \right)^2 \right] \qquad (1)$$

where w is the weight of the hypothesis function; Disp the distance that separates the function feature in the molecule from the centroid of the hypothesis function; Tol is the tolerance of the hypothesis function, which determines the radius of the function sphere.

A relationship between log (activities) and the corresponding Fit-values for all training set molecules is computed using linear regression after mapping of each molecule to the hypothesis. The equation derived from the linear regression is then used to calculate log (activities) values for both training and test set molecules, and the internal correlation coefficient q^2 for training sets (or external R^2 for test sets) are determined with Eq. (2):

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \qquad (2)$$

where y_i and $\hat{y}_i$ are the actual and predicted activities of the ith compound, respectively; and $\bar{y}$ is the average activity of all compounds in the training set (or test set for R^2).

4. *k*NN QSAR algorithm

The variable selection *k*NN QSAR method [23] was described previously [17]. The original method was recently enhanced using weighted molecular similarity to give a higher weight to the neighbor in the training set with the higher similarity to the compound in the test set (for $k > 1$) [24].

4.1. Generation of molecular descriptors

All chemical structures were generated using Catalyst software [15]. Molecular topological indices [25,26] were obtained with the MolConnZ program (MCI descriptors) [19], and molecular operating environment (MOE) descriptors were generated with the QuaSAR-descriptor option implemented in the MOE molecular modeling software [18].

4.2. Selection of training and test sets and model validation

4.2.1. HypoGen modeling

The HypoGen module within Catalyst requires both biological (e.g. activity) and structural data of the training set molecules to develop predictive pharmacophore models. Therefore, while selecting a training set for developing a pharmacophore model in Catalyst, it is critical to consider both biological and structural diversity, thus covering sufficient activity and structural spaces. Generally, to develop a meaningful HypoGen model, SAR data of the training set are first inspected to avoid biological and structural redundancy. This process of manually selecting training set compounds can often be a daunting task, especially for a large amount of SAR data resulting from a homologous series. For this purpose, an automated method for selecting training set for pharmacophore modeling was developed and HypoGen models created to test if indeed this automated training set selection protocol can be used for making predictive pharmacophore models in Catalyst [16]. Functionalities within both Cerius^2 [27] and Catalyst were used to develop this training set selection protocol. This protocol involves selection of the most diverse set of compounds on the basis of their (a) multi-conformer 3D pharmacophoric fingerprints, (b) multi-conformer shape indices, and (c) activity.

In order to calculate the 3D fingerprints, we first created a feature file containing the coordinates of all features present in each conformer. This was generated from a multi-conformer Catalyst .bdb file using the Cerius^2 3D Keys [27]. The features (e.g. HBA, HBD, RING AROMATIC, HYDROPHOBIC, NEG CHARGE, NEG IONIZABLE, POS CHARGE, and POS IONIZABLE) are predefined by Catalyst and are surface accessible. The shape descriptors consist of volume descriptors and x, y, and z components of principal axes. The highly correlated and high-dimensional fingerprint data were then subjected to multidimensional scaling [28]. MDS coordinates, together with shape descriptors and activity data, were then analyzed by principal component analysis (PCA). Finally, Cerius^2 diversity tools were used to select the most diverse compounds for the training set (Fig. 3). It should be noted that this is just one suggested protocol for an automated training set selection and it can be modified to perform diversity sampling directly on 3D fingerprints or any other meaningful variables.

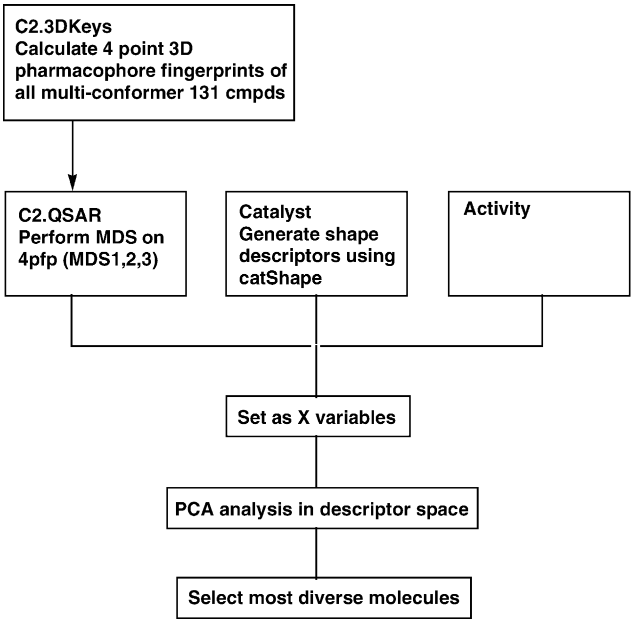

C2.3DKeys
Calculate 4 point 3D pharmacophore fingerprints of all multi-conformer 131 cmpds
C2.QSAR
Perform MDS on 4pfp (MDS1,2,3)
Catalyst
Generate shape descriptors using catShape
Activity
Set as X variables
PCA analysis in descriptor space
Select most diverse molecules

Fig. 3. The flow-chart of the automated training set selection protocol for Catalyst HypoGen model 1.

4.2.2. kNN QSAR modeling

An algorithm similar to the stochastic cluster analysis (SCA) method developed earlier by Reynolds et al. [29] was used for diversity sampling in both MCI and MOE descriptor spaces to select the training and test sets. The volume of the original descriptor space is normalized to 1, and then the volume corresponding to an individual point representing a compound is defined as $1/N$, where N is the number of representative points in the descriptor space. A random compound is selected and included into the training set. A sphere is constructed with its center at the representative point of this compound and with radius $R = c(1/N)^{1/K\cdot}$. Here, K denotes the number of descriptors, and c refers to the dissimilarity level. Compounds corresponding to representative points within this sphere, other than the center, are included in the test set and all the points within this sphere are then excluded from the initial set. The point with smallest distance from the first representative point is selected as the second center and the above procedure is repeated until all the compounds are assigned to either training or test set. This algorithm allows construction of training sets that cover the whole descriptor space occupied by representative points. Additional details of this algorithm are reported elsewhere [30,31]. As anticipated, the distribution of compounds between training and test sets is sensitive to the types of descriptors used in the calculations.

Multiple training and test sets were selected from the original dataset and models were developed for different training sets. Models were considered acceptable if their leave-one-out cross-validated R^2 (q^2) values (for *k*NN QSAR) or internal correlation coefficients (for Catalyst HypoGen) for the training set exceeded 0.4 and the corresponding predictive R^2 for the test set exceeded 0.6.

5. Results and discussion

5.1. Training and test set selection and Catalyst HypoGen pharmacophore modeling

According to the HypoGen approach [22], Catalyst QSAR pharmacophore modeling requires (i) training sets that cover a wide range of activities to have effective subtractive phases and differentiate the most active from the least active compounds, and (ii) selection of active compounds as diverse as possible to maximize the number of pharmacophore hypotheses. It is unnecessary to simultaneously include several very similar compounds in the training sets for the hypothesis generation, since it may only provide redundant information and bias the hypotheses toward similar structures. Therefore, a novel training set selection methodology described above was applied prior to the construction of HypoGen pharmacophore models to select 25 most diverse compounds as the training set from the original dataset of 131 compounds. Nineteen compounds were then selected as the test set using the same procedure.

We have also examined the possible influence of different types of descriptors on the selection of training and test sets and the performance of both Catalyst and *k*NN QSAR modeling. To this end, the diversity sampling was performed in multidimensional descriptor space including both the activity values and MCI or MOE descriptors. The original dataset was first divided into three subsets based on the activity data, with 13 compounds in the most active group, 5 compounds in the least active group and the remaining 113 compounds in the moderately active group; a small set (about 20) of most diverse compounds was additionally selected as representative subset of the latter group. Each of these three subgroups were further divided into training and test sets, and the final training and test sets were generated by combining the three selected training and test sets together.

In summary, multiple training and test sets were selected from the original dataset in the activity and MCI or MOE descriptor spaces and Catalyst HypoGen pharmacophore models were generated accordingly (Table 1). The model generated from the training set chosen in the activity and MCI descriptor spaces (model 2) had internal q^2 and external R^2 values of 0.684 and 0.558, respectively. These values were slightly lower than those of the earlier model [16] (model 1) where training and test sets were selected in activity, four point 3D pharmacophore fingerprint, and shape descriptors. Three different models (models 3–5) were generated with different training sets derived from the activity and MOE descriptor spaces. These models were validated; the internal q^2 and external R^2 values of the corresponding training and test sets are reported in Table 1. All these models provided consistently high values of q^2 and R^2 (with q^2 values ranging from 0.62 to 0.79, and R^2 values ranging from 0.56 to 0.90). Although the performance of the Catalyst HypoGen method varied slightly with the selection and the size of the training versus test sets, the models obtained with datasets selected by diversity sampling in descriptor spaces of four-point 3D pharmacophore fingerprints, topological indices or MOE descriptors provided consistently favorable q^2 and R^2 values.

A closer examination of the pharmacophore models derived from different training sets revealed high homology among models (Fig. 4). Both models 1 and 2 identified four functional features (two hydrogen bond donors and two hydrophobic aromatic features) in similar spatial orientations. Models 3–5 identified three functional features (two hydrogen bond donors and one hydrophobic, aromatic feature). These features match with three of the four features in models 1 and 2 in a highly consistent spatial arrangement. The hypotheses were readily mapped onto specific molecular areas of the active compounds **11** and **111**. The

Table 1
Catalyst HypoGen pharmacophore modeling with multiple representative training and test set selected in different descriptor spaces

	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[c]	Model 5[c]
Size of training set	25	24	23	26	21
Size of test set	19	19	26	14	19
q^2	0.76	0.68	0.79	0.62	0.76
R^2	0.74	0.56	0.56	0.90	0.71

[a] Derived from the training set selected in the activity, four-point 3D pharmacophore fingerprint and shape descriptor spaces.
[b] Derived from the training set selected in activity and MCI descriptor spaces.
[c] Derived from training sets selected in activity and MOE descriptor spaces.

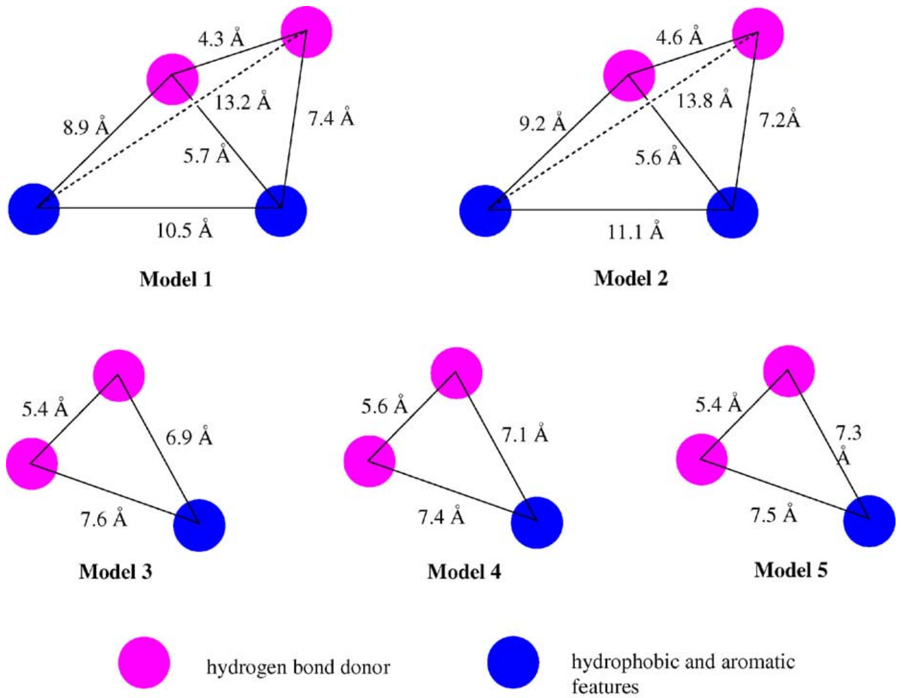


Fig. 4. Catalyst HypoGen pharmacophore models derived from multiple training sets selected in different descriptor spaces (cf. Table 1 for models 1–5. The distances among functional features are illustrated with an error range of ±1 Å.).

functional features identified by the pharmacophore models as critical for p38 MAP kinase inhibitory activity are illustrated in Fig. 5. A hydrophobic aromatic ring at C_4, a hydrogen bond donor at the C_5 side chain and the hydrogen bond donor of N_1 were recognized as structural features important to the activity of the pyridinylimidazole inhibitors, which was consistent with those characterized in the pyridinylimidazole inhibitor-p38 kinase complex structure and the previous pharmacophore model derived from pyridinylimidazole inhibitors. One of the HypoGen models (model 1) developed from the automated training set was mapped back onto the receptor active site (1A9U.pdb) with the bound ligand (Sb203580) one of the most active compound from the SAR series. The pharmacophore model (model 1) derived from this automated selection of the training set correctly identified the two well characterized hydrophobic pockets and the hydrogen bond donors [16].

The plot of observed vs. predicted activities for both training and test sets derived from model 5 (cf. Table 1) is shown in Fig. 6. Although both internal q^2 and external R^2 values of model 5 were high, the predicted activity values of the inactive compounds in both training and test sets showed a tendency to converge toward identical values, in contrast with dissimilar experimental activity data for these compounds.

To expand the earlier Catalyst HypoGen modeling of p38 MAP kinase inhibitors, additional calculations were performed. The entire original dataset was rationally divided into three subsets, training and test sets were generated, and QSAR models were built for each subset respectively. Briefly, successive diversity sampling of the most active, moderately active and least active subgroups (see Section 3) in the MOE descriptor space was performed. After each sampling, about one-third of the original 131 compounds were included into each subset, and then diversity sampling

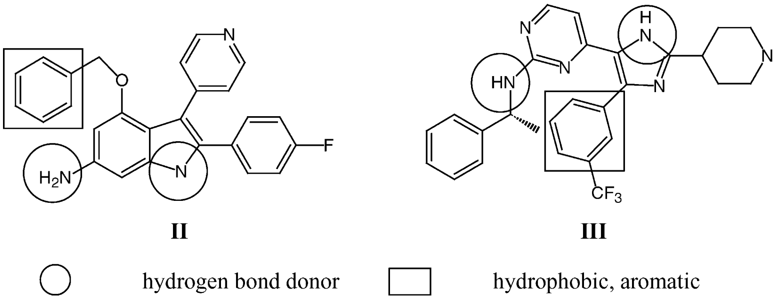


Fig. 5. Functional features identified by catalyst pharmacophore models as critical for p38 MAP kinase inhibitory activity.

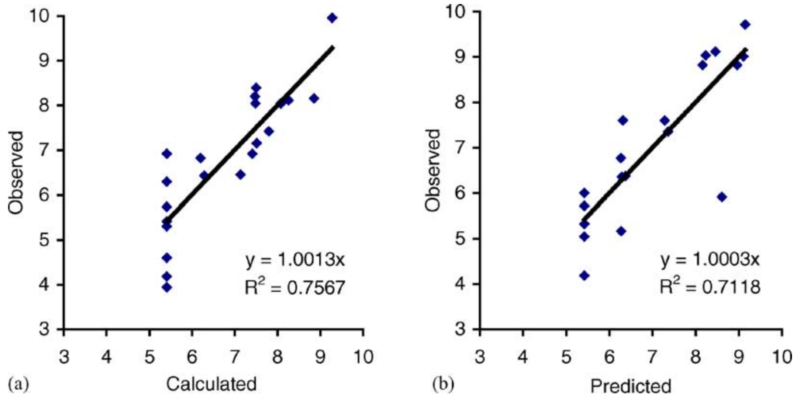


Fig. 6. The correlation between observed vs. calculated [9-log(IC_{50})] values of training (a) and test (b) sets molecules based on model 5.

Table 2
Catalyst HypoGen pharmacophore modeling with training and tests sets selected from three subsets of the original dataset in multidimensional descriptor spaces

	Subset 1 (model 6)		Subset 2 (model 7)		Subset 3 (model 8)	
	Training	Test	Training	Test	Training	Test
Size	24	19	21	19	25	17
q^2	0.559	–	0.717	–	0.644	–
R^2	–	0.328	–	0.021	–	0.629

was carried out again in each subset to split it into training and test sets. The QSAR models derived from each subset are summarized in Table 2. Although each subset afforded acceptable q^2 values, only the subset with the lowest diversity (subset 3) showed satisfactory predictive power (with R^2 value above 0.6). The poor predictive ability of subsets 1 and 2 may be due to the selection of insufficiently representative training sets.

In addition, model 1 was applied to predict activities of those compounds in the original datasets not included among the 44 molecules selected for the model generation [16]. The predictive R^2 was only 0.337. This result implies that QSAR pharmacophore models generated from selected compounds may still have limited applicability outside of the representative dataset that the models were derived from.

5.2. *k*NN QSAR modeling

Two sets of calculations similar to those done with the Catalyst HypoGen pharmacophore modeling were carried out using *k*NN QSAR method with MOE descriptors. Initially, *k*NN QSAR models were built with the same training and test sets used previously [16] and the statistics for resulting models is summarized in Table 3. These models were comparable to the HypoGen model (model 1). This indicated that alignment-free MOE descriptors combined with a statistically robust, non-linear variable selection technique afforded models with at least similar accuracy. The plot of observed versus predicted activities for both training and test sets derived from model 11 (cf. Table 3) are shown in Fig. 7. Although the q^2 and R^2 values of model 11 were statistically significant, when applied to the remaining 87 compounds in the original dataset left outside of the training and test sets, predictions derived from model 11 were unsuccessful (with predictive R^2 of only 0.115). Again, as discussed above for the Catalyst HypoGen modeling, this result implies that the relatively small diverse dataset of 44 compounds selected from the original dataset for structural and biological diversity was in fact insufficiently representative to afford accurate activity prediction of all remaining compounds.

Using MOE descriptors, the entire original data set was divided into a training set with 88 compounds and a test set with 43 compounds as described in Methods. The number of variables (n_{var}) was set to 10, 20, 30, 40, and with each predefined n_{var}, five models were generated. The best models obtained for each number of variables are described in Table 4 in comparison with model 1. To evaluate the robustness of these models, both leave-one-out cross-validated R^2 (q^2) for the training sets and predictive R^2 for the test sets

Table 3
*k*NN QSAR modeling using MOE descriptors with the training and test sets used for model 1

Model no.	n_{var}	*k* value	q^2	R^2
9	10	2	0.91	0.68
10	20	2	0.9	0.67
11	**30**	**2**	**0.9**	**0.76**
12	40	2	0.88	0.75
13	50	2	0.89	0.59
1			0.764	0.738

The results for the best model are in bold.

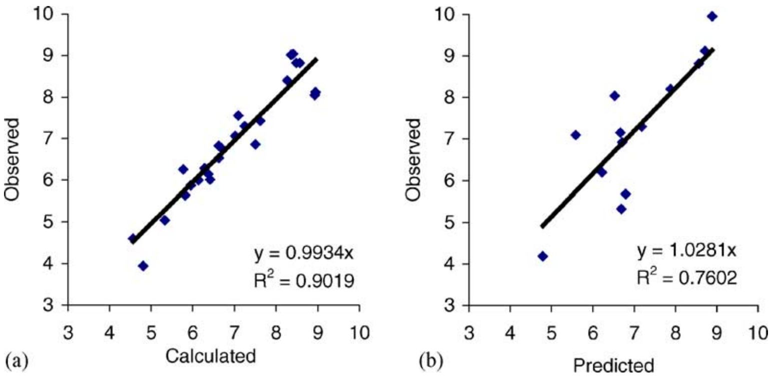


Fig. 7. The correlation between observed vs. calculated [9-log(IC_{50})] of training (a) and test (b) sets molecules based on model 11.

Table 4
*k*NN QSAR modeling using MOE descriptors with training and test sets selected from the whole dataset

Model no.	n_{var}	k value	q^2	R^2
19	10	3	0.68	0.46
20	20	2	0.67	0.59
21	30	3	0.7	0.52
22	**40**	**2**	**0.73**	**0.69**
23	50	2	0.68	0.67
1			0.764	0.738

The results for the best model are in bold.

were calculated. At a lower number of variables, the predictive ability of the models was relatively poor; however, as the n_{var} increased to 40, predictive models were obtained with q^2 and R^2 values comparable to those of model 1. The observed versus calculated activities generated with the best *k*NN/MOE model (model 22) for both training and test sets are plotted in Fig. 8.

It is important to stress at this point that unlike the earlier HypoGen model which used only representative compounds selected for diversity from the entire dataset of 131 compounds, these results were generated for training and test sets obtained by dividing the entire original dataset with the sphere exclusion algorithm. Apparently, due to the greater structural and biological diversity of compounds, the *k*NN QSAR models derived for expanded training sets have greater applicability than those derived for “representative” but small training sets. These results illustrate that an adequate applicability domain [32] must be established for any QSAR model to afford reliable and accurate prediction of the activity for external compounds.

6. Summary

In this study, we have applied both Catalyst HypoGen pharmacophore and *k*NN QSAR modeling approaches to obtain QSAR models for a data set of 131 p38 MAP kinase inhibitors. Statistically significant models were produced with both approaches, but there were substantial differences in computational interpretability and external prediction power of the underlying models.

The Catalyst pharmacophore models identify functional features important for p38 MAP kinase binding, designate possible molecular alignments for the active conformers of the diverse inhibitors and exhibit certain predictive capacity over compounds outside of the training set. These models are readily interpretable and can be used for the rational design of novel p38 MAP kinase inhibitors. However, extensive calculations associated with the generation of multiple conformers exploited in hypothesis generation makes it

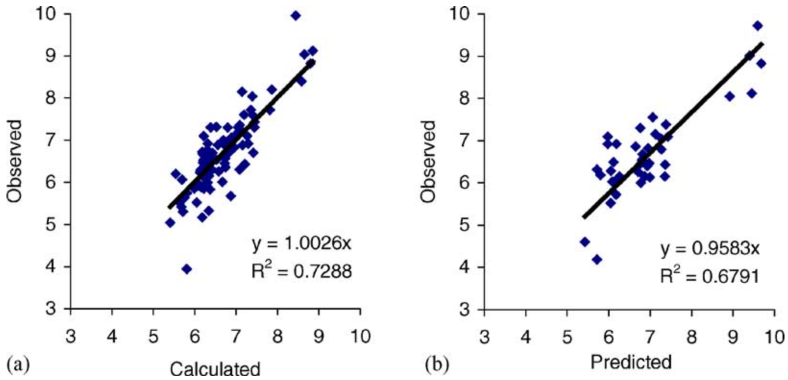


Fig. 8. The correlation between observed vs. calculated [9-log(IC_{50})] values of the training (a) and test (b) sets molecules based on model 22.

impractical for Catalyst HypoGen modeling to handle large datasets. Therefore, special algorithms should be applied to select representative training sets for pharmacophore generation, which raises the critical issue of rational diversity sampling. We have demonstrated that Catalyst QSAR models developed for rationally automated selected diverse training subsets of the entire available dataset do have certain predictive power. However, we caution that even in this case of the rationally selected representative training set the resulting models may not be sufficiently robust to guarantee their external predictive power even when applied to the remaining compounds from the original dataset. Similar conclusions have been obtained using completely different QSAR modeling principles (*k*NN QSAR modeling) and 2D topology based descriptors.

In contrast with Catalyst approach, the ambiguity of physico-chemical interpretation of topological descriptors makes the *k*NN QSAR models reported in this paper not applicable to direct specific chemical modifications of existing molecules. However, the high predictive ability of the models allows efficient virtual screening of chemical databases or virtual libraries determined by either synthetic feasibility or commercial availability of starting materials to prioritize synthesis of the most promising candidates. Models obtained in this study should facilitate the rational design of novel p38 MAP kinase inhibitors, guide the design of focused libraries based on the pyridinylimidazole or 3-(4-pyridyl)-1*H*-pyrrolo[2,3-*b*]pyridine skeletons, and facilitate the search for related structures with similar biological activity from large databases.

The results of this study suggest that it may be beneficial to apply Catalyst pharmacophore and *k*NN QSAR modeling techniques concomitantly to screening external databases for p38 MAP kinase inhibitors. *k*NN based models using alignment free descriptors can be used to mine large collections of chemical compounds for potential candidates with high predicted activities [33]. Catalyst models can be used concurrently to identify compounds with pharmacophoric features characteristic of the active molecules. Those molecules identified as active by both methods should be viewed as more likely candidates for further experimental validation.

Acknowledgements

This investigation was supported by the NIH grant GM66940-01 awarded to A. Tropsha. The authors at UNC acknowledge Accelrys, Inc. for the Catalyst and Cerius2 software grants.

References

[1] A. Paul, S. Wilson, C.M. Belham, C.J.M. Robinson, P.H. Scott, G.W. Gould, R. Plevin, Stress activated protein kinases: activation, regulation and function, Cell Signal 9 (1997) 403–410.

[2] P. Cohen, The search for physiological substrates of MAP and SAP kinases in mammalian cells, Trends Cell Biol. 7 (1997) 353–361.

[3] F.M. Brennan, M. Feldman, Cytokines in autoimmunity, Curr. Opin. Immunol. 8 (1996) 872–877.

[4] G. Camussi, E. Lupia, The future of antitumor necrosis factor (TNF) products in the treatment of rheumatoid arthritis, Drugs 55 (1998) 613–620.

[5] A. Cuenda, J. Rouse, Y.N. Doza, R. Meier, P. Cohen, T.F. Gallagher, P.R. Young, J.C. Lee, SB 203580 is a specific inhibitor of a MAP kinase homologue which is stimulated by cellular stresses and interleukin-1, FEBS Lett. 364 (1995) 229–233.

[6] J.C. Lee, A.M. Badger, D.E. Griswold, D. Dunnington, A. Truneh, B. Votta, J.R. White, P.R. Young, P.E. Bender, Bicyclic imidazoles as a novel class of cytokine biosynthesis inhibitors, Ann. NY Acad. Sci. 696 (1993) 149–170.

[7] T.F. Gallagher, G.L. Seibel, S. Kassis, J.T. Laydon, M.J. Blumenthal, J.C. Lee, D. Lee, J.C. Boehm, S.M. Fier-Thompson, J.W. Abt, M.E. Soreson, J.M. Smietana, R.F. Hall, R.S. Garigipati, P.E. Bender, K.F. Erhard, A.J. Krog, G.A. Hofmann, P.L. Sheldrake, P.C. McDonnell, S. Kumar, P.R. Young, J.L. Adams, Regulation of stress-induced cytokine production by pyridinylimidazoles inhibition of CSBP kinase, Bioorg. Med. Chem. 5 (1997) 49–64.

[8] T.F. Gallagher, S.M. Fier-Thompson, R.S. Garigipati, M.E. Sorenson, J.M. Smietana, D. Lee, P.E. Bender, J.C. Lee, J.T. Laydon, D.E. Griswold, M.C. Chabot-Fletcher, J.J. Breton, J.L. Adams, 2,4,5-Triarylimidazole inhibitors of IL-1 biosynthesis, Bioorg. Med. Chem. Lett. 5 (1995) 1171–1176.

[9] J.C. Boehm, J.M. Smietana, M.E. Sorenson, R.S. Garigipati, T.F. Gallagher, P.L. Sheldrake, J. Bradbeer, A.M. Badger, J.T. Laydon, J.C. Lee, L.M. Hillegass, D.E. Griswold, J.J. Breton, M.C. Chabot-Fletcher, J.L. Adams, 1-Substituted 4-Aryl-5-pyridinylimidazoles: a new class of cytokine suppressive drugs with low 5-lipoxygenase and cyclooxygenase inhibitory potency, J. Med. Chem. 39 (1996) 3929–3937.

[10] J.R. Henry, K.C. Rupert, J.H. Dodd, I.J. Turchi, S.A. Wadsworth, D.E. Cavender, B. Fahmy, G.C. Olini, J.E. Davis, J.L. Pellegrino-Gensey, P.H. Schafer, J.J. Siekierka, 6-Amino-2-(4-fluorophenyl)-4-methoxy-3-(4-pyridyl)-1*H*-pyrrolo[2,3-*b*]pyridine (RWJ 68354): a potent and selective p38 kinase inhibitor, J. Med. Chem. 41 (1998) 4196–4198.

[11] N.J. Liverton, J.W. Butcher, C.F. Claiborne, D.A. Claremon, B.E. Libby, K.T. Nguyen, S.M. Pitzenberger, H.G. Selnick, G.R. Smith, A. Tebben, J.P. Vacca, S.L. Varga, L. Agarwal, K. Dancheck, A.J. Forsyth, D.S. Fletcher, B. Frantz, W.A. Hanlon, C.F. Harper, S.J. Hofsess, M.L.J. Kostura, S. Luell, E.A. O'Neill, C.J. Orevillo, M. Pang, J. Parsons, A. Rolando, Y. Sahly, D.M. Visco, S.J. O'Keefe, Design and synthesis of potent, selective, and orally bioavailable tetrasubstituted imidazole inhibitors of p38 mitogen-activated protein kinase, J. Med. Chem. 42 (1999) 2180–2190.

[12] Z. Wang, B.J. Canagarajah, J.C. Boehm, S. Kassisa, M.H. Cobb, P.R. Young, S. Abdel-Meguid, J.L. Adams, E.J. Goldsmith, Structural basis of inhibitor selectivity in MAP kinases, Structure 6 (1998) 1117–1128.

[13] K.P. Wilson, P.G. McCaffrey, K. Hsiao, S. Pazhanisamy, V. Galullo, G.W. Bemis, M.J. Fitzgibbon, P.R. Caron, M.A. Murcko, M.S. Su, The structural basis for the specificity of pyridinylimidazole inhibitors of p38 MAP kinase, Chem. Biol. 4 (1997) 423–431.

[14] L. Tong, S. Pav, D.M. White, S. Rogers, K.M. Crane, C.L. Cywin, M.L. Brown, C.A. Pargellis, A highly specific inhibitor of human p38 MAP kinase binds in the ATP pocket, Nat. Struct. Biol. 4 (1997) 311–316.

[15] Accelrys, Inc., Catalyst 4.6, San Diego, 2001.

[16] S. Varma, R. Hoffmann, Novel methodologies in training-set selection for pharmacophore modeling: a pharmacophore model of p38 MAP kinase inhibitors, abstracts of papers, in: Proceedings of the 224th ACS National Meeting, Boston, MA, 2002.

[17] W. Zheng, A. Tropsha, Novel variable selection quantitative structure-property relationship approach based on the *k*-nearest-neighbor principle, J. Chem. Inf. Comput. Sci. 40 (2000) 185–194.

[18] Information on MOE is available at: http://www.chemcomp.com/.

[19] Molconn-Z Version 3.5, Hall Associates Consulting. Quincy, MA, Information on Molconn Z is Available at http://www.eslc.vabiotech.com/molconn/.

[20] Z. Wang, B.J. Canagarajah, J.C. Boehm, S. Kassisa, M.H. Cobb, P.R. Young, S. Abdel-Meguid, J.L. Adams, E.J. Goldsmith, Structural basis of inhibitor selectivity in MAP kinases, Structure 6 (1998) 1117–1128.

[21] D.C. Underwood, R.R. Osborn, C.J. Kotzer, J.L. Adams, J.C. Lee, E.F. Webb, D.C. Carpenter, S. Bochnowicz, H.C. Thomas, D.W.P. Hay, Griswold SB239063, a potent p38 MAP kinase inhibitor, reduces inflammatory cytokine production, airways eosinophil infiltration, and persistence, J. Phar. Expt. Therap. 293 (2000) 281–288.

[22] Y. Kurogi, O.F. Güner, Pharmacophore modeling and three-dimensional database searching for drug design using catalyst, Curr. Med. Chem. 8 (2001) 1035–1055.

[23] M.A. Sharaf, D.L. Illman, B.R. Kowalski, Chemometrics, John Wiley & Sons, New York, 1986.

[24] Z. Xiao, Y. Xiao, J. Feng, A. Golbraikh, A. Tropsha, K.H. Lee, Antitumor agents. 213. Modeling of epipodophyllotoxin derivatives using variable selection k nearest neighbor QSAR method, J. Med. Chem. 45 (2002) 2294–2309.

[25] L.B. Kier, L.H. Hall, Molecular Structure Description—The Electrotopological State, Academic Press, San Diego, 1999.

[26] L.H. Hall, L.B. Kier, The molecular connectivity chi indexes and kappa shape indexes in structure-property modeling, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry II, VCH Publishers, 1991, p. 367.

[27] Accelrys. Inc., Cerius2 4.6, San Diego, 2001.

[28] K.V. Mardia, J.T. Kent, J.M. Bibby, Multivariate Analysis, Academic Press, London, 1979.

[29] C.H. Reynolds, R. Druker, L.B. Pfahler, Lead discovery using stochastic cluster analysis (SCA): a new method for clustering structurally similar compounds, J. Chem. Inf. Comput. Sci. 38 (1998) 305–312.

[30] A. Golbraikh, A. Tropsha, Predictive QSAR modeling based on diversity sampling of experimental datasets for the test and training set selection, J. Comp. Aid. Mol. Design 16 (2002) 357–369.

[31] A. Golbraikh, M. Shen, Z. Xiao, Y. Xiao, K.-H. Lee, A. Tropsha, Rational selection of training and test sets for the development of validated QSAR models, J. Comp. Aid. Mol. Design 17 (2003) 241–253.

[32] A. Tropsha, P. Gombar, V.K. Gramatica, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, Quant. Struct. Act. Relat. Comb. Sci. 22 (2003) 69–77.

[33] A. Tropsha, S.J. Cho, W. Zheng, New tricks for an old dog: development and application of novel QSAR methods for rational design of combinatorial chemical libraries and database mining, in: A.L. Parrill, M.R. Reddy (Eds.), Rational Drug Design: Novel Methodology and Practical Applications, ACS Symposium Series No. 719, 1999, pp. 198–211.