

Article

Insight into the Structural Determinants of Imidazole Scaffold-Based Derivatives as TNF- α Release Inhibitors by *in Silico* Explorations

Yuan Wang ^{1,†}, Mingwei Wu ^{1,†}, Chunzhi Ai ² and Yonghua Wang ^{1,*}

¹ Lab of Systems Pharmacology, College of Life Sciences, Northwest A&F (Agriculture and Forestry) University, Yangling 712100, China; E-Mails: yuan-w-09@nwsuaf.edu.cn (Y.W.); audi.lg@163.com (M.W.)

² Lab of Pharmaceutical Resource Discovery, Dalian Institute of Chemical Physics, Graduate School of the Chinese Academy of Sciences, Dalian 116023, China; E-Mail: yling@dicp.ac.cn

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: yh_wang@nwsuaf.edu.cn; Tel.: +86-187-0080-1229; Fax: +86-029-8709-2262.

Academic Editor: Jesus Vicente De Julián Ortiz

Received: 28 May 2015 / Accepted: 30 July 2015 / Published: 25 August 2015

Abstract: Presently, 151 widely-diverse pyridinylimidazole-based compounds that show inhibitory activities at the TNF- α release were investigated. By using the distance comparison technique (DISCOtech), comparative molecular field analysis (CoMFA), and comparative molecular similarity index analysis (CoMSIA) methods, the pharmacophore models and the three-dimensional quantitative structure-activity relationships (3D-QSAR) of the compounds were explored. The proposed pharmacophore model, including two hydrophobic sites, two aromatic centers, two H-bond donor atoms, two H-bond acceptor atoms, and two H-bond donor sites characterizes the necessary structural features of TNF- α release inhibitors. Both the resultant CoMFA and CoMSIA models exhibited satisfactory predictability (with Q^2 (cross-validated correlation coefficient) = 0.557, R^2_{ncv} (non-cross-validated correlation coefficient) = 0.740, R^2_{pre} (predicted correlation coefficient) = 0.749 and Q^2 = 0.598, R^2_{ncv} = 0.767, R^2_{pre} = 0.860, respectively). Good consistency was observed between the 3D-QSAR models and the pharmacophore model that the hydrophobic interaction and hydrogen bonds play crucial roles in the mechanism

of actions. The corresponding contour maps generated by these models provide more diverse information about the key intermolecular interactions of inhibitors with the surrounding environment. All these models have extended the understanding of imidazole-based compounds in the structure-activity relationship, and are useful for rational design and screening of novel 2-thioimidazole-based TNF- α release inhibitors.

Keywords: imidazoles; TNF- α ; inhibitor; 3D-QSAR; CoMFA; CoMSIA; DISCOtech; pharmacophore

1. Introduction

Chronic inflammatory diseases, such as rheumatoid arthritis (RA) and psoriatic arthritis, are highly debilitating sicknesses affecting a large segment of the population. Recently, it has become obvious that even metabolic diseases (such as type II diabetes) and cardiovascular disease (such as atherosclerosis) should also be considered to be inflammatory in nature [1]. Thus, it is not surprising that huge efforts are constantly focused on the development of anti-inflammatory drugs. The progress of chronic inflammation is driven by an amplified systemic occurrence of several proinflammatory cytokines, like tumor necrosis factor- α (TNF- α) and interleukin-1 β (IL-1 β), which are produced in response to infection and other cellular stresses [2]. Although an appropriate amount of TNF- α plays an important role in the host immune response, excess levels are thought to underlie a number of serious inflammatory diseases [3,4]. Recent clinical data, obtained with either chimeric TNF- α antibodies or soluble TNF- α receptor in the treatment of rheumatoid arthritis and Crohn's disease, have confirmed the key role of TNF- α in these inflammatory disorders [5–9]. Antagonism of these proinflammatory cytokines has been recognized as an effective possibility for the treatment of inflammatory conditions [10].

Though the complete biochemical pathway for the production of TNF- α and IL-1 β in response to inflammatory stimuli has not been elucidated yet, p38 mitogen-activated protein kinase (p38, p38 MAPK) has been demonstrated to play a central role in the inflammatory process [1,11]. P38 MAPKs, first identified and cloned by Han *et al.* in 1994 as an important moiety of the immune response machinery activated by cytokines [12,13], were triggered by multiple extracellular stimuli, including stress signals, such as lipopolysaccharide (LPS), osmotic or heat shock, and proinflammatory cytokines, such as TNF- α or IL-1 β [14,15]. Research results have shown that after the activation of p38 MAPK, pro-inflammatory cytokines, such as TNF- α and IL-1 β , raised in both production and release [15]. Inversely, p38 can also be phosphorylated upon TNF- α or IL-1 β receptor binding and functions in the cell signaling network responsible for the up-regulation of these inflammatory mediators, both at the transcriptional and translational level, forming a vicious cycle in the development of inflammatory response [16,17]. Importantly, previous work has demonstrated that several small-molecular inhibitors of p38 MAP kinase, which compete with ATP for access to the catalytic site, have been shown to effectively block the activation of the pro-inflammatory transcription factor AP-1, which inhibits the transcriptional induction of TNF- α , IL-1 β and other cytokines and then blocks the production and release of these pro-inflammatory cytokines [1,2,11]. Thus, the biological importance of p38 MAP kinase, related to the release of pro-inflammatory cytokines, has aroused many studies aiming at

the development of selective inhibitors of p38 MAP kinase for the treatment of inflammatory conditions resulting from excess cytokine production.

Most of these small-molecular inhibitors of p38 MAP kinase are derived from the prototypical pyridin-4-ylimidazole SB203580, which definitely contributes to the identification and characterization of the p38 signaling pathway as a valid therapeutic target in inflammatory conditions [1,2,11]. Presently, based on the pyridin-4-ylimidazole SB203580, Laufer *et al.* [2,10,18,19] have reported newer pyridinylimidazole inhibitors, which possess not only high p38 α inhibitory activity, but also excellent inhibitory potency in suppression of the TNF- α release. In addition, these inhibitors have many other decisive advantages over prototype SB203580-like 2-arylimidazoles, e.g., higher selectivity, better kinetic and metabolic properties, and fewer interactions with metabolic enzymes like CYP450. Compared with the prototype SB203580, there are two major improvements in these newly-synthesized imidazole inhibitors: firstly, introduce additional substituents at the ortho position of the pyridin-4-yl moiety, which is sterically demanding or electronically shielding, and secondly, incorporate different substituents on the imidazole core, particularly the substituents at N1 and C2 positions of the imidazole, which might lead to improved physicochemical properties and reduced toxicity.

During the past decades, numerous new p38 inhibitors have been synthesized and reported, and to explore the structure-activity relationships of the p38 inhibitors, several groups have done excellent work on the imidazole or pyridinylimidazole inhibitors [20–26]. However, to our best knowledge, almost no *in silico* studies on those imidazoles inhibitors involving the inhibitory activity of TNF- α release in the human whole blood model has been reported until now. As we know, in recent years 3D-QSAR methods, like comparative molecular field analysis (CoMFA) [27] and comparative molecular similarity indices analysis (CoMSIA) [28], have been increasingly employed in rational drug discovery processes to understand the drug-receptor interaction and to design new molecules [20], due to their outstanding advantages of time-saving, cost-reducing, as well as high efficiency *in silico* screening and prediction of candidate drugs [29–31].

In the present work, 151 newly-synthesized imidazoles derivatives reported [2,10,18,19] as potent and selective TNF- α release inhibitors were employed as a data set to carry out a series of QSAR studies using a combination of CoMFA, CoMSIA, and pharmacophore modeling computational methods. The obtained CoMFA and CoMSIA studies not only illustrate the conformation or spatial orientation of those imidazole derivatives, but also provide useful indicators for the design of new drug candidates for inflammation diseases. These results are applicable to the prediction of the activities of new TNF- α release inhibitors and would be of help in providing structural implications for designing potent and selective TNF- α release inhibitors. Furthermore, the pharmacophore model was established to understand the essential features required for p38 binding using DISCOtech, which could provide important information for understanding of the mechanism of p38 enzyme catalysis. Based on the molecular field information of 3D-QSAR tools and pharmacophore modeling protocols, a few strategies were proposed to design new molecules with improved activities.

2. Results and Discussion

2.1. Split the Training and Test Sets

In order to validate the predicting ability of the produced 3D-QSAR models, an available dataset should be split into the training and test sets [32]. For the prediction statistics to be reliable, the selection of training and test sets should satisfy the following three rules [33,34]: (1) at least five molecules must be included in the test set; (2) the whole molecule-points of the test set in the descriptor space should be close to those in the training set; and (3) the points of the training set should be distributed evenly in the whole space possessed by the available dataset. In the present work, the Kohonen's self-organizing maps (SOM) was employed to divide the dataset. SOM, first developed by Kohonen in the 1980s, is a class of unsupervised neural networks widely employed in classification and similarity perception [35,36]. It is a remarkable tool in exploring phase of data mining and has a special advantage of calculating various features with spatial organization to represent input signals and abstractions [37]. SOM generates a set of prototype vectors to represent the dataset and then maps these prototypes from the high-dimensional [37,38]. Taking advantage of clustering capability, SOM makes it possible that training and test sets are homogeneously distributed in the whole descriptor space and the representative molecules in both sets can describe the depth of distribution of the whole molecules [39].

In the present work, a total of 1664 molecular descriptors were calculated for each TNF- α release inhibitor by Dragon (version 5.4). Then, based on these descriptors as input vectors, a SOM with 6×6 neurons was generated for the dataset. Figure 1 shows the SOM for TNF- α release inhibitors, in which the test set is labeled in red and the training set in black, respectively. It is clear that the entire distribution of the molecules in the map is satisfactory and both sets present a uniform spread in the whole chemical space. Furthermore, the representative points in the test set are close to those in the training set. The above results indicate that the division of the dataset is reliable and rational.

23 66 49		5 83 82 81	29 26 ³² 30 31 35 136 36	90 91	95 134 40 94 39 132 50
13 8 10 87 9 14 12 80	131 137	141 140	28 25 3	92 130 93 67	6 98 24 37 97 7 21 11 96
	99 100	146 147	119 101	133 125 129 135 38	89 128 79 78 68 127 22 126
122 123 124	114 110	150 138 139	148 151 145 143 142 144	109 116 111	120 121 16 118 117
20 18 15 17 19	69 88 70 75 76 4 1 2 72	73 77 60 74 59	63 65 86 64 62 61 33	112	113 108 105 107 115 102 103 104
	43 84 41 45 42 71 44 85	46 48 34	58 57 55 54 51 56 52 53		149 106

Figure 1. Self-organizing map showing the distribution of the training and test sets. The test set is labeled in red and the training set in black, respectively. The number equals to the series number of the molecules of the TNF- α release inhibitors.

2.2. 3D-QSAR Statistical Results

In the present work, the ligand-based alignment was employed to overlay the whole 151 compounds. All subsequent CoMFA and CoMSIA models were derived using the same training (118 molecules) and test (33 molecules) sets. In the present study, the most potent compound 3v was chosen as a template to fit the remaining compounds on the common substructure (Figure 2A, shown in bold) by using the “align database” command in Sybyl 6.9. Figure 2A also shows three important substituents of compound 3v, including the R¹, R², and R³ substituents. Figure 2B shows the resulting alignment model.

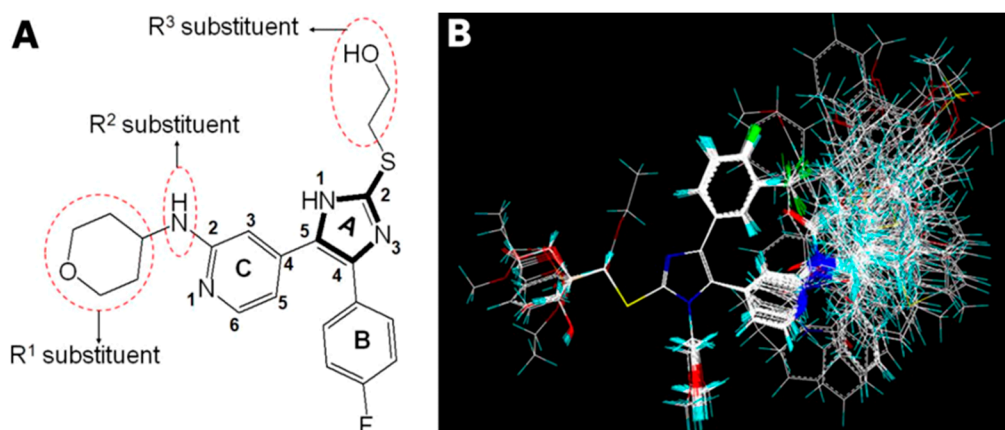


Figure 2. Molecular alignment of all compounds in the data set. (A) Compound 3v was used as a template for alignment, with the common substructure shown in bold, and three important substituents including the R¹, R², and R³ substituents shown in the dashed red border; and (B) Ligand-based alignment model of all the compounds. Molecules are colored in white for common C, blue for N, red for O, yellow for S, cyan for H, and green for F, respectively.

To determine the reliability of these models, all crucial statistical parameters were analyzed here, including the cross-validated correlation coefficient (Q^2), non-cross-validated correlation coefficient (R^2_{ncv}), standard error of estimate (SEE), F-statistic values, the optimum number of components (OPN), as well as the predicted correlation coefficient (R^2_{pre}). $ClogP$ (calculated logarithm) and hydrophobic fields are both important parameters describing the hydrophobic property of a molecule and the $ClogP$ parameter always plays a crucial role in building appropriate 3D-QSAR models, which has been confirmed by many earlier studies [26,40–43]. In all of these studies, the inclusion of $ClogP$ as an additional descriptor added to the QSAR models led to significantly improved statistical results of the *in silico* models. Therefore, in the present study, a 3D-QSAR analysis including $ClogP$ as an additional descriptor has been carried out. Table 1 summarizes the statistical results of the CoMFA and CoMSIA analyses.

Table 1. Summary of CoMFA and CoMSIA results.

PLS Statistics	Model A		Model B	
	CoMFA	CoMSIA	CoMFA	CoMSIA
Q^2	0.524	0.593	0.557	0.598
R^2_{ncv}	0.856	0.778	0.740	0.767
SEE	0.323	0.399	0.432	0.409

Table 1. Cont.

PLS Statistics	Model A		Model B	
	CoMFA	CoMSIA	CoMFA	CoMSIA
<i>F</i>	133.655	99.117	80.388	93.091
R^2_{pre}	0.730	0.876	0.748	0.860
<i>SEP</i>	0.588	0.541	0.564	0.538
<i>OPN</i>	5	4	4	4
Contribution				
<i>S</i>	0.574	0.121	0.492	0.135
<i>E</i>	0.426	0.294	0.373	0.279
<i>H</i>	–	0.195	–	0.158
<i>D</i>	–	0.208	–	0.195
<i>A</i>	–	0.181	–	0.146
<i>ClogP</i>	–	–	0.135	0.086

Q^2 , cross-validated correlation coefficient; R^2_{ncv} , non-cross-validated correlation coefficient; *SEE*, standard error of estimate; *F*, *F*-test value; R^2_{pre} , predicted correlation coefficient; *SEP*, smallest predicted error; *OPN*, optimum number of components; *S*, steric hindrance; *E*, electric charge; *H*, hydrogen bond; *D*, hydrogen-bond donor; *A*, hydrogen bond acceptor; *ClogP*, calculated logarithm.

2.2.1. CoMFA Details

For CoMFA analysis, the steric, electrostatic, and *ClogP* field descriptors were fitted together in every possible form to build appropriate CoMFA mathematical models. Finally, the optimal model (model A), employing both the steric and electrostatic field descriptors, obtains a LOO cross-validated Q^2 of 0.524 with five components, indicating a proper internal predictive capacity of the model. A high correlation coefficient (R^2_{ncv}) of 0.856 for the final non-cross-validated model shows self-consistency. In addition, other statistical results including *SEE* value of 0.323, and an *F*-test value of 133.655 all suggest that the CoMFA model is a reliable predictor. While a more statistically significant model (model B) was obtained, including *ClogP* as a third parameter, which has a higher cross-validated Q^2 of 0.557 with four optimum components, a non-cross-validated R^2_{ncv} of 0.740, a standard error of estimation of 0.432, and *F* ratio of 80.388, proving the reliability of the model. In terms of the relative contributions in model A, the steric and electrostatic fields account for 0.574 and 0.426, respectively. In model B, the steric and electrostatic fields contribute 0.492 and 0.373, respectively, while the *ClogP* field contributes 0.135, indicating that the steric property contributes a majority to the antagonist activity.

2.2.2. CoMSIA Details

Combined with *ClogP*, a total of six parameters (steric, electrostatic, hydrophobic, H-Bond donor, H-Bond acceptor, and *ClogP*) were fitted together in every possible form to build appropriate CoMSIA models. Using the same training set as in the CoMFA method, two superior models, out of all the CoMSIA models established, were obtained with high Q^2 values using all the five fields (steric, electrostatic, hydrophobic, H-bond donor, and H-Bond acceptor field) (Table 1). The optimal CoMSIA model (model B), employing the *ClogP* descriptor in statistical results, showed $Q^2 = 0.598$, $R^2_{ncv} = 0.767$, *SEE* = 0.409, and *F* = 93.091 with four optimum components. However, another superior model without a *ClogP*

index (model A) has a Q^2 value of 0.593 with four optimum components, an R^2_{ncv} value of 0.778, a SEE value of 0.399, and an F value of 99.117. The optimal CoMSIA model reveals that steric, electrostatic, hydrophobic, hydrogen bond-donor, and hydrogen bond-acceptor fields all have major influences in the inhibitory activity of imidazole derivatives. The corresponding field contributions of steric, electrostatic, hydrophobic, hydrogen bond-donor, and hydrogen bond-acceptor fields in model A were 0.121, 0.294, 0.195, 0.208, and 0.181, respectively, and in model B were 0.135, 0.279, 0.158, 0.195, and 0.146, respectively. Once again, the steric and electrostatic fields are demonstrated to contribute a lot to model B (with a sum of about 0.414 contribution). However, at this time, the hydrogen bond (H-Bond) descriptors, especially the H-bond fields, outstand their roles for correlation with the inhibitory activity, where the donor field possesses about 0.195 and the acceptor field occupies about 0.146 of relative contributions. Moreover, the contribution of the hydrophobic field is found to be about 0.158, while $ClogP$ demonstrates a slight contribution of 0.086, suggesting a good correlation between the hydrophobic descriptor and their TNF- α release inhibitory activities. All of these conclusions can be verified in other studies on the p38 MAPK complexes that the effect of hydrogen bonding, steric and hydrophobic fields could be important for the inhibitory activity of imidazole and quinazoline functionalized inhibitors [21,26,44].

According to the above analysis, the predictive ability of both 3D-QSAR models for A are not good enough when compared with that of model B. Hence, the models including the $ClogP$ parameter were selected as the optimal models and were utilized for further discussion.

2.2.3. Validation of the 3D QSAR Models

To test the predictive ability of the models, the most appropriate method is to predict the activities of molecules excluded from the training set. For this purpose, the test set (33 molecules) which accounts for 28% of the training set, was used here to validate the accuracy of models A and B. We should identify possible outliers first before the final validation by the test set. Outliers from a QSAR are compounds that do not fit the model or are poorly predicted [45]. QSAR models that have very few or no outliers are known as good models. On the other hand, those QSAR models that have a large number of outliers are bad models. Many reasons may account for the presence of outliers in the dataset used for *in silico* modeling, including unique structural differences, different binding conformation, or a higher residual between the observed and predicted biological activity of an inhibitor [46,47].

However, the removal of outliers will be allowed for the development of stronger and more significant models and the outlier test is, therefore, reasonable and necessary in the derived models. There are a variety of methods to highlight outliers, including identifying those compounds with significantly high residuals from regression-based techniques. At present, the proposed models (CoMFA and CoMSIA) are checked to identify possible outliers. It can be observed that no residuals in both training and test sets are more than 1.5 log units, illustrating that the CoMFA and CoMSIA models are both robust and predictive. Thus, it is reasonable to consider that there are no outliers in the present models.

After outlier testing, both CoMFA and CoMSIA models in model B exhibit good prediction ability, yielding R^2_{pre} of 0.748 for CoMFA and 0.860 for CoMSIA, respectively. The plot of actual *versus* predicted activities of all compounds for the optimal CoMFA and CoMSIA models is depicted in Figure 3A,B, respectively, where the data points are rather uniformly distributed around the regression line, indicating the reasonability of the obtained models.

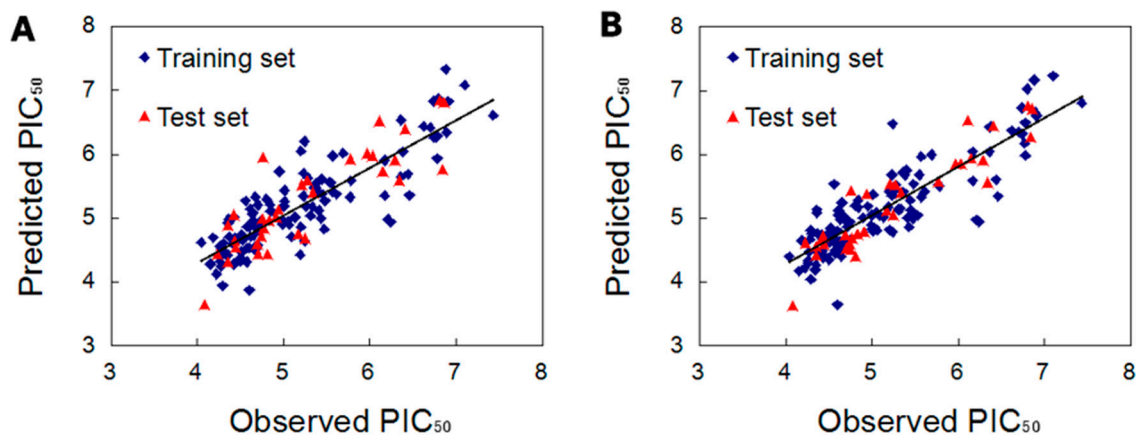


Figure 3. The correlation plots of the predicted pIC_{50} values *versus* the observed pIC_{50} values using the training set and the test set based on (A) CoMFA model and (B) CoMSIA model. The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively. (more detail in Table S1).

2.3. Interpretation of 3D-QSAR Contour Maps

After consideration of both the internal and external predictive powers of all the derived models, the best CoMFA and CoMSIA models are selected to construct the standard deviation \times coefficient (StDev \times Coeff) contour maps to view the field effects on the target features. All contour maps obtained from the models are illustrated together with the most potent compound 3v (Figure 4A). In order to select appropriate contour levels for each feature, the resulting histograms of actual field values were analyzed, and a contour level was chosen interactively as those that produce the best interpretable contour plot.

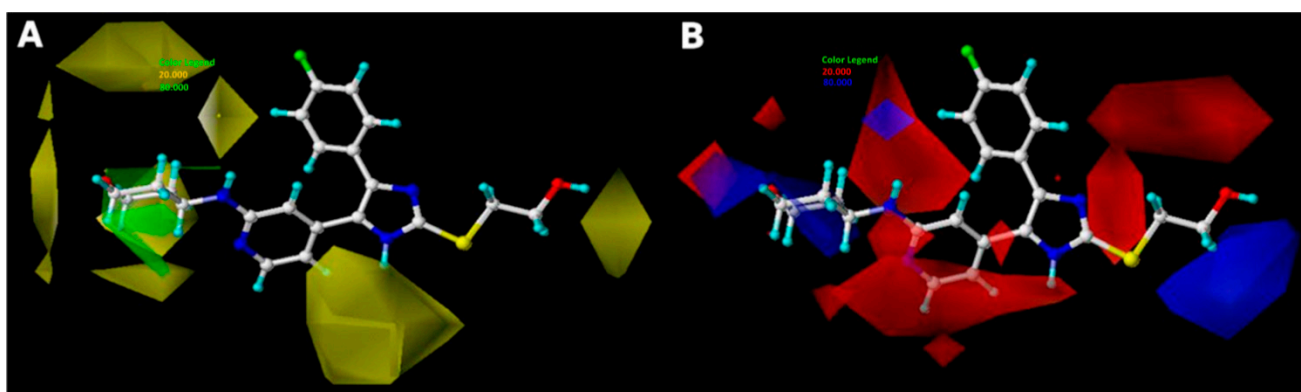


Figure 4. CoMFA StDev \times Coeff contour plots. (A) Steric (green/yellow) contour map in combination with compound 3v in ball and stick. Green contours indicate regions where bulky groups increase the activity; yellow contours indicate regions where bulky groups decrease the activity; and (B) Electrostatic contour map (red/blue) in combination with compound 3v. Red contours indicate regions where negative charges increase the activity; blue contours indicate regions where positive charges increase the activity.

Figure 4A depicts the CoMFA steric contour map of the optimal model with compound 3v overlaid. In this map, the green (sterically-favorable) and yellow (sterically-unfavorable) contours represent 80% and 20% level contributions, respectively. The green contour region around ring C (R^1 substituent) indicates that a bulky substituent is preferred in this position to produce higher inhibitory activity. Compounds with high potency such as 30h, 3v, and 3w have sterically-favoured groups at the R^1 position occupying the big green contour. However, due to the absence of such groups in the sterically-favoured green areas, compounds 5k, 5i, and 5f show much lower activity. There are large yellow contours near the green contour region, which suggests that substitution of too bulky a group at R^1 substituent would lead to a decrease in activity, such as compounds 22j and 46 (with $C_{11}H_{16}$ and $C_{14}H_{20}O$ at the R^1 substituent position, respectively). In addition, a medium-sized yellow contour is seen in the vicinity of the OH (R^3 substituent), which suggests that occupancy of this sterically-unfavorable region with bulky substituent would have a detrimental effect on the inhibitory activity. This discovery is well illustrated by the example that compound 1b with $-\text{CH}(\text{COOCH}_2\text{CH}_3)_2$, 1e with $-\text{CH}_2\text{CH}_2\text{CH}_2\text{COOCH}_3$, and 1f with $-\text{CH}_2\text{COOCH}_3$ at the R^3 substituent have much lower activity than any other compound with smaller substituent at the same location, like 1a with $-\text{CH}_2\text{CH}(\text{OH})\text{CH}_2\text{OH}$, 1c with $-\text{CH}_2\text{CH}_2\text{CH}_2\text{OH}$, and 1d with $-\text{CH}_2\text{CH}_2\text{OH}$.

The CoMFA electrostatic contour map is shown in Figure 4B. The blue (electropositive groups are favorable) and red (electronegative groups are favorable) contours represent 80% and 20% level contributions, respectively. The blue contour around the 2-position of ring C (R^1 and R^2 substituents) indicates that substitutions with positively-charged groups are favored for inhibitory activity. This is in agreement with the experimental findings, such as the order of activity for those compounds: 22b ($C_9H_{13}N$) > 5g (C_6H_7N) > 33b ($C_9H_{12}O$). Compound 1d has a higher activity when compared to 1e, explaining why a large blue contour exists near the 2-position of ring A. In addition, some red contours are observed around ring A, ring B, and ring C, which indicates that electronegative groups would be favorable.

The electrostatic contour map of the CoMSIA model is displayed in Figure 5B. The blue (electropositive groups are favorable) and red (electronegative groups are favorable) contours represent 80% and 20% level contributions, respectively. A red region is observed near the 2-position of ring A (R^3 substituent), suggesting that substitution with negatively-charged group would increase the inhibitory activity. Due to the presence of electronegative groups ($-\text{CH}_2\text{CH}_2\text{OH}$) at the 2-position of ring A, compounds 3v, 3w, and 3u show dramatic increases in activity, which implies that there may exist important hydrogen bond interactions with those groups. In addition, a large red contour near the R^2 substituent indicates substitution with a negatively-charged group is favored for inhibitory activity. This discovery is well-illustrated by the example that compound 1d has a higher activity than 1c. The location of blue contours is similar to the previous CoMFA model, which is thus not specified here.

The CoMSIA contour map of hydrophobic contribution is described in Figure 5C. In this figure, the yellow (hydrophobic groups are favorable) and white (hydrophobic groups are unfavorable) contours represent 80% and 20% level contributions, respectively. A big yellow contour is observed around the 2-position of ring C (R^1 substituent), suggesting this region's taste for hydrophobic substituents. The fact that compounds with high activity, like 22f, 22g, 28l, 35b, and 35d, all have bulky hydrophobic substituents at R^1 position (occupying the yellow contours) well illustrates this conclusion. Whereas, a medium-sized white contour at the 2-position of ring C (R^2 substituent) and another two at the left of R^1 substituent indicate hydrophilic substituents groups (like hydroxy or amido) are favorable for activity.

Hence, substitution of bulky hydrophobic moieties with small polar groups on the left extended to the large yellow and medium-sized white regions resulting in a higher inhibitory activity, which can be exemplified by the higher activity of chemical 30 (with $C_{12}H_{16}O_2$), 20 (with $C_{11}H_{14}O_2$), 31 (with $C_{10}H_{12}O_2$), and 5 (with $C_9H_{10}O_2$) than their analog compound 18 (with $C_{11}H_{14}O$).

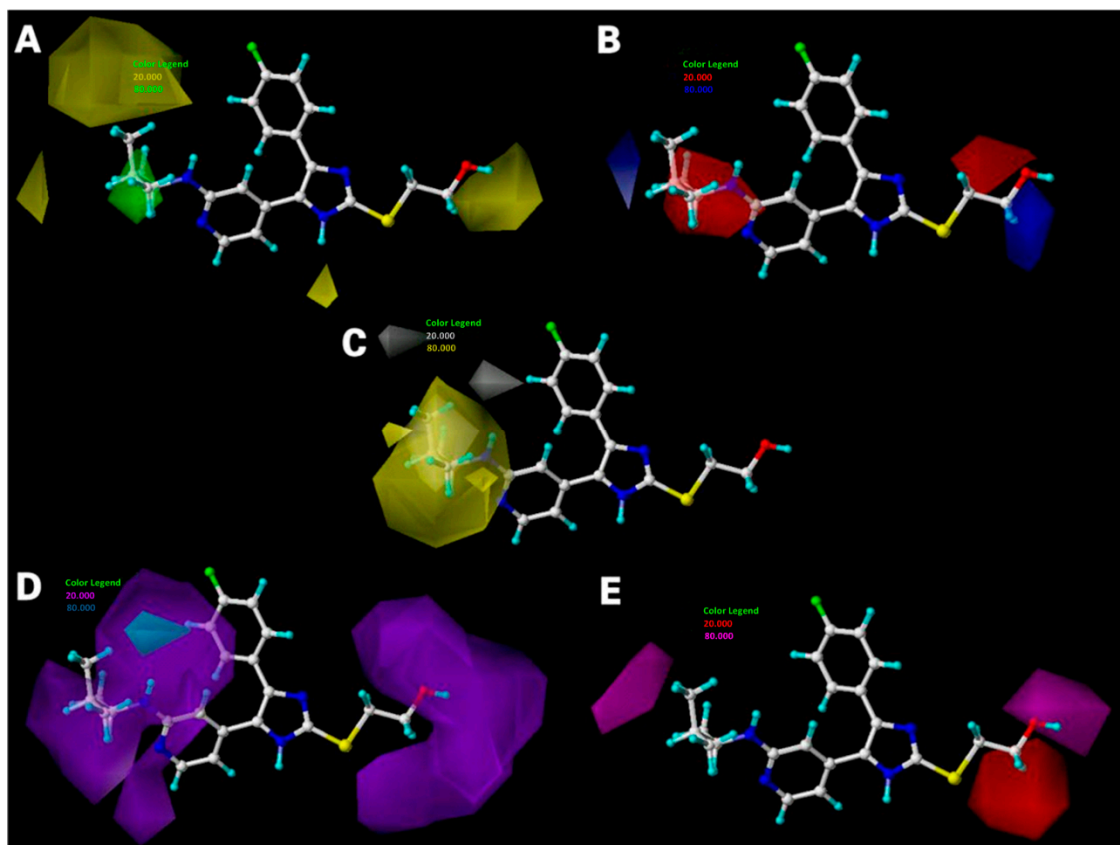


Figure 5. CoMSIA StDev*Coeff (A) steric; (B) electrostatic; (C) hydrophobic; (D) H-bond donor and (E) H-bond acceptor contour maps. The color code is as follows: (A) green and yellow contours indicate favorable and unfavorable bulky groups, respectively; (B) blue and red contours indicate favorable and unfavorable electropositive groups, respectively; (C) yellow and white contours indicate favorable and unfavorable hydrophobic groups, respectively; (D) cyan and purple contours indicate favorable and unfavorable H-bond donor groups, respectively; (E) magenta and red contours indicate favorable and unfavorable H-bond acceptor groups, respectively. The compound 3v in ball and stick is displayed as a reference.

The CoMSIA contour map of H-bond donor contribution is described in Figure 5D. In this figure, the cyan (H-bond donor groups are favorable) and purple (H-bond donor groups are unfavorable) contours represent 80% and 20% level contributions, respectively. A small cyan contour map near the R^1 substituent suggests that the presence of strong electropositive H-bond donors in these regions is favorable for the biological activity. There is a large purple contour map observed around ring C, suggesting its unfavorable effect on the H-bond donor groups. There is another purple contour observed near the R^3 substituent suggesting an H-bond donor moiety located near these regions will result in impaired biological activity.

The H-bond acceptor contour map of the CoMSIA model is displayed in Figure 5E. The magenta (H-bond acceptor groups are favorable) and red (H-bond acceptor groups are unfavorable) contours represent 80% and 20% level contributions, respectively. A large magenta contour is seen in the vicinity of the OH, which suggests that substitution with a H-bond acceptor group would increase the inhibitory activity. Furthermore, a medium-sized magenta contour map appears near the R¹ substituent, which is consistent with the findings revealed by the H-bond donor contour map. A relatively larger red contour map near the 2-position of ring A implies that H-bond donor groups could have a positive influence on the inhibitory activity.

2.4. Pharmacophore Modeling

Presently, 100 highly active compounds participated in the establishment of the pharmacophore model in order to identify the features required for TNF- α release inhibitors because of their structural diversity and high activity. The most active compound, 3v, was selected as the reference molecule for the pharmacophore model. The maximum number of conformers generated for each compound was 50, and seven conformers were then selected. In total, eight models were generated using the most active 100 antagonists, and the statistic results of pharmacophore modeling by Distance Comparisons technique (DISCOtech) are listed in Table 2. Ten features were searched for all eight obtained models, and all molecules matched with the obtained models. In general, the DISCOtech pharmacophore model with a relatively high score, more useful features, and moderate pairwise tolerance would be selected as the best model. As depicted in Figure 6A,B, Model_006 was selected as the best model for subsequent studies because of its highest score value (4.1075) and useful features of a higher diversity compared with others.

Table 2. Number of models obtained along with the pharmacophoric features and tolerance values for each of the DISCO pharmacophoric.

Model	Size ^a	Hits ^b	Score ^c	Tolerance ^d	Dmean ^e
MODEL_006	10	100	4.1075	0.25	4.5615
MODEL_003	10	100	4.1015	0.50	4.5264
MODEL_008	10	100	4.1015	0.50	4.5263
MODEL_007	10	100	3.4537	0.50	3.7971
MODEL_005	10	100	3.4526	0.25	3.7925
MODEL_001	10	100	3.4517	0.25	3.789
MODEL_004	10	100	3.4517	0.25	3.789
MODEL_002	10	100	3.4516	0.25	3.7885

^a, SIZE, number of features in the model; ^b, HITS, number of molecules that matched during the research; ^c, SCORE, an overall measure of fit and of overlap for the entire collection of structure; ^d, TOLERANCE, initial tolerance setting (from 0.25 to 2.5); ^e, DMEAN, average inter-point distance.

The DISCOtech model with the alignment of 100 molecules is presented in Figure 6B. All aligned conformers represent stochastic conformations of the molecules and the final alignment shows a satisfactory superimposition of the pharmacophoric points. Figure 6A shows the features arrangement of the optimal pharmacophore model (model_006) obtained with a highest score of 4.1075 and tolerance distance of 0.25 Å. Clearly, the model consists of ten essential features required for high receptor binding affinity,

defined as two hydrophobic sites (HP1 and HP2), two aromatic centers (AR1 and AR2), two H-bond donor atoms (DA1 and DA1), two H-bond acceptor atoms (AA1 and AA2), and two H-bond donor sites (DS1 and DS2).

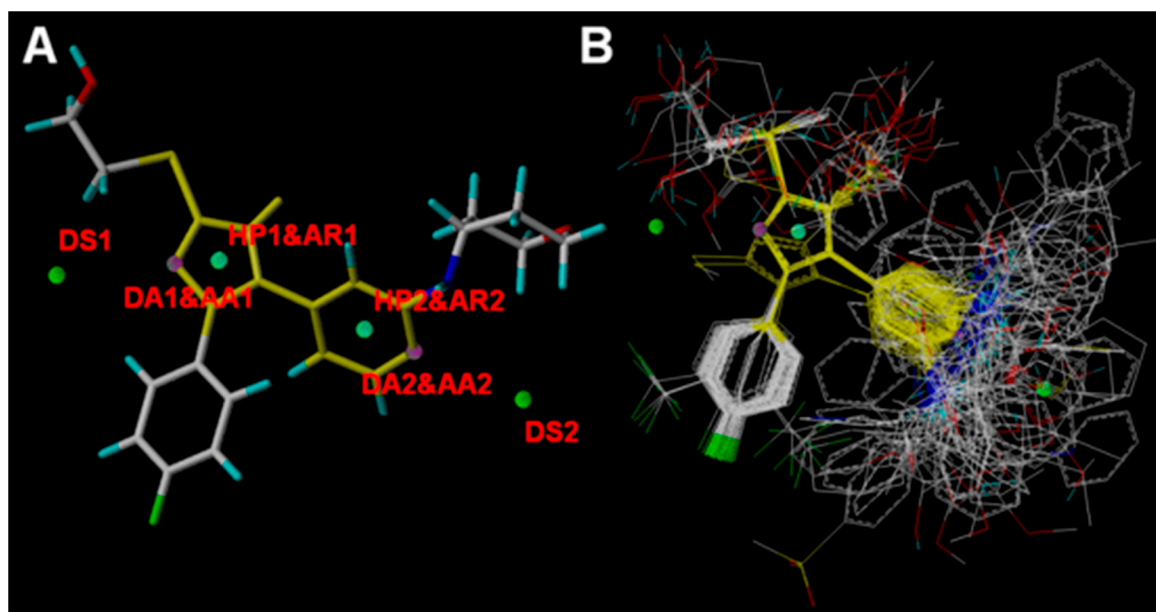


Figure 6. Pharmacophore model of DISCOtech. (A) Model_006 with ten features generated presented in template molecule 3v (B) Alignment of 100 molecules. AR represents aromatic center; HP refers to hydrophobic center; DA and AA are short for H-bond donor and acceptor atom, respectively; DS and AS refer to H-bond donor site and H-bond acceptor site, respectively. Molecules are colored in yellow and white for common C, blue for N, red for O, cyan for H, and green for F, respectively.

The pharmacophoric feature points and the distances between pharmacophore features are presented in Figure 7 and Table 3, respectively. The distance between HP1&AR1 in ring A and HP2&AR2 in ring C is $5.05 \pm 0.25 \text{ \AA}$, and the hydrophobic ring centers or aromatic centers revealed by this model indicate that hydrophobic interaction has important roles in the action mechanism between inhibitors and the TNF- α release inhibitory activity. Moreover, the nitrogen atom at the 3-position of ring A corresponds to the DA1 and AA1 features of the model, with the DS1 representing its counterpart on the putative receptor. Similarly, the nitrogen atom at the 1-position of ring C also corresponds to the DA2 and AA2 features, with the DS2 representing its counterpart on the putative receptor. The distance between those features are: $12.16 \pm 0.25 \text{ \AA}$ (DS1 and DS2), $3.00 \pm 0.25 \text{ \AA}$ (DS1 and DA1&AA1), $9.20 \pm 0.25 \text{ \AA}$ (DS1 and DA2&AA2), $16.41 \pm 0.25 \text{ \AA}$ (DA1&AA1 and DA2&AA2), $9.41 \pm 0.25 \text{ \AA}$ (DA1&AA1 and DS2), $3.00 \pm 0.25 \text{ \AA}$ (DA2&AA2 and DS2). Those important hydrogen bond-donor and hydrogen bond-acceptor features corresponding to H-bond interactions also play an important role in the improvements of the activity of TNF- α release inhibitors. Closer inspection of the pharmacophore model reveals that the hydrogen bond-donor and hydrogen bond-acceptor results are in agreement with the results from the previously reported study results that three important H-bonds were formed with N atom of imidazole (ring A) and 4-pyridyl (ring C): H-bonds (1) between Lys53 and imidazole N3; (2) between the 4-pyridyl N (ring C) and Met109N; and (3) between the 4-pyridyl N (ring C) and Gly110N [26]. All of the features are

mapped perfectly for the highly active compounds, whereas the features of inactive compounds are not mapped well. In conclusion, our pharmacophore results reveal that hydrophobic interaction and hydrogen bonds are the crucial factors acting on TNF- α release inhibitory activity, which is also consistent with the 3D-QSAR models.

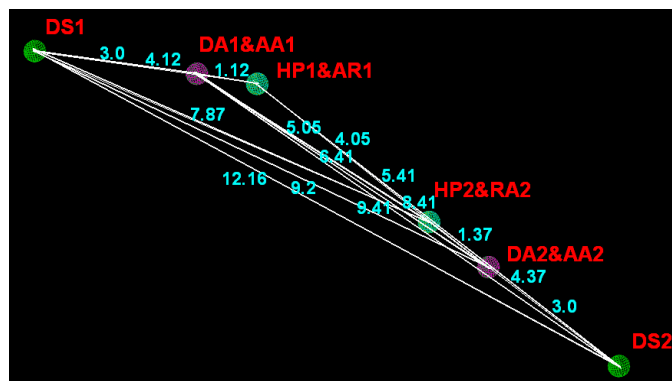


Figure 7. Pharmacophoric features and their distance relation generated by DISCOtech module, AR represents aromatic center; HP refers to hydrophobic center; DA and AA are short for H-bond donor and acceptor atom, respectively; DS and AS refer to H-bond donor site and H-bond acceptor site, respectively.

Table 3. Relative intramolecular distances between pharmacophoric feature points for model_006 (Å).

Domain	DS1	DA1/AA1	HP1/AR1	HP2/AR2	AA2/DA2
DA1/AA1	3.00 ± 0.25	–	–	–	–
HP1/AR1	4.12 ± 0.25	1.12 ± 0.25	–	–	–
HP2/AR2	7.87 ± 0.25	5.05 ± 0.25	4.05 ± 0.25	–	–
AA2/DA2	9.20 ± 0.25	6.41 ± 0.25	5.41 ± 0.25	1.37 ± 0.25	–
DS2	12.16 ± 0.25	9.41 ± 0.25	8.41 ± 0.25	4.37 ± 0.25	3.00 ± 0.25

AR represents aromatic center; HP refers to hydrophobic center; DA and AA are short for H-bond donor and acceptor atom, respectively; DS and AS refer to H-bond donor site and H-bond acceptor site, respectively.

3. Experimental Section

3.1. Dataset and Biological Activity

A total of 151 2-thioimidazoles were used as the data set to derive the 3D-QSAR models in this work, with a wide spectrum of activity values against TNF- α release collected from the work of Stefan A. Laufer and co-workers [2,10,18,19]. The *in vitro* biological activities of these compounds were converted into the corresponding pIC₅₀ (–log IC₅₀) values, which were used as dependent variables in the *in silico* analyses. All the structures and biological activities of the data set are listed in Table S2. In an approximate ratio of 4:1 the whole data set was divided into training (118) and test (33 molecules) sets by using Kohonen's self-organizing maps (SOM). In order to classify the dataset, a neural network mapped all the points in a high-dimensional space into a two dimensional space. In this two dimensional Kohonen map, the same neurons or neurons near to each other were occupied by data with similar input.

3.2. Molecular Modeling and Alignment Procedure

All molecular modeling and 3D-QSAR studies were performed using the Sybyl 6.9 molecular modeling software package (Tripos Associates, St. Louis, MO, USA). Partial atomic charges were calculated by the Gasteiger-Huckel method [48], while energy minimization and conformational search were performed using Tripos molecular mechanics force field [49] by conjugating method with a convergence criterion of 0.001 kcal/mol. The energy gradient limit was set at 0.05 kcal/mol·Å to obtain the most stable conformation. The predicting ability of the produced models can be done by applying the statistical tests for the continuous QSAR models[50,51].

The structural alignment of the compounds is one of the most important and challenging steps in the development of a successful 3D-QSAR model [52]. Since all the compounds share a common scaffold, it was assumed that each molecule binds into the integrin active site in a similar mode [53]. In the present study, the most potent compound, 3v, was chosen as a template to fit the remaining compounds on the common substructure by using the “align database” command in Sybyl 6.9 (Tripos Associates, St Louis, MO,USA).

3.3. CoMFA and CoMSIA

Several approaches have been developed for the 3D-QSAR study and CoMFA and CoMSIA are two of the most popular methods. CoMFA/CoMSIA were anticipated by MTD/Minimal Topological Difference methods [54,55]. These two methods were recently reported in modeling degenerative diseases [56–58].

In order to obtain the CoMFA and CoMSIA descriptor fields, a 3D cubic lattice intersection of a regularly-spaced grid of 2.0 Å was generated to encompass the aligned molecules. The grid box dimensions were determined automatically in such a way that region boundaries were extended beyond 4 Å in each direction from coordinates of each molecule. In CoMFA, the steric and electrostatic fields were calculated separately for each molecule using a sp³ carbon atom probe with a charge of +1.0 and energy cut-off values of 30 kcal/mol for both steric and electrostatic fields. The probe atom was placed at each lattice point, and the steric and electrostatic interactions with each atom in the molecule were computed using the CoMFA standard scaling.

The same grid was used for the CoMSIA field calculation as that constructed for the CoMFA field calculation. Five descriptors (steric, electrostatic, hydrophobic, hydrogen-bond donor and acceptor) were evaluated using a common probe atom with a radius of 1.0 Å, a +1 charge, hydrophobicity of +1.0, and H-bond donor and acceptor properties of +1.0. The contributions from these descriptors were truncated at 0.3 kcal/mol. Validation of the CoMSIA analysis was performed as described for CoMFA. Due to the different shape of the Gaussian function, CoMSIA similarity indices (A_F) for molecule j with atom i at grid point q are calculated by Equation (1):

$$A_{F,K}^q(j) = - \sum \omega_{\text{probe},k} \omega_{ik} e^{-\alpha r_{iq}^2} \quad (1)$$

where k represents the steric, electrostatic, hydrophobic, or H-bond donor or acceptor descriptor. $\omega_{\text{probe},k}$ is the probe atom with radius 1.0 Å, charge +1, hydrophobicity +1, H-Bond donating +1, H-bond accepting +1; ω_{ik} is the actual value of the physicochemical property k of atom i ; r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule. The attenuation factor α was set to 0.3.

3.4. Partial Least Square Analysis and Statistical Validation

To quantify the relationship between the structural parameters (CoMFA/CoMSIA interaction energies) and the biological activities, the Partial Least Squares (PLS) algorithm [59,60] was used. The CoMFA/CoMSIA descriptors were used as independent variables, and pIC₅₀ values as dependent variables in partial least square regression analysis. PLS is a statistical approach that generalizes and combines features from principal component analysis (PCA) and multiple regressions. It is particularly useful to predict a set of dependent variables from a large set of independent variables when the matrix of predictors has more variables than observations (multicollinearity).

To evaluate the reliability of the models and determine the optimal number of components for further non-cross-validated analysis, cross-validations were carried out using the “leave-one-out” option (LOO), wherein one compound was moved away from the dataset and its activity was predicted by using the model derived from the rest of the dataset. A cross-validated coefficient, Q^2 , was subsequently obtained and provided as a statistical index of predictive power. The highest Q^2 , corresponding to the smallest predicted error (*SEP*) and the optimal number of component, was used to obtain the final QSAR model. Then, the non-cross-validated models were evaluated by the Pearson coefficient (R^2_{ncv}), standard error of estimate (*SEE*), and *F* test ratio. Finally, the CoMFA/CoMSIA results were graphically represented by field contour maps, where the coefficients were generated using the field type “StDev*Coeff”.

In order to evaluate the real predictive ability of the best models generated by the CoMFA/CoMSIA analyses using the same training set, the pIC₅₀ values of test compounds are treated as the external validation set. The predictive ability of the model was evaluated by defining R^2_{pre} , which was then obtained with the following formula:

$$R^2_{pre} = \sqrt{(SD - PRESS)/SD} \quad (2)$$

where *SD* denotes the sum of squared deviation between the biological activities of the test set molecule and the mean activity of the training set molecules, *PRESS* represents the sum of squared deviations between the experimental and predicted activities of the test molecules.

3.5. DISCOtech Analysis

To identify the general pharmacophoric features the dataset was studied, at first, by the Distance Comparisons technique, a distance constraint pharmacophore building method. DISCOtech, a well-established module for designing pharmacophoric maps and frequently used in the process of virtual screening to discover new leads, is an enhanced, faster version of DISCO (DIStance Comparison) [61]. DISCOtech identifies features that could be elements in a pharmacophore model from a set of molecules that are related by their ability to bind to a common binding site [62]. DISCOtech, operating in distance space, can carry out clique detection to generate pharmacophore hypotheses for up to 300 conformers per molecule [63–65]. Elements considered in developing the pharmacophore model include hydrogen bond donor atoms, hydrogen bond acceptor atoms, hydrogen donor and acceptor site, charge centers, centers of mass of hydrophobic rings, aromatic rings, and positive NA (numerical aperture) stochastic search method was run to generate a maximum of 50 conformers for each molecule on the basis of maximum diversity to cover as many probable conformers as possible, and seven

conformers were then selected. Min 4 and Max 16 features were allowed to be found during the analysis. Tanimoto threshold was set as 0.85. All other parameters were retained as default values.

4. Conclusions

In the present work, a series of CoMFA, CoMSIA, and pharmacophore studies were performed on 151 imidazole-based TNF- α release inhibitors (see Figure 8). Both the reliability and predictivity of the resultant optimal CoMFA and CoMSIA models (with $Q^2 = 0.557$, $R^2_{ncv} = 0.740$, $R^2_{pre} = 0.748$ and $Q^2 = 0.598$, $R^2_{ncv} = 0.767$, $R^2_{pre} = 0.860$, respectively) were validated by their high Q^2 , R^2_{ncv} , and R^2_{pre} values. The corresponding contour maps generated by these models provide more diverse information about the key intermolecular interactions of inhibitors with the surrounding environment. A good consistency was also observed between the QSAR models and pharmacophore modeling studies. To sum up, our findings are:

1. Bulky substituents at R¹ position may improve the inhibitory activity.
2. Hydrophobic groups around R¹ substituent are helpful to enhance the potency of the inhibitors.
3. Electropositive groups at R³ position are beneficial to improve the biological activity of inhibitors.
4. H-bond donor groups around ring C and acceptor groups at R³ substituent promote the inhibitory activity, respectively.
5. Hydrophobic interaction and hydrogen bonds were the crucial factors acting on the inhibitory activity of TNF- α release.

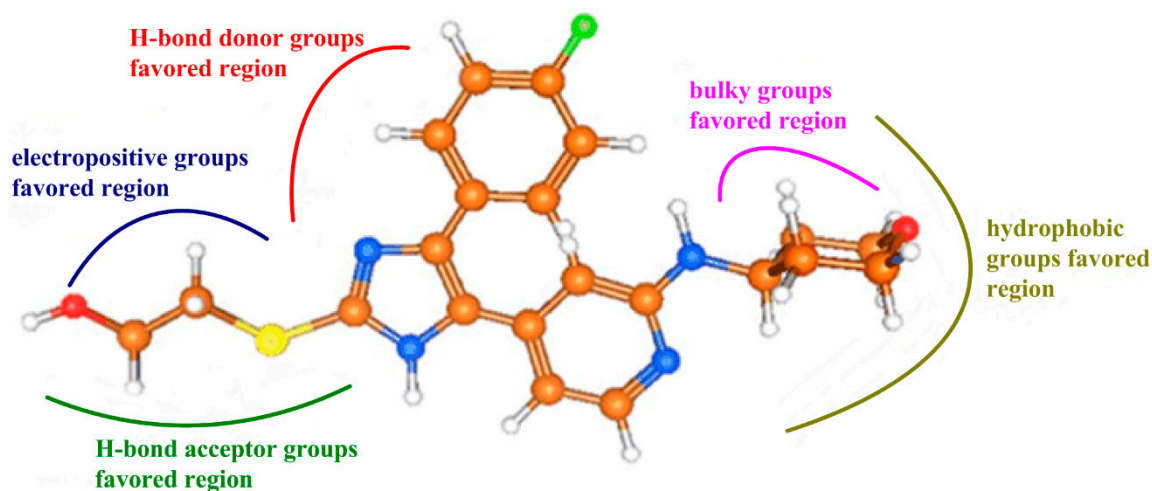


Figure 8. The interaction features of compound 3v impacting the antagonistic activity obtained from our present work.

We do hope the results of the present study may provide further support to the design of imidazoles as potential inhibitors of TNF- α release.

Supplementary Materials

Supplementary materials can be accessed at: <http://www.mdpi.com/1422-0067/16/09/20118/s1>.

Acknowledgments

Thanks for the support given by the high-performance computing platform of Northwest A&F (Agriculture and Forestry) University. This work is supported by grants from Northwest A&F University, National Natural Science Foundation of China (11201049 and 31170796), Program for New Century Excellent Talents in University of Ministry of Education of China, Interdisciplinary Research Matching Scheme (IRMS) of Hong Kong Baptist University (RC-IRMS/12-13/02) and Hong Kong Baptist University Strategic Development Fund (SDF13-1209-P01).

Author Contributions

Yuan Wang contributed to the analysis of the study and manuscript writing; Mingwei Wu contributed to the conception and instruction of the study; Chunzhi Ai performed the data analysis; Yonghua Wang helped perform DISCOtech Analysis.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Karin, M. Mitogen activated protein kinases as targets for development of novel anti-inflammatory drugs. *Anna. Rheum. Dis.* **2004**, *63*, ii62–ii64.
2. Laufer, S.A.; Hauser, D.R.; Domeyer, D.M.; Kinkel, K.; Liedtke, A.J. Design, synthesis, and biological evaluation of novel Tri- and tetrasubstituted imidazoles as highly potent and specific ATP-mimetic inhibitors of p38 MAP kinase: Focus on optimized interactions with the enzyme's surface-exposed front region. *J. Med. Chem.* **2008**, *51*, 4122–4149.
3. Tracey, K.J.; Cerami, A. Tumor necrosis factor, other cytokines and disease. *Ann. Rev. Cell Boil.* **1993**, *9*, 317–343.
4. Dinarello, C.A. Inflammatory cytokines: Interleukin-1 and tumor necrosis factor as effector molecules in autoimmune diseases. *Curr. Opin. Immunol.* **1991**, *3*, 941–948.
5. Elliott, M.J.; Maini, R.N.; Feldmann, M.; Kalden, J.R.; Antoni, C.; Smolen, J.S.; Leeb, B.; Breedveld, F.C.; Macfarlane, J.D.; Bijl, H.; *et al.* Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor α (cA2) *versus* placebo in rheumatoid arthritis. *The Lancet* **1994**, *344*, 1105–1110.
6. Elliott, M.J.; Maini, R.N.; Feldmann, M.; Long-Fox, A.; Charles, P.; Bijl, H.; Woody, J.N. Repeated therapy with monoclonal antibody to tumour necrosis factor α (cA2) in patients with rheumatoid arthritis. *The Lancet* **1994**, *344*, 1125–1127.
7. Rankin, E.C.; Choy, E.H.; Kassimos, D.; Kingsley, G.H.; Sopwith, A.M.; Isenberg, D.A.; Panayi, G.S. The therapeutic effects of an engineered human anti-tumour necrosis factor α antibody (CDP571) in rheumatoid arthritis. *Brit. J. Rheumatol.* **1995**, *34*, 334–342.
8. Van Dullemen, H.M.; van Deventer, S.J.; Hommes, D.W.; Bijl, H.A.; Jansen, J.; Tytgat, G.N.; Woody, J. Treatment of Crohn's disease with anti-tumor necrosis factor chimeric monoclonal antibody (cA2). *Gastroenterology* **1995**, *109*, 129–135.

9. Moreland, L.W.; Baumgartner, S.W.; Schiff, M.H.; Tindall, E.A.; Fleischmann, R.M.; Weaver, A.L.; Ettliger, R.E.; Cohen, S.; Koopman, W.J.; Mohler, K.; *et al.* Treatment of rheumatoid arthritis with a recombinant human tumor necrosis factor receptor (p75)-Fc fusion protein. *N. Engl. J. Med.* **1997**, *337*, 141–147.
10. Pierre, K.; Christiane, B.; Hartmut, J.; Stefan, L. Targeting the ribose and phosphate binding site of p38 mitogen-activated protein (MAP) kinase: Synthesis and biological testing of 2-alkylsulfanyl-, 4(5)-aryl-, 5(4)-heteroaryl-substituted imidazoles. *J. Med. Chem.* **2008**, *51*, 5630–5640.
11. Waetzig, G.H.; Schreiber, S. Review article: Mitogen-activated protein kinases in chronic intestinal inflammation—Targeting ancient pathways to treat modern diseases. *Aliment. Pharmacol. Ther.* **2003**, *18*, 17–32.
12. Fukami, Y.; Tokmakov, A.A.; Konaka, K.; Sato, K. Peptide inhibitors of the mitogen-activated protein kinase pathway: A structure-mimetic peptide corresponding to the conserved inter-DFG-APE region in the kinase domain. *Pharmacol. Ther.* **1999**, *82*, 399–407.
13. Engelberg, D. Stress-activated protein kinases-tumor suppressors or tumor initiators? *Semin. Cancer Biol.* **2004**, *14*, 271–282.
14. Dalrymple, S.A. p38 mitogen activated protein kinase as a therapeutic target for Alzheimer's disease. *J. Mol. Neurosci.* **2002**, *19*, 295–299.
15. Feng, Y.J.; Li, Y.Y. The role of p38 mitogen-activated protein kinase in the pathogenesis of inflammatory bowel disease. *J. Dig. Dis.* **2011**, *12*, 327–332.
16. Wagner, G.; Laufer, S. Small molecular anti-cytokine agents. *Med. Res. Rev.* **2006**, *26*, 1–62.
17. Saklatvala, J. The p38 MAP kinase pathway as a therapeutic target in inflammatory disease. *Curr. Opin. Pharmacol.* **2004**, *4*, 372–377.
18. Bracht, C.; Hauser, D.R.; Schattel, V.; Albrecht, W.; Laufer, S.A. Synthesis and biological testing of *N*-aminoimidazole-based p38 α MAP kinase inhibitors. *Chemmedchem* **2010**, *5*, 1134–1142.
19. Ziegler, K.; Hauser, D.R.; Unger, A.; Albrecht, W.; Laufer, S.A. 2-Acylaminopyridin-4-ylimidazoles as p38 MAP kinase inhibitors: Design, synthesis, and biological and metabolic evaluations. *Chemmedchem* **2009**, *4*, 1939–1948.
20. Ravindra, G.K.; Achaiah, G.; Sastry, G.N. Molecular modeling studies of phenoxyimidazole as p38 kinase inhibitors using QSAR and docking. *Eur. J. Med. Chem.* **2008**, *43*, 830–838.
21. Sperandio da Silva, G.M.; Sant'Anna, C.M.; Barreiro, E.J. A novel 3D-QSAR comparative molecular field analysis (CoMFA) model of imidazole and quinazolinone functionalized p38 MAP kinase inhibitors. *Bioorg. Med. Chem.* **2004**, *12*, 3159–3166.
22. Da Silva, G.M.S.; Lima, L.M.; Fraga, C.A.; Sant'Anna, C.M.; Barreiro, E.J. The molecular basis for coxib inhibition of p38 α MAP kinase. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3506–3509.
23. Xiao, Z.; Varma, S.; Xiao, Y.D.; Tropsha, A. Modeling of p38 mitogen-activated protein kinase inhibitors using the Catalyst™ HypoGen and *k*-nearest neighbor QSAR methods. *J. Mol. Graph. Model.* **2004**, *23*, 129–138.
24. Romeiro, N.C.; Albuquerque, M.G.; de Alencastro, R.B.; Ravi, M.; Hopfinger, A.J. Construction of 4D-QSAR models for use in the design of novel p38-MAPK inhibitors. *J. Comput. Aided. Mol. Des.* **2005**, *19*, 385–400.

25. Edraki, N.; Hemmateenejad, B.; Miri, R.; Khoshneviszade, M. QSAR study of phenoxyprymidine derivatives as potent inhibitors of p38 kinase using different chemometric tools. *Chem. Biol. Drug Des.* **2007**, *70*, 530–539.
26. Huang, C.; Li, Y.; Ren, H.; Wang, J.; Shao, L.; Zhang, S.; Li, G.; Yang, L. Insight into the structural determinants of imidazole scaffold-based derivatives as p38 MAP kinase inhibitors by computational explorations. *Curr. Med. Chem.* **2012**, *19*, 4024–4037.
27. Cramer, R.D.; Patterson, D.E.; Bunce, J.D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
28. Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130–4146.
29. Wei, S.P.; Ji, Z.Q.; Zhang, H.X.; Zhang, J.W.; Wang, Y.H.; Wu, W.J. Isolation, biological evaluation and 3D-QSAR studies of insecticidal/narcotic sesquiterpene polyol esters. *J. Mol. Model.* **2011**, *17*, 681–693.
30. Wang, X.; Xu, X.; Ma, M.; Zhou, W.; Wang, Y.; Yang, L. pH-Dependent channel gating in connexin26 hemichannels involves conformational changes in N-terminus. *Biochim. Biophys. Acta.* **2012**, *1818*, 1148–1157.
31. Xu, X.; Wang, X.; Li, Y.; Wang, Y.; Yang, L. A large-scale association study for nanoparticle C60 uncovers mechanisms of nanotoxicity disrupting the native conformations of DNA/RNA. *Nucleic Acids Res.* **2012**, *40*, 7622–7632.
32. Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.; Lee, K.; Tropsha, A. Rational selection of training and test sets for the development of validated QSAR models. *J. Comput. Aided Mol. Des.* **2003**, *17*, 241–253.
33. Golbraikh, A.; Tropsha, A. Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *Mol. Divers.* **2000**, *5*, 231–243.
34. Hemmateenejad, B.; Javadnia, K.; Elyasi, M. Quantitative structure-retention relationship for the Kovats retention indices of a large set of terpenes: A combined data splitting-feature selection strategy. *Anal. Chim. Acta* **2007**, *592*, 72–81.
35. Teuvo, K. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **1982**, *43*, 59–69.
36. Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*, 2nd ed.; Wiley: Weinheim, Germany; 1999.
37. Juha, V.; Esa, A. Clustering of the self-organizing map. *Neural Netw. IEEE Trans.* **2000**, *11*, 586–600.
38. Juha, V. SOM-based data visualization methods. *Intell. Data Analys.* **1999**, *3*, 111–126.
39. Liu, H.; Gramatica, P. QSAR study of selective ligands for the thyroid hormone receptor β . *Bioorg. Med. Chem.* **2007**, *15*, 5251–5261.
40. Fortin, S.; Wei, L.; Moreau, E.; Lacroix, J.; Cote, M.F.; Petitclerc, E.; Kotra, L.P.; R, C.G. Design, synthesis, biological evaluation, and structure-activity relationships of substituted phenyl 4-(2-oxoimidazolidin-1-yl)benzenesulfonates as new tubulin inhibitors mimicking combretastatin A-4. *J. Med. Chem.* **2011**, *54*, 4559–4580.
41. Hevener, K.E.; Ball, D.M.; Buolamwini, J.K.; Lee, R.E. Quantitative structure-activity relationship studies on nitrofuranyl anti-tubercular agents. *Bioorg. Med. Chem.* **2008**, *16*, 8042–8053.

42. Juvale, D.C.; Kulkarni, V.V.; Deokar, H.S.; Wagh, N.K.; Padhye, S.B.; Kulkarni, V.M. 3D-QSAR of histone deacetylase inhibitors: Hydroxamate analogues. *Org. Biomol. Chem.* **2006**, *4*, 2858–2868.
43. Raichurkar, A.V.; Kulkarni, V.M. Understanding the antitumor activity of novel hydroxysemicarbazide derivatives as ribonucleotide reductase inhibitors using CoMFA and CoMSIA. *J. Med. Chem.* **2003**, *46*, 4419–4427.
44. Cheeseright, T.J.; Holm, M.; Lehmann, F.; Luik, S.; Gottert, M.; Melville, J.L.; Laufer, S. Novel lead structures for p38 MAP kinase via FieldScreen virtual screening. *J. Med. Chem.* **2009**, *52*, 4200–4209.
45. Egan, W.J.; Morgan, S.L. Outlier detection in multivariate analytical chemical data. *Anal. Chem.* **1998**, *70*, 2372–2379.
46. Patel, P.D.; Patel, M.R.; Kaushik-Basu, N.; Talele, T.T. 3D QSAR and molecular docking studies of benzimidazole derivatives as hepatitis C virus NS5B polymerase inhibitors. *J. Chem. Inf. Model* **2008**, *48*, 42–55.
47. Verma, R.P.; Hansch, C. An approach toward the problem of outliers in QSAR. *Bioorg. Med. Chem.* **2005**, *13*, 4597–4621.
48. Barken, F.M.; Gasteiger, E.L. Excitability of a penicillin-induced cortical epileptic focus. *Exp. Neurol.* **1980**, *70*, 539–547.
49. O’Neil, D.W.; Clark, M.V.; Lowe, J.W.; Harrington, M.S. Oral trauma in children: A hospital survey. *Oral Surg. Oral Med. Oral Pathol.* **1989**, *68*, 691–696.
50. Alexander, G.; Alexander, T. Beware of q^2 ! *J. Mol. Graph. Model.* **2002**, *20*, 269–276.
51. Melagraki, G.; Afantitis, A. Enalos KNIME nodes: Exploring corrosion inhibition of steel in acidic medium. *Chemometr. Intell. Lab.* **2013**, *123*, 9–14.
52. AbdulHameed, M.D.M.; Hamza, A.; Liu, J.; Zhan, C.G. Combined 3D-QSAR modeling and molecular docking study on indolinone derivatives as inhibitors of 3-phosphoinositide-dependent protein kinase-1. *J. Chem. Inf. Model.* **2008**, *48*, 1760–1772.
53. Da Cunha, E.F.F.; Sippl, W.; de Castro Ramalho, T.; Ceva Antunes, O.A.; de Alencastro, R.B.; Albuquerque, M.G. 3D-QSAR CoMFA/CoMSIA models based on theoretical active conformers of HOE/BAY-793 analogs derived from HIV-1 protease inhibitor complexes. *Eur. J. Med. Chem.* **2009**, *44*, 4344–4352.
54. Duda-Seiman, C.; Duda-Seiman, D.; Dragos, D.; Medeleanu, M.; Careja, V.; Putz, M.V.; Lacrama, A.-M.; Chiriac, A.; Nutiu, R.; Ciubotariu, D. Design of anti-HIV ligands by means of minimal topological difference (MTD) method. *Int. J. Mol. Sci.* **2006**, *7*, 537–555.
55. Duda-Seiman, D.M.; Speranța, A.; Mancaș, S.; Careja, V.; Duda-Seiman, C.; Putz, M.V.; Ciubotariu, D. MTD-CoMSIA modelling of HMG-CoA reductase inhibitors. *J. Serb. Chem. Soc.* **2011**, *76*, 85–99.
56. Avram, S.; Milac, A.; Mernea, M.; Mihailescu, D.; Putz, M.V.; Buiu, C. Structure–Biological function relationship extended to mitotic arrest-deficient 2-like protein Mad2 native and mutants-new opportunity for genetic disorder control. *Int. J. Mol. Sci.* **2014**, *15*, 21381–21400.
57. Speranta, A.; Maria, M.; Dan, F.M.; Corina, D.S.; Daniel, D.S.; Mihai V.P. Mitotic checkpoint proteins Mad1 and Mad2 structural and functional relationship with implication in genetic diseases. *Curr. Comput. Aided Drug Des.* **2013**, *2*, 1573–4099.

58. Putz, M.V.; Duda-Seiman, C.; Duda-Seiman, D.M.; Putz, A.M. Turning SPECTRAL-SAR into 3D-QSAR analysis. Application on H⁺K⁺-ATPase inhibitory activity. *Int. J. Chem. Model.* **2008**, *1*, 45–62.
59. Jain, A.N. Surfex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46*, 499–511.
60. Sprous, D.G.; Palmer, R.K.; Swanson, J.T.; Lawless, M. QSAR in the pharmaceutical research setting: QSAR models for broad, large problems. *Curr. Top Med. Chem.* **2010**, *10*, 619–637.
61. Zhang, M.Q.; Zhang, X.L.; Li, Y.; Fan, W.J.; Wang, Y.H.; Hao, M.; Zhang, S.W.; Ai, C.Z. Investigation on Quantitative Structure Activity Relationships and Pharmacophore Modeling of a Series of mGluR2 Antagonists. *Int. J. Mol. Sci.* **2011**, *12*, 5999–6023.
62. Bondavalli, F.; Botta, M.; Bruno, O.; Ciacci, A.; Corelli, F.; Fossa, P.; Lucacchini, A.; Manetti, F.; Martini, C.; Menozzi, G.; *et al.* Synthesis, molecular modeling studies, and pharmacological activity of selective A1 receptor antagonists. *J. Med. Chem.* **2002**, *45*, 4875–4887.
63. Liu, G.Y.; Ju, X.L.; Cheng, J.; Liu, Z.Q. 3D-QSAR studies of insecticidal anthranilic diamides as ryanodine receptor activators using CoMFA, CoMSIA and DISCOtech. *Chemosphere* **2010**, *78*, 300–306.
64. Ju, X.L.; Hao, Y.L.; Pei, J.F.; Ozoe, Y. Investigation of structural requirements for inhibitory activity at the rat and housefly picrotoxinin binding sites in ionotropic GABA receptors using DISCOtech and CoMFA. *Chemosphere* **2007**, *69*, 864–871.
65. Jung, D.; Floyd, J.; Gund, T.M. A comparative molecular field analysis (CoMFA) study using semiempirical, density functional, *ab initio* methods and pharmacophore derivation using DISCOtech on sigma 1 ligands. *J. Comput. Chem.* **2004**, *25*, 1385–1399.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).