Available online at www.sciencedirect.com
ScienceDirect
European Journal of Medicinal Chemistry 43 (2008) 2724–2734
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
http://www.elsevier.com/locate/ejmech

Original article

Three-dimensional QSAR analyses of 1,3,4-trisubstituted pyrrolidine-based CCR5 receptor inhibitors

Ya Zhuo[1], Ren Kong[1], Xiao-jing Cong, Wei-zu Chen, Cun-xin Wang*

College of Life Science and Bioengineering, Beijing University of Technology, Ping Le Yuan 100, Chao Yang District, Beijing 100124, China

Received 13 November 2007; received in revised form 14 January 2008; accepted 14 January 2008
Available online 8 February 2008

Abstract

In this study, a series of 1,3,4-trisubstituted pyrrolidine-based CCR5 receptor inhibitors were taken as our target with the method of the three-dimensional quantitative structure-activity relationship (3D-QSAR) analyses in order to investigate the interactions between CCR5 receptor and their inhibitors. For a comparison, Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) were, respectively, used to build predictive models, which were generated from a training set consisting of 72 selected molecules, derived from literatures. Two alignment rules, including rigid body rms (root mean square) fit and field fit, were performed in the superimposition of inhibitors structures. As a result, a better CoMFA model based on common structure alignment obtains a conventional correlation coefficient r^2 of 0.952 and a leave-one-out cross-validated coefficient q^2 of 0.637, while the desirable CoMSIA model based on the same alignment rule acquires the r^2 of 0.958 and the q^2 of 0.677. To further validate the reliability of the models, we also investigated into the externally test set composed of 39 molecules under the criterions of squared correlation coefficient between experimental and predicted activities with intercept R^2 and without intercept R_0^2, along with R_m^2 as the modified R^2 with a penalty function due to difference between R^2 and R_0^2. At last, the contour map also provides a visual representation of contributions of steric, electrostatic, hydrogen bond and hydrophobic fields, as well as the prospective binding modes. These results may provide meaningful guidance to the further work including the similar lead compounds' structure modification and activity prediction.

Keywords: Computer-aided drug design; 3D-QSAR analysis; CoMFA and CoMSIA methods; Inhibitors of CCR5 receptor

1. Introduction

With rapid progress in exploration of HIV infection processes, it was found by recent studies that in addition to the CD4 receptor, a new class of seven-membrane-domain receptors called chemokine receptors was proved to play a crucial role in the membrane-fusion stage of HIV infection [1–3]. In the early stage of HIV infection, the virus tends to attack the immune cells by sequentially binding to the CD4 receptor and chemokine receptors on the cell surface, then the membrane-infusion can be achieved [4,5]. Recognized as a member of the chemokine receptor family, CCR5 was discovered to be utilized in the early stage of the replication cycle by the most commonly transmitted M-tropic strains of HIV-1. Notable findings showed that a few individuals genetically bearing a defective CCR5 allele were protected from HIV-1 infection without any unhealthy consequence [6–8]. Hence the idea of setting CCR5 as a possible target for therapeutic intervention was brought up and well supported by evidence that blocking the function of CCR5 could strongly inactivate HIV virus, resulting in effectively prevention of HIV-1 entering into cells while exhibiting few side effects [7,8].

On the basis of these studies mentioned above, extensive exploration into this potential target for anti-HIV treatments has motivated the development of some CCR5 inhibitors as a new group of anti-HIV therapeutics. It is reported that monoclonal antibodies [9] and endogenous ligands of CCR5 receptor, including RANTES, MIP-1α and MIP-1β [10], as well as

* Corresponding author. Tel.: +86 10 67392724; fax: +86 10 67392837.
E-mail address: cxwang@bjut.edu.cn (C.-x. Wang).
[1] Both authors contributed equally.

0223-5234/$ - see front matter © 2008 Published by Elsevier Masson SAS.
doi:10.1016/j.ejmech.2008.01.040

their modified analogues [11,12], could exhibit anti-HIV activity. Furthermore, the development of small molecule CCR5 inhibitors has drawn much attention and efforts for several advantages. First, instead of dealing with the HIV virus itself which is widely used as the major drug target, CCR5 inhibitors will aim to work on the structure binding to the human cells, probably avoiding the problem of viral mutation and drug resistance. Second, small molecule inhibitors can be produced as oral medicines more easily and can overcome the bioavailability problems associated with large molecules [13–16]. Some of these small molecular antagonists are still under clinical research [17,18]; however, none of them yet appear on the market. It is obvious that this promising class of anti-HIV drugs is in great demand while facing the serious situation of AIDS epidemics.

By means of rational drug design methods, this paper aims to work on a series of active CCR5 inhibitors binding to the receptor and provide better understanding of structure-activity relationships of CCR5 receptor and its inhibitors, which may offer some practical guidelines for further modification of CCR5 antagonists lead compounds.

Recent years, computer-aided drug design has contributed to elucidating active relationships between drug molecules and their receptors, gaining deep insight into the mechanism of receptor–antagonist interaction, which is used to find modification requirements for highly active drug molecules. With computer-aided ligand-based design methods advanced, some 3D-QSAR methods can closely delineate the structural features and binding characteristics of the CCR5 receptor active sites. Actually, previous studies using different QSAR methods including Molecular Docking, Pharmacophore Analysis, CoMFA (Comparative Molecular Field Analysis) and CoMSIA (Comparative Similarity Indices Analysis) method, MSA (Molecular Shape Analysis), RSA (Receptor Surface Analysis) and MFA (Molecular Field Analysis) have also shed some light on the related types of CCR5 receptor inhibitors and have sort of inspired our research in terms of methodology and discussion [19–25]. In this paper, the 3D-QSAR methods, Comparative Molecular Field Analysis (CoMFA) [26] and its extension Comparative Similarity Indices Analysis (CoMSIA) [27] were applied to generate a few predictive models, which may identify the physicochemical features of CCR5 inhibitors and provide essential information to increase their binding affinity.

2. Materials and methodology

2.1. Biological data

In the present study, a total of 111 pyrrolidine-based CCR5 inhibitors were utilized to build 3D-QSAR models with the biological activity data collected from the literatures [15,16,28–45]. The binding affinity was represented as IC_{50} (nmol L^{-1}), α binding to the CCR5 expressing Chinese hamster ovary (CHO) cell in a high-throughput assay against [^{125}I]-MIP-1α [46]. All the compounds developed by the Merck Research Laboratories, were confirmed the consistency of the biological assay [15,16,28–45]. Initially, the compounds were partitioned into a training set with 72 compounds and a test set consisting of 39 compounds, with the training set mainly used to develop the QSAR models and the test set employed to assess the predictive quality of the constructed models. As for the division rule for training and test set, previous studies have also provided some available strategies such as random selection, activity-range algorithm and *K*-means-cluster based selection [19,20,47]. Of the investigation they have made, the separation rule was emphasized to the key point that how to make the training set to represent the entire data set to the most. Hence here, we adopted the method to focus on the diversity of structures and biological activity of the compounds so that the training set and test set could mostly cover the variances of structures and activities. To start with, we roughly separated all the compounds into 10 classes according to their structural features. Afterwards, we sorted the compounds of each class based on the biological activity. And the selection was conducted from each class to ensure enough and essential structural variations of the compounds to be studied, furthermore the selection of each class also aimed to cover the range of activity as wide as possible. Finally, the IC_{50} value of the training set molecules covered a range from 0.06 to 200 nmol L^{-1}. For the benefit of the following calculation, all biological activities were converted into the corresponding $-\log IC_{50}$ values (pIC_{50}). The chemical structures and biological activities of the training set and test set are listed in the Supplementary material.

2.2. Molecular modeling

All computational studies were performed in the molecular modeling package SYBYL 7.0 on a SGI Octane workstation. To start with, all molecules were constructed in the state of neutral form using the SYBYL sketch program. It is of great importance to determine the suitable active conformation of all molecules, which would greatly affect the interaction between drug molecules and the receptor taking a crucial role in CoMFA operation as well. Due to the limited structural information of the receptor, it is impossible to obtain the exactly practical conformation of each molecule, but the determination of optimum conformation was strived in many effective methods ranging from the molecular docking [23,48], systematic search [19,20] and pharmacophore model selection [49] for the best fit to the reality. In this work, a systematic conformational search was adopted concerning the flexibility of the substituents. Because of the highest potency, the compound **310** was set as the template. Every rotatable bond in the compound **310** was identified and rotated at the step of 10° through the whole space. Then, the lowest energy conformation calculated in this search was selected as a template to build the other molecular structures. Partial atomic charges were assigned to each atom using the Gasteiger–Hückel method. All structures of compounds were optimized using a conjugate gradient minimizer and the standard Tripos force field with a distance-dependent dielectric function. The minimization was terminated when the energy gradient convergence criterion of 0.001 kcal/mol was achieved or when the 5000-step minimization cycle limit was exceeded.

2.3. Structural alignment

Since there is critical requirement of structure alignment in 3D-QSAR, the alignment rule remains to be a crucial and controversial process in 3D-QSAR analyses. In this study, two widely used alignment rules were adopted for comparison to ensure a successfully reasonable model. Alignment I was recognized as the common substructure-based alignment in which all molecules were superimposed on the template molecule according to the selected common backbone. And the best alignment module was obtained through the optimization of atom/centroid root mean square (RMS). The common backbone used for alignment is illustrated in Fig. 1. Apparently, this method would generate a better result from compounds with resembling structures. Alignment II was known as a field fit alignment. It attempted to adjust the geometry of the molecule in terms of steric and electrostatic fields to match the fields of the template molecule as close as possible. The resultant alignments of training set molecules employing these rules are shown in Fig. 2. In fact, it should be noted that variations occurring to conformations of drug molecules exhibited by the receptor were not taken into account by both alignment rules. Since it is not sensible to determine which method would be more superior, we were able to make sure the alignment rule comparably suitable to molecules in this study.

3. CoMFA and CoMSIA analysis

After the alignment of training set molecules, the 3D-QSAR interaction fields would be established to perform the CoMFA [26] and CoMSIA [27] calculations. In general, the following standard parameters were used to calculate the CoMFA fields: a 3D cubic lattice with a grid spacing of 2.0 Å and extending 4.0 Å beyond the volumes of all studied molecules; a probe atom represented by a sp3-hybrized carbon atom charged of +1.0 and a distance-dependent dielectric constant. Both standard CoMFA fields including steric field and electrostatic field were calculated, respectively, under the rules of Lennard-Jones 6-12 and Coulomb potentials with the energy cutoff of 30 kcal/mol. As to the CoMSIA analysis, which was improved on the basis of CoMFA, it is supposed to consider more parameters by adding steric, electrostatic, hydrogen bond (donor and acceptor) and hydrophobic field, as well as a sp3-hybrized carbon atom probe with a radius of 1.0 Å, a charge of +1, hydrophobicity of +1 and hydrogen bond properties of +1. When the atom probe was placed at each grid point of the same lattice with CoMFA to measure the related physicochemical properties, the CoMSIA similarity indices for a molecule were calculated by the following Gaussian-type expression:

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

where $A_{F,K}$ stands for the similarity indices of a molecule j and ω_{ik} represents the actual value of the physicochemical property while $\omega_{\text{probe},k}$ is the probe atom and r_{iq} is the mutual distance between the probe atom at the grid point q and the atom i of the molecule. A default value 0.3 was selected for the attenuation factor α.

Consequently, a spreadsheet, comprising of conventional CoMFA or CoMSIA descriptors of each molecule and the corresponding bioactivity value was built to perform the multiple regression analysis. Because thousands of actual parameters were estimated by CoMFA and CoMSIA, which made the dependent data much smaller, it was appropriate to select the Partial Least Squares (PLS) regression as the critical tool. In PLS regression analysis, the CoMFA and CoMSIA descriptors were set as explanatory variables and pIC_{50} values were set as dependent variables. It usually started with the cross-validation with Leave-One-Out (LOO), which was carried out under the condition of column filtering and a predicted number of components. After the optimal number of components was determined, a non-cross-validated analysis was performed without column filtering to build a final 3D-QSAR model. The optimal number of components tends to produce the smallest root mean predictive sum of squared errors, usually corresponding to the highest cross-validated coefficient q^2. In addition, a set of parameters including q^2, Spress (cross-validated standard error of prediction), r^2 (non-cross-validated coefficient), standard error of estimate (SEE) and F values were, respectively, computed according to the definitions in the PLS procedures. Those parameters are useful indicators of the quality of the constructed models. It is acceptable to gain the higher q^2 value more than 0.5 and an r^2 value more than 0.9, which may confirm a model predictable to some degree. In fact, the cross-validated coefficient q^2 was calculated to indicate the linear relevant relationship of independent and dependent variables by the following equation:

$$q^2 = 1 - \frac{\sum\left(Y_{\text{pred}} - Y_{\text{exp}}\right)^2}{\sum\left(Y_{\text{pred}} - Y_{\text{mean}}\right)^2}$$

where Y_{pred}, Y_{exp}, Y_{mean} are predicted, experimental and mean values of the target property (pIC_{50}), respectively. As for the values of Spress and SE, the well-built models are more preferable to have small Spress and SEE.

Notably, after all the parameters mentioned above that could sort of reveal the predictive ability of the resultant model, the acceptable results of them may not sufficient to guarantee a reliable model. According to some incisive studies on validation of the QSAR modeling [50,51], the high value of the LOO cross-validation q^2 alone do not automatically reflect

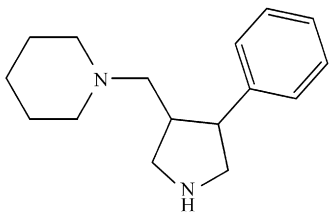

Fig. 1. Common substructure alignment.

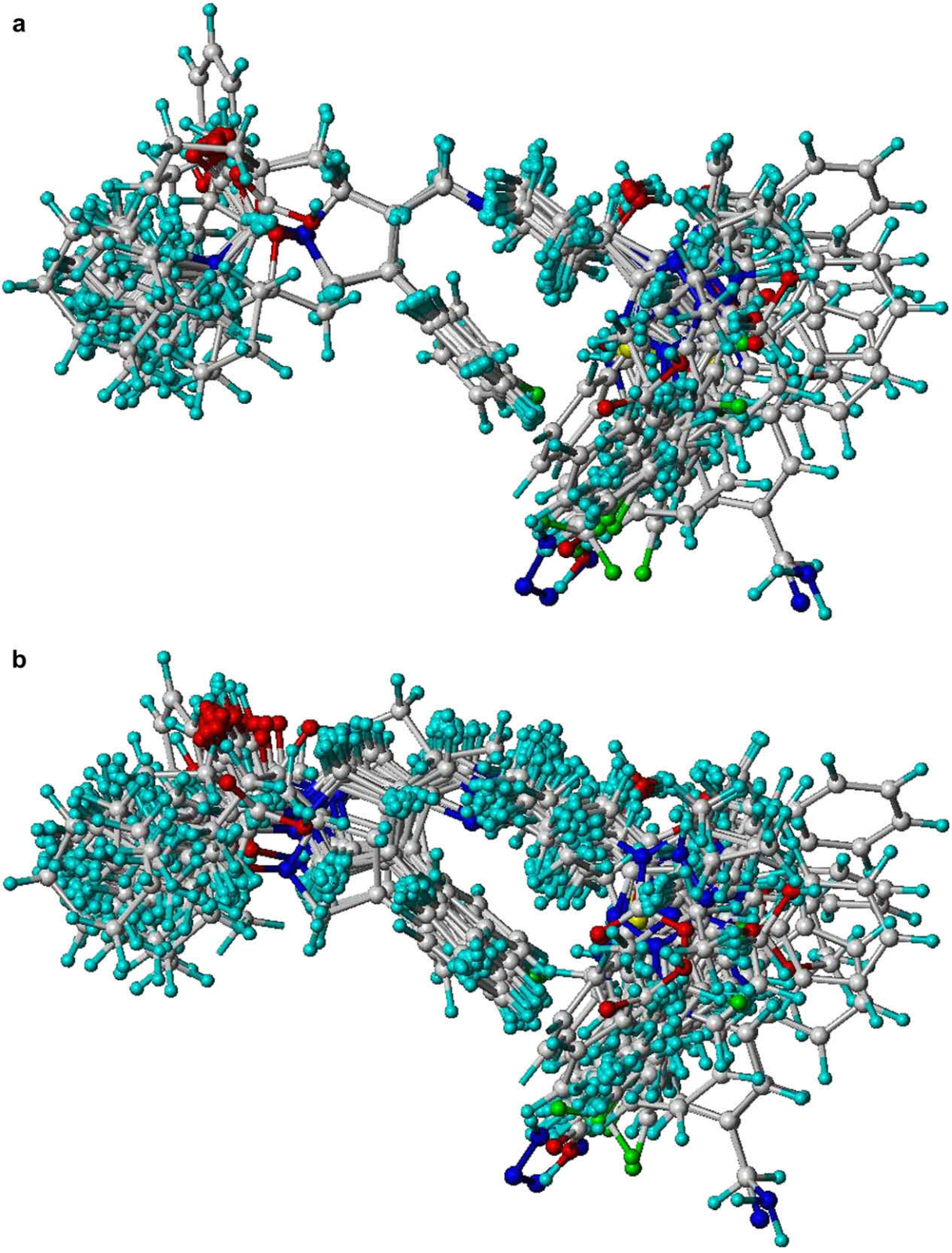

Fig. 2. Superposition of compounds in the training set: (a) common substructure-based alignment; (b) field fit alignment.

a general predictability of the models, the only effective way to judge the true predictive power of derived models is to test them on the external test set. Hence, in addition to q^2 and r^2, we also implemented the calculation of R^2 (squared correlation coefficient between observed and predicted activities), R_0^2 (R^2 with intercept set at zero) and R_m^2 (a modified R^2 with a penalty function due to difference between R^2 and R_0^2) [50,51] from the test set predication in order to better validate

the predictive power of QSAR models. The R_0^2 and R_m^2 are calculated as follows:

$$R_0^2 = 1 - \frac{\sum(y_i - y_i^{r_0})^2}{\sum(y_i - \overline{y_i})^2}$$

$$R_m^2 = R^2\left(1 - \left|\sqrt{R^2 - R_0^2}\right|\right)$$

Here, y_i refers to the observed activity of each compound from the test set, $y_i^{r_0}$ indicates the corresponding value of the predicted activity of each compound based on the regression line with the intercept at zero, $\overline{y_i}$ stands for the mean value of all the observed activities from the test set compounds.

Specifically, both applied to correlate the observed and predicted values of the test set compounds, R^2 and R_0^2 were used to describe the regression line of the test set prediction with and without intercept, respectively. Based on the expressions shown above, if the angle between the two regression lines is small, the closer R^2 must be to R_0^2 and also to the corresponding R_m^2, which are necessary to judge the predictive results from the test set. In conclusion, when the values of each parameter (R^2, R_0^2 and R_m^2) exhibited approximate and also greater than 0.5, it can exactly imply the predicted values of test set compounds are close to the corresponding experimental values and thus prove a highly predictive QSAR model.

4. Results and discussion

4.1. CoMFA and CoMSIA statistical results

The statistical results of CoMFA and CoMSIA studies are summarized in Table 1. Actually it is quite common to generate some poor models at the beginning. In this work, some initiatives have been tried to improve the results, including choosing a different conformation of the poorly predicted molecules and changing parameters in the QSAR exercise. The resultant q^2 is ultimately enhanced close to 0.6 while r^2 is around 0.95. To make a further initiate, another way accomplished in the Advanced CoMFA employing Region Focusing to increase q^2 through selective re-weight of the grid points in a region and reducing the random but cross-correlated brown noise in the data matrix [52,53].

According to Table 1, the CoMFA1 model ($q^2 = 0.653$, $r^2 = 0.855$, with three components) resulted in much higher quality against the CoMFA3 model ($q^2 = 0.588$, $r^2 = 0.802$, with two components). After improvement of these two models performed by focusing regions, the CoMFA2 model ($q^2 = 0.637$, $r^2 = 0.952$, with five components) was still superior to the CoMFA4 model ($q^2 = 0.622$, $r^2 = 0.794$, with two components). With respect to predictions of 39 test compounds, the CoMFA2 model also achieved more desirable correlated coefficient R^2 of 0.616 and standard deviation (SD) of 0.384 than the CoMFA4 model ($R^2 = 0.531$, SD = 0.357). Meanwhile, the R_0^2 of 0.615 and R_m^2 of 0.597 generated from the CoMFA2 model further strengthened its good predictive power by showing acceptable results both in closeness and size. In comparison, the CoMFA4 model exhibited a little poorer with $R_0^2 = 0.528$, $R_m^2 = 0.501$. In the case of CoMSIA

Table 1
Summary of CoMFA and CoMSIA results for various models

	CoMFA1	CoMFA2	CoMFA3	CoMFA4	CoMSIA1	CoMSIA2	CoMSIA3	CoMSIA4
Alignment[a]	I	I	II	II	I	I	II	II
Region focusing[b]	N	Y	N	Y	N	Y	N	Y
q^2	0.653	0.637	0.588	0.622	0.618	0.677	0.649	0.673
Spress	0.490	0.509	0.530	0.508	0.522	0.487	0.501	0.483
r^2	0.855	0.952	0.802	0.794	0.936	0.958	0.941	0.939
SEE	0.317	0.186	0.367	0.375	0.214	0.176	0.205	0.209
Components	3	5	2	2	5	7	5	5
F value	133.835	259.873	140.097	133.165	192.035	207.208	210.172	202.338
Fraction								
Steric	0.458	0.373	0.371	0.431	0.089	0.146	0.060	0.079
Electrostatic	0.542	0.627	0.629	0.569	0.284	0.280	0.311	0.291
Donor	–	–	–	–	0.249	0.195	0.220	0.206
Acceptor	–	–	–	–	0.189	0.161	0.225	0.229
Hydrophobic	–	–	–	–	0.189	0.218	0.184	0.194
Test set prediction								
R^2	–	0.616	–	0.531	–	0.526	–	0.482
SD	–	0.384	–	0.357	–	0.467	–	0.431
R_0^2	–	0.615	–	0.528	–	0.524	–	0.469
R_m^2	–	0.597	–	0.501	–	0.503	–	0.427

R^2 stands for squared correlation coefficient of the regression line with intercept in the test set prediction; SD represents standard deviation of test set prediction; R_0^2 stands for squared correlation coefficient of the regression line without intercept in the test set prediction; R_m^2 stands for difference between values of R^2 and R_0^2.

[a] Indicates the alignment rules adopted in the CoMFA and CoMSIA models.

[b] Indicates whether the process of region focusing was used in the CoMFA and CoMSIA models, while N means not and Y means yes.

analyses, the results turned a little different, where the CoMSIA3 model produced a higher q^2 of 0.649 and an r^2 of 0.941, with five components than the CoMSIA1 model ($q^2 = 0.618$, $r^2 = 0.936$, with five components). It was also observed that the process of region focusing has greatly enhanced the CoMSIA models, which was reflected by the CoMSIA2 model with the q^2 of 0.677 and the r^2 of 0.958 bearing seven components as well as the CoMSIA4 model ($q^2 = 0.673$, $r^2 = 0.939$, with five components). The corresponding Spress was also reduced from 0.522 and 0.501 to 0.487 and 0.483, respectively. Additionally the SEE was cut down from 0.214 and 0.205 to 0.176 and 0.209, respectively. In spite of those CoMSIA results in which Alignment I exhibited more advantageous than Alignment II, it is necessary to extend further validation on the test set. In fact, the CoMSIA2 model acquired more favorable predictability of compounds in the test set, with R^2 of 0.526 and SD of 0.467 than the CoMSIA4 model with R^2 of 0.482 and SD of 0.431. Additionally, the CoMSIA2 model was also confirmed to be highly reliable by achieving greater R_0^2 and R_m^2 of 0.524 and 0.503, respectively, compared with the poorer results from CoMSIA4 model ($R_0^2 = 0.469$, $R_m^2 = 0.427$). As described before, statistical results from PLS analyses can reflect the predictable quality to some extent, however, the validation of test set compounds weighed much more on the basis of comprehensive judgment. Hence, it is revealed that the common substructure alignment exhibits more fitness to this group of compounds to be studied due to their relative similarities. Actually in the CoMSIA studies, different fields have been included to find out the best combination and it turned out that using exactly five fields could provide an adequate description for the receptor–ligand interaction for this series of CCR5 inhibitors. It can be found that there is no significant difference between the best model referred to CoMFA2 and the best model of CoMSIA2 and both models are selected to make further analyses and conclusions.

In addition, the performance of prediction on the molecules in training set derived from both the CoMFA2 and the CoMSIA2 model suggests a good internal consistency in the training set with relevant coefficient of 0.975 and 0.979, standard deviation of 0.176 and 0.165, respectively. The detailed experimental and predicted pIC_{50} built on the CoMFA2 and CoMSIA2 models for the training compounds are shown in Table 2, while the test compounds results are shown in Table 3. To

Table 2
Experimental pIC_{50} and the corresponding predicted values for the training compounds of the CoMFA and CoMSIA models through region focusing

Compound	pIC_{50}	CoMFA2	CoMSIA2	Compound	pIC_{50}	CoMFA2	CoMSIA2
240	−0.778	−0.748	−0.848	**437**	−1.258	−1.441	−1.303
241	−1.079	−1.042	−1.179	**438**	−1.431	−1.374	−1.339
242	−2	−2.123	−2.043	**439**	−1.362	−1.148	−1.309
243	−0.477	−0.596	−0.456	**440**	−0.792	−0.869	−0.675
244	−1	−0.724	−1.007	**444**	−0.301	−0.256	−0.423
248	0.699	0.799	0.758	**445**	0.538	0.417	0.383
260	0.097	0.176	−0.014	**448**	0.137	0.327	0.235
261	−0.301	−0.153	−0.157	**449**	0.420	0.373	0.369
262	−0.477	−0.285	−0.468	**451**	−0.204	0.028	0.19
263	−0.778	−0.656	−0.780	**452**	0.469	0.084	0.218
265	0.155	−0.104	−0.154	**454**	0.638	0.322	0.383
266	−0.602	−0.696	−0.566	**458**	0.051	−0.165	0.083
268	−0.477	−0.206	0.096	**459**	0.119	0.098	0.318
270	1	0.847	0.802	**461**	0	−0.25	−0.013
274	−0.954	−1.131	−1.127	**499**	−1.799	−1.820	−1.684
275	−1	−1.177	−1.231	**500**	−1.362	−1.065	−1.125
276	−1.079	−0.998	−1.163	**501**	−0.447	−0.660	−0.555
279	−1.362	−1.374	−1.295	**506**	−1.114	−1.252	−1.28
280	−1.820	−1.694	−1.668	**510**	−0.544	−0.457	−0.564
286	−0.820	−0.756	−0.909	**514**	−1.462	−1.382	−1.385
299	0.097	0.053	0.142	**515**	−2.301	−2.004	−2.387
301	0.699	0.861	0.528	**523**	−0.079	0.086	−0.133
303	0.046	0.222	0.198	**526**	−0.079	−0.048	−0.009
304	1	0.843	1.196	**527**	−0.176	0.012	0.012
309	0.301	0.671	0.624	**535**	−0.477	−0.514	−0.521
310	1.222	0.807	0.844	**539**	−1.301	−1.232	−1.15
312	0.523	0.464	0.588	**540**	−0.362	−0.416	−0.469
314	0.398	0.221	−0.071	**542**	−0.342	−0.139	−0.268
316	−0.681	−0.473	−0.528	**543**	−0.623	−0.698	−0.688
317	−0.204	−0.038	−0.029	**549**	0.097	−0.174	−0.003
319	−0.176	−0.028	−0.326	**550**	−0.079	−0.13	−0.097
323	−1.477	−1.886	−1.413	**551**	−1.342	−1.349	−1.182
325	−1.663	−1.799	−1.75	**552**	−1.204	−1.418	−1.306
331	−1.519	−1.454	−1.365	**553**	−2.086	−2.086	−2.09
332	−1.935	−1.798	−1.893	**554**	−0.322	−0.273	−0.361
333	−1.146	−0.974	−1.100	**558**	−0.839	−0.94	−0.873

Table 3
Experimental pIC_{50} and the corresponding predicted values for the test compounds of the CoMFA and CoMSIA models through region focusing

Compound	pIC_{50}	CoMFA2	CoMSIA2	Compound	pIC_{50}	CoMFA2	CoMSIA2
239	−0.845	−0.792	−1.138	**447**	−0.845	−0.796	−0.176
245	−0.778	−1.079	−1.044	**450**	−0.580	0.012	0.399
246	−0.699	−0.763	−1.261	**453**	0.292	−0.178	0.084
251	0.398	−0.034	−0.690	**455**	0.638	0.179	0.406
253	−0.556	−0.249	−0.845	**456**	0.244	0.153	0.058
254	0	−0.410	−0.923	**457**	0.222	0.223	0.549
258	−0.602	−0.201	−0.727	**495**	−0.398	−1.241	−0.881
259	−0.699	−0.789	−0.766	**508**	−1.079	−1.372	−1.166
267	−1.556	−1.159	−0.856	**511**	−0.477	−0.944	−1.059
272	0.699	−0.041	−0.103	**513**	−1.813	−1.573	−1.651
281	−1.079	−1.811	−1.604	**525**	−0.079	0.004	0.022
285	−0.699	−1.044	−0.310	**528**	−0.204	0.043	−0.049
288	−0.255	−0.665	−1.060	**531**	0.097	−0.129	0.103
302	0.097	0.190	0.205	**532**	0.046	0.067	0.013
311	0.699	0.682	0.948	**537**	−0.681	−1.352	−1.013
313	0.523	0.483	0.650	**541**	−0.322	−0.419	−0.162
315	−0.568	−0.171	−0.199	**544**	−0.663	−0.016	−0.121
442	0.174	0.453	0.398	**545**	−0.477	−0.789	−0.709
443	−0.114	−0.428	−0.532	**557**	−1.061	−0.580	−0.884
446	−0.176	−0.972	−0.794				

visually view the predictions, the scatter plot of experimental versus predicted binding affinity pIC_{50} for the training set and the test set are shown in Fig. 3. It is clear that compounds both in the training set and test set from selected CoMFA and CoMSIA models are distributed evenly across the line of $y = x$ with little distracted ones, which emphasized their favorable ability of prediction again.

4.2. Comparison of CoMFA and CoMSIA contour maps

One of the most pleasant advantages of CoMFA and CoMSIA is that the field effect on the target property can be mapped as 3D coefficient contour plots. As shown in Figs. 4–6, the coefficient contour maps provide a distinguished observation of important regions where any change in the steric,

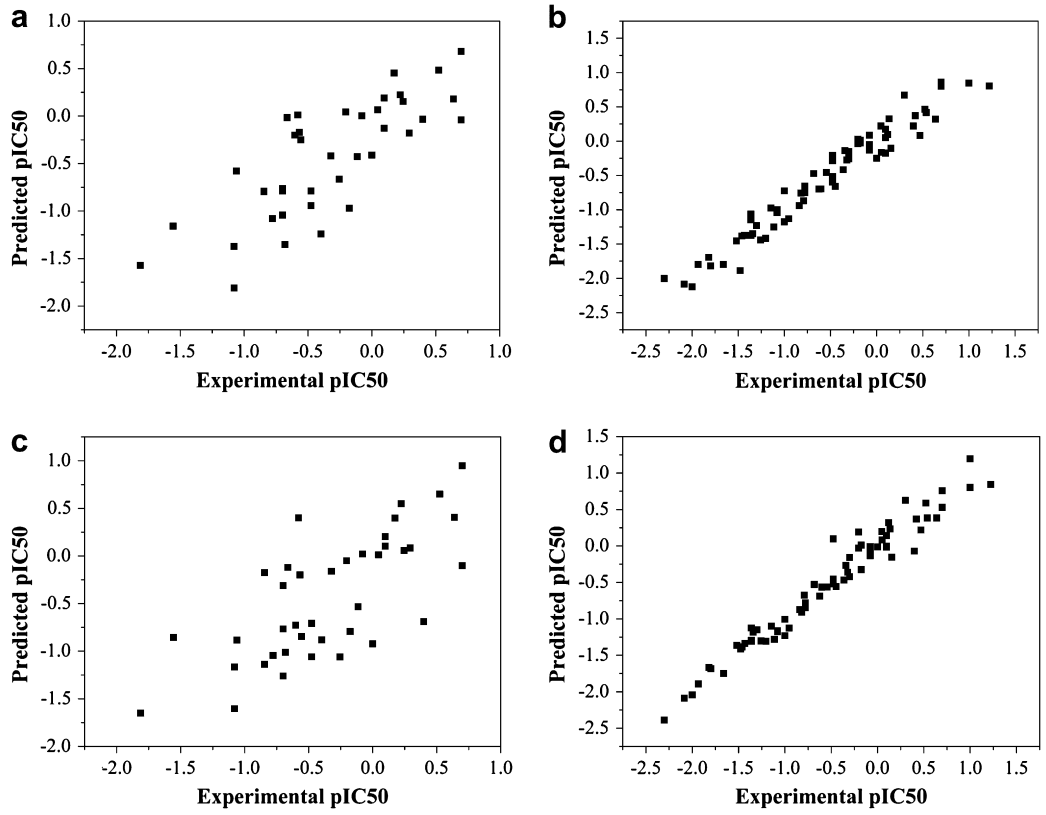

Fig. 3. Predicted versus experimental activity (pIC_{50}, nM) of compounds. The test set (a) and the training set (b) for the CoMFA2 model. The test set (c) and the training set (d) for the CoMSIA2 model.

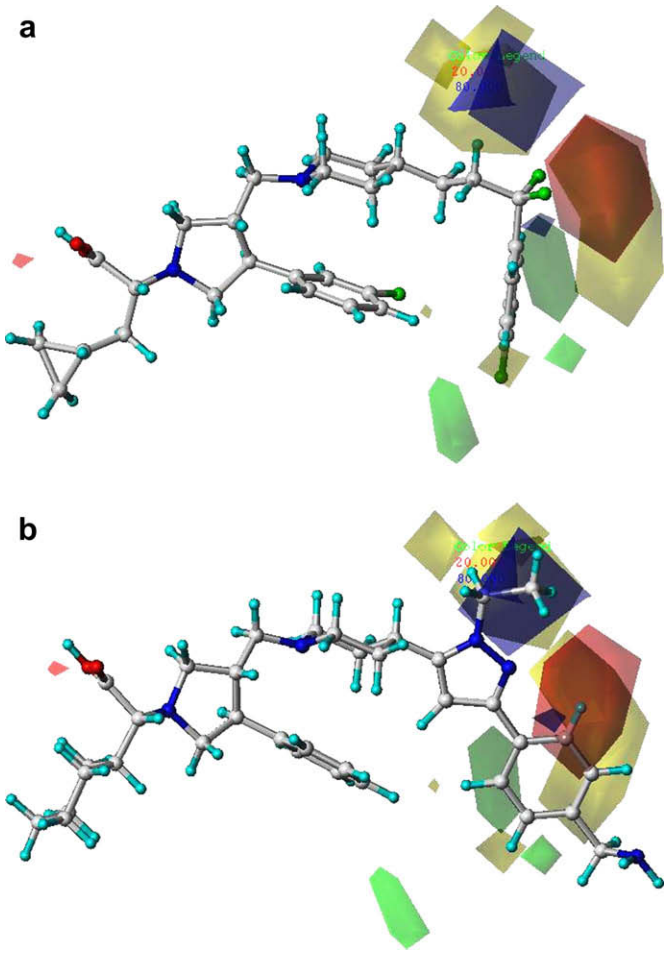

Fig. 4. Steric and electrostatic maps from the CoMFA model of compound **310** (a) and **553** (b) are displayed as comparison. Sterically favored areas (contribution level of 80%) and disfavored areas (contribution level of 20%) are represented by green polyhedron and yellow polyhedron, respectively. Blue and red polyhedron depicts the favorable site for positively (80%) and negatively (20%) charged groups, respectively.

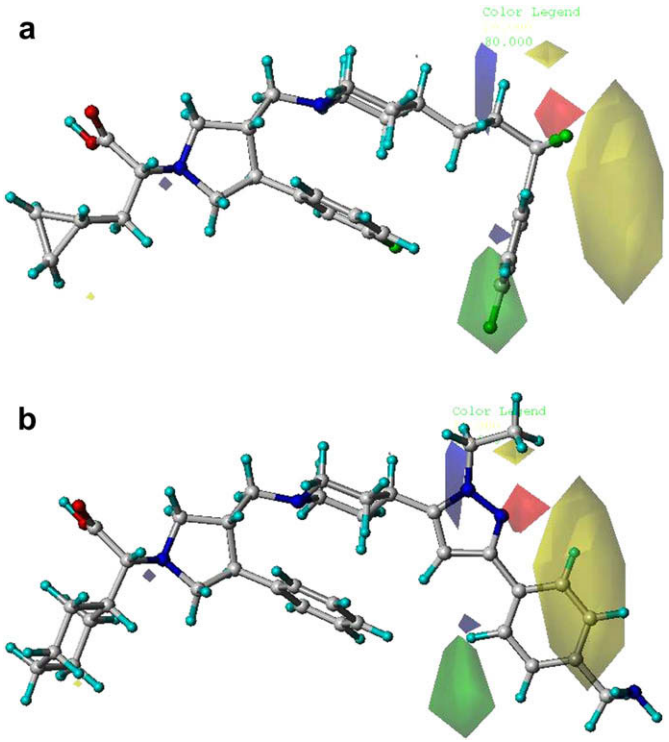

Fig. 5. Steric and electrostatic maps from the CoMSIA model of compound **310** (a) and **553** (b) are exhibited above, respectively. Sterically favored areas (contribution level of 80%) and disfavored areas (contribution level of 20%) are represented by green polyhedron and yellow polyhedron, respectively. Blue and red polyhedron depicts the favorable site for positively (80%) and negatively (20%) charged groups, respectively.

electrostatic, hydrophobic and hydrogen bond fields may affect the biological activity. It also can make a qualitative judgment and identify the possible interaction sites, which may offer instructive guidance for the structural modification of lead compounds. While the CoMFA contribution maps highlight the 'pseudo-receptor' region where aligned molecules would interact favorably or unfavorably. The CoMSIA method can give a diagram that allows each field contribution to be mapped back directly to regions within molecular structures so that information regarding the contribution of particular functional groups in certain areas may be evaluated [54]. Yet, the combination of two different approaches can be applied to verify the convergence of the results and complement conclusions of each other as well [48]. In Figs. 4–6, the contour diagrams of the field contributions determined by different properties and obtained from analyses are illustrated with the exemplary ligands.

The most potent compound **310** is shown superimposed on the important steric and electrostatic field regions in Fig. 4a, from which the green polyhedron characterize the regions where bulky substituents would increase the binding affinity, whereas yellow contours depict regions where steric bulk would not be tolerated. The blue and red polyhedron denotes the favorable sites for positively and negatively charged groups, respectively. It can be observed that most interaction revealed by the contour maps tends to focus at the right side of the molecular structures, left the other side of chains lack of information. Meanwhile, the distribution of steric and electrostatic fields is found to integrate together to some degree. However, taking a fully consideration of selected molecules in the study, the disproportion of fields distribution may result from more intricate variation at the right side chains rather than the limited structural changes on the left side. Furthermore, the steric and electrostatic contour maps from the CoMSIA model are demonstrated to be consistent with the CoMFA results, which obviously display the serious requirement around the area of right side chains. Despite the similarity in steric and electrostatic fields of the CoMFA and CoMSIA model, yet their polyhedra have appeared some difference in size. The rather smaller contour maps of electrostatic fields can be attributed to the more comprehensive concerns of hydrogen bond and hydrophobic fields introduced in CoMSIA approach. As displayed in Fig. 5a and Fig. 6a, the color scheme for steric and electrostatic fields is represented in the same manner as used for CoMFA. Additionally, the donor field contour shows regions where hydrogen bond donors are predicted to enhance (cyan) or disfavor (purple) binding. As

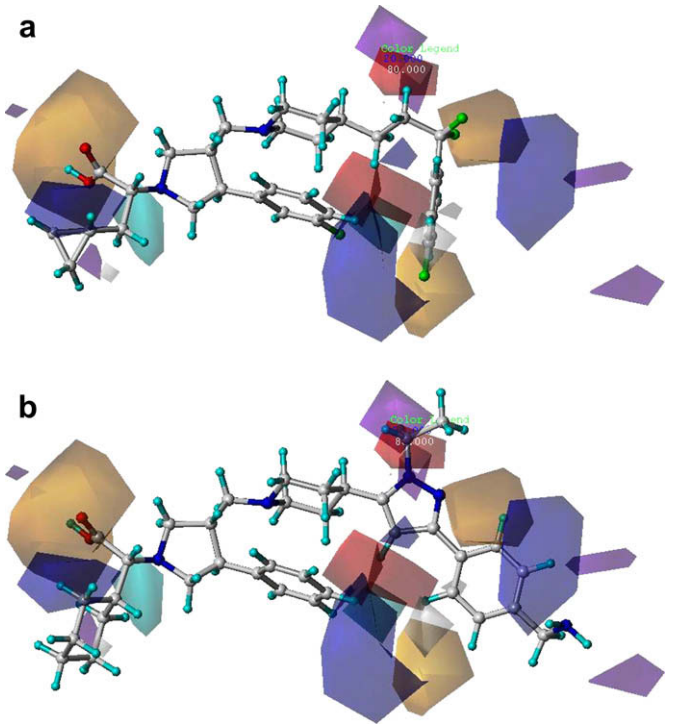

Fig. 6. Hydrogen bond and hydrophobic maps from the CoMSIA model of compound **310** (a) and **553** (b) are exhibited above, respectively. The hydrogen bond donor field represented by cyan (80%)/purple (20%) contours shows regions where donors are predicted to enhance or disfavor binding. The hydrogen bond acceptor field represented by orange (80%)/red (20%) contours encloses areas where acceptors are expected to increase affinity or reduce it. White/blue contours indicate the regions where hydrophobic (80%)/hydrophilic (20%) groups would enhance the binding affinity.

to the hydrogen bond acceptor field, the orange and red contours enclose areas where acceptors are expected to enhance affinity or reduce it, respectively. Blue and white contours of the CoMSIA model indicate the area where hydrophilic and hydrophobic properties are preferred, respectively. With its advantages, the CoMSIA results were able to reflect the hydrogen bond and hydrophobic properties of molecules, which has provided some helpful notes at the left area. As proved by the PLS analysis and test set prediction above, the CoMSIA model actually exhibited much higher predictive capacity than the CoMFA model. On the whole, the message generated from the predictive CoMFA and CoMSIA contour maps will be of great importance and help.

4.3. Structure-activity analyses based on contour maps

Within the CoMFA contour map shown in Fig. 4, two medium-sized yellow polyhedra lying in the right region include an upper one mainly indicating that bulky branches attached to substituents at the vicinity of 4-position of the piperidine ring would not be favorable, and a lower steric disfavorable diagram with restriction of substituents on the side and end chains. Therefore the substituents along the side chains of 4-position of the piperidine ring would prefer small branches and flexible ones over the huge substituents, as can be seen in a series of highly active compounds like compound **310** ($IC_{50} = 0.06$ nM) presented in Fig. 4a as well as compound **298** ($IC_{50} = 0.5$ nM), **301** ($IC_{50} = 0.2$ nM) and **304** ($IC_{50} = 0.1$ nM). By contrast, compound **553** ($IC_{50} = 122$ nM) shown in Fig. 4b on behalf of the poorly active compounds bears an ethyl rightly intruding the upper yellow region and a phenyl ring conflicting with the other steric disfavorable region. The improper orientation of redundancy branches such as ethyl and allyl growing at the position of disfavorable steric interaction may lead to the low binding affinities of compounds such as **280** ($IC_{50} = 66$ nM) and **332** ($IC_{50} = 86$ nM). Additionally, there is a blue polyhedron favorable for positive charges under the forbidden steric region related above, which suggests a strict requirement for substitution in the vicinity of the 4-position of the piperidine ring. The result may also explain the poor activity of compound **553** on account of the negative nitrogen on the pyrazole ring closely approaching the positive favorable region. To enhance binding activities, this region can be modified by moderate groups bearing fewer branches and positive partial charges. Notably, there is a red polyhedron favorably for negative charges embedded in the lower yellow polyhedron, which implies that this region should be limited to substituents with small bulk and negative partial charges. The hypothesis is in agreement with the comparison of compound **310** and **553** as mentioned above, for the presence of two fluorine atoms in the first carbon attached to the aromatic ring effectively improve the potent of compound **310** while compound **553** hardly meets this requirement. In the case of a series of molecules like compound **311** ($IC_{50} = 0.2$ nM), **313** ($IC_{50} = 0.3$ nM) and **312** ($IC_{50} = 0.3$ nM), the introduction of halogen over hydrogen in this area is more favorable for binding affinity. The observation of better binding affinity in compound **315** ($IC_{50} = 3.7$ nM) and **319** ($IC_{50} = 1.5$ nM) also demonstrates these regions would prefer small groups rich in electrons.

Turning to the CoMSIA resultant contour maps in Fig. 5, the steric distribution represented by a huge yellow polyhedron and a green region lying below, is found to be in agreement with the CoMFA results. Specifically, the green polyhedra characterize the regions where bulky substituents would enhance the binding affinity. The observation is well proved in compounds **301** ($IC_{50} = 0.2$ nM), **270** ($IC_{50} = 0.1$ nM) and **454** ($IC_{50} = 0.23$ nM), which exhibit higher activity than compounds **438** ($IC_{50} = 27$ nM) and **513** ($IC_{50} = 65$ nM). Besides, the red polyhedron located on the right side most highlight electron-rich substituents in the molecular structure, which can be observed in another series of molecules such as compound **445** ($IC_{50} = 0.29$ nM), **454** ($IC_{50} = 0.23$ nM) and **442** ($IC_{50} = 0.67$ nM). Hence, the occupation of nitrogen atom on the ring at this area may have a positive effect on the activity. From the cyan contours around substituents at 1-position of pyrrolidine ring, it can be concluded that hydrogen bond donors only enhance the activity when located near this area, likewise the orange contours largely lying in the left and underside define the appearance of acceptors in this area are expected to enhance binding. The purple and red contours probably identify these

regions as poor place for hydrogen bond donor groups or acceptors to exist. According to the comparison of compound **310** and **553** shown in Fig. 6a and b, the introduction of electron-rich fluorine atoms of compound **310** in the underside and right orange regions is probably proved to enhance its potent to some extent. Seemly, the upper orange contour is in accordance with the negatively favorable red region, which suggests the space in the vicinity of substituents at 4-position of piperidine ring is restricted to small branches bearing negative charges or hydrogen bond acceptor features for the purpose of better binding affinity.

Inspection of the hydrophobic field contribution map illustrates that it is in accordance with the analysis of the steric and hydrogen bond fields. One blue contour representing a zone of favorable hydrophilic interactions appears in the proximity of substituents at 1-position of pyrrolidine ring where hydrogen-bonding interactions are found to enhance affinity. This result could be supported by the general existence of the −COOH group in this area, which may provide the oxygen atom conducive to the interaction. The hydrophilic blue polyhedron that appears near the upper-right corner of the map tends to more or less explain the poor activity of compound **553** due to the interference of a phenyl ring without little polarity. Besides, there is a rather small hydrophobic white polyhedron locates near the region of the green contour, which implies what matters to the activity at this region is not the polarity of the substituent but the bulk. Last but not the least, the lower-right green polyhedron surrounded by an orange and a blue contours indicates this region is favorable for a bulky substituents bearing the hydrophilic and hydrogen bond receptor quality to increase the interaction between CCR5 receptor and inhibitors. As shown in compound **310**, the fluorine atom in the *ortho*-position of phenyl ring instead of a hydrogen definitely increases the size of the group as well as the capacity of hydrophilic and hydrogen bond receptor, so that promote its formation with the corresponding receptor.

5. Conclusions

The 3D-QSAR CoMFA and CoMSIA methods were utilized to study a series of 1,3,4-trisubstituted pyrrolidine-based CCR5 receptor inhibitors comprising a training set of 72 molecules and a test set of 39 molecules. To derive an effective and reasonable model, different approaches were taken involving the optimum conformation, alignment rules and region focusing. Moreover, the resulting models have demonstrated potential predictive capability through PLS procedure and test set validation. High statistical results of q^2 and optimal examination of external test set were observed in both CoMFA and CoMSIA models based on Alignment I, supporting much confidence to these analyses. On the basis of CoMFA and CoMSIA contour maps, a combination of steric, electrostatic, hydrogen bond and hydrophobic fields was investigated to realize some important structural features for inhibitors behavior. The maps also identified crucial structural requirements for inhibitors to enhance the activity. Referred as a rational outline, this helpful information should contribute to the structural optimization of potent CCR5 inhibitors lead compounds and activity prediction of selective molecules belonging to this series. In the future, there is no doubt that more efforts should be taken to find optimum solutions for the development of new CCR5 inhibitors.

Acknowledgements

Project supported by the National Natural Science Foundation of China (No. 30500429 and No. 30670497) and the Beijing Natural Science Foundation (No. 5072002). We also extend our great thanks to Dr. Chang Shan of Beijing University of Technology for the discussion about some statistical work in the paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejmech.2008.01.040.

References

[1] H.K. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. DiMarzio, S. Marmon, R.E. Sutton, C.M. Hill, C.B. Davis, S.C. Peiper, T.J. Schall, D.R. Littman, N.R. Landau, Nature 381 (1996) 661–666.
[2] T. Dragic, V. Litwin, G.P. Allaway, S.R. Martin, Y.X. Huang, K.A. Nagashima, C. Cayanan, P.J. Maddon, R.A. Koup, J.P. Moore, W.A. Paxton, Nature 381 (1996) 667–673.
[3] A.S. Fauci, Nature 384 (1996) 529–534.
[4] J.P. Moore, Curr. Opin. Immunol. 9 (1997) 551–562.
[5] W.A. Paxton, R. Liu, S. Kang, L.J. Wu, T.R. Gingeras, N.R. Landau, C.R. Mackay, R.A. Koup, Virology 244 (1998) 66–73.
[6] R. Liu, W.A. Paxton, S. Choe, D. Ceradini, S.R. Martin, R. Horuk, M.E. MacDonald, H. Stuhlmann, R.A. Koup, N.R. Landau, Cell 86 (1996) 367–377.
[7] M. Marmor, H.W. Sheppard, D. Donnell, S. Bozeman, C. Celum, S. Buchbinder, B. Koblin, G.R. Seage, J. Acquir. Immune Defic. Syndr. 27 (2001) 472–481.
[8] M. Samson, F. Libert, B.J. Doranz, J. Rucker, C. Liesnard, C.M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y.J. Yi, R.J. Smyth, R.G. Collman, R.W. Doms, G. Vassart, M. Parmentier, Nature 382 (1996) 722–725.
[9] W.C. Olson, G.E.E. Rabut, K.A. Nagashima, D.N.H. Tran, D.J. Anselma, S.P. Monard, J.P. Segal, D.A.D. Thompson, F. Kajumo, Y. Guo, J.P. Moore, P.J. Maddon, T. Dragic, J. Virol. 73 (1999) 4145–4155.
[10] F. Cocchi, A.L. Devico, A. Garzinodemo, S.K. Arya, R.C. Gallo, P. Lusso, Science 270 (1995) 1811–1815.
[11] G. Simmons, P.R. Clapham, L. Picard, R.E. Offord, M.M. Rosenkilde, T.W. Schwartz, R. Buser, T.N.C. Wells, A.E.I. Proudfoot, Science 276 (1997) 276–279.
[12] N. Heveker, M. Montes, L. Germeroth, A. Amara, A. Trautmann, M. Alizon, J. Schneider-Mergener, Curr. Biol. 8 (1998) 369–376.
[13] O.M.Z. Howard, J.J. Oppenheim, M.G. Hollingshead, J.M. Covey, J. Bigelow, J.J. McCormack, R.W. Buckheit, D.J. Clanton, J.A. Turpin, W.G. Rice, J. Med. Chem. 41 (1998) 2184–2193.
[14] J.M. Strizki, S. Xu, N.E. Wagner, L. Wojcik, J. Liu, Y. Hou, M. Endres, A. Palani, S. Shapiro, J.W. Clader, W.J. Greenlee, J.R. Tagat, S. McCombie, K. Cox, A.B. Fawzi, C.C. Chou, C. Pugliese-Sivo, L. Davies, M.E. Moreno, D.D. Ho, A. Trkola, C.A. Stoddart,

J.P. Moore, G.R. Reyes, B.M. Baroudy, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 12718–12723.
[15] P.E. Finke, L.C. Meurer, B. Oates, S.G. Mills, M. MacCoss, L. Malkowitz, M.S. Springer, B.L. Daugherty, S.L. Gould, J.A. DeMartino, S.J. Siciliano, A. Carella, G. Carver, K. Holmes, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, W.A. Schleif, E.A. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 265–270.
[16] C.P. Dorn, P.E. Finke, B. Oates, R.J. Budhu, S.G. Mills, M. MacCoss, L. Malkowitz, M.S. Springer, B.L. Daugherty, S.L. Gould, J.A. DeMartino, S.J. Siciliano, A. Carella, G. Carver, K. Holmes, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, W.A. Schleif, E.A. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 259–264.
[17] O.M.Z. Howard, T. Korte, N.I. Tarasova, M. Grimm, J.A. Turpin, W.G. Rice, C.J. Michejda, R. Blumenthal, J.J. Oppenheim, J. Leukocyte Biol. 64 (1998) 6–13.
[18] M. Baba, O. Nishimura, N. Kanzaki, M. Okamoto, H. Sawada, Y. Iizawa, M. Shiraishi, Y. Aramaki, K. Okonogi, Y. Ogawa, K. Meguro, M. Fujino, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 5698–5703.
[19] Y.D. Aher, A. Agrawal, P.V. Bharatam, P. Garg, J. Mol. Model. 13 (2007) 519–529.
[20] M.S. Song, C.M. Breneman, N. Sukumar, Bioorg. Med. Chem. 12 (2004) 489–499.
[21] J. Thomas Leonard, K. Roy, Bioorg. Med. Chem. Lett. 16 (2006) 4467–4474.
[22] A. Afantitis, G. Melagraki, H. Sarimveis, P.A. Koutentis, J. Markopoulos, O. Igglessi-Markopoulou, J. Comput. Aided Mol. Des. 20 (2006) 83–95.
[23] Yong Xu, Hong Liu, Chunying Niu, C. Luo, X. Luo, J. Shen, K. Chen, H. Jiang, Bioorg. Med. Chem. 12 (2004) 6193–6208.
[24] K. Roy, J.T. Leonard, J. Chem. Inf. Model. 45 (2005) 1352–1368.
[25] J.T. Leonard, K. Roy, QSAR Comb. Sci. 23 (2004) 387–398.
[26] R.D. Cramer, D.E. Patterson, J.D. Bunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[27] G. Klebe, U. Abraham, T. Mietzner, J. Med. Chem. 37 (1994) 4130–4146.
[28] D.-M. Shen, M. Shu, S.G. Mills, K.T. Chapman, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, G.Y. Kwei, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M.D. Miller, E.A. Emini, Bioorg. Med. Chem. Lett. 14 (2004) 935–939.
[29] K. Shankaran, K.L. Donnelly, S.K. Shah, C.G. Caldwell, P. Chen, P.E. Finke, B. Oates, M. MacCoss, S.G. Mills, J.A. DeMartino, S.L. Gould, L. Malkowitz, S.J. Siciliano, M.S. Springer, G. Kwei, A. Carella, G. Carver, R. Danzeisen, D. Hazuda, K. Holmes, J. Kessler, J. Lineberger, M.D. Miller, E.A. Emini, W.A. Schleif, Bioorg. Med. Chem. Lett. 14 (2004) 3589–3593.
[30] D.-M. Shen, M. Shu, C.A. Willoughby, S. Shah, C.L. Lynch, J.J. Hale, S.G. Mills, K.T. Chapman, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, K. Lyons, J.V. Pivnichny, G.Y. Kwei, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M.D. Miller, E.A. Emini, Bioorg. Med. Chem. Lett. 14 (2004) 941–945.
[31] M. Shu, J.L. Loebach, K.A. Parker, S.G. Mills, K.T. Chapman, D.-M. Shen, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, J.D. Salvo, K. Lyons, J.V. Pivnichny, G.Y. Kwei, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M.D. Miller, E.A. Emini, Bioorg. Med. Chem. Lett. 14 (2004) 947–952.
[32] P.E. Finke, L.C. Meurer, B. Oates, S.K. Shah, J.L. Loebach, S.G. Mills, M. MacCoss, L. Castonguay, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, Bioorg. Med. Chem. Lett. 11 (2001) 2469–2473.
[33] P.E. Finke, B. Oates, S.G. Mills, M. MacCoss, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, A. Carella, G. Carver, K. Holmes, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, W.A. Schleif, E.A. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 2475–2479.
[34] C.L. Lynch, A.L. Gentry, J.J. Hale, S.G. Mills, M. MacCoss, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, M.A. Cascieri, G. Doss, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, E.A. Emini, Bioorg. Med. Chem. Lett. 12 (2002) 677–679.
[35] C.A. Willoughby, S.C. Berk, K.G. Rosauer, S. Degrado, K.T. Chapman, S.L. Gould, M.S. Springer, L. Malkowitz, W.A. Schleif, D. Hazuda, M. Miller, J. Kessler, R. Danzeisen, K. Holmes, J. Lineberger, A. Carella, G. Carver, E.A. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 3137–3141.
[36] D. Kim, L. Wang, C.G. Caldwell, P. Chen, P.E. Finke, B. Oates, M. MacCoss, S.G. Mills, L. Malkowitz, S.L. Gould, J.A. DeMartino, M.S. Springer, D. Hazuda, M. Miller, J. Kessler, R. Danzeisen, G. Carver, A. Carella, K. Holmes, J. Lineberger, W.A. Schleif, E.A. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 3103–3106.
[37] D. Kim, L. Wang, C.G. Caldwell, P. Chen, P.E. Finke, B. Oates, M. MacCoss, S.G. Mills, L. Malkowitz, S.L. Gould, J.A. DeMartino, M.S. Springer, D. Hazuda, M. Miller, J. Kessler, R. Danzeisen, G. Carver, A. Carella, K. Holmes, J. Lineberger, W.A. Schleif, E.A. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 3099–3102.
[38] J.N. Burrows, J.G. Cumming, S.M. Fillery, G.A. Hamlin, J.A. Hudson, R.J. Jackson, S. McLaughlin, J.S. Shaw, Bioorg. Med. Chem. Lett. 15 (2005) 25–28.
[39] J.G. Cumming, S.J. Brown, A.E. Cooper, A.W. Faull, A.P. Flynn, K. Grime, J. Oldfield, J.S. Shaw, E. Shepherd, H. Tucker, D. Whittaker, Bioorg. Med. Chem. Lett. 16 (2006) 3533–3536.
[40] D. Kim, L. Wang, J.J. Hale, C.L. Lynch, R.J. Budhu, M. MacCoss, S.G. Mills, L. Malkowitz, S.L. Gould, J.A. DeMartino, M.S. Springer, D. Hazuda, M. Miller, J. Kessler, R.C. Hrin, G. Carver, A. Carella, K. Henry, J. Lineberger, W.A. Schleif, E.A. Emini, Bioorg. Med. Chem. Lett. 15 (2005) 2129–2134.
[41] C.A. Willoughby, K.G. Rosauer, J.J. Hale, R.J. Budhu, S.G. Mills, K.T. Chapman, M. MacCoss, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, M.A. Cascieri, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, E.A. Emini, Bioorg. Med. Chem. Lett. 13 (2003) 427–431.
[42] J.J. Hale, R.J. Budhu, S.G. Mills, M. MacCoss, L. Malkowitz, S. Siciliano, S.L. Gould, J.A. DeMartino, M.S. Springer, Bioorg. Med. Chem. Lett. 11 (2001) 1437–1440.
[43] J.J. Hale, R.J. Budhu, E.B. Holson, P.E. Finke, B. Oates, S.G. Mills, M. MacCoss, S.L. Gould, J.A. DeMartino, M.S. Springer, S. Siciliano, L. Malkowitz, W.A. Schleif, D. Hazuda, M. Miller, J. Kessler, R. Danzeisen, K. Holmes, J. Lineberger, A. Carella, G. Carver, E. Emini, Bioorg. Med. Chem. Lett. 11 (2001) 2741–2745.
[44] J.J. Hale, R.J. Budhu, S.G. Mills, M. MacCoss, S.L. Gould, J.A. DeMartino, M.S. Springer, S.J. Siciliano, L. Malkowitz, W.A. Schleif, D. Hazuda, M. Miller, J. Kessler, R. Danzeisen, K. Holmes, J. Lineberger, A. Carella, G. Carver, E.A. Emini, Bioorg. Med. Chem. Lett. 12 (2002) 2997–3000.
[45] C.L. Lynch, J.J. Hale, R.J. Budhu, A.L. Gentry, S.G. Mills, K.T. Chapman, M. MacCoss, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, M.A. Cascieri, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, E.A. Emini, Bioorg. Med. Chem. Lett. 12 (2002) 3001–3004.
[46] S.J. Siciliano, S.E. Kuhmann, Y. Weng, N. Madani, M.S. Springer, J.E. Lineberger, R. Danzeisen, M.D. Miller, M.P. Kavanaugh, J.A. DeMartino, D. Kabat, J. Biol. Chem. 274 (1999) 1905.
[47] K. Roy, J.T. Leonard, QSAR Comb. Sci. 25 (2006) 235–251.
[48] J.K. Buolamwini, H. Assefa, J. Med. Chem. 45 (2002) 841–852.
[49] H. Yuan, A.P. Kozikowski, P.A. Petukhov, J. Med. Chem. 47 (2004) 6137–6143.
[50] K. Roy, P.P. Roy, QSAR Comb. Sci. 9999 (2007). doi:10.1002/qsar.200710043.
[51] A. Golbraikh, A. Tropsha, J. Mol. Graph. Model. 20 (2002) 269–276.
[52] M. Forina, C. Casolino, C.P. Millan, J. Chemometr 13 (1999) 165–184.
[53] F. Lindgren, P. Geladi, S. Rannar, S. Wold, J. Chemometr. 8 (1994) 349–363.
[54] G. Klebe, Perspect. Drug Discov. Design 12 (1998) 87–104.