J. Chosun Natural Sci.
Vol. 10, No. 2 (2017) pp. 78 – 84
https://doi.org/10.13160/ricns.2017.10.2.78

Hologram Based QSAR Analysis of CXCR-2 Inhibitors

Sathya. B†

Abstract

CXC chemokine receptor 2 (CXCR2) is a prominent chemokine receptor on neutrophils. CXCR2 antagonist may reduce the neutrophil chemotaxis and alter the inflammatory response because the neutrophilic inflammation in the lung diseases is found to be largely regulated through CXCR2 receptor. Hence, in the present study, Hologram based Quantitative Structure Activity Relationship Study was performed on a series of CXCR2 antagonist named pyrimidine-5-carbonitrile-6-alkyl derivatives. The best HQSAR model was obtained using atoms, bonds, and chirality as fragment distinction parameter using hologram length 151 and 6 components with fragment size of minimum 4 and maximum 7. Significant cross-validated correlation coefficient (q^2=0.774) and non cross-validated correlation coefficients (r^2=0.977) were obtained. The model was then used to evaluate the six external test compounds and its r^2_{pred} was found to be 0.614. Contribution map show that presence of cyclopropyl ring and its bulkier substituent's makes big contributions for improving the biological activities of the compounds. We hope that our HQSAR model and analysis will be helpful for future design of novel and structurally related CXCR2 antagonists.

Keywords: CXCR2, HQSAR

1. Introduction

Chemokines are G-protein-coupled receptors (GPCRs) which have long been implicated in the initiation and amplification of inflammatory responses by their role in leukocyte chemotaxis[1,2]. Chemokines are small 8-10 kDa proteins which act to regulate a variety of effects, including cell migration and inflammatory events. They are currently seven known CXCR receptor found in mammals named CXCR1-CXCR7. CXCR2 (also called CD182, IL8) plays a critical role in the regulation of neutrophil homeostasis[3] and is found on many cells including leukocytes, endothelial and epithelial cells[4,5]. CXCR2 plays an important role in chronic obstructive pulmonary disease (COPD), asthma, fibrotic pulmonary disorders[6-8]. CXCR2 receptor can be released from a number of inflammatory cell types and may have a broad functional role in number of acute and chronic diseases. It was found that neutrophilic inflammation in the lung diseases is found to be largely regulated through CXCR2[9,2]. Therefore blockade of CXCR2 substantially reduces leukocyte recruitment, tissue damage and mortality. An antagonist of CXCR2 reduces neutrophilic chemotaxis and may alter the airway inflammation. To date, there are no CXCR2 receptor antagonists approved for use in humans. However, several pharmaceutical companies have disclosed CXCR2 antagonists and amongst these, navarixin and AZD-5069 are noteworthy.

Quantitative Structural Activity relationship (QSAR) models attempts to relate the chemical structure to biological activity computationally or mathematically and to discover new compounds with improved biological activity. Hologram Quantitative Structure Activity Relationship (HQSAR) is the novel 2D fragment-based QSAR method that employs specialized molecular fingerprints[11,12] and eliminates the need for 3D structure, molecular alignment and conformational search[13,14]. In HQSAR, each molecule in the training set is divided into several structural fragments, which are arranged to form a molecular hologram, assigned by a cyclic redundancy check (CRC) algorithm. Although HQSAR uses two dimensional information of a molecule, it also utilizes some three dimensional information such as chirality and molecular hybridization. In addition, HQSAR

Department of Bioinformatics, School of Bioengineering, SRM University, SRM Nagar, Kattankulathur, Chennai 603203, India

†Corresponding author : sathyainfo26@gmail.com
(Received : May 8, 2017, Revised : June 16, 2017, Accepted : June 25, 2017)

models interpret positive and negative contributions based on various atoms and structural units. With HQSAR technique we can easily and rapidly generate QSAR models for both small and large data set with high predictive value compared to other QSAR models[11]. The limitation is that it could not make biological activity predictions accurately to molecules lacking fragments or structural units included in the training set which are used to set up the model. In the present study, HQSAR has been employed to study the activity of pyrimidine-5-carbonitrile-6-alkyl derivatives as CXCR2 antagonist. Many HQSAR models were generated with different combinations of parameters and based on statistical values of the model, the best model was selected and its contribution map was also analyzed. We also identified the important features of the compounds for improving the activity.

2. Materials and Methods

2.1. Data Set

The structure of the pyrimidine-5-carbonitrile-6-alkyl derivatives and their biological activities of 26 compounds were taken from the literature[15]. Biological activities i.e. IC_{50} values of each inhibitor was converted into pIC_{50} ($-\log IC_{50}$) and the dataset (26 compounds) were segregated into test (6 compounds) and training set (20 compounds). The training and test sets were classified to ensure that both sets could completely cover the whole range of biological activity and structural diversity. The structures and their activity values are displayed in Table 1.

Table 1. Structures and biological activities (pIC_{50}) of CXCR2 inhibitors

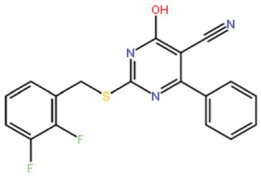
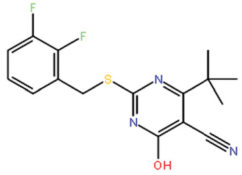
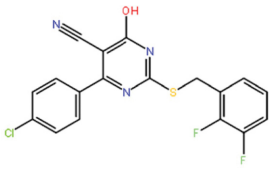
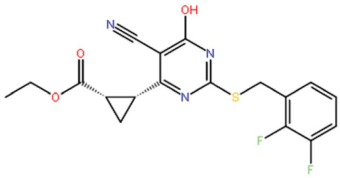
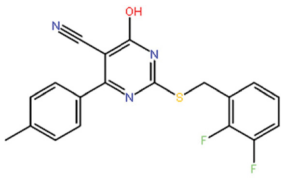
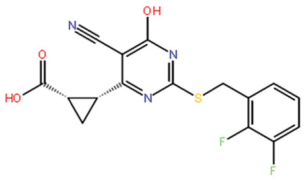
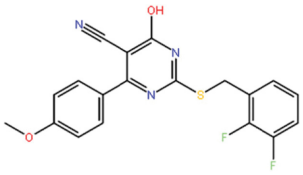
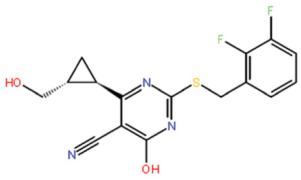
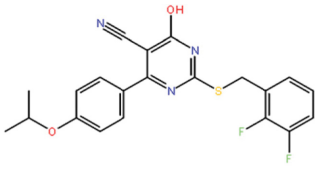
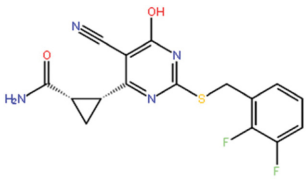
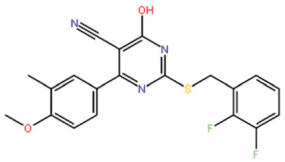
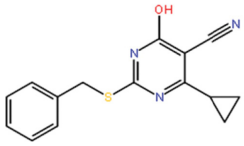
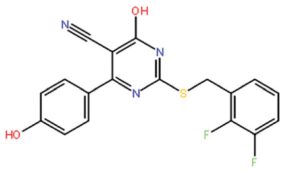
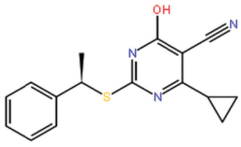
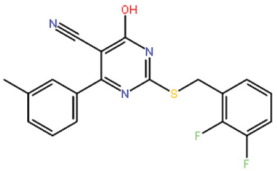
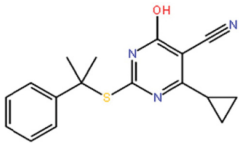
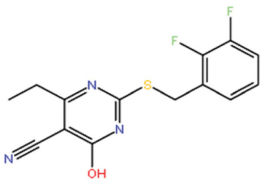
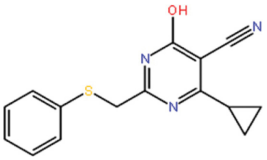
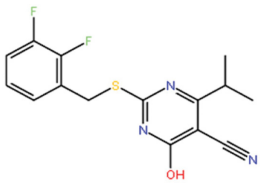
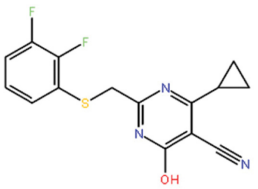
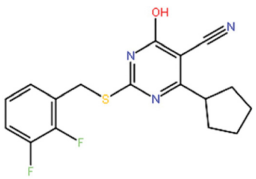
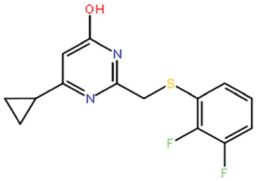
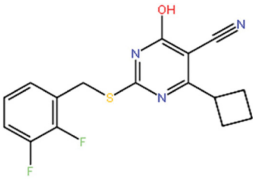
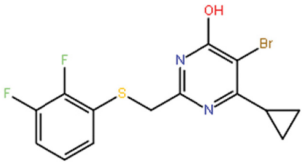
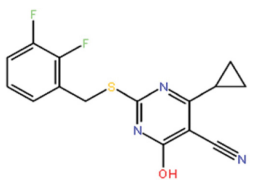
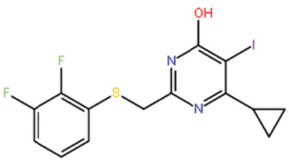
Cmpd no	Structure	pIC_{50} values	Cmpd no	Structure	pIC_{50} values
1		5.432	14		5.432
2		5.130	15		5.824
3		5.854	16		8.000
4		6.148	17		8.222
5		5.795	18		8.155

Table 1. Continued

Cmpd no	Structure	pIC_{50} values	Cmpd no	Structure	pIC_{50} values
6		5.337	19		6.292
7		6.107	20		5.318
8		5.327	21		5.193
9		5.366	22		4.522
10		6.045	23		5.000
11		5.309	24		6.495
12		5.769	25		5.854
13		6.853	26		5.495

2.2. HQSAR

HQSAR is a two dimensional computational technique that uses a fragmenting approaches that relates substructural components of compounds to their biological activity. In this method, each molecule is divided into a series of unique structural fragments that are counted in the bins of a fixed length array to form the molecular hologram[16]. The parameters such as hologram length, fragment size and fragment distinction affect the HQSAR model. The hologram length (HL) determines the number of bins in the hologram into which the fragments are hashed. The optimal HQSAR model was derived from screening through the default HL values, which were set of prime numbers ranging from 53 to 401 to avoid fragment collisions. Fragment size controls the minimum and maximum length of the fragments to be included on the hologram fingerprint with the default values as 4 and 7 respectively. Molecular fragment generation utilizes the following fragment distinctions: atoms (A), bonds (B), connections (C), chirality (Ch), hydrogen atoms (H) and donor/acceptor (DA). To evaluate the hologram generation, numerous models with the various combinations of the parameters were developed. The validity of the model depends on the statistical parameters such as cross-validated r^2 (q^2), non cross-validated r^2 by Leave-One-Out (LOO), r^2_{pred} and standard error.

2.3. Predictive Correlation Coefficient (r^2_{pred})

The predictive power of CoMFA model was determined from six test molecules which were excluded during model development. The predictive correlation coefficient (r^2_{pred}) based on the test set molecules, is defined as:

$$r^2_{pred} = \frac{(SD - PRESS)}{SD}$$

where PRESS is the sum of the squared deviation between the predicted and actual activity of the test set molecules, and SD is defined as the sum of the square deviation between the biological activity of the test set compounds and the mean activity of the training set molecules[17].

3. Results and Discussion

3.1. HQSAR Analysis

HQSAR model generation was performed on 26 pyrimidine-5-carbonitrile-6-alkyl derivatives using three distinct parameters namely fragment size, hologram length and fragment distinction. 15 HQSAR models were generated using the different fragment distinction with the fragment size 4-7. The models generated using the combination of atoms, bonds and chirality gave better results compared to others. The statistical results of the generated HQSAR models are shown in Table 2.

Table 2. HQSAR analysis for various fragment distinctions using default fragment size (4-7)

Model no	Fragment Distinction	q^2	r^2	SEE	N	HL
1	A/B	0.733	0.970	0.190	6	151
2	A/B/C	0.360	0.949	0.236	5	253
3	A/B/C/H	0.473	0.798	0.456	4	53
4	A/B/C/Ch	0.501	0.988	0.117	6	353
5	A/B/C/H/Ch	0.643	0.915	0.296	4	97
6	A/C/DA	0.191	0.951	0.241	6	353
7	A/B/C/H/DA	0.493	0.948	0.231	4	307
8	A/B/H	0.654	0.966	0.195	5	199
9	A/B/H/DA	0.610	0.932	0.264	4	307
10	A/B/C/DA	0.103	0.939	0.270	6	53
11	A/B/Ch/DA	0.614	0.962	0.211	6	307
12	A/B/H/Ch	0.645	0.877	0.356	4	199
13	A/B/DA	0.624	0.917	0.302	5	151
14	**A/B/Ch**	**0.774**	**0.977**	**0.164**	**6**	**151**
15	A/B/C/H/Ch/DA	0.585	0.933	0.262	4	307

Table 3. Influence of various fragment size using the best fragment distinction combination (A/B/Ch)

Model no	Fragment Size	q^2	r^2	SEE	N	HL	r^2_{pred}
	2-5	0.742	0.958	0.224	6	61	-
	3-6	0.655	0.957	0.219	5	401	-
	4-7	**0.774**	**0.977**	**0.164**	**6**	**151**	**0.614**
14	5-8	0.748	0.983	0.141	6	151	-
	6-9	0.710	0.980	0.154	6	97	-
	7-10	0.741	0.982	0.146	6	97	-
	8-11	0.750	0.981	0.149	6	97	-

Table 4. Experimental and predicted pIC_{50} values of training and test set compounds

Compound no	Actual pIC_{50}	HQSAR	
		Predicted (pIC_{50})	Residual
1	5.432	5.739	-0.307
2*	5.131	5.678	-0.547
3	5.854	5.628	0.226
4	6.149	5.983	0.166
5	5.796	5.957	-0.161
6	5.337	5.419	-0.082
7	6.108	6.005	0.103
8	5.328	5.291	0.037
9*	5.367	6.551	-1.185
10	6.046	6.106	-0.060
11	5.310	5.120	0.190
12	5.770	5.786	-0.016
13*	6.854	6.937	-0.083
14	5.432	5.474	-0.042
15*	5.824	5.824	0.000
16	8.000	8.182	-0.182
17	8.222	8.080	0.142
18*	8.155	7.469	0.686
19	6.292	6.181	0.111
20	5.319	5.406	-0.087
21	5.194	5.209	-0.015
22	4.523	4.542	-0.019
23	5.000	5.078	-0.078
24	6.495	6.393	0.102
25	5.854	5.880	-0.026
26*	5.495	6.081	-0.586

The model selected was further investigated to see the influence of length of fragment sizes (2-5, 3-6, 4-7, 5-8, 6-9, 7-10 and 8-11) and its results are summarized in Table 3. The statistical parameters showed that there is no significant improvement by changing the fragment size for this dataset. We had chosen the best model with higher q^2 values as summarized in Table 3 for examining the predictive ability r^2_{pred} of test set molecules. Based on better q^2 and r^2_{pred} values the final model was selected (q^2=0.774, r^2=0.977, SEE=0.164, r^2_{pred}=0.614) which was built using parameters A/B/Ch as fragment distinction, fragment size set to min 4 and max 7 with hologram length 151 and 6 components. The detailed predicted versus actual activities along with the residual values for training and test set was depicted in Table 4. Low residual values obtained for developed HQSAR model indicates its reliability and can be used to predict the biological activity of novel compounds.

3.2. HQSAR Contribution Map Analysis

The HQSAR results gave direct evidence about the individual atomic contributions to the biological activity through the use of different color codes. The contributions of the different fragments for the activity of the molecules are displayed in Fig. 1. The colors at the red end of the spectrum indicates the poor contributions (red, red orange and orange), while colors at the green end reflect favorable contributions (yellow, green blue and green). Atoms with intermediate contributions are colored in white. In the contribution map we found that the scaffold pyrimidine carbonitrile are represented in white in all compounds which depicts the intermediate contribution to the activity of all molecules. The generated HQSAR model for few compounds is shown in Fig. 1 where the atoms colored in cyan color indicates the common substructure and it contributes to the inhibitory activity of the compound. In the highly active compounds (17, 18 and 16), the cyclopropyl ring is covered by green and yellow color and it its combination

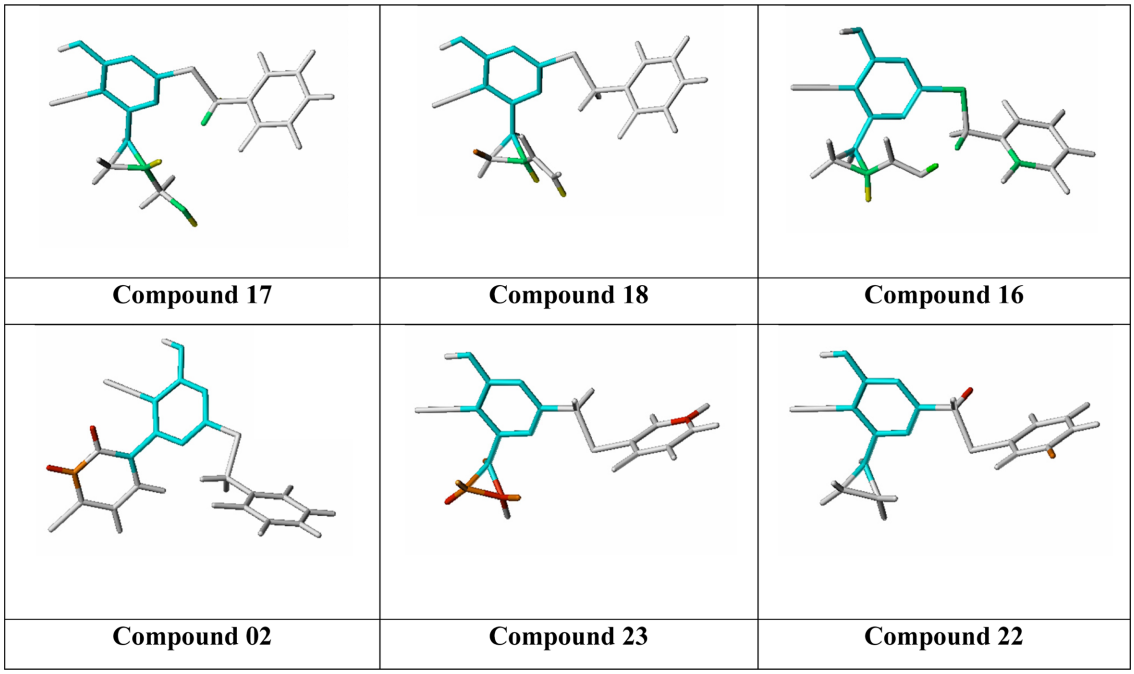


Fig. 1. HQSAR contribution map of highly active compound (compound 17, 18 and 16) and least active molecule (compound 02, 23 and 22).

with other substituent's like CH_2OH, $CONH_2$ and CO_2H helps in improving the activity of the molecule. Thus we observe that the presence of bulky substituents to the cyclopropyl ring makes the compound more potent. In the least active compounds (21, 22 and 23), we found there is the lack of bulky substituents to the cyclopropyl ring. In compound 02 the presence of chlorine atom attached to pyrimidine ring is covered by red and brown color which shows that this substituent is not suitable at that position and might be responsible for the low activity of that molecule. In compound 11 the presence of red and brown color around the cyclopentyl group attached to pyrimidine ring indicates that cyclopentyl is not favourable at that position for improving the activity.

4. Conclusion

This study was conducted to rationalize the pyrimidine-5-carbonitrile-6-alkyl derivatives by HQSAR analysis. All the generated models showed different statistical results in terms of q^2 and r^2 values when different combinations of fragment distinction were used. The best model was selected based on high q^2 (0.774) and r^2_{pred} (0.614) values. Contribution map show that presence of cyclopropyl ring along with the bulkier substituent's makes favorable contributions in the highly active compounds. Hence, this study is useful for the discovery of new antagonists for CXCR2 receptor.

References

[1] M. Uguccioni, M. D'Apuzzo, M. Loetscher, B. Dewald, and M. Baggiolini, "Actions of the chemotactic cytokines MCP-1, MCP-2, MCP-3, RANTES, MIP-1 alpha and MIP-1 beta on human monocytes", Eur. J. Immunol., Vol. 25, pp. 64-68, 1995.

[2] R. W. Chapman, J. E. Phillips, R. W. Hipkin, A. K. Curran, D. Lundell, and J. S. Fine, "CXCR2 antagonists for the treatment of pulmonary disease", Pharmacol. Ther., Vol. 121, pp. 55-68, 2009.

[3] K. J. Eash, A. M. Greenbaum, P. K. Gopalan, and D. C. Link, "CXCR2 and CXCR4 antagonistically regulate neutrophil trafficking from murine bone marrow", J. Clin. Invest., Vol. 120, pp. 2423-2431, 2010.

[4] J. Reutershan, A. Basit, E. V. Galkina, and K. Ley,

"Sequential recruitment of neutrophils into lung and bronchoalveolar lavage fluid in LPS-induced lung injury", Am. J. Physiol. Lung Cell. Mol. Physiol., Vol. 289, pp. L807-L815, 2005.

[5] S. L. Traves, S. J. Smith, P. J. Barnes, and L. E. Donnelly, "Specific CXC but not CC chemokines cause elevated monocyte migration in COPD: a role for CXCR2", J. Leukoc. Biol., Vol. 76, pp. 441-450, 2004.

[6] R. A. Pauwels and K. F. Rabe, "Burden and clinical features of chronic obstructive pulmonary disease (COPD)", Lancet, Vol. 364, pp. 613-620, 2004.

[7] E. H. Bel, A. Sousa, L. Fleming, A. Bush, K. F Chung, J. Versnel, A. H. Wagener, S. S. Wagers, P. J. Sterk, and C. H. Compton, "Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI)", Thorax, Vol. 66, pp. 910-917, 2011.

[8] P. Anderson, "Emerging therapies in cystic fibrosis", Ther. Adv. Respir. Dis., Vol. 4, pp. 177-185, 2010.

[9] P. M. Murphy, "Neutrophil receptors for interleukin-8 and related CXC chemokines", Semin. Hematol., Vol. 34, pp. 311-318, 1997.

[10] T. Puzyn, J. Lesczynski, and M. T. Cronin, "Recent advances in QSAR Studies" in Challenges and Advances in Computational Chemistry and Physics, Vol. 8, Germany: Springer, pp. 103-125, 2010.

[11] Tripos Sybyl, HQSAR manual.

[12] C. L. Waller, "A comparative QSAR study using CoMFA, HQSAR, and FRED/SKEYS paradigms for estrogen receptor binding affinities of structurally diverse compounds", J. Chem. Inf. Model., Vol. 44, pp. 758-765, 2004.

[13] W. Tong, D. R. Lowis, R. Perkins, Y. Chen, W. J. Welsh, D. W. Goddette, T. W. Heritage, and D. M. Sheehan, "Evaluation of quantitative structure-activity relationship methods for large-scale prediction of chemicals binding to the estrogen receptor", J. Chem. Inf. Model., Vol. 38, pp. 669-677, 1998.

[14] T. W Heritage and D. R. Lowis, "Molecular hologram QSAR", In Rational Drug design: Novel methodology and Practical Applications, eds. A. L. Parrill, M. R. Reddy, Washington: American Chemical Society, Symposium Series, Vol. 719, pp. 212-225, 1999.

[15] D. W. Porter , M. Bradley, Z. Brown, S. J. Charlton, B. Cox, P. Hunt, D. Janus, S. Lewis, P. Oakley, D. O'Connor, J. Reilly, N. Smith, and N. J. Press, "The discovery of potent, orally bioavailable pyrimidine-5-carbonitrile-6-alkyl CXCR2 receptor antagonists", Bioorg. Med. Chem. Lett., Vol. 24, pp. 3285-3290, 2014.

[16] D. A Winker and F. R. Burden, "Holographic QSAR of benzodiazepines", Mol. Inform., Vol. 17, pp. 224-231, 1998.

[17] S. Wold, "Cross-validatory estimation of the number of components in factor and principal component model", Technometrics, Vol. 20, pp. 397-405, 1978.