*molecules*

MDPI

*Article*

# Quantitative Structure–Activity Relationship (QSAR) Modeling of Chiral CCR2 Antagonists with a Multidimensional Space of Novel Chirality Descriptors

**Ramanathan Natarajan [1,*], Ganapathy S. Natarajan [2] and Subhash C. Basak [3]**

[1] Department of Research and Development, Saranathan College of Engineering, Panjappur, Tiruchirappalli 620 012, Tamil Nadu, India
[2] Department of Mechanical Engineering and Industrial Engineering, University of Wisconsin-Platteville, Platteville, WI 53818, USA; natarajang@uwplatt.edu
[3] Independent Researcher, 1802 Stanford Avenue, Duluth, MN 55811, USA; sbasak@d.umn.edu
* Correspondence: natarajan-rd@saranathan.ac.in

**Abstract:** The development of chirality descriptors for quantitative chirality structure–activity relationship (QCSAR) modeling has always attracted attention, owing to the importance of chiral molecules in pharmaceutical, agriculture, food, and fragrance industries, and environmental toxicology. The utility of a multidimensional space of novel relative chirality indices (RCIs) in the QCSAR modeling of twenty CCR2 antagonists is reported upon in this paper. The numerical characterization of chirality by the RCI approach gives a large pool of chirality descriptors with different degrees of mutual correlation (the correlation coefficient among the computed descriptors varied from 0.02 to 0.99). In the present study, the final data set contains 198 chirality descriptors for each of the twenty CCR2 antagonist molecules, providing a multidimensional space for modeling. The data reduction using principal component analysis resulted in the extraction of eight principal components (PCs). The linear regression using the principal component scores (PCSs) resulted in a three-predictor prediction model with good statistics: $R^2$ = 0.823; Adj $R^2$ = 0.790. The regression models were rebuilt using the chirality descriptors (RCIs) that are most correlated with each of the scores (PCSs) of the three principal components. The $R^2$ value for the regression models with three RCIs as the predictors is 0.742 and the five-fold cross validation, $R_{cv}^2$, is 0.839. The new chirality descriptors, namely, the RCIs calculated using a different weighting scheme, provide a multidimensional space of chirality descriptors for a set of chiral molecules, and such a multidimensional chirality space is a powerful tool to build quantitative chiral structure–activity relationship (QCSAR) models.

**Keywords:** quantitative chirality structure–activity relationship (QCSAR); QSAR; chirality; chirality indices; CCR2 antagonists

check for updates

Academic Editor: Alla P. Toropova

Received: 10 December 2024
Revised: 1 January 2025
Accepted: 8 January 2025
Published: 14 January 2025

**Citation:** Natarajan, R.; Natarajan, G.S.; Basak, S.C. Quantitative Structure–Activity Relationship (QSAR) Modeling of Chiral CCR2 Antagonists with a Multidimensional Space of Novel Chirality Descriptors. *Molecules* **2025**, *30*, 307. https://doi.org/10.3390/molecules30020307

**Copyright:** © 2025 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 (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

The term chirality was introduced by Lord Kelvin to explain the relation between the optical isomers or enantiomers that differ in their three-dimensional disposition around an asymmetric carbon atom, which are related as the object and the non-superimposable mirror image. In chemistry, there is a solubility rule to explain the solubility of solutes and it states that "Like dissolves like," which means polar solvents dissolve polar solutes and non-polar solvents dissolve non-polar solutes. The same is true for the chiral distinction. Only a chiral environment can distinguish chiral isomers. Biological systems are chiral in

nature, owing to the homochirality of the building blocks of the macromolecules. Thus, chirality is the biosignature present in all biological macromolecules [1]. The homochirality that is prevailing among the biological molecules is considered to have a cosmic origin [2] in support of these propylene oxides, since a chiral molecule has been identified in interstellar space [3]. The biological systems are enantio-selective, and the enantio-distinction plays a vital role in drug action, i.e., pharmacology [4,5]. The chiral distinction or enantio-selectivity is not a phenomenon relevant only to the pharmaceutical industry [6] but extends to other areas such as the agriculture [7], fragrance, and food industries [8], and environmental toxicology [9].

The QSAR approach has been extensively used to model various physiochemical properties, biological activities, and toxicological end points [10–18]. In the QSAR approach for modeling and predicting the property/activity under consideration, computed molecular descriptors are preferred and have been successfully used over the years. The large set of molecular descriptors used in the QSAR approach under the name of topological indices are derived from their molecular graphs. In a molecular graph, the atoms form the vertices, and the bonds form the edges. The topological indices are therefore computed based on the connectivity of atoms in the molecular graphs. In the case of enantiomers, though they have the object and the non-superimposable mirror image relation, the molecular connectivity remains the same. That is, enantiomers have identical molecular graphs; therefore, *R*- and *S*- isomers will have the same numerical values for the molecular descriptors and fail to distinguish them. To overcome this limitation in modifications, which are generally regarded as chiral modifications, chiral corrections are applied in the computation of the topological indices. These modifications may be classified into two major types:

(1) including a chiral correction into the connectivity matrix and computing the chiral descriptors.
(2) applying a chiral correction factor to the topological indices computed based on a well-defined algorithm, such that new Chiral TI = conventional TI × correction factor.

The second approach modifies the existing topological index; therefore, they must be computed as the first step. The different methods on the numerical characterization of chirality have resulted in chirality descriptors that could distinguish enantiomers and diastereomers [19–24]. In most of these approaches, only one type of index is derived and of course they satisfy the primary objective in differentiating the enantiomers. The numerical characterization approach that provides just a pair of descriptors for a set of enantiomers has a limitation in its application to different biological responses for the same set of chiral molecules. The stereospecific recognition of the same set of molecules by two different receptors may be completely uncorrelated. The chiral descriptors might be successful in modeling one of these biological activities of a set of enantiomers but fail to model other properties related to a different receptor. In such a scenario, a single set of descriptors may be able to model one of the two responses. This was illustrated in the dopamine sigma receptor and D2 receptor affinities for seven pairs of 3HPPs, 3-(3-hydroxyphynyl)piperidine [19]. These two biological responses are mutually uncorrelated (correlation coefficient r = 0.195). When a single set of chiral descriptors for these seven pairs of enantiomers were used, it could model the sigma receptor affinities but failed in modeling the D2 receptor affinities. In order to handle such a situation, a family of chiral descriptors was proposed [19]. Among these descriptors, some were mutually perfectly correlated (r = 0.999), while some were uncorrelated (r < 0.1). This approach was like that of Kier and Hall [16,17,19] in extending Randić connectivity indices [20] to calculate a family of topological indices for a given set of molecules. Similar to Kier and Hall's approach, this approach gave the Randić connectivity index new dimensions and extended its applicability. The QSAR approach, which is a new approach of calculating a large pool

of chirality indices for chiral molecules, was expected to have wider applicability to many different properties, because the chiral descriptors provide a multidimensional space. The new type of chirality indices proposed to differentiate enantiomers and diastereomers were called the relative chirality indices (RCIs).

Readers may wonder how chirality, an intrinsically three-dimensional (3D) phenomenon, is characterized by a graph theoretic approach which is only two-dimensional (2D). Our approach of computing relative chirality indices may be looked upon as a method of transforming the 3D disposition of the four substituents attached to a chiral center into two directed graphs. The transition of CIP rules of ordering the four substituents attached to a chiral carbon into two directed graphs to represent *R*- and *S*- isomers is outlined in Figure 1. The four groups A, B, C, and D are attached to the chiral carbon, and the order of priority according to Cahn–Ingold–Prolog (CIP) rules is A > B > C > D. While assigning the R/S configuration, group D (least priority) is kept away from the viewer and the other three groups are placed in the front (Figure 1). This disposition is transferred into two directed graphs:

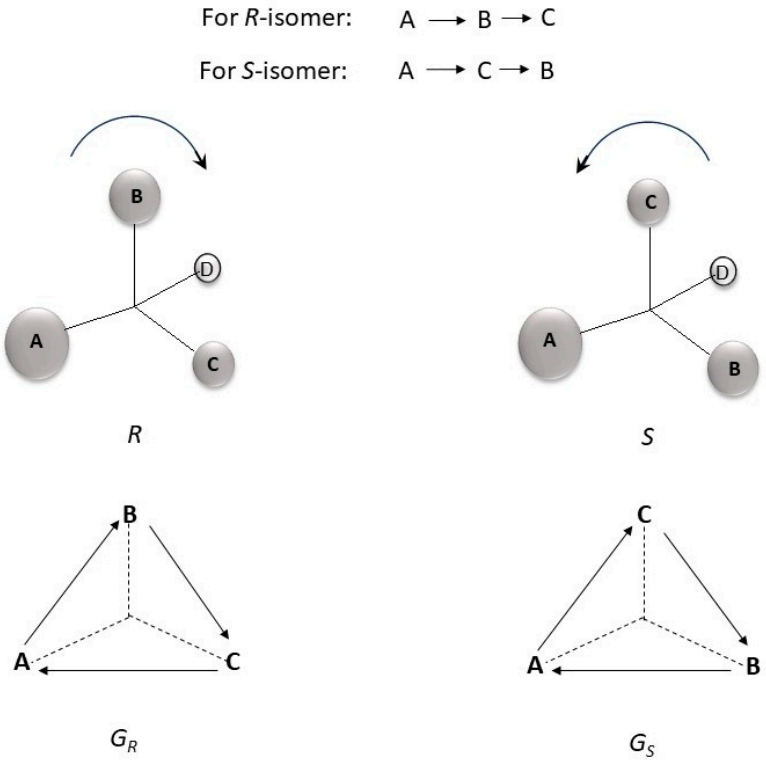



**Figure 1.** 3D structure to 2D directed graph.

The two directed graphs differ in their adjacency matrix. The adjacency matrices for the two directed graphs are as follows:

$$\text{Adjacency matrix for the } R \text{ isomer } A(G_R) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\text{Adjacency matrix for the } S \text{isomer } A(G_S) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

In order to derive indices from the directed graph, a set of formulas were proposed to calculate chirality indices that could distinguish the enantiomers. The initial formu-

lation [25] was recently modified [22] to compute a pack of several chirality indices for a given set of enantiomers. The formulas to calculate the chirality descriptors, relative chirality indices (RCI), $^{R}$RCI, and $^{S}$RCI are given below:

$$^{R}\mathrm{RCI} = \delta_a + \{\delta_a + (\delta_a \times \delta_b)\} + \{\delta_a + (\delta_a \times \delta_b) + (\delta_a \times \delta_b \times \delta_c)\} + (\delta_a \times \delta_b \times \delta_c \times \delta_d) \quad (1)$$

$$^{S}\mathrm{RCI} = \delta_a + \{\delta_a + (\delta_a \times \delta_c)\} + \{\delta_a + (\delta_a \times \delta_c) + (\delta_a \times \delta_b \times \delta_c)\} + (\delta_a \times \delta_b \times \delta_c \times \delta_d) \quad (2)$$

where $\delta_a$, $\delta_b$, $\delta_c$, and $\delta_d$ are the weights assigned to the four substituents A, B, C, and D attached to the chiral carbon, while the CIP priority order is A > B > C > D.

This new approach enables the extension the of computation of RCIs to molecules with more than one chiral center. When a molecule has more than one chiral center, the RCI is computed for each center based on substituents attached, and the individual RCI values thus computed for each chiral center in a molecule is combined by taking a root mean square value to derive a single RCI for the chiral molecule (Equation (3)). Thus, RCIs can be calculated not only for enantiomers but also for diastereomers.

$$\mathrm{RCI} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \mathrm{RCI}_i^2} \quad (3)$$

The calculation of the RCI and its application to model the biological activity of diastereomers was illustrated for the differential repellency (mosquito) of *SS* 220 [21], and recently, Natarajan et al. [19] extended the application of the large pool of RCIs calculated to model two biological responses for a set of seven pairs of enantiomers of 3-(3-hydroxyphenyl)piperidines. The present paper is an extension of this approach to model the biological activities of a diverse data set of chiral molecules.

Physiological and pathological processes in the human body are regulated by a system of chemokine and chemokine receptors, a subfamily of human Class A G-protein coupled receptors (GPCRs) [22,23]. The chemokine receptors play a significant role in the migration and localization of leukocytes. To date, the protein data bank (PDB) has a repository of the structures of 22 chemokine receptors. Urvas and Kellenberger [24] analyzed and compared their structures in a recent review. Out of the chemokine receptors, the CC-chemokine receptor 2 (CCR2) is the second most studied receptor. CCR2 is involved in various neurodegenerative disorders including Alzheimer's disease, multiple sclerosis, and ischemic brain injury [25–28]. During the SARS-CoV-2 pandemic, the involvement of CCR2 in fighting the inflammation of lungs was extensively studied [29,30]. Hence, CCR2 has attracted attention as a therapeutic target for autoimmune diseases such as rheumatoid arthritis [27], cancer [25,29], traumatic brain injury [31], etc. In most of these, the overexpression of CCR2 is the main cause; therefore, the suppression or dampening by CCR2 antagonists is one of the therapeutic strategies. Owing to the therapeutic importance of CCR2, we considered developing QSAR models for CCR2 antagonists that are chiral using a large pool of RCIs computed based on various algorithms to assign weights to the four substituents attached to the chiral carbon.

## 2. Materials and Methods

### *2.1. Biological Data*

The in vitro inhibitory activities of CCR2 antagonists have been reported upon by Merck Laboratories [32–34]. Nair et al. [35] selected only 50 molecules that were reported with definite $IC_{50}$ (nM) values for the QSAR studies using chiral sensitive hologram descriptors. All 50 molecules used by Nair et al. [35] are not chiral and only 20 of these are chiral molecules. As the purpose of the current study is to build a QSAR model with chirality descriptors, we used the inhibitory activity, $IC_{50}$ (nM), of the 20 chiral CCR2 antagonists.

The selected molecules belong to two structural classes, namely, arylglycinamide and α-aminothiazole-γ-aminobutanoic amide. Among these structural classes, the substituents vary from substituted-phenyl, substituted-thiophenyl, to substituted-piperidinyl. The structures of these molecules are given in Figure 2. The $IC_{50}$ (nM) values are converted to $pIC_{50}$ values by taking $-\log_{10}[IC_{50}]$ (values are given in Table 1).

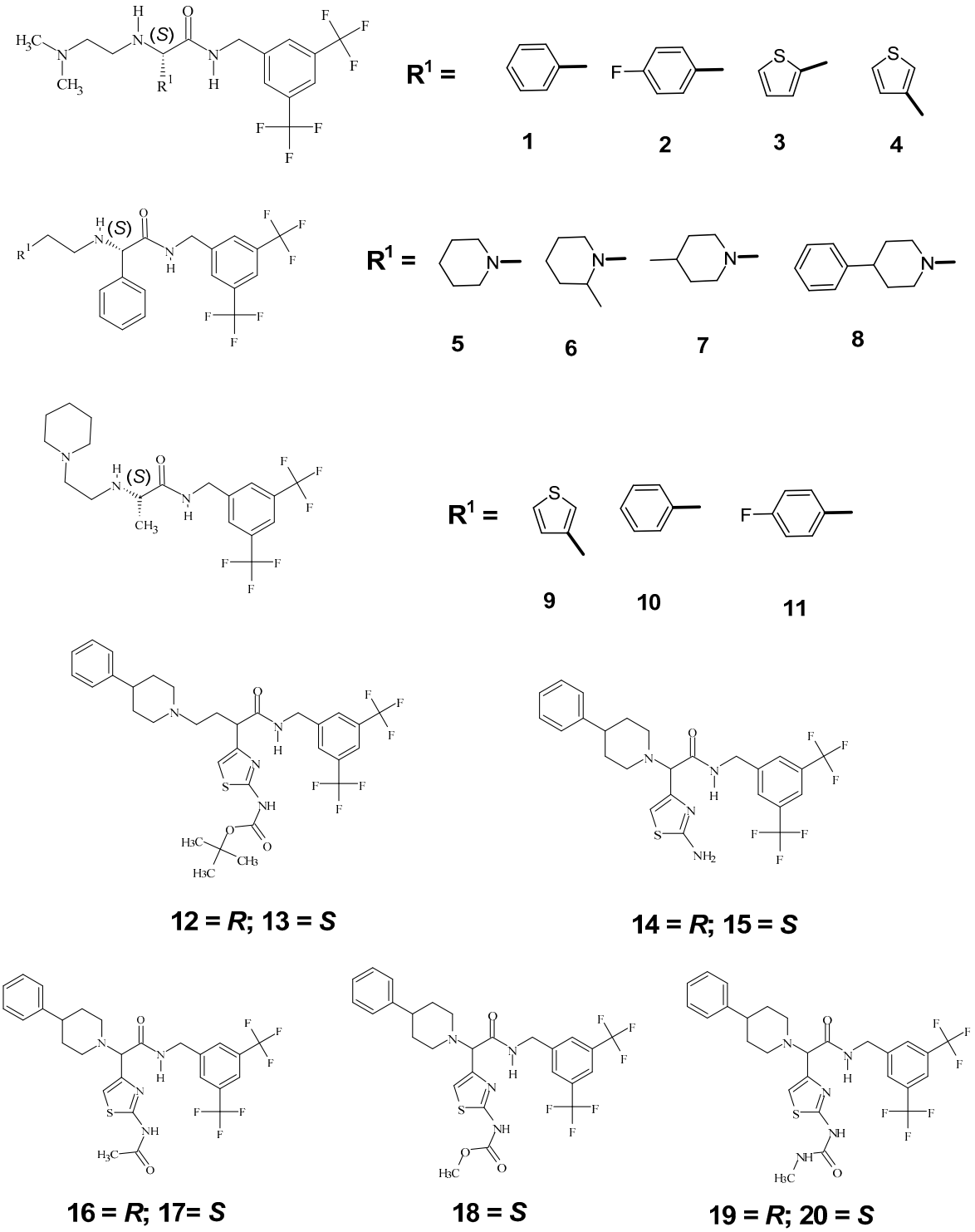



**Figure 2.** Structures of CCR2 (From 1 to 11, all the compounds are *S* isomers. Among compounds 12 to 20, except 18, the other 8 molecules are 4 pairs of enantiomers. The solid bond indicates the position of attachment of the substituent to the main structure).

**Table 1.** CCR2 antagonist inhibitory activity $pIC_{50}$.

| Molecule ID | $pIC_{50}$ (nM) | Molecule ID | $pIC_{50}$ (nM) |
|---|---|---|---|
| 1 | 6.00 | 11 | 6.76 |
| 2 | 6.11 | 12 | 7.88 |
| 3 | 6.06 | 13 | 7.84 |
| 4 | 6.37 | 14 | 6.64 |
| 5 | 6.40 | 15 | 7.88 |
| 6 | 6.08 | 16 | 7.38 |
| 7 | 6.12 | 17 | 8.74 |
| 8 | 6.08 | 18 | 8.46 |
| 9 | 7.41 | 19 | 7.52 |
| 10 | 6.66 | 20 | 9.37 |

Nair et al. [35] used hologram descriptors to model the inhibitory activity of CCR2 antagonist activities of 50 compounds comprising 20 chiral compounds. In the hologram analysis, the quality of the model improved only after adding chirality features and the $R^2$ value was 0.945 for the 50-molecule data set. As we are using only the 20 chiral molecules, we calculated the $R^2$ value for only the 20 molecules as a benchmark to compare the results of our regression analysis. As we used only a subset of 20 molecules, the predicted values using hologram descriptors for this subset of chiral molecules alone were considered, $R^2_{pred} = 0.897$.

### *2.2. Calculation of Relative Chirality Indices for R and S Isomers*

The steps followed in calculating the relative chirality indices (RCIs) for the *R* and the *S* isomers denoted as $^{R}$RCI and $^{S}$RCI, respectively, are given below:

1. The four groups attached to the chiral carbon are assigned priority following the Cahn–Ingold–Prelog (CIP) system as a, b, c, and d, where the order of priority is a > b > c > d.
2. SMILES notation is written for each of the four groups and to maintain the connectivity of the atom (vertex) connected to the chiral carbon, the chiral carbon is also included to the groups. This is illustrated for Compound 1, as shown in Figure 3.
3. For each of the four groups (a, b, c, and d), weights are calculated in terms of various topological indices using the SMILES [36–38] notations as the input. The topological indices, namely, topo-structural, topo-chemical, triplet, overall connectivity indices, and shape indices $^{1}\kappa$ and $^{2}\kappa$, were calculated as the weights ($\delta$) for each of the four substituents (groups) attached to the chiral carbon as $\delta_a$, $\delta_b$, $\delta_c$, and $\delta_d$.
4. These weights were then used to compute $^{R}$RCI and $^{S}$RCI for all the chiral molecules in the data set using Equations (1) and (2).

Invariably, all the twenty molecules have hydrogen (H) as the fourth and the least priority substituent (d) and its $\delta$-value ($\delta_d$) is zero. If the $\delta$-value of any other substituent computed as its descriptor was zero, then that descriptor was deleted from the list before computing RCIs. The list of various descriptors that could be calculated using the computer programs *POLLY* [39], *INDCAL* [40], and *TRPLET* [41,42] are given in the Supporting Information S1. In addition to the above topological indices, shape indices [43], otherwise called as kappa indices ($^{1}\kappa$, $^{2}\kappa$) of the first order and the second order, were also calculated for the groups. The RCI values on various formalisms are generated for one set of enantiomers by substituting these $\delta$-values in the equations for $^{R}$RCI and $^{S}$RCI, using Equations (1) and (2), respectively. For the sake of illustration $\delta_a$, $\delta_b$, $\delta_c$, and $\delta_d$, computed based on only five different formalisms, are shown in Table 2. This could be extended to various orders of the connectivity indices, information contents, and triplet indices.

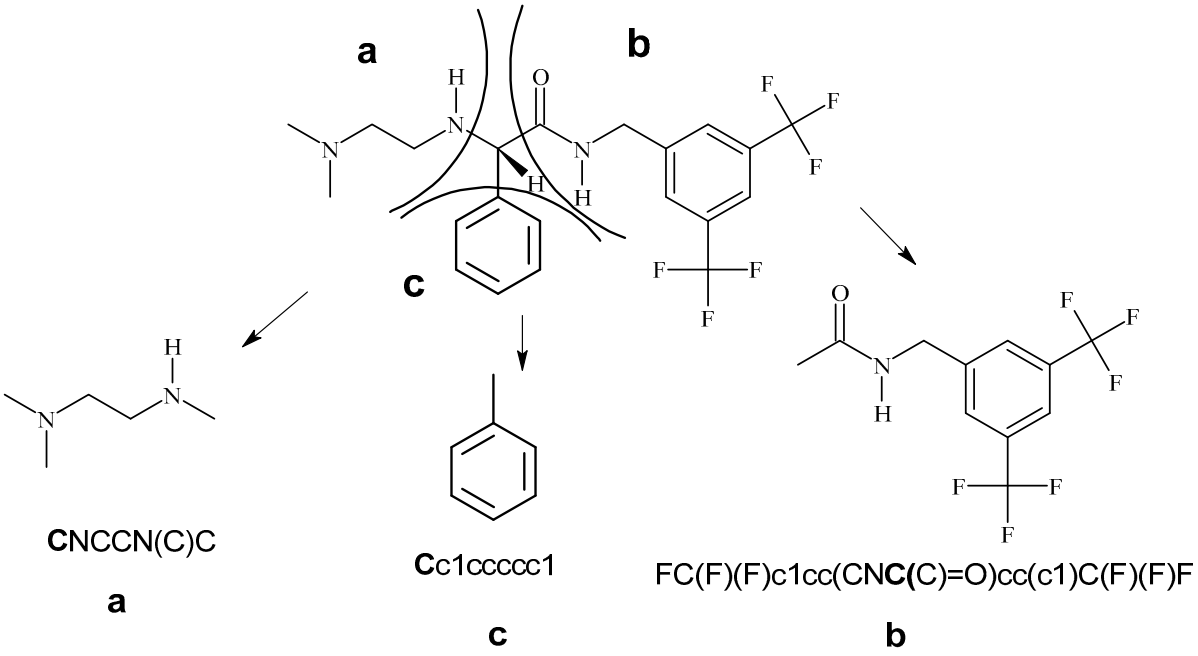



**Figure 3.** Illustration for application of sequence rule writing of SMILES notation for the substituents (chiral carbon included into the SMILES notation is shown by bold faces).

**Table 2.** SMILES notation and group weights δ for the examples shown in Figure 3 (several δ are computed for the group but two samples are shown for each compound).

| Sub | SMILES | Group Weights $\delta_a$, $\delta_b$, $\delta_c$, $\delta_d$ | | | | |
|---|---|---|---|---|---|---|
| | | J | $^2\kappa$ | $^1\chi^v$ | IC1 | AZV1 |
| a | CNCCN(C)C | 2.678 | 2.083 | 2.564 | 1.7799 | 1.422 |
| b | FC(F)(F)c1cc(CNC(C)=O)cc(c1)C(F)(F)F | 3.298 | 3.093 | 5.027 | 2.9852 | 4.180 |
| c | Cc1ccccc1 | 3.033 | 1.172 | 2.411 | 1.5329 | 1.711 |
| d | C (d = H) | 0 | 0 | 0 | 0 | 0 |

Note: J = Balaban index; $^2\kappa$ = shape index; $^1\chi^v$ = first order valence connectivity index; IC1 = dirst order information content; AZV1 = triplet index from adjacency matrix, graph order, and vertex degree; operation y = 1.

The RCI values calculated based on the five types of δ-values listed for Compound 1 in Table 2 are given below:

$$^R\text{RCI}^J = 52.482;\ ^S\text{RCI}^J = 51.064$$

$$^R\text{RCI}^{2\kappa} = 26.687;\ ^S\text{RCI}^{2\kappa} = 18.683$$

$$^R\text{RCI}^{1\chi v} = 64.550;\ ^S\text{RCI}^{1\chi v} = 51.131$$

$$^R\text{RCI}^{IC1} = 24.112;\ ^S\text{RCI}^{IC1} = 18.942$$

$$^R\text{RCI}^{AZV1} = 26.330;\ ^S\text{RCI}^{AZV1} = 19.307$$

### *2.3. Statistical Analysis*

The number of descriptors available are much more than the number of biological responses; therefore, a dimension reduction was applied using principal component analysis (PCA). Before performing the PCA, the descriptors were standardized by a logarithmic transformation, where the transformed descriptor = $\log_e$(Descriptor + 1). This was necessary because the descriptors differ very much in their dimensions. The PCA and stepwise regression were performed using SPSS 25, while cross validation was carried out using the Statsmodels 0.13.2 module in Python 3.9.13.

## 3. Results and Discussion

Several chirality descriptors based on different theories or concepts of treating molecular connectivity shape attributes were computed. Some of the higher-order connectivity-based descriptors have zero values and the values of higher-order information content become degenerate; therefore, they were removed from the data set. The final set of descriptors contains a lower number of chirality descriptors for each enantiomer (*R* or *S*) than the original number of descriptors computed. The final list of RCIs calculated has 198 different chiral descriptors for each chiral molecule. The number descriptors and the classification according to the formalisms are given in Table 3.

**Table 3.** Number of RCIs and their classification.

| S. No. | Descriptor Type | Number |
|---|---|---|
| 1 | Triplet indices | 100 |
| 2 | Topo structural indices (TSIs) | 24 |
| 3 | Topochemical indices (TCIs) | 20 |
| 4 | Overall connectivity indices | 10 |
| 5 | Information content (IC) | 37 |
| 6 | ICs for H-suppressed chemical graphs | 34 |
| 7 | Shape indices | 2 |

The complete set of chirality descriptors computed for the CCR2 antagonists (20 compounds) is given in the Supporting Information (S1) and the descriptor values are given in the Supporting Information (S2) as an Excel Book.

The intercorrelations of the chirality descriptors are presented in the Supporting Information S2 as one of the Excel worksheets. The RCI with highest correlation with the $pI_{50}$ values is $RCI^{HIC4}$ (r = 0.838) and the least correlated RCI is based on the complementary information content of the zeroth order, $RCI^{CIC0}$ (r = 0.028). Among the chirality descriptors computed, RCIs based on information contents are highly correlated with the bio-efficacy. In the case of connectivity-based chirality descriptors, the cluster and path cluster type of indices are among the top 10 and the overall valence connectivity index-based RCI has the highest negative *r* value of 0.792. The top ten most correlated descriptors from each class of RCI are presented in Table 4.

**Table 4.** Top ten chirality indices (RCIs) highly correlated with CCR2 antagonist affinities.

| No. | $RCI_{Information\ content}$ | | $RCI_{Triplet}$ | | $RCI_{Connectivity}$ | |
|---|---|---|---|---|---|---|
| | **Desc** | **r** | **Desc** | **r** | **Desc** | **r** |
| 1 | HIC4 | 0.838 | AZN4 | 0.804 | V_MPC | −0.792 |
| 2 | IC2 | 0.830 | DN2S3 | −0.801 | SC3 | 0.778 |
| 3 | HIC2 | 0.825 | DN2S4 | −0.801 | SPC4 | 0.776 |
| 4 | HTIC0 | 0.825 | DN2S5 | −0.798 | BC3 | 0.766 |
| 5 | IC1 | 0.818 | ANZ5 | 0.789 | BPC4 | 0.765 |
| 6 | IC0 | 0.814 | ANZ1 | 0.781 | SPC5 | 0.756 |
| 7 | HTIC1 | 0.811 | AZV4 | 0.760 | SPC6 | 0.755 |
| 8 | HIC3 | 0.809 | ANZ3 | 0.740 | B_OPM | 0.741 |
| 9 | HIC5 | 0.809 | ANN4 | 0.718 | M2 | 0.739 |
| 10 | HIORB | 0.809 | ANN3 | 0.701 | M1 | 0.738 |

### *3.1. Regression Analysis*

Since the number of descriptors (198) far exceeded the available data points or observations (20), it is necessary to perform a dimension reduction. Principal component analysis (PCA) was used for the dimension reduction and principal components with eigen values > 1 were extracted. A total of six principal components were extracted, and they

represented 98.7% of the data variance. The six PCs and their cumulative variance are presented in Table 5. The six principal component scores were used for predicating the inhibitor activity of the chiral CCR2 antagonists.

**Table 5.** Eigen values of principal components and their cumulative variances.

| Component | Eigen Values | % of Variance | Cumulative % of Variance |
|---|---|---|---|
| 1 | 153.09 | 77.3 | 77.3 |
| 2 | 24.83 | 12.5 | 89.9 |
| 3 | 8.79 | 4.4 | 94.3 |
| 4 | 4.39 | 2.2 | 96.5 |
| 5 | 2.67 | 1.4 | 97.9 |
| 6 | 1.71 | 0.9 | 98.7 |

Regression models using computed molecular descriptors suffer from overfitting, owing to collinearity of predictors. However, when principal component scores (PCSs) are used as predictors, the collinearity problem does not arise because the principal components are orthogonal to each other. Hence, we used the six principal component scores to model the $pI_{50}$ values by stepwise linear regression. The final regression model has only three predictors and the regression equation for the prediction of the inhibitor activity of the chiral CCR2 antagonists is as follows:

$$\begin{gathered} pI_{50} = 7.088 + (0.728 \times PC_1) + (0.469 \times PC_2) + (0.3219 \times PC_3) \\ \text{n} = 20;\ \text{R}^2 = 0.823;\ \text{Adj R}^2 = 0.790 \end{gathered} \quad (4)$$

In the case of a regression model developed with PCS, the utility in future prediction is very limited because for a new molecule, the principal component analysis must be repeated to extract the factor scores. Hence, it is always better to develop prediction models with the actual computed descriptors. In the present case, the relative chirality descriptors that correlate with each of the three principal component scores were selected; if more descriptors have the same correlation, the least degenerate descriptor, i.e., the descriptor that discriminates the chiral molecules effectively, was selected. The three relative chirality descriptors thus selected are $\text{RCI}^{OPM}$, $\text{RCI}^{BICO}$, and $\text{RCI}^{J}$. $\text{RCI}^{OPM}$ is the relative chirality index calculated using the overall path multiplicity of the substituents attached to the chiral center, and similarly, $\text{RCI}^{BICO}$ and $\text{RCI}^{J}$ are the relative chirality descriptors computed using the bond information content of order zero and the Balaban *J* index, respectively. The overall path multiplicity index and Balaban *J* index are related to the branching pattern in the chiral molecule, while the information content on the bond order encodes the bond multiplicity in the molecule. Hence, the branching pattern and the types of bonds in the chiral molecules seem to influence the CCR2 antagonist inhibition of the molecules. However, it should be remembered that correlation is not a proof of causality. The regression equation with the three RCIs is as follows:

$$\begin{gathered} pI_{50} = \left(0.553 \times \text{RCI}^{OPM}\right) + (5.763 \times \text{RCI}^{BIC0} - \left(1.639 \times \text{RCI}^{J}\right) - 7.460 \\ \text{n} = 20;\ \text{R}^2 = 0.742;\ \text{Adj R}^2 = 0.677 \end{gathered} \quad (5)$$

The benchmark model [35], which is considered the best predictive model, contains four predictors comprising hologram descriptors for atom counts from four to seven atoms, the nature of bonds, the connectivity of atoms, and a chirality parameter. The three-predictor model obtained in this work has molecular descriptors that encode almost identical information about the molecules. Hence, there is good cross talk between the present model and the model developed with hologram descriptors.

A five-fold cross validation was performed, and the results are presented in Table 6. The best $R_{cv}^2$ of 0.839 was obtained for the second fold. The results of the five-fold cross validation for each fold are given in Supporting Information S3.

**Table 6.** Results of five-fold cross validation.

| Fold No. | Training Set (16 Data) | | Test Set (4 Data) |
|---|---|---|---|
| | $R^2$ | Adj $R^2$ | $R^2$ |
| 1 | 0.743 | 0.679 | 0.778 |
| 2 | 0.742 | 0.677 | 0.839 |
| 3 | 0.813 | 0.767 | 0.512 |
| 4 | 0.798 | 0.748 | 0.556 |
| 5 | 0.750 | 0.688 | 0.629 |

### *3.2. Comparison of the Models*

The predicted ($pI_{50}$) values and the residuals of the prediction by each of the models are presented in Table 7 along with that of Nair et al. [35] as the benchmark. It may be noted that the predicted values taken from the literature [35] are part of a set of 50 molecules and the prediction model is not for the 20 chiral molecules only. The plots of the predicted $pI_{50}$ values from each of the regression equations (4) and (5) versus the experimental $pI_{50}$ values are given in Figure 4, and the comparison of residuals of each model is presented in Figure 5. From Figures 4 and 5, it is evident that the predicted $pI_{50}$ values deviate greatly from the experimental values when $pI_{50}$ is > 8.5, indicating that the model did not capture the correct molecular features for molecules with high inhibitory activity. Even in the case of the model developed with hologram descriptors, the predicted value (8.88) for molecule 20 (experimental $pI_{50}$ = 9.37) has a higher deviation from the experimental value. When the $pI_{50}$ value is in the range from 6.0 to 7.0, the predicted values are reasonably closer to the experimental values.

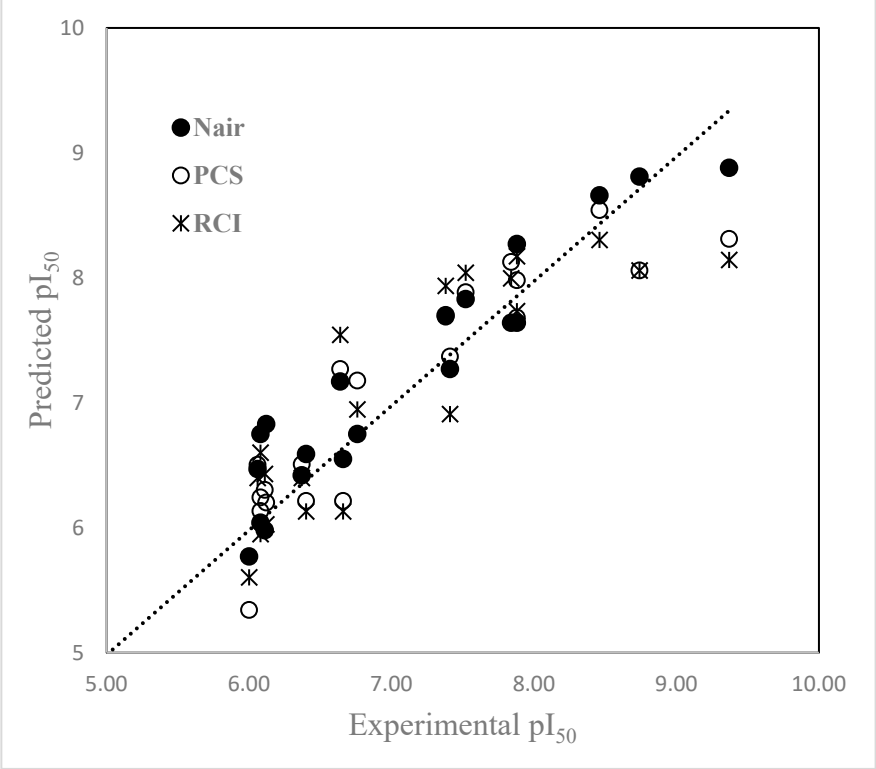



**Figure 4.** Comparison of predicted $pI_{50}$ vs. actual (experimental) $pI_{50}$ values.

**Table 7.** The experimental (actual) and predicted $pI_{50}$ values from different models.

| ID | $pI_{50}$ Values | | | | Residuals | | |
|---|---|---|---|---|---|---|---|
| | Exptl | PCS | RCI | Nair [30] | PCS | RCI | Nair |
| 1 | 6.00 | 5.34 | 5.60 | 5.77 | 0.66 | 0.40 | 0.23 |
| 2 | 6.11 | 6.30 | 6.43 | 5.98 | −0.19 | −0.32 | 0.13 |
| 3 | 6.06 | 6.50 | 6.40 | 6.47 | −0.44 | −0.34 | −0.41 |
| 4 | 6.37 | 6.51 | 6.40 | 6.42 | −0.14 | −0.03 | −0.05 |
| 5 | 6.40 | 6.21 | 6.13 | 6.59 | 0.19 | 0.27 | −0.19 |
| 6 | 6.08 | 6.24 | 5.95 | 6.75 | −0.16 | 0.13 | −0.67 |
| 7 | 6.12 | 6.20 | 6.03 | 6.83 | −0.08 | 0.09 | −0.71 |
| 8 | 6.08 | 6.13 | 6.60 | 6.04 | −0.05 | −0.52 | 0.04 |
| 9 | 7.41 | 7.37 | 6.91 | 7.27 | 0.04 | 0.50 | 0.14 |
| 10 | 6.66 | 6.21 | 6.13 | 6.55 | 0.45 | 0.53 | 0.11 |
| 11 | 6.76 | 7.18 | 6.95 | 6.75 | −0.42 | −0.19 | 0.01 |
| 12 | 7.88 | 7.98 | 8.17 | 7.64 | −0.10 | −0.29 | 0.24 |
| 13 | 7.84 | 8.13 | 8.00 | 7.64 | −0.29 | −0.16 | 0.20 |
| 14 | 6.64 | 7.27 | 7.54 | 7.17 | −0.63 | −0.90 | −0.53 |
| 15 | 7.88 | 7.68 | 7.73 | 8.27 | 0.20 | 0.15 | −0.39 |
| 16 | 7.38 | 7.69 | 7.94 | 7.70 | −0.31 | −0.56 | −0.32 |
| 17 | 8.74 | 8.06 | 8.06 | 8.81 | 0.68 | 0.68 | −0.07 |
| 18 | 8.46 | 8.54 | 8.30 | 8.66 | −0.08 | 0.16 | −0.20 |
| 19 | 7.52 | 7.89 | 8.04 | 7.83 | −0.37 | −0.52 | −0.31 |
| 20 | 9.37 | 8.31 | 8.14 | 8.88 | 1.06 | 1.23 | 0.49 |

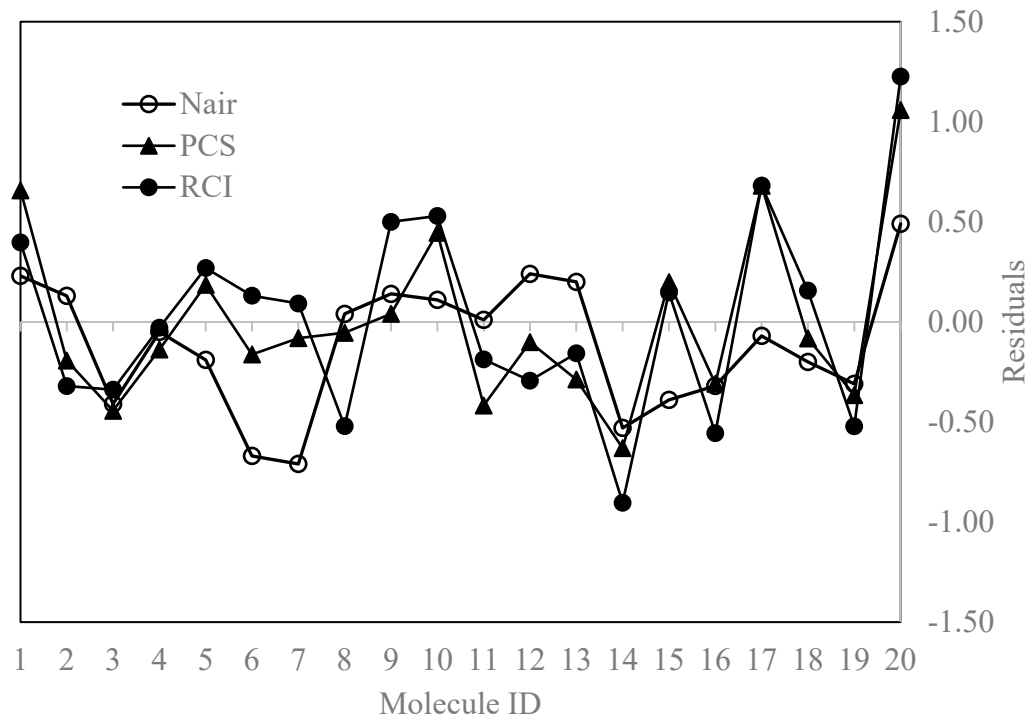



**Figure 5.** Comparison of residuals of predictions of the various models.

Figure 5 compares the residual for each observation, calculated as the difference between the experimental $pI_{50}$ values and those calculated by the regression models and the benchmark values.

## 4. Conclusions

The results show the application of the relative chirality indices (RCIs) in quantitative chiral structure–activity relationship (QCSAR) modeling. When using RCIs for a given set of chiral molecules, a multidimensional space of chirality descriptors with varying levels of correlations is available. Hence, the RCI approach is a very powerful tool.

I. Because nobody can know a priori what type of interactions will happen between chiral ligands on the one hand and the biotarget on the other, a diverse set of descriptors will have a better chance of success as compared to one single chirality index, no matter how it is derived.
II. The weighting scheme of deriving RCIs is based on calculated properties that have a good physicochemical basis in terms of sigma, pi bonds, count of lone pair, etc., which can characterize many physical and biological properties. For instance, the information content-based RCIs are not only varied, but also take care of neighborhoods of different orders, which are important for properties like the inductive effect, electron donation or withdrawal, etc.

The RCI approach has expanded the descriptor landscape from one single number to a larger set of mutually distinct/uncorrelated indices; thus, a lager net is spread to capture the panorama of ligand–biotarget interactions in diverse situations.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/molecules30020307/s1, S1: Brief description of molecular descriptors used to compute the weight of the substituents attached to a chiral carbon; S2: Excel workbook: RCI calculated for the 20 chiral CCR2 antagonists and the mutual intercorrelation of RCIs; S3: Results of 5-fold cross validation of each fold.

**Author Contributions:** Conceptualization, R.N. and S.C.B.; methodology, R.N.; validation, G.S.N.; writing—original draft preparation, R.N.; writing—review and editing, S.C.B. and G.S.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available upon request to the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Guijarro, A. *The Origin of Chirality in the Molecules of Life*; Royal Society of Chemistry: London, UK, 2022. [CrossRef]
2. Pizzarello, S.; Groy, T.L. Molecular asymmetry in extraterrestrial organic chemistry: An analytical perspective. *Geochim. Cosmochim. Acta* **2011**, *75*, 645–656. [CrossRef]
3. McGuire, B.A.; Carroll, P.B.; Loomis, R.A.; Finneran, I.A.; Jewell, P.R.; Remijan, A.J.; Blake, G.A. Discovery of the interstellar chiral molecule propylene oxide ($CH_3CHCH_2O$). *Science (1979)* **2016**, *352*, 1449–1452. [CrossRef]
4. Hutt, A.J. Drug Chirality and its Pharmacological Consequences. In *Introduction to the Principles of Drug Design and Action*, 1998th ed.; Smith, J.H., Williams, H., Eds.; CRC Press: Boca Raton, FL, USA, 1998.
5. Hutt, A.; Patel, B. Stereoselectivity in Drug Action and Disposition. In *Chirality in Drug Design and Development*; CRC Press: Boca Raton, FL, USA, 2004. [CrossRef]
6. Francotte, E.; Lindner, W. *Chirality in Drug Research*; Francotte, E., Lindner, W., Eds.; Wiley: New York, NY, USA, 2006; Volume 33. [CrossRef]

7. Kurihara, N.; Miyamoto, J.; Paulson, G.D.; Zeeh, B.; Skidmore, M.W.; Hollingworth, R.M.; Kuiper, H.A. Pesticides Report 37: Chirality in Synthetic Agrochemicals: Bioactivity and Safety Consideration. *IUPAC Stand. Online* **1997**, *69*, 1335–1348. [CrossRef]
8. Gal, J. The Discovery of Stereoselectivity at Biological Receptors: Arnaldo Piutti and the Taste of the Asparagine Enantiomers—History and Analysis on the 125th Anniversary. *Chirality* **2012**, *24*, 959–976. [CrossRef]
9. Kallenborn, R.; Hühnerfuss, H.; Aboul-Enein, H.Y.; Ali, I. Chirality in Environmental Toxicity and Fate Assessments. In *Chiral Environmental Pollutants*; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 279–305. [CrossRef]
10. Dudek, A.; Arodz, T.; Galvez, J. Computational Methods in Developing Quantitative Structure-Activity Relationships (QSAR): A Review. *Comb. Chem. High Throughput Screen* **2006**, *9*, 213–228. [CrossRef] [PubMed]
11. Roncaglioni, A.; Toropov, A.A.; Toropova, A.P.; Benfenati, E. In silico methods to predict drug toxicity. *Curr. Opin. Pharmacol.* **2013**, *13*, 802–806. [CrossRef]
12. Raies, A.B.; Bajic, V.B. In silico toxicology: Computational methods for the prediction of chemical toxicity. *WIREs Comput. Mol. Sci.* **2016**, *6*, 147–172. [CrossRef]
13. Basak, S.C.; Niemi, G.J.; Veith, G.D. Optimal characterization of structure for prediction of properties. *J. Math. Chem.* **1990**, *4*, 185–205. [CrossRef]
14. Basak, S.C.; Vracko, M. (Eds.) *Big Data Analytics in Chemoinformatics and Bioinformatics (With Applications to Computer-Aided Drug Design, Cancer Biology, Emerging Pathogens and Computational Toxicology*; Elsevier: Amsterdam, The Netherlands, 2022.
15. Devillers, J.; Balaban, A.T. (Eds.) *Topological Indices and Related Descriptors in QSAR and QSPR*; CRC Press: Boca Raton, FL, USA, 2000. [CrossRef]
16. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Chemistry and Drug Research*, 1976th ed.; Academic Press: Cambridge, MA, USA, 1976.
17. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Structure-Activity Analysis*; John Wiley: Hoboken, NJ, USA, 1986.
18. Basak, S.C.; Natarajan, R.; Sinha, D.K. Graph theoretical invariants of chemical and biological systems: Development and applications. In *Springer Proceedings in Mathematics and Statistics*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 146. [CrossRef]
19. Natarajan, R.; Lungu, C.N.; Basak, S.C. Chirality descriptors for structure–activity relationship modeling of bioactive molecules. *J. Math. Chem.* **2024**, *62*, 2388–2406. [CrossRef]
20. Randic, M. Characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615. [CrossRef]
21. Natarajan, R.; Basak, S.C.; Neumann, T.S. Novel approach for the numerical characterization of molecular chirality. *J. Chem. Inf. Model.* **2007**, *47*, 771–775. [CrossRef] [PubMed]
22. Bakos, E.; Thaiss, C.A.; Kramer, M.P.; Cohen, S.; Radomir, L.; Orr, I.; Kaushansky, N.; Ben-Nun, A.; Becker-Herman, S.; Shachar, I. CCR2 Regulates the Immune Response by Modulating the Interconversion and Function of Effector and Regulatory T Cells. *J. Immunol.* **2017**, *198*, 4659–4671. [CrossRef]
23. Shroka, T.M.; Kufareva, I.; Salanga, C.L.; Handel, T.M. The dual-function chemokine receptor CCR2 drives migration and chemokine scavenging through distinct mechanisms. *Sci. Signal* **2023**, *16*, eabo4314. [CrossRef]
24. Urvas, L.; Kellenberger, E. Structural Insights into Molecular Recognition and Receptor Activation in Chemokine–Chemokine Receptor Complexes. *J. Med. Chem.* **2023**, *66*, 7070–7085. [CrossRef]
25. Pozzi, S.; Satchi-Fainaro, R. The role of CCL2/CCR2 axis in cancer and inflammation: The next frontier in nanomedicine. *Adv. Drug Deliv. Rev.* **2024**, *209*, 115318. [CrossRef]
26. Živković, L.; Asare, Y.; Bernhagen, J.; Dichgans, M.; Georgakis, M.K. Pharmacological Targeting of the CCL2/CCR2 Axis for Atheroprotection: A Meta-Analysis of Preclinical Studies. *Arterioscler. Thromb. Vasc. Biol.* **2022**, *42*, e131–e144. [CrossRef] [PubMed]
27. Quinones, M.P.; Estrada, C.A.; Kalkonde, Y.; Ahuja, S.K.; Kuziel, W.A.; Mack, M.; Ahuja, S.S. The complex role of the chemokine receptor CCR2 in collagen-induced arthritis: Implications for therapeutic targeting of CCR2 in rheumatoid arthritis. *J. Mol. Med.* **2005**, *83*, 672–681. [CrossRef] [PubMed]
28. Charo, I.F.; Peter, S.W. Chemokine Receptor 2 (CCR2) in Atherosclerosis, Infectious Diseases, and Regulation of T-Cell Polarization. *Microcirculation* **2003**, *10*, 259–264. [CrossRef] [PubMed]
29. Modak, R.V.; de Oliveira Rebola, K.G.; McClatchy, J.; Mohammadhosseini, M.; Damnernsawad, A.; Kurtz, S.E.; Eide, C.A.; Wu, G.; Laderas, T.; Nechiporuk, T.; et al. Targeting CCL2/CCR2 Signaling Overcomes MEK Inhibitor Resistance in Acute Myeloid Leukemia. *Clin. Cancer Res.* **2024**, *30*, 2245–2259. [CrossRef] [PubMed]
30. Vanderheiden, A.; Thomas, J.; Soung, A.L.; Davis-Gardner, M.E.; Floyd, K.; Jin, F.; Cowan, D.A.; Pellegrini, K.; Shi, P.Y.; Grakoui, A.; et al. CCR2 Signaling Restricts SARS-CoV-2 Infection. *mBio* **2021**, *12*, e0274921. [CrossRef]
31. Sachdev, K.R.; Lynch, K.J.; Barreto, G.E. Exploration of novel ligands to target C-C Motif Chemokine Receptor 2 (CCR2) as a promising pharmacological treatment against traumatic brain injury. *Biomed. Pharmacother.* **2022**, *151*, 113155. [CrossRef]
32. Zhou, C.; Guo, L.; Parsons, W.H.; Mills, S.G.; MacCoss, M.; Vicario, P.P.; Zweerink, H.; Cascieri, M.A.; Springer, M.S.; Yang, L. α-Aminothiazole-γ-aminobutanoic amides as potent, small molecule CCR2 receptor antagonists. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 309–314. [CrossRef]

33. Xia, M.; Hou, C.; DeMong, D.; Pollack, S.; Pan, M.; Singer, M.; Matheis, M.; Murray, W.; Cavender, D.; Wachter, M. Synthesis and structure–activity relationship of 7-azaindole piperidine derivatives as CCR2 antagonists. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6468–6470. [CrossRef] [PubMed]
34. Butora, G.; Morriello, G.J.; Kothandaraman, S.; Guiadeen, D.; Pasternak, A.; Parsons, W.H.; MacCoss, M.; Vicario, P.P.; Cascieri, M.A.; Yang, L. 4-Amino-2-alkyl-butyramides as small molecule CCR2 antagonists with favorable pharmacokinetic properties. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4715–4722. [CrossRef]
35. Nair, P.C.; Srikanth, K.; Sobhia, M.E. QSAR studies on CCR2 antagonists with chiral sensitive hologram descriptors. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1323–1330. [CrossRef]
36. Weininger, D. SMILES. 3. DEPICT. Graphical depiction of chemical structures. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 237–243. [CrossRef]
37. Weininger, D.; Weininger, A.; Weininger, J.L. SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 97–101. [CrossRef]
38. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36. [CrossRef]
39. Basak, S.C.; Harriss, D.K.; Magnuson, V.R. POLLY, Copyright of University of Minnesota, USA 1988.
40. Natarajan, R.; Anbazhagan, T.M.; Murali, T.M. INDCAL Topological index calculator, In-house software 2002.
41. Balaban, A.T.; Mills, D.; Kodali, V.; Basak, S.C. Complexity of chemical graphs in terms of size, branching, and cyclicity. *SAR QSAR Environ. Res.* **2006**, *17*, 429–466. [CrossRef] [PubMed]
42. Filip, P.A.; Balaban, T.S.; Balaban, A.T. A new approach for devising local graph invariants: Derived topological indices with low degeneracy and good correlation ability. *J. Math. Chem.* **1987**, *1*, 61–83. [CrossRef]
43. Kier, L.B.; Hall, L.H. The kappa indices for modeling molecular shape and flexibility. In *Topological Indices and Related Descriptors in QSAR and QSPR*; Balaban, A.T., Devillers, J., Eds.; CRC Press: London, UK, 1999; pp. 455–489.


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.