QSAR Modeling Using Chirality Descriptors Derived from Molecular Topology

Alexander Golbraikh and Alexander Tropsha*

Laboratory for Molecular Modeling, Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7360

Received March 22, 2002

Topological descriptors of chemical structures (such as molecular connectivity indices) are widely used in Quantitative Structure—Activity Relationships (QSAR) studies. Unfortunately, these descriptors lack the ability to discriminate between stereoisomers, which limits their application in QSAR. To circumvent this problem, we recently introduced chirality descriptors derived from molecular graphs and applied them in QSAR studies of ecdysteroids (Golbraikh A.; Bonchev, D.; Tropsha, A. J. Chem. Inf. Comput. Sci. 2001, 41, 147−158). In this paper, we extend our earlier work by applying chirality descriptors to four data sets containing chiral compounds. All models were derived with the k-nearest neighbors (KNN) QSAR method developed in our laboratory (Zheng, W.; Tropsha, A. J. Chem. Inf. Comput. Sci. 2000, 40, 185−194). They were validated using the same training and test sets that were employed in various, mostly 3D-QSAR, investigations published by other authors. We show that for all data sets 2D-QSAR models that use a combination of chirality descriptors with conventional (chirality insensitive) topological descriptors afford better or similar predictive ability as compared to models generated with 3D-QSAR approaches. The results presented in this paper reassure that 2D-QSAR modeling provides a powerful alternative to 3D-QSAR.

1. INTRODUCTION

It is well-known that many biological molecules are asymmetric (or chiral), i.e., that they contain atoms in one of the two possible spatial configurations, which are mirror images of each other. Amino acids, carbohydrates, and lipids as well as many other natural and artificial receptor ligands are chiral.2,3 Many biochemical processes and phenomena are stereospecific. For instance, L− and D−enantiomers of amino acids have different tastes,4,4 enantiomers of some compounds have different odors,5,6 many medicinal preparations have physiological properties different from those of their enantiomers,7−9 many insecticides are stereospecific,10 etc. In 1999, the worldwide annual sales of chiral drugs exceeded $100 billion that constituted almost one-third of all drug sales,11 and in 2000, these numbers were $133 billion and 40%, respectively.12 If a drug candidate is a racemate, the regulations of Food & Drug Administration (FDA) require a detailed study of both enantiomers.11

To establish the mathematical relations between biological activities of molecules and their physicochemical properties, quantitative structure−activity relationship (QSAR) methods are used. Due to stereospecificity of biological effects, QSAR methods must be capable of taking into account atomic chiralities. Indeed, one of the most popular 3D-QSAR methods, Comparative Molecular Field Analysis (CoMFA), developed in mid-1980s13,14 and other CoMFA-like methods (some of them are discussed in refs 16−19) take into account chirality by default, since molecular fields of chiral isomers are different. However, as was outlined in our previous paper,19 CoMFA and many other 3D-QSAR methods have several shortcomings. In many cases, it is impossible to

* Corresponding author phone: (919)966-2955; fax: (919)966-6919; e-mail: tropsha@email.unc.edu.
studies of several data sets of compounds. One of these data sets, a series of ecdy steroid analogues of 20-hydroxyecdysone (20E), the steroid hormone responsible for onset and regulation of molting in almost all arthropods, was used in independent CoMFA studies and was already considered in our previous publications. Here we analyze this data set in more detail. Additional data sets include a "Q SAR benchmark" data set of 31 Cramer's steroids, and a set of 66 histamine H1 receptor ligands, and a set of 49 HIV-1 protease inhibitors. In all of these studies chirality descriptors were used in combination with conventional (chirality insensitive) topological descriptors. We demonstrate that our models are characterized by similar or better statistics and predictive power as compared with CoMFA and/or other 3D-QSAR models reported in the literature for the same data sets. The success of our modeling studies expands the range of applicability of 2D-QSAR modeling and reassures that this approach provides a powerful alternative to more popular 3D-QSAR methods.

2. METHODS

2.1. Chirality Descriptors. We give a rather brief definition of chirality descriptors used in this work (cf. ref 19 for a complete discussion). These descriptors include modified overall Zagreb indices, molecular connectivity indices, extended connectivity indices, and overall connectivity indices. All of the indices make use of so-called chirality correction, which can be a real or imaginary number added to or subtracted from vertex degrees of a hydrogen-depleted molecular graph corresponding to atoms in R- and S-configurations, respectively. For example, conventional index is defined as , where and are vertex degrees of adjacent atoms and , respectively. Chirality index is defined as , where is the chirality correction for atom . The plus sign is used, if atom is in R-configuration, and the minus sign is used, if atom is in S-configuration. For achiral atoms, chirality correction is zero.

The chirality correction can be a real or imaginary number. Chirality descriptors based on the real chirality correction are real numbers. These descriptors are referred to as class I of chirality descriptors. Chirality descriptors based on the imaginary chirality correction are referred to as class II of descriptors. Four subclasses of class II of chirality descriptors, which are complex numbers, were defined earlier. Subclass IIA descriptors are equal to these complex numbers. Subclass IIC descriptors are real and imaginary parts of these complex numbers. Subclass IIB descriptors are defined according to the following formula: , where is a subclass IIC descriptor, is the corresponding complex number. Subclass IID descriptors are defined according to the formula: , where is arctangent dependent on signs of both arguments, and thus it can be defined in segment . In this work, all subclasses of chirality descriptors have been used, except for subclass IIA, since QSAR software available is not adapted to complex descriptors.

2.2. Conventional Descriptors. Chirality descriptors do not take into account atom and bond types. To overcome this limitation, the chirality descriptors were combined with conventional topological descriptors such as overall Zagreb indices, molecular connectivity indices, extended connectivity indices, and overall connectivity indices, and descriptors obtained using Molconn-Z program. All descriptors were normalized by range-scaling, so that all normalized descriptors had values within the interval [0,1]. The total volume occupied by the representative points in the normalized descriptor space is equal to one.

2.3. Training and Test Set Compounds Selection. Descriptors defined in the previous two sections have been used for the development of QSAR models for examples described in the Introduction (refs 15, 23, 34, 37-40). To compare predictive ability of the models developed herein with respective original QSAR models for the same data sets, the same training and test sets of compounds as in the original reports were used.

For two examples, additional QSAR models were developed with other training and test sets. The following sphere-exclusion algorithm similar to that described in refs 50 and 51 for the selection of a representative subset of compounds from the whole data set was used to divide a set of compounds into training and test sets.

1. A total volume occupied by the representative points of compounds in the descriptor space was estimated as in ref 52. (In this paper, ).
2. Select a compound with the highest activity.
3. Include this compound into the training set.
4. Construct a sphere with the center in the representative point of this compound with radius .
5. Include compounds, corresponding to representative points within this sphere, except for the center of it, in the test set.
6. Exclude all points within this sphere from the initial set of compounds.
7. Let be the number of remaining compounds. If , go to step 11; otherwise go to step 8.
8. Let be the number of spheres already constructed. Calculate the distances from to , .
9. Select a compound with the smallest .
10. Go to step 3.
11. Stop.

This algorithm allows constructing training sets covering all descriptor space areas occupied by representative points. The higher the dissimilarity level is, the smaller the training set is and the larger the test set is. It is expected that the predictive ability of QSAR models generally decreases when the dissimilarity level increases.

2.4. k-Nearest Neighbors QSAR. K-nearest neighbors (kNN) QSAR method uses leave-one-out (LOO) cross-validation procedure and an evolutionary simulated-annealing algorithm for descriptor selection. The procedure starts with the random selection of a predefined number of descriptors out of all descriptors. Activities of compounds excluded in LOO procedure are estimated using the following formula.
\[
\tilde{y} = \frac{\sum_{\text{nearest neighbors}} y_i \exp(-d_i)}{\sum_{\text{nearest neighbors}} \exp(-d_i)}
\]

(1)

where \( d_i \) are the distances between nearest neighbors and this compound. After each run, cross-validated \( R^2 (q^2) \) is calculated

\[
q^2 = \frac{\sum (y_i - \tilde{y}_i)^2}{\sum (y_i - \bar{y})^2}
\]

(2)

where \( y_i, \tilde{y}_i \) and \( \bar{y} \) are the actual, predicted and mean values of activity. The summation in (2) is performed over all compounds. After each run, a predefined number \( M \) of descriptors are randomly replaced by other descriptors from the original pool, and the new value of \( q^2 \) is obtained. If \( q^2(\text{new}) = q^2(\text{old}) \), the new set of descriptors is accepted with probability \( p = \exp(q^2(\text{new}) - q^2(\text{old}))/T \) and rejected with probability \( 1-p \), where \( T \) is a simulated “temperature” annealing parameter. During this process, \( T \) is gradually decreasing from \( T_{\text{max}} \) to \( T_{\text{min}} = d \times T_{\text{curr}} \), where \( d \) is a parameter of the algorithm. Thus, \( q^2 \) is optimized (see ref 30 for additional details). In the prediction process, the final set of descriptors selected is used, and expression (1) is applied to predict activities of the test set compounds.

In all calculations in this work, the maximum number of nearest neighbors was equal to five, \( T_{\text{max}} = 100 \), \( T_{\text{min}} = 10^{-9} \), \( d = 0.9 \), and \( M = 3 \). The number of selected descriptors was varied from 10 to 40 with step two.

### 2.5. Estimation of the Predictive Ability of a Model.

We have shown earlier that a QSAR model with high predictive power must satisfy the following three conditions:\(^{54}\)

1. High value of cross-validated \( R^2(q^2) \).
2. Correlation coefficient \( R \) between the predicted and observed activities of compounds from an external test set close to one. At least one (but better both) of the correlation coefficients for regressions through the origin (in the literature they are also referred to as coefficients of determination\(^{55}\)) comparing predicted versus observed activities, (or observed versus predicted activities) \( R_0^2 \) (or \( R_0^{25} \)) close to \( R^2 \).
3. At least one slope \( k \) or \( k' \) of regression lines through the origin close to one. (It will correspond to \( R_0^2 \) or \( R_0^{25} \) that is more similar to \( R^2 \)).

To find the boundary significance level \( \alpha \), the following equation must be solved

\[
F_{1,n-2,\alpha} = F
\]

(3)

where \( F \) is the F-ratio, \( F_{1,n-2,\alpha} \) is the F-distribution function with one and \( n-2 \) degrees of freedom. The higher \( \alpha \), the better is the model. \( \alpha \) values were obtained using MATLAB\(^{60}\) cdf function. Frequently, \( p \)-values (defined as \( p \)-value = \( 1 - \alpha \)-value) are used instead of \( \alpha \)-values.

Additionally, our models were tested for robustness. This test implies comparison of models built with real data with models built with randomized activities (i.e., with shuffled target properties). The original model is considered acceptable if (i) it has significantly higher \( q^2 \) value than models developed with randomized activities or (ii) if random models have no predictive ability. Sometimes the first of the above conditions cannot be satisfied, particularly in the case of structural redundancy of the training set, or if the total number of descriptors is higher than or comparable with the number of compounds.\(^{36}\)

### 2.6. Calculations. Descriptor Sets.

For each example, the models were performed with the following values of chirality correction (see Section 2.1): 0.5, 1.0, 1.5, 2.0, and 2.5. The following subsets of chirality descriptors were used to develop QSAR models (see Section 2.1).

1. Only class I descriptors.
2. Only subclass IIB descriptors.
3. Only subclass IIC descriptors.
4. Only subclass IID descriptors.
5. Class I with subclass IIB descriptors.
6. Class I with subclass IIC descriptors.
7. Class I with subclass IID descriptors.
8. Class I and class II descriptors except for subclass IIA descriptors.

All descriptors with zero variance and duplicate descriptors were excluded.

### Statistical Characteristics of Models.

Using the \( k \)-NN-QSAR procedure, for each of the eight subsets of descriptors and for each of the 16 sizes of descriptor sets (see Section 2.4), 10 QSAR models were built. Thus, for each example for each chirality correction value, the total number of models built was \( 16 \times 10 = 1280 \). Models with \( q^2 > 0.5 \) were selected for further validation using external test sets of compounds. The models were considered statistically significant if they satisfied the following criteria (see section 2.5): \( q^2 > 0.5; R^2 > 0.6; R_0^2 \) or \( R_0^{25} \) close to \( R^2 \), i.e., \((R^2 - R_0^2)/R^2 < 0.1 \) or \((R^2 - R_0^{25})/R^2 < 0.1 \); \( k \) or \( k' \) lie within the interval [0.85, 1.15]. We sought models that satisfied these conditions for each chirality correction value. The models with the highest correlation coefficient \( R \) were considered as best. All satisfactory models were also characterized by the F-ratio and the \( \alpha- \) or \( p \)-values.

For comparison, we also generated additional models using nonchiral descriptors only. In this case, for each data set and each set of descriptors \( 16 \times 10 = 160 \) models were built.

### Division into Training and Test Sets.

The procedure for dividing a data set into training and test sets\(^{51}\) (Section 2.3) was repeated with different dissimilarity levels. The starting dissimilarity level was equal to 0.5 and then increased with step 0.5. Consequently, the number of compounds in subsequent training sets was gradually decreasing, and the number of compounds in the corresponding test sets was gradually increasing. It is expected that as training sets become smaller the predictive ability of QSAR models must generally decrease. The minimum dissimilarity level was defined such that the test set would contain at least five compounds. The maximum dissimilarity level was defined such that the test set would contain at least half of all compounds yet no good QSAR model would be obtained.

RS configurations for chiral atoms were assigned using a SYBYL Programming Language (SPL) script. Following the input data preparation, the process of QSAR model development was completely automated. Calculation time depended on the number of compounds in a data set. For the largest data set (78 ecdysteroids) all calculations for the same
Initially, QSAR models were developed using the same training and test sets as in ref 34. For each value of chirality correction and each subset of descriptors many predictive models were found, which satisfied all conditions considered in Sections 2.5 and 2.6. As many as 703 highly predictive models were generated; building one predictive model required on average slightly more than four minutes of computational time. Models with the highest predictive power for each value of chirality correction and each subset of descriptors are included in Table 1. All these models also contain asymmetric atoms; (2) the data set contains 19 pairs of \( \alpha \)-diastereomeric pairs have been excluded from the training set. (3) The values in parentheses are obtained after excluding one compound from the test set as an outlier. For each value of chirality correction and each descriptor subset \( k \)NN QSAR model with the highest predictive power is included. In the last two rows, statistics of CoMFA models from ref 34 is presented.

### Table 1. Ecdysteroids: Training and Test Sets Contained 71 and 7 Compounds, Respectively

<table>
<thead>
<tr>
<th>Model</th>
<th>Chir. Corr.</th>
<th>Chir. Descr</th>
<th>( q^2 )</th>
<th>( R )</th>
<th>( R^2 )</th>
<th>( R^2_{adj} )</th>
<th>( k^c )</th>
<th>RMSE</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>0.62</td>
<td>0.93</td>
<td>0.86</td>
<td>0.85</td>
<td>0.89</td>
<td>0.02</td>
<td>31.7</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>1,1b</td>
<td>0.57</td>
<td>0.94</td>
<td>0.89</td>
<td>0.82</td>
<td>0.89</td>
<td>0.16</td>
<td>38.8</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>1,1b,1c,1d</td>
<td>0.66</td>
<td>0.94</td>
<td>0.88</td>
<td>0.88</td>
<td>0.89</td>
<td>0.28</td>
<td>38.0</td>
</tr>
<tr>
<td>4</td>
<td>0.5</td>
<td>1,1c</td>
<td>0.59</td>
<td>0.93</td>
<td>0.86</td>
<td>0.85</td>
<td>0.93</td>
<td>0.27</td>
<td>29.5</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>1,1d</td>
<td>0.66</td>
<td>0.91</td>
<td>0.82</td>
<td>0.82</td>
<td>0.87</td>
<td>0.43</td>
<td>23.5</td>
</tr>
<tr>
<td>6</td>
<td>0.5</td>
<td>1b</td>
<td>0.71</td>
<td>0.97</td>
<td>0.94</td>
<td>0.88</td>
<td>0.93</td>
<td>0.09</td>
<td>72.8</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>1c</td>
<td>0.69</td>
<td>0.88</td>
<td>0.78</td>
<td>0.71</td>
<td>0.92</td>
<td>0.24</td>
<td>17.2</td>
</tr>
<tr>
<td>8</td>
<td>0.5</td>
<td>1d</td>
<td>0.66</td>
<td>0.92</td>
<td>0.84</td>
<td>0.84</td>
<td>0.87</td>
<td>0.34</td>
<td>27.1</td>
</tr>
<tr>
<td>9</td>
<td>1.0</td>
<td>1</td>
<td>0.62</td>
<td>0.92</td>
<td>0.84</td>
<td>0.84</td>
<td>0.87</td>
<td>0.36</td>
<td>27.0</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
<td>1,1b</td>
<td>0.61</td>
<td>0.88</td>
<td>0.77</td>
<td>0.77</td>
<td>1.07</td>
<td>0.13</td>
<td>16.7</td>
</tr>
<tr>
<td>11</td>
<td>1.0</td>
<td>1,1b,1c,1d</td>
<td>0.74</td>
<td>0.85</td>
<td>0.73</td>
<td>0.72</td>
<td>1.13</td>
<td>0.90</td>
<td>13.5</td>
</tr>
<tr>
<td>12</td>
<td>1.0</td>
<td>1,1c</td>
<td>0.63</td>
<td>0.90</td>
<td>0.80</td>
<td>0.81</td>
<td>0.90</td>
<td>0.48</td>
<td>20.1</td>
</tr>
<tr>
<td>13</td>
<td>1.0</td>
<td>1,1d</td>
<td>0.62</td>
<td>0.92</td>
<td>0.85</td>
<td>0.85</td>
<td>0.89</td>
<td>0.31</td>
<td>29.1</td>
</tr>
<tr>
<td>14</td>
<td>1.0</td>
<td>1b</td>
<td>0.72</td>
<td>0.90</td>
<td>0.82</td>
<td>0.82</td>
<td>0.90</td>
<td>0.40</td>
<td>22.3</td>
</tr>
<tr>
<td>15</td>
<td>1.0</td>
<td>1c</td>
<td>0.67</td>
<td>0.87</td>
<td>0.76</td>
<td>0.76</td>
<td>1.09</td>
<td>0.88</td>
<td>15.8</td>
</tr>
<tr>
<td>16</td>
<td>1.0</td>
<td>1d</td>
<td>0.60</td>
<td>0.90</td>
<td>0.81</td>
<td>0.81</td>
<td>0.92</td>
<td>0.34</td>
<td>21.5</td>
</tr>
<tr>
<td>17</td>
<td>1.5</td>
<td>1</td>
<td>0.62</td>
<td>0.91</td>
<td>0.83</td>
<td>0.82</td>
<td>0.88</td>
<td>0.35</td>
<td>23.6</td>
</tr>
<tr>
<td>18</td>
<td>1.5</td>
<td>1,1b</td>
<td>0.63</td>
<td>0.89</td>
<td>0.79</td>
<td>0.76</td>
<td>0.91</td>
<td>0.30</td>
<td>19.3</td>
</tr>
<tr>
<td>19</td>
<td>1.5</td>
<td>1,1b,1c,1d</td>
<td>0.53</td>
<td>0.89</td>
<td>0.78</td>
<td>0.78</td>
<td>0.88</td>
<td>0.40</td>
<td>18.2</td>
</tr>
</tbody>
</table>

\( \alpha \) The highest of \( R^2 \) or \( R^2_{adj} \) is given; the corresponding value of \( k \) or \( k^c \) is given in the next column. \( \beta \) RMSE – residual mean square of error of prediction. 13 compounds from \( \alpha \)-diastereomeric pairs have been excluded from the training set. The values in parentheses are obtained after excluding one compound from the test set as an outlier. For each value of chirality correction and each descriptor subset \( k \)NN QSAR model with the highest predictive power is included. In the last two rows, statistics of CoMFA models from ref 34 is presented.

#### 3. RESULTS AND DISCUSSION

#### 3.1. Ecdysteroids

Preliminary QSAR studies of this data set using chirality descriptors were reported elsewhere. This data set was selected as a good first example to implement our chirality descriptors in QSAR studies where. This data set was selected as a good first example to implement our chirality descriptors in QSAR studies where. 19,35

Table 1 contains a list of compounds used in this study, along with their properties. The training and test sets as in ref 34 took two days on a Pentium III 500 MHz PC.
correction of 2.0 was used to obtain this model. For comparison with our best models, the corresponding statistics for two CoMFA models A and B from ref 34 are also included in Table 1 (see Models 42 and 43). The latter models also have low \( R \) values of 0.7039 and 0.5497 for Models A and B, respectively. After deleting one outlier, the corresponding \( R \) values are 0.8958 and 0.8145, respectively. It means that even after deleting the outlier, these models cannot be accepted even with significance level 10%.

Additional calculations were performed using Molconn-Z\(^{31}\) descriptors only. The same kNN-QSAR procedure was applied. For these calculations, only more active \( \sigma \)-diastereomers from each pair have been retained in the training set. Thus, 13 compounds had to be excluded from the training set. Interestingly, one model was found to have a higher predictive power than the best models obtained with chirality descriptors (see Table 1). However, this model was apparently unable to distinguish between \( \sigma \)-diastereomeric compounds. For three pairs of \( \sigma \)-diastereomers with only one chiral atom in the opposite configurations, the differences between \( ED_{50} \) values were particularly large. In Table 2, both estimated (Model 27, Table 1) and observed \( ED_{50} \) values for these compounds are presented.

To demonstrate that QSAR models developed with chirality descriptors are capable of predicting the activities of \( \sigma \)-diastereomers in the test set, additional calculations were performed. For these calculations, the test set included six compounds (three pairs of \( \sigma \)-diastereomers, Table 2), and the training set consisted of the remaining 72 compounds. All five chirality correction values and eight sets of descriptors (see Section 2.6) were used in these calculations. For 20 out of these 40 cases, models satisfying conditions of Sections 2.5 and 2.6 were found. As in Table 1, the models generated with different subsets of descriptors even using
For chirality correction 0.5, a predictive model was built that for smaller training sets (and, therefore, the conditions considered in Sections 2.5 and 2.6. We found, as we shall see, this conclusion is also true for other data

<table>
<thead>
<tr>
<th>no. of the compd</th>
<th>−log(ED₅₀) estimated</th>
<th>no. of the compd</th>
<th>−log(ED₅₀) estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>7.85</td>
<td>7.35</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>6.18</td>
<td>6.66</td>
</tr>
<tr>
<td>3</td>
<td>66</td>
<td>7.85</td>
<td>7.52</td>
</tr>
</tbody>
</table>

Table 2. Ecdysteroids: ED₅₀ Values for Three Pairs of α-Diastereomers Estimated in Cross-Validation Procedure (Model 27, Table 1) and Observed Experimentally

The 2D-QSAR models used for this data set contain chiral atoms, and binding affinities of these compounds to corticosteroid binding globulin are available. Different methods were used to develop 3D-QSAR models for this data set, including CoMFA, CoMSIA, Comparative Molecular Similarity Indices Analysis (CoMSIA), Quantitative Similarity-Activity Relationships (QSiAR) Analysis, Pseudo-Atomic Receptor Model QSAR (PARM), and Comparative Molecular Mo-

3.2. 31 Cramer’s Steroids. This data set was introduced by Cramer et al. in 1988 and since then has become a benchmark for the assessment of novel QSAR methods. All compounds in this data set contain chiral atoms, and binding affinities of these compounds to corticosteroid binding globulin are available. Different methods were used to develop 3D-QSAR models for this data set, including CoMFA, CoMSIA, Comparative Molecular Similarity Indices Analysis (CoMSIA), Quantitative Similarity-Activity Relationships (QSiAR) Analysis, Pseudo-Atomic Receptor Model QSAR (PARM), and Comparative Molecular Mo-

We have employed the same training and test sets as in the above-mentioned publications: the training set included compounds 1 to 21 and the test set contained compounds 22 to 31 (the first training and test set selection). Additionally, following ref 18 the models were constructed based on compounds 1 to 12 and 23 to 31, while compounds 13 to 22 were included into the test set (the second training and test set selection). QSAR models for each of the five chirality correction values and each of the eight subsets of descriptors (see Section 2.6) were generated. For the first training and test set selection, for 11 out of these 40 combinations of descriptor sets and chirality corrections satisfactory (Sections 2.5 and 2.6) were found, while for the second training and test set selection predictive models were found for 15 combinations. The total number of models with high predictive ability was 37 and 35, for the first and second training and test set selection, respectively. Best models for the first training and test set selection had the following statistics: \( q^2 = 0.83, R^2 = 0.89, R_{c}^2 = 0.81, k = 0.93, F = 67.7 \) (Model 1) and \( \alpha = 1 - 3.5656 \times 10^{-3} \) and \( q^2 = 0.74, R^2 = 0.86, R_{c}^2 = 0.82, k = 0.98, F = 49.3 \) and \( \alpha = 0.9999 \) (Model 2). (The first model has higher \( R^2 \) than the second one, whereas \( k \) for the second model is closer to one.) The first model was built using nonchiral, and class I and subclass IIc chirality descriptors were calculated with the chirality
The second model was built using all descriptors (nonchiral, I, IIb, Ic and IId) calculated with the chirality correction 1.5. In Model 1, all nearest neighbors of all 10 compounds of the test set included only five different compounds, two of them having the same binding affinity. Predictive abilities of these two models are demonstrated in Figure 2a,b.

For both training and test set selections, calculations were performed using Molconn-Z descriptors only and Molconn-Z descriptors along with chirality-insensitive overall Zagreb indices, molecular connectivity indices, extended connectivity indices and overall connectivity indices. No one model was found to satisfy conditions considered in Sections 2.5 and 2.6. We note that models based on different subsets of descriptors even for one chirality correction have different predictive ability, and models based on all subsets of descriptors are not necessarily the best ones. Thus, we come to the same conclusion as for the previous example. Indeed, it is relatively fast to obtain one QSAR model with high predictive ability. However, extensive calculations with different values of chirality correction and selections of subsets of chirality descriptors should be performed in order to find the QSAR model with the highest predictive power.

In Tables 3 and 4 our best 2D-QSAR models are compared with those obtained by other authors using various 3D-QSAR methods. Obviously, our best models built for the first training and test set selection are superior to almost all other models with respect to their predictive ability (Table 3). They appear to be comparable (but still better) to the models based on QSiAR approach (Table 3). In case of the second training and test selection, our best models are almost as good as ones obtained with the QSiAR approach (Table 4).

In summary, we have demonstrated that in the case of Cramer’s steroids chirality descriptors implemented in 2D-QSAR appear to be useful. Our models compared favorably with most of other QSAR models obtained with different methods.

### 3.3. 66 Histamine H1 Receptor Ligands

This data set includes 35 analogues of 1-phenyl-3-amino-1,2,3,4-tetrahydronaphthalenes [1-phenyl-3-aminotetralins (PATs)] and 31 non-PATs. Of 66 molecules in this data set contain chiral
the compounds are available. Standard CoMFA and CoMFA/q validated using the predictions of K1 -test set consisted of the remaining 16 compounds. The highest predictive ability had the following statistics: models satisfying these conditions were found. Model with atom. Binding affinities K0.5 values for the test set. The model was built using the training set (50 compounds) and validated using the test set (16 compounds) obtained using the sphere-exclusion algorithm (Section 2.3). (c) Histamine H1 receptor ligands. Observed vs predicted K0.5 values for the test set. The model was built using the training set (36 compounds) and validated using the test set (30 compounds) obtained using the sphere-exclusion algorithm (Section 2.3).

Figure 3. (a) Histamine H1 receptor ligands. Observed vs predicted K0.5 values for the test set including 16 compounds (both training and test sets were the same as in ref 40). (b) Histamine H1 receptor ligands. Observed vs predicted K0.5 values for the test set. The model was built using the training set (16 compounds) obtained using the sphere-exclusion algorithm (Section 2.3). (c) Histamine H1 receptor ligands. Observed vs predicted K0.5 values for the test set. The model was built using the training set (36 compounds) and validated using the test set (30 compounds) obtained using the sphere-exclusion algorithm (Section 2.3).

Table 5. Histamine H1 Receptor Ligands

<table>
<thead>
<tr>
<th>method</th>
<th>alignment</th>
<th>q2</th>
<th>R2 (pred)</th>
<th>ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoMFA</td>
<td>rigid</td>
<td>0.46</td>
<td>0.89</td>
<td>40</td>
</tr>
<tr>
<td>CoMFA/q2-GRS</td>
<td>rigid</td>
<td>0.51</td>
<td>0.77</td>
<td>40</td>
</tr>
<tr>
<td>Our model</td>
<td></td>
<td>0.69</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>

*Comparison of our best model with models obtained with other methods.

Chirality descriptors of subclass IId calculated with the chirality correction 0.5. Its predictive ability is demonstrated in Figure 3a. Again, the results show that the best value of the chirality correction must be established experimentally for each particular data set.

QSAR models were also built using Molconn-Z3 descriptors only. One of these models was found almost as good as the best models in which chirality descriptors were used (q2 = 0.65, R2 = 0.70, R2 = 0.68, k = 0.87, F = 33.0 and α = 0.9996). For the development of this model, the less active compound from each enantiomeric and α-diastereomeric pair was excluded from the training set. Of course, this model was unable to distinguish between enantiomers and α-diastereomers. Comparison with other 3D-QSAR models is given in Table 5. Apparently, the predictive power of 3D-QSAR models appears to be slightly better than that of our 2D-QSAR models. On the other hand, 3D-QSAR models have much lower q2 than our models. The model based on standard CoMFA had q2 < 0.5 which implies that this model is not robust (in this work, we did not even try to make predictions using models with q2 less than 0.5).

Additional QSAR models were developed with the sphere-exclusion algorithm described in Sections 2.3 and 2.6. The models were obtained for each value of the chirality correction, each descriptor subset, and each training and test set. Several models built for the training set containing about 50 compounds have much better predictive ability than the models built using the training set from ref 40. Predictive power of one of these models is characterized by q2 = 0.55, R2 = 0.83, R2 = 0.83, k = 0.97, F = 69.1, and α = 1−8.7402 × 10−7 (see also Figure 3b). All descriptor subsets (nonchiral, I, IIb, IIc, and IId) were used to build this model; chirality correction was equal to 1.5. Better predictive ability of these models can be explained by better (rational) division of all compounds into the training and test sets used in these calculations. The rational methodology for the selection of training and test sets affords predictive QSAR models for relatively large test sets. Thus, one model built with the training set containing only 36 molecules and validated using the test set containing 30 compounds had the following statistics: q2 = 0.67, R2 = 0.61, R2 = 0.60, k = 0.97, F = 43.2, and α = 1−3.9767 × 10−7 (see also Figure 3c). Nonchiral and subclass IIa chirality descriptors were used to develop this model; chirality correction was equal to 0.5. As in the previous examples, we emphasize the dependence of the best values of the chirality correction on the choice of compounds for the training and test sets and the descriptor subset used to build the model.

Thus, in the case of Histamine H1 receptor ligands, we were able to build relatively good 2D-QSAR models for the prediction of K0.5 values. We also emphasize that in this case, the data set included both chiral and achiral compounds. Chiral compounds from this data set contained one to four
chiral atoms, while Cramer’s steroids contained four to eight, and ecdysteroids contained eight to fifteen chiral atoms. Our descriptors appeared to be applicable to QSAR studies of all these examples.

3.4. 49 HIV-1 Protease Inhibitors. All molecules in this data set contain chiral atoms. Activity data (IC₅₀ values) for these compounds were published.²⁻²³ The correlation between the calculated binding energy and IC₅₀ values has been established and regression analysis was used to predict activities of compounds from the test set.²³ Comparative Binding Energy Analysis (COMBINE) model was built and used to predict activities of compounds from the test set.¹⁵ The same training set (33 compounds) and test set (16 compounds) as in refs 15 and 23 were used for building QSAR models. For each chirality correction value and each subset of descriptors the models were found that satisfied the conditions considered in Sections 2.5 and 2.6. The total number of models satisfying these conditions was as high as 606. Two models with the highest predictive ability had the following statistics: $q^2 = 0.77$, $R^2 = 0.85$, $R_0^2 = 0.79$, $k = 0.95$, $F = 81.1$ and $\alpha = 1 - 3.3633 \times 10^{-7}$ (Model 1) and $q^2 = 0.79$, $R^2 = 0.85$, $R_0^2 = 0.77$, $k = 0.92$, $F = 76.4$ and $\alpha = 1 - 4.8142 \times 10^{-7}$ (Model 2). The first of these models was generated using nonchiral and 1d chirality descriptors calculated with the chirality correction 2.5. For the second model nonchiral and all classes and subclasses of chirality descriptors were used; the chirality correction was equal to 0.5. Predictive power of these models is demonstrated in Figure 4a,b. Models using only Molconn-Z descriptors were constructed as well. (There was only one stereoisomeric pair in this data set, and the compound with the lower activity from this pair was excluded from the test set.) In this case, one model based on Molconn-Z descriptors was found, which was only slightly worse than the best models built using chirality descriptors ($q^2 = 0.79$, $R^2 = 0.71^{18}$, $R_0^2 = 0.77$, $k = 0.96$, $F = 47.4$, and $\alpha = 1 - 7.5045 \times 10^{-6}$). Apparently, this model cannot discriminate between stereoisomers. In ref 23, the average absolute error for the prediction was 1.01 log units. After deleting one outlier, the absolute error of prediction was 0.79 log units. In contrast, the average absolute error for prediction using our best Models 1 and 2 were 0.79 log units (without outliers) and 0.85 log units (without outliers), respectively. The COMBINE model in ref 15 gave SDEP = 0.59, while our models gave SDEP = 0.61 and SDEP = 0.51, respectively.

Thus, in this case as in all previous cases we were able to obtain predictive 2D-QSAR models using chirality descriptors. Predictive ability of our best models appeared to be at least comparable with those of two different 3D-QSAR models.

4. CONCLUSIONS

We have employed chirality descriptors developed in our laboratory earlier for 2D-QSAR studies of four data sets of ligands to different receptors. In all cases, the kNN QSAR procedure with leave-one-out cross-validation was used. We compared our models with those reported in the literature for the same data sets using different QSAR approaches. The main criterion of comparison was the predictive power of the models estimated by their ability to predict accurately the activities of the external test set compounds. The training and test sets for all examples were the same as in the original publications. In all cases, our best models were either comparable to other (mostly, 3D) QSAR models or outperformed them. Thus, we have demonstrated that topological chirality descriptors can be successfully used to generate 2D-QSAR models for data sets containing stereoisomers.

We have established that for each particular data set, different chirality correction values should be used to obtain models with the highest predictive power. No preferable chirality correction value was found for all data sets. The best values of chirality correction were generally different even when models were built for different training and test sets derived from the same data set of compounds. This result suggests that this value has to be found experimentally for any given data set.

Our studies demonstrate that chirality descriptors calculated on the basis of chemical graph theory augmented by the knowledge of the absolute configurations of chiral atoms can be successfully applied in 2D-QSAR studies of data sets, which include chiral molecules. 2D-QSAR methods have several important advantages as compared to 3D-QSAR methods, i.e., they do not require extensive conformational analysis and spatial alignment of molecules; they can be easily automated and are much more computationally efficient. Therefore, we suggest that 2D-QSAR methods enhanced by chirality descriptors present a powerful alternative to popular 3D-QSAR approaches. A program to calculate chirality descriptors is available from authors upon request.

ACKNOWLEDGMENT

The support for this project was provided in part by the National Institutes of Health (grant number MH60328). The
authors are grateful to Prof. Danail Bonchev who has inspired and contributed to the development of chirality descriptors.

REFERENCES AND NOTES


CI025516B