TI-MTD MODEL
APPLICATIONS IN MOLECULAR DESIGN
Ovidiu M. Minailiuc and Mircea V. Diudea
Faculty of Chemistry and Chemical Engineering, 'Babes-Bolyai' University
3400 Cluj, Romania
in: M. V. Diudea, Ed., QSPR/QSAR Studies by Molecular Descriptors.

1. Introduction
2. TI – MTD Model
3. Line Graphs
4. Hypermolecule: Construction and Optimization
5. Parameters of TI-MTD
6. Applications of TI-MTD Model in QSAR and Molecular Design
7. Conclusions

1. Introduction

A bioactive molecule interacts with a biological receptor by means of an effector-receptor complex, in which the two partners mutually accommodate their structures such that the resulting complex exists sufficiently long for generating a biological response. The structural features of the bioactive molecule, responsible for a given biological activity (i.e. pharmacological profile), are called the pharmacophore. It plays an essential role in the recognition process preceding the complex formation. The receptor (i.e. a situs in a biomacromolecule) must contain complementary features to the pharmacophore for the recognition (complexation and ultimately biological response) to occur. In addition to the pharmacophore and its complementary points, the effector and biomacromolecule fix by areas of nonspecific interaction, which modulate the biological response.

The relative arrangement of all those atoms in the biomacromolecule (sensitive to the shape, size and chemical identity of the bioactive molecule), interested in the receptor-effector interaction, is called the receptor map. Usually the receptor structure is not known. Excepting the fortunate cases in which crystallographic data are available, the receptor map is drawn as complementary to a leading bioactive structure.

Several 2D- and 3D-QSAR approaches are performed for modeling the receptor-drug interaction, with ultimate aim the drug design. A quantitative structure-activity relationship (QSAR) correlates the changes in the observed biological activity with the changes in the chemical structure of a series of molecules. Such correlating studies work usually on sets of congeneric structures but, in the last time, non-congeneric and conformationally flexible molecules were considered.
Refined models using some 3D-databases include MSD, MTD, SIBIS, CoMFA, Genetic Algorithms. These models are, however, time-consuming particularly because of the huge number of molecular geometries needed to be calculated, for which they clearly require supercomputers.

In this chapter, a variant of MSD-MTD-SIBIS models, called TI-MTD, is presented. It makes use of graph-theoretical (i.e. topological) molecular descriptors in describing and validating a receptor map.

2. TI – MTD Model

As SIBIS, TI-MTD uses a virtual topological Receptor Space RS as a coordinate system. It is a hypermolecule constructed by superimposing the candidate molecules \( M_i \) over the reference molecule \( M_r \). The set of \( M_1, M_2, \ldots, M_m \) molecules (i.e. the learning set) is assumed to act by the same mechanism upon the bioreceptor and possess the measured activities \( Y_1, Y_2, \ldots, Y_m \). The lead molecule \( M_r \) is the most active structure within the considered set: \( Y_r = \max(Y_i; i = 1, 2, \ldots, m) \). In words, the lead molecule is considered the best negative copy of the bioreceptor.

RS reflects the topology of the space explored by the bioactive molecules \( M_i \). It provides a topological basis for some grid-based numerical description of the investigated set of chemical structures.

In construction of RS, a row-vector \( X_i \) of dimension \( N_R \) is attached to each candidate \( M_i \)

\[
X_i = X_j; \quad j = 1, 2, \ldots, N_R
\]

where \( N_R \) represents the number of vertices in RS; \( X_j = 1 \) if the point \( j \in RS \) is occupied by an atom (different from the hydrogen) of molecule \( M_i \), and zero if it is empty.

Chemical nature of the points (i.e. atoms) is indicated by a property \( P_j \) of atoms \( j \in M_i \) (e.g. the Van der Waals volume, molar refraction, the local value of a topological index TI, etc.) associated to the above vector

\[
P_jX_i = P_jX_j; \quad j = 1, 2, \ldots, N_R
\]

With the above notions, the Receptor Space RS can be constructed as:

(1) Draw the hypermolecule (i.e. the \([RS]_{init}\)) by superimposing the structures \( M_i \) over the reference \( M_r \). Consider additional connectivities in \([RS]_{init}\) if the edge resulted by joining points \( p \) and \( q \) may represent a covalent bond.

(2) Share \([RS]_{init}\) in three vertex classes:
   - points of the receptor cavity (C);
- points of the receptor walls ($W$) and
- irrelevant points ($I$).

Basically, the cavity points $C$ correspond to the atoms of the reference structure $M_r$, while the wall points $W$ occur in structures $M_i$ in the zone of their non-overlap over $M_r$. For convenience, an additional point of $I$ type (i.e. an empty point) is introduced to be bonded with those points in $RS$ whose relevance is tested.

The initial map looks as

$$[RS]_{init} = [C(m_1,...); W(n_1,...); I(p_1,...)]$$

where $m$, $n$ and $p$ are points belonging to the classes $C$, $W$ and $I$, respectively, in $[RS]_{init}$.

3) Describe numerically the hypermolecule by using some molecular topological descriptors $TD$:

$$TD(Hyp) = [TD_C(Hyp); TD_W(Hyp); TD_I(Hyp)]$$

The candidates $M_i$ are described by using the local values of $TI$ (above calculated for the whole hypermolecule) according to their own $X_i$ vector:

$$TD_C(i) = \sum_{j\in C-type} P_{ij} X_{ij}$$

$$TD_W(i) = \sum_{j\in W-type} P_{ij} X_{ij}$$

where $P_{ij}$ and $X_{ij}$ have the meaning above mentioned.

4) Optimize $[RS]_{init}$ by following the scenario:

(a) calculate the regression (7) within $[RS]_{init}$:

$$Y_{calc}(i) = a + b_1 TD_C(i) + b_2 TD_W(i) + F(i)$$

that provides the correlation coefficient $r_0$. $F$ is a linear function of non-steric factors (e.g. electronegativities) that may influence the bioactivity.

(b) change the initial attribute of vertices $j \in [RS]_{init}$, (e.g. $C$ changes to $W$ or $W$ changes to $C$ or $I$) if:

(i) eq 7 gives a better correlation index $r_1 \geq r_0 + \Delta r$, where $\Delta r$ is the desired improvement;

(ii) the novel set of $C$ is connected.

(c) an improved map $[RS]_k$ is thus obtained; the attribute changing is iterated until no improvement of the correlation coefficient is recorded.

(d) when the self-consistency is reached, the final optimal map $[RS]_{opt}$ is obtained and the algorithm is canceled.
If condition \((b(ii))\) and \(F(i)\) are ignored, the model TI-MTD (and SIBIS, of course) works as a FREE - WILSON bi-parametric variant.\(^{19}\)

The vertex set of type \(C\in[RS]\) represents the most probable molecular profile complementary to the active center of the bioreceptor and it is reliable for improving the leading structure \(M_r\).

Condition \((b(ii))\) imposes the \(C\) set be connected, that is, a true fragment. This is the reason for which a procedure able to select a desired fragment within a molecular structure was performed. Such a fragmentation algorithm will be presented in the next section.

3. Line Graphs

The line graph\(^{30,31}\) of a graph is a natural concept, independently discovered by many authors. By this reason it was called interchange graph,\(^{32}\) derivative,\(^{33}\) derived graph,\(^{34}\) edge-to-vertex dual,\(^{35}\) covering graph\(^{36}\) or also adjoint.\(^{37}\)

Several properties of line graph were discovered so far.

3.1. Definitions and Properties

The line graph \(L(G)\) of a given graph \(G\) is obtained by representing the lines in \(G = L_0\) by points and then joining two such points by a line if the corresponding lines in \(G\) are incident to a common point.\(^{38}\) By repeating this procedure \(n\) times, the iterated line graph \(L_n\); \(n = 0, 1, 2, \ldots\) (with \(L_n(G) = L_1(L_{n-1}(G))\) and \(n = 0\) for the original graph, \(G\)) is obtained. Figure 1 illustrates the line graphs \(L_n\) of \(G_1\) (2-Methylbutane); \(n = 0 - 3\), by edge evolution.

5 4 3 2 1

\[ L_0(G_1) \]
1 1 = (1,2)
2 2 = (2,3)
3 3 = (2,5)
4 4 = (3,4)
5 5

4 3

\[ L_1(G_1) \]
1 = (1,2)-(2,3)
2 = (1,2)-(2,5)
3 = (2,3)-(3,4)
4 = (2,3)-(2,5)
5 = (2,3)-(3,4)

3 2 4

\[ L_2(G_1) \]
1 = ((1,2)-(2,3))-((1,2)-(2,5))
2 = ((1,2)-(2,3))-((2,3)-(3,4))
3 = ((1,2)-(2,3))-((2,3)-(2,5))
4 = ((1,2)-(2,5))-((2,3)-(2,5))
5 = ((2,3)-(3,4))-((2,3)-(2,5))

2 3 5

\[ L_3(G_1) \]
1 = ((1,2)-(2,3))-((1,2)-(2,5))
2 = ((1,2)-(2,3))-((2,3)-(3,4))
3 = ((1,2)-(2,3))-((2,3)-(2,5))
4 = ((1,2)-(2,5))-((2,3)-(2,5))
5 = ((2,3)-(3,4))-((2,3)-(2,5))

Figure 1. The line graphs \(L_n(G_1)\); \(n = 0\).
The number of vertices, \( N \) and edges \( Q \) in \( L_{n+1} \) is given by relations \(^{39-41} \)

\[
N(L_{n+1}) = Q(L_n) \\
Q(L_{n+1}) = -Q(L_n) + (1/2) \sum_{i \in L_n} (k_i)^2 \\
Q(L_{n+1}) = \sum_{i \in L_n} \left( \frac{k_i}{2} \right) = \sum_{i \in L_n} k_i(k_i - 1)/2 = B_n
\]

where \( k_i \) is the vertex degree and \( B_n \) - Bertz' s branching index, \(^{41} \) which is the exact number of edges in the \( L_{n+1} \) line graph.

In regular graphs (i.e. graphs with all vertices having the same degree), the number of edges \( Q(L_{n+1}) \) can be calculated by a recursion, derived from eq 9 or eq 10 by substituting the value for the vertex degree

\[
k(L_n) = 2Q(L_n) / N(L_n) = 2Q(L_n) / Q(L_{n+1}) \\
Q(L_{n+1}) = -Q(L_n) + 2Q^2(L_n) / Q(L_{n+1})
\]

The number of edges in \( L_{n+1} \) can also be expressed as a function of \( k \) and \( N \)

\[
Q(L_{n+1}) = (1/2) k(L_{n+1}) N(L_{n+1})
\]

Since in regular graphs

\[
k(L_{n+1}) = 2 (k(L_n) - 1)
\]

and taking into account eq 8, eq 13 becomes

\[
Q(L_{n+1}) = Q(L_n) (k(L_n) - 1)
\]

From relations (14) and (15), \( k(L_n) \) and \( Q(L_n) \) can be expressed in terms of the starting parameters, \( k(L_0) \) and \( Q(L_0) \)

\[
k(L_n) = 1 + 2^n k(L_0) - \sum_{i=0}^{n} 2^i = 2^n k(L_0) - 2^{(n+1)} + 2 \\
Q(L_n) = Q(L_0) \prod_{i=0}^{n-1} (k(L_i) - 1) = Q(L_0) \prod_{i=0}^{n-1} (2^i k(L_0) - 2^{(i+1)} + 1)
\]

### 3.2. Reduced Line Graph

By definition, the line graph \( L_1(G) \) provides information on lines (i.e. edges) in \( G \). Each point in \( L_2(G) \) corresponds to a fragment of two lines (or three points) in \( G \). Iteratively, a point in \( L_n(G) \) will represent a fragment of \( n \) lines in the original graph \( L_0 \). A natural limit of iteration could
be \( n = Q(L_0) \) (i.e. the number of lines in \( G \)). In such a line graph each point would represent the whole original graph \( L_0 \). The above statement is true for any graph except the simple cycles \( C_N \) and the star \( S_{1,3} \), for which \( Q(L_n) = Q(L_0) = \text{constant.} \)

The iterated line graphs can be written by edge evolution.\(^4\) Figure 2 shows the fragments corresponding to \( L_3 \) in the original graph.

\[
L_3(G_1) = (1) \left( (1, 2) - (2, 3) \right) - \left( (1, 2) - (2, 5) \right) \\
(2) \left( (1, 2) - (2, 3) \right) - \left( (2, 3) - (3, 4) \right) \\
(3) \left( (1, 2) - (2, 3) \right) - \left( (2, 3) - (2, 5) \right) \\
(4) \left( (1, 2) - (2, 5) \right) - \left( (2, 3) - (2, 5) \right) \\
(5) \left( (2, 3) - (3, 4) \right) - \left( (2, 3) - (2, 5) \right)
\]

Figure 2. Edge evolution for \( L_n(G_1); n = 3 \) and the corresponding fragments in \( L_0(G_1) \).

From Figure 2 it can be seen that some lines occur twice (boldface):

\[
(1) \left( (1, 2) - (2, 3) \right) - \left( (1, 2) - (2, 5) \right) \\
(3) \left( (1, 2) - (2, 3) \right) - \left( (2, 3) - (2, 5) \right)
\]

If the multiplicity of lines is ignored in the iterated line graphs of rank \( n \geq 3 \) and keeping in mind the procedure for generating a line graph, a reduced line graph \( R_n(G) \) is obtained.\(^4\) For \( L_3(G_1) \) the corresponding \( R_n(G_1) \) looks as in Figure 3.

\[
(1) (1, 2) (2, 3) (2, 5) \\
(2) (1, 2) (2, 3) (3, 4) \\
(3) (1, 2) (2, 3) (2, 5) \\
(4) (1, 2) (2, 3) (2, 5) \\
(5) (2, 3) (3, 4) (2, 5)
\]

\[
(1) (1, 2) (2, 3) (2, 5) \\
(2) (1, 2) (2, 3) (3, 4) \\
(3) (2, 3) (3, 4) (2, 5)
\]

\( R_n(G_1) \)

Figure 3. Reduced line graph \( R_n(G_1); n = 3 \) and the corresponding fragments.

It is easily seen that the resulting fragments consist of three lines, equalling the rank of \( R_n \). It is also seen that the fragments 1, 3 and 4 are identical. By eliminating the redundant fragments, (e.g. one of the points (1), (3) and (4) in Figure 3) all the distinct fragments (consisting of \( n \) lines) in \( G \) are obtained. Two fragments are called connected in \( R_n \) if they have \( n - 1 \) common edges.
The iteration in $R_n$ is cancelled at $n = Q(L_0)$ when the resulting fragment is just the original graph. It is obvious that, in the reduced line graph

$$N(R_{n+1}) \leq Q(R_n)$$

By comparing eqs 8 and 18, on a hand, and $L_3$ (Figure 1) and $R_3$ (Figure 3), on the other hand, the difference between $L_n$ and $R_n$ is highlighted.

3.3. The FRAGGEN Algorithm.

In a set of congeneric structures, the substituents are linked in various positions of the basic chemical structure (e.g. an aromatic ring). In such a case it is useful to mark the docking positions. The FRAGGEN program allows the automatic generation of all fragments or, only of those fragments of a choice docking. The fragments thus generated are stored on disk in the view of optimizing the investigated receptor space $RS$. Figure 4 illustrates the fragmentation of the structure $G_2$, with the marked points 1, 2 and 3.

![Diagram of $G_2$ and its fragmentation](image)

$R_n; n = 6$

(1) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(4, 5)  
(2) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(6, 7)  
(3) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(8, 9)  
(4) (1, 2)(1, 4)(2, 3)(2, 6)(4, 5)(6, 7)  
(5) (1, 2)(1, 4)(2, 3)(3, 8)(4, 5)(8, 9)  
(6) (1, 2)(2, 3)(2, 6)(3, 8)(6, 7)(8, 9)

$R_n; n = 7$

(1) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(4, 5)(6, 7)  
(2) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(4, 5)(8, 9)  
(3) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(6, 7)(8, 9)

$R_n; n = 8$

(1) (1, 2)(1, 4)(2, 3)(2, 6)(3, 8)(4, 5)(6, 7)(8, 9)

Figure 4. A structure ($G_2$) and its fragmentation: reduced line graph $R_n; n = 6-8$. 
4. Hypermolecule: Alignment Rule

The bioreceptor structure is most often unknown. It is believed that, within a set of candidates, the lead structure $M_r$ is most complementary to the receptor situs. The receptor space $[RS]_{mr}$ is assimilated to a hypermolecule $Hyp$. It is constructed as a labeled grid (i.e. graph) in a pure topological space, disregarding the bond lengths and angles.

As a preliminary alignment rule, the lead structure $M_r$ is used as template. The candidates $M_i$ are aligned by docking them to the reference according to a pharmacophore hypothesis or to the maximum common substructure approach. Additional connectivities are considered if they may represent covalent bonds.

Figure 5 illustrates the docking procedure for the set of trimethoprim congeners\textsuperscript{18} (see Table 3, Sect. 6.1)

![Docking procedure for the trimethoprim congeners.](image)

It is known that the alignment rule is the key input for CoMFA.\textsuperscript{20-28} The alignment can be achieved either by manual atom-to-atom mapping technique or by an automatic mapping procedure that analyzes the topological equivalence of each atom of both the reference and the candidate. More elaborated (and more time consuming) molecular field-matching approaches are also used.

In TI-MTD the alignment modes are identified as follows: the maximum number of edges $n$ of the $M_i$ structures (or substructures) interested in the interaction with $RS$ is taken as the highest rank for generating the iterated reduced line graphs $R_n(Hyp)$ of hypermolecule. The procedure supplies subgraphs consisting of 1, 2,...$n$ edges. All distinct subgraphs (of size according to each candidate) containing the marked docking points, are selected as alignment modes of the series.

The selection of alignment modes for $M_{14} = 5$-(3'-OC\textsubscript{4}H\textsubscript{9}-benzyl)-2,4-diamino-pyrimidine and a hypothetical $M_{xx} = 5$-(4'-OC\textsubscript{4}H\textsubscript{9}-benzyl)-2,4-diamino-pyrimidine (see the set of the
trimethoprim congeners,\textsuperscript{18} Sect. 5.1) is illustrated in Figure 6. The focused fragments are 3'-\text{OC}_4\text{H}_9 and 4'-\text{OC}_4\text{H}_9. The reduced line graph $R(Hyp)$ of hypermolecules I and II, (with marked docking points) is generated thus resulting five substructures of 7 edges ($a$, $b$, $c$, $d$ and $e$) for fragment 4'-\text{OC}_4\text{H}_9 and two ones ($f$ and $g$) for 3'-\text{OC}_4\text{H}_9. They represent distinct (labeled) subgraphs of the hypermolecule and consequently distinct modes of alignment of the corresponding candidates.

![Diagram of hypermolecules I and II with alignment modes](image)

Figure 6. Alignment modes of 4'-\text{OC}_4\text{H}_9 (a–e) and 3'-\text{OC}_4\text{H}_9 (f, g) fragments within the hypermolecule of trimethoprim congeners.

In the selection of the above substructures (i.e. alignment modes) a highly discriminating $TI$ is used as a first criterion. In case of degenerate values of $TI$, permutations in the adjacency matrix are performed, within a routine procedure.\textsuperscript{44,45}

The alignment procedure is a substructure mapping procedure, useful in any approach requiring structure superposition.

At this point, the TI-MTD algorithm follows the steps 2 to 4. The optimization step 4 is iterated until the self-consistency provides the optimal map $[RS]_{opt}$. The steps 2 to 4 are performed
for all alignment modes and the best statistically validated $[RS]_{opt}$ is selected. It represents the TI-MTD map that can be used to predict the biological activity of new compounds.

## 5. Parametrization of TI-MTD

TI-MTD is basically a topological approach. Its ground parameters are topological matrices and/or topological indices.

Any TI (locally defined) can be a property $P_{ij}$ in eq 2, or in other words, a parameter of TI-MTD model.

### 5.1. Connectivity Indices

Among topological indices, frequently used in correlating studies is the Randić index\textsuperscript{46} $\chi$, defined by

$$\chi = \sum_{(j \in E(G))} (\delta_i \delta_j)^{-1/2}$$

where $\delta_i$ and $\delta_j$ are the degrees of two adjacent vertices. Two extensions of $\chi$ index, proposed by Diudea et al.\textsuperscript{47,48} namely $DSI$ and $EC$ indices, were particularly used in TI-MTD.

$DSI$\textsuperscript{47} is constructed by using some group electronegativity valences $EVG$ (defined in Sect. 5.2)

$$DSI_i = \sum_{j \in E(G)} (EVG_i EVG_j)^{-1/2}$$

$$DSI = \sum_i DSI_i$$

The summation in (20) runs over all $j$ vertices adjacent to $i$.

$EC_{P/N}$ were built up by using the $EC$ electronegativities\textsuperscript{48} (see below)

$$EC_{P/N} = \sum_{j \in E(G)} (EC_i EC_j)^{-1/2}$$

$$EC_{P/N} = \sum_i EC_{P/N}$$

The subscript symbol is $P$ for $+1/2$ and $N$ for $-1/2$. Excellent correlations of $DS$ and $EC$ indices with some physico-chemical and biological properties of aliphatic alcohols, amines and halogeno-derivatives were reported.\textsuperscript{47,48} This was a promise for using these descriptors as $P_{ij}$ local parameters in TI-MTD model.
5.2. Electronegativity Valences of Groups

For characterizing the chemical nature of vertices in a molecular graph, Diudea and Silaghi\(^4^7\) have proposed the *electronegativity valences of groups EVG*

\[ EVG_i = \left( \frac{ESA_i ESH_i}{h_i} \right)^{1/(1+h_i)} \]  
(24)

\[ h_i = (8 - GA_i) - k_i \]  
(25)

\[ ESG_i = \left( \frac{ESA_i}{ESH_i} \right)^{1/(1+k_i)} \]  
(26)

where \(GA_i\) is the number of column in the Periodic Table for the atom \(A\) belonging to the vertex (i.e. group) \(i\). \(ESA\) and \(ESH\) denote the Sanderson\(^4^9\) electronegativities for the atom \(A\) and hydrogen, respectively. The number of hydrogen atoms attached to the group \(i\) is denoted by \(h_i\) while \(k_i\) stands for the degree of \(i\). When \(k_i > (8 - GA_i)\), then \(h_i = 0\). In case of multiple bonds \(k_i = \sum b_j\), with \(b_j\) being the conventional bond orders around the vertex \(i\).

Note that these group electronegativities obey the electronegativity equalizing principle within the group \(i\) (see eq 24) and per molecule, each group being considered bonded to neighbors with electronegativity 1.\(^4^7\)

The \(EVG_i\) values were used in construction of DS index as well as parameters of TI-MTD model.

The \(EC\) valence electronegativities\(^4^8\) were developed as a variant of \(EVG\) parameters. They are based on the idea of modification of covalent radius of an atom by its hybridization state.\(^5^0\) Such a modification is reflected in the electronegativity values corresponding to the considered state. The following scenario defined the \(EC\) parameters:

(i) - covalent radii, relative to carbon atom (0.772 ANG) are calculated by eqs 27-29

\[ rc_{ni} = rc_{1i} + \Delta rc_{ni} \]  
(27)

\[ rc_{1i} = r_{1i} / 0.772 \]  
(28)

\[ \Delta rc_{ni} = (r_{ni} - r_{1i}) / 7.72 \]  
(29)

where : \(rc\) is the atomic radius relative to the carbon atom; \(n\) is the row and \(i\) is the column in the Periodic Table; \(\Delta rc\) stand for the *excess of relative radius*.

(ii) - values \(EC\), for the atoms in the \(n^{th}\) row of Periodic Table are calculated by dividing the group electronegativities \(ESG_i\) to the mean relative length, \(mlc\), of the bonds around the considered vertex/group \(i\) :

\[ EC_{ni} = \left( \frac{ESG_{ni}}{mlc_{ni}} \right) / EC_C \]  
(30)

\[ EC_C = 2.746 / 1.4996 \]  
(31)

\[ mlc_{ni} = mlc \cdot rc_{ni} \]  
(32)
EC values are listed in Table 1. They were used in construction of EC indices as well as parameters of TI-MTD model.

Table 1. EC Electronegativities.

<table>
<thead>
<tr>
<th>-Br</th>
<th>1.2447</th>
<th>-NH₂</th>
<th>1.0644</th>
</tr>
</thead>
<tbody>
<tr>
<td>-CBr₃</td>
<td>1.1266</td>
<td>-NO</td>
<td>1.4063</td>
</tr>
<tr>
<td>-CCl₃</td>
<td>1.1932</td>
<td>-NO₂</td>
<td>1.4861</td>
</tr>
<tr>
<td>-CF₃</td>
<td>1.3260</td>
<td>-O</td>
<td>1.4634</td>
</tr>
<tr>
<td>-CH₂⁻</td>
<td>0.9622</td>
<td>-OCH₃</td>
<td>1.1248</td>
</tr>
<tr>
<td>-CH₂Br</td>
<td>1.0110</td>
<td>-OH</td>
<td>1.2325</td>
</tr>
<tr>
<td>-CH₂Cl</td>
<td>1.0305</td>
<td>-P(CH₃)₂</td>
<td>0.9351</td>
</tr>
<tr>
<td>-CH₂F</td>
<td>1.0674</td>
<td>-P&lt;</td>
<td>0.8988</td>
</tr>
<tr>
<td>-CH₂I</td>
<td>0.9744</td>
<td>-PCH₃⁻</td>
<td>0.9314</td>
</tr>
<tr>
<td>-CH₂OH</td>
<td>1.0228</td>
<td>-PH⁻</td>
<td>0.9124</td>
</tr>
<tr>
<td>-CH₂SH</td>
<td>0.9804</td>
<td>-PH₂</td>
<td>0.9170</td>
</tr>
<tr>
<td>-CH₃</td>
<td>0.9575</td>
<td>-PHCH₃</td>
<td>0.93053</td>
</tr>
<tr>
<td>-CH⁻</td>
<td>0.9716</td>
<td>-S</td>
<td>1.1064</td>
</tr>
<tr>
<td>-CH=CH₂</td>
<td>1.0381</td>
<td>-SCH₃</td>
<td>1.0073</td>
</tr>
<tr>
<td>-CH=O</td>
<td>1.1596</td>
<td>2PO</td>
<td>0.1222</td>
</tr>
<tr>
<td>-CHBr₂</td>
<td>1.0672</td>
<td>3P=O</td>
<td>1.3333</td>
</tr>
<tr>
<td>-CHCl₂</td>
<td>1.1089</td>
<td>=C=</td>
<td>1.1581</td>
</tr>
<tr>
<td>-CHF₂</td>
<td>1.1897</td>
<td>=CH⁻</td>
<td>1.0441</td>
</tr>
<tr>
<td>-CH₂⁻</td>
<td>0.9914</td>
<td>=CH₂</td>
<td>1.0891</td>
</tr>
<tr>
<td>-Cl₁</td>
<td>1.0088</td>
<td>=N⁻</td>
<td>1.3147</td>
</tr>
<tr>
<td>-COOH</td>
<td>1.2220</td>
<td>=NH</td>
<td>1.2474</td>
</tr>
<tr>
<td>-Cl</td>
<td>1.3717</td>
<td>=O</td>
<td>1.6564</td>
</tr>
<tr>
<td>-C≡</td>
<td>1.1476</td>
<td>=P⁻</td>
<td>0.9658</td>
</tr>
<tr>
<td>-C≡N</td>
<td>1.2377</td>
<td>=S</td>
<td>1.2523</td>
</tr>
<tr>
<td>-F</td>
<td>1.6514</td>
<td>&gt;C&lt;</td>
<td>1.0000</td>
</tr>
<tr>
<td>-H</td>
<td>0.9175</td>
<td>&gt;C≡</td>
<td>1.0747</td>
</tr>
<tr>
<td>-I</td>
<td>1.0262</td>
<td>&gt;C=O</td>
<td>1.2397</td>
</tr>
<tr>
<td>-N(CH₃)₂</td>
<td>1.0292</td>
<td>-NHCH₃</td>
<td>1.0379</td>
</tr>
<tr>
<td>-N&lt;</td>
<td>1.2234</td>
<td>≡CH</td>
<td>1.2142</td>
</tr>
<tr>
<td>-NH⁻</td>
<td>1.1021</td>
<td>≡N</td>
<td>1.5288</td>
</tr>
</tbody>
</table>

5.3. Indices Based on Layer Matrices

The most used TI within TI-MTD are the indices of centrality $C(LM)$ and centrocomplexity $X(LM)^{51,52}$ constructed on some layer matrices.

Indices of centrality$^{51,52}$ look for the center of the graph and are defined as
\[ C(LM)_i = \left( \sum_{j=1}^{\text{deg}(LM)} (|LM|)^{d} \right)^{-1} \]  

(33)

\[ C(LM) = \sum_i C(LM)_i \]  

(34)

where \( d \) is a specified topological distance (e.g. \( d = 10 \)). Most frequently used layer matrices were \( L^1W \) and \( LDS \).

Indices of centrocomplexity express the location vs. a vertex of high complexity (e.g. a vertex of high degree or a heteroatom). They are defined as

\[ X(LM)_i = \left( \sum_{j=0}^{\text{deg}(LM)} |LM|_{ij} 10^{-z} \right)^{t_i} \]  

(35)

\[ X(LM) = \sum_i X(LM)_i \]  

(36)

where \( z \) denotes the number of bits (of the integer part) of \( \max(|LM|_{ij}) \) in \( G \) and \( t_i \) is a weighting factor for heteroatom specification. Within TI-MTD model \( t_i \) was either \( EVG_i \) or \( EC_i \) parameters. Matrices \( L^1W \) and \( LDS \) and the corresponding indices for the graph \( G_j \), are shown in Table 2.

<table>
<thead>
<tr>
<th>( L^1W )</th>
<th>( C(L^1W)_i )</th>
<th>( X(L^1W)_i )</th>
<th>( LDS )</th>
<th>( C(LDS)_i )</th>
<th>( X(LDS)_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 3 3 1</td>
<td>0.2975</td>
<td>1.331</td>
<td>8 5 14 9</td>
<td>0.2082</td>
<td>8.051409</td>
</tr>
<tr>
<td>2 3 4 1 0</td>
<td>0.4654</td>
<td>3.410</td>
<td>5 22 9 0</td>
<td>0.3432</td>
<td>5.220900</td>
</tr>
<tr>
<td>3 2 4 2 0</td>
<td>0.4353</td>
<td>2.420</td>
<td>6 14 16 0</td>
<td>0.3286</td>
<td>6.141600</td>
</tr>
<tr>
<td>4 1 2 3 2</td>
<td>0.2818</td>
<td>1.232</td>
<td>9 6 5 16</td>
<td>0.2052</td>
<td>9.060516</td>
</tr>
<tr>
<td>5 1 3 3 1</td>
<td>0.2975</td>
<td>1.331</td>
<td>8 5 14 9</td>
<td>0.2082</td>
<td>8.051409</td>
</tr>
</tbody>
</table>

Global Index: \( 1.7774 \) 9.724 1.2934 36.525834

These indices showed good sensitivity in discriminating isomeric structures. They were used in studies on intramolecular ordering of subgraphs of various sizes as well as intermolecular ordering.
6. Applications of TI-MTD Model in QSAR and Molecular Design

6.1. Inhibitory Activity of the Trimethoprim Congeners

An application of TI-MTD on a set of trimethoprim (5-(3',4',5'-trimethoxy-benzyl)-2, 4-diaminopyrimidine) congeners18 (see Table 3) in presented. They show inhibitory activity upon dihydrofolate reductase of *Escherichia coli*.

Figure 7 shows the basic structure of this set while Table 3 lists the substituents for all structures $M_i$, $i = 1, ..., 22$ along with the observed inhibition constant (as $-\log K_{app}$).

![Figure 7. Basic structure of the set of trimethoprim derivatives](image)

**Table 3. 5-(Benzyl-Substituted)-2,4-Diaminopyrimidines, Inhibition Constant (as $-\log K_{app}$) and Topological Data for the Active Fragment.**

<table>
<thead>
<tr>
<th>No</th>
<th>$X$</th>
<th>$-\log K_{app}$</th>
<th>$D_{SC,S}$</th>
<th>$D_{SW,S}$</th>
<th>$D_{SC,N}$</th>
<th>$EC_{3'}$</th>
<th>$EC_{4'}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3',4',5'-(OCH$_3$)$_3$</td>
<td>8.87</td>
<td>317</td>
<td>0</td>
<td>252</td>
<td>1.1248</td>
<td>1.1248</td>
</tr>
<tr>
<td>2</td>
<td>3'-CF$_3$</td>
<td>7.02</td>
<td>113</td>
<td>148</td>
<td>82</td>
<td>1.3260</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3'-CH$_3$</td>
<td>6.70</td>
<td>57</td>
<td>0</td>
<td>52</td>
<td>0.9575</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>4'-CH$_3$</td>
<td>6.48</td>
<td>44</td>
<td>0</td>
<td>34</td>
<td>0</td>
<td>0.9575</td>
</tr>
<tr>
<td>5</td>
<td>3'-Cl</td>
<td>6.65</td>
<td>57</td>
<td>0</td>
<td>52</td>
<td>1.3717</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>4'-Cl</td>
<td>6.45</td>
<td>44</td>
<td>0</td>
<td>34</td>
<td>0</td>
<td>1.3717</td>
</tr>
<tr>
<td>7</td>
<td>3'-F</td>
<td>6.23</td>
<td>57</td>
<td>0</td>
<td>42</td>
<td>1.6514</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>4'-F</td>
<td>6.35</td>
<td>44</td>
<td>0</td>
<td>34</td>
<td>0</td>
<td>1.6514</td>
</tr>
<tr>
<td>9</td>
<td>H</td>
<td>6.18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>4'-N(CH$_3$)$_2$</td>
<td>6.78</td>
<td>91</td>
<td>47</td>
<td>68</td>
<td>0</td>
<td>1.0292</td>
</tr>
<tr>
<td>11</td>
<td>4'-NH$_2$</td>
<td>6.30</td>
<td>44</td>
<td>0</td>
<td>34</td>
<td>0</td>
<td>1.0644</td>
</tr>
<tr>
<td>12</td>
<td>4'-NCOCH$_3$</td>
<td>6.89</td>
<td>91</td>
<td>124</td>
<td>68</td>
<td>0</td>
<td>1.1021</td>
</tr>
<tr>
<td>13</td>
<td>4'-NO$_2$</td>
<td>6.20</td>
<td>91</td>
<td>47</td>
<td>34</td>
<td>0</td>
<td>1.4861</td>
</tr>
<tr>
<td>14</td>
<td>3'-OC$_6$H$_5$</td>
<td>6.89</td>
<td>113</td>
<td>230</td>
<td>82</td>
<td>1.4634</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>4'-OCF$_3$</td>
<td>6.57</td>
<td>91</td>
<td>180</td>
<td>51</td>
<td>0</td>
<td>1.4634</td>
</tr>
<tr>
<td>16</td>
<td>3'-OCH$_2$C$_6$H$_5$</td>
<td>6.93</td>
<td>113</td>
<td>321</td>
<td>82</td>
<td>1.4634</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>3'-OCH$_3$</td>
<td>6.93</td>
<td>113</td>
<td>0</td>
<td>82</td>
<td>1.1248</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>3',4'-(OCH$_3$)$_2$</td>
<td>7.72</td>
<td>204</td>
<td>0</td>
<td>170</td>
<td>1.1248</td>
<td>1.1248</td>
</tr>
<tr>
<td>19</td>
<td>3',5'-(OCH$_3$)$_2$</td>
<td>8.38</td>
<td>226</td>
<td>0</td>
<td>184</td>
<td>1.1248</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>4'-OCH$_3$</td>
<td>6.82</td>
<td>91</td>
<td>0</td>
<td>68</td>
<td>0</td>
<td>1.1248</td>
</tr>
<tr>
<td>21</td>
<td>3'-OH</td>
<td>6.47</td>
<td>57</td>
<td>0</td>
<td>52</td>
<td>1.2325</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>3',4'-(OH)$_2$</td>
<td>6.46</td>
<td>101</td>
<td>0</td>
<td>76</td>
<td>1.2325</td>
<td>1.2325</td>
</tr>
</tbody>
</table>
In building the hypermolecule, (i.e. \([RS]_{init}\)), the leading structure \(M_r\) was taken the trimethoprime, having the substituents 3',4',5'-(OCH\(_3\))\(_3\). For first, a symmetric \([RS]_{S,init}\), was considers. It is illustrated in Figure 8, a.

![Figure 8](image_url)

Figure 8. Symmetric (a) and non-symmetric (b) hypermolecules of trimethoprime derivatives

Selected \(TI's\) are calculated for the whole hypermolecule and subsequently the descriptors \(TD_C\) and \(TD_W\), that represent the predictor variables to be introduced in the regression equation (7).

By the hypermolecule hypothesis, the cavity points (i.e. the \(C\) subset) are the atoms of the leading structure (i.e. trimethoprime) having the vector \(X_1 = (0001111011000000000)\). In words, the \(C\) set includes the points: 4, 5, 6, 7, 9 and 10 of the hypermolecule (Figure 8, a), while the remainders are considered wall points (i.e. the \(W\) subset) or irrelevant ones.

The corresponding descriptors \(TD(i)\) for candidates are calculated by summing the local values of the \(TI\) according to their \(X_i\) vector.

Thus, the cavity descriptor for the reference is

\[
TD_C(1) = TI_4 + TI_5 + TI_6 + TI_7 + TI_9 + TI_{10}
\]  

(37)

Analogously one calculates for all candidates \(M_i\). For the lead the wall descriptor is \(TD_W(1) = 0\), since all its points are considered, by the above hypothesis, entering in the cavity. The descriptors \(TD_C\) and \(TD_W\) thus obtained are used in the regression analysis.

As \(TD\), the distance sum \(DS\) (i.e. the sum of topological distances from \(i\) to all other vertices in a molecular graph) was here used. For the trimethoprime set, the descriptors included in Table 3 led to the following regression equation
In eq 38, the coefficient of the wall descriptor is one order of magnitude lower than the coefficient of the cavity descriptor (at the same magnitude of descriptor values). It suggests that, in this case, the wall interaction could be neglected. Indeed, the statistics of the monovariate equation (39) is at least as good as the bivariate regression equation (38)

\[ Y_{\text{calc}} = 5.9317 + 0.0092DS_{c,s} \]
\[ n = 22; r = 0.9506; s = 0.217; F = 187.635 \]  

(39)

The good quality of the regression equations (38 and 39) indicates that $DS$ is an appropriate topological descriptor. It encloses the steric dimension of these molecular structures.

Since for the characterization of the receptor-effector system, the steric factor is, however, not sufficient, the electronegativity valences $EC$ (see Sect. 5.2) were introduced as an additional predictor variable. By the viewpoint of electronegativity, the attachment sites of substituents to the benzyl are important. A new regression equation is thus obtained

\[ Y_{\text{calc}} = 6.2548 + 0.01049DS_{c,S} - 0.2466EC_3 - 0.3056EC_4 + 0.1473EC_5 \]
\[ n = 22, r = 0.9780, s = 0.1584, F = 93.69 \]  

(40)

The Student test for $EC_5$ suggests the possibility of neglecting this variable

\[ Y_{\text{calc}} = 6.2477 + 0.01124DS_{c,S} - 0.2807EC_3 - 0.3374EC_4 \]
\[ n = 22, r = 0.9774, s = 0.2563, F = 128.16 \]  

(41)

Note that in calculating the above descriptors, in case of large substituents (such as -NO$_2$, -NHCOCH$_3$, -OCF$_3$) only the atom attached to the benzyl is considered to be of C-type.

The map $[RS]_{\text{opt}}$ above calculated is based on a symmetric hypermolecule (Figure 8, (a)). It is also conceivable a non-symmetric hypermolecule (Figure 8, b). Indeed, in the new hypothesis, the compounds $M_3$, $M_5$, and $M_7$ show a residual value $\Delta = Y_{\text{obs}} - Y_{\text{calc}}$ lower that in case of symmetric hypermolecule.

With the non-symmetric $[RS]_{\text{opt}}$, the correlation equation becomes

\[ Y_{\text{calc}} = -6.288081 + 0.012941DS_{C,N} - 0.304293EC_3 - 0.294274EC_4 \]
\[ n = 22, r = 0.9931, s = 0.086, F = 433.20 \]  

(42)
The statistics of eq 42 are far more improved and it can be used for predicting novel bioactive structures. The plot of \( Y_{\text{obs}} \) vs \( Y_{\text{calc}} \) (cf. eq 42) is shown in Figure 9.

![Figure 9. The plot \( Y_{\text{obs}} = f(Y_{\text{calc}}) \).](image)

The values \( Y_{\text{obs}}, Y_{\text{calc}} \) as well as the residuals \( \Delta \) for the whole set of trimethoprim congeners are listed in Table 4.

<table>
<thead>
<tr>
<th>( M_i )</th>
<th>( Y_{\text{obs}} )</th>
<th>( Y_{\text{calc}} )</th>
<th>( \Delta = Y_{\text{obs}} - Y_{\text{calc}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>8.8700</td>
<td>8.8803</td>
<td>-0.0103</td>
</tr>
<tr>
<td>2:</td>
<td>7.0200</td>
<td>6.9328</td>
<td>0.0872</td>
</tr>
<tr>
<td>3:</td>
<td>6.7000</td>
<td>6.6557</td>
<td>0.0443</td>
</tr>
<tr>
<td>4:</td>
<td>6.4800</td>
<td>6.4388</td>
<td>0.0412</td>
</tr>
<tr>
<td>5:</td>
<td>6.6500</td>
<td>6.5287</td>
<td>0.1213</td>
</tr>
<tr>
<td>6:</td>
<td>6.4500</td>
<td>6.3193</td>
<td>0.1307</td>
</tr>
<tr>
<td>7:</td>
<td>6.2300</td>
<td>6.3130</td>
<td>-0.0830</td>
</tr>
<tr>
<td>8:</td>
<td>6.3500</td>
<td>6.2386</td>
<td>0.1114</td>
</tr>
<tr>
<td>9:</td>
<td>6.1800</td>
<td>6.2730</td>
<td>-0.0930</td>
</tr>
<tr>
<td>10:</td>
<td>6.7800</td>
<td>6.8602</td>
<td>-0.0802</td>
</tr>
<tr>
<td>11:</td>
<td>6.3000</td>
<td>6.4080</td>
<td>-0.1080</td>
</tr>
<tr>
<td>12:</td>
<td>6.8900</td>
<td>6.8392</td>
<td>0.0508</td>
</tr>
<tr>
<td>13:</td>
<td>6.2000</td>
<td>6.2863</td>
<td>-0.0863</td>
</tr>
<tr>
<td>14:</td>
<td>6.8900</td>
<td>6.8907</td>
<td>-0.0007</td>
</tr>
<tr>
<td>15:</td>
<td>6.5700</td>
<td>6.5139</td>
<td>0.0561</td>
</tr>
<tr>
<td>16:</td>
<td>6.9900</td>
<td>6.8907</td>
<td>0.0993</td>
</tr>
</tbody>
</table>
The predicting ability of eq 42 was tested by a cross validation (leave 5% out) procedure. The test supplied eq 43

\[
Y_{\text{calc}} = -0.00345 + 0.99992 Y_{\text{cv}}
\]

\[n = 22, \quad r_{cv} = 0.989, \quad s = 0.105, \quad F = 478.0\]  \hspace{1cm} (43)

for which the plot \(Y_{\text{obs}} = f(Y_{\text{cv}})\) is shown in Figure 10.

Figure 10. The plot \(Y_{\text{obs}} = f(Y_{\text{cv}})\) cf eq 43.

The good predicting ability of eq 42 allowed to estimate the activity of a test structure \(M_{23}\), having the substituents \(3',5'-\text{OCH}_3-4'\)-\(\text{O(CH}_2\text{)OCH}_3\): \(Y_{\text{calc}} = 8.336\) vs \(Y_{\text{obs}} = 8.35\), \(\Delta = 0.014\). The substituent -\(\text{O(CH}_2\text{)OCH}_3\) is considered a voluminous one and, by this reason, it overlap the cavity set only in position 7 of\([RS]_{N_{opt}}\), cf Figure 11. Note that this structure was dropped out of the training set.
Figure 11. The optimal RS map for the trimethoprime series

The parameters used in calculating the activity of the test structure $M_{23}$ are: $DSC = 218; EC_3 = 1.1248; EC_4 = 1.4634$.

The same eq 42 allowed the prediction of three novel trimethoprime derivatives, for which good activity is expected:

$M_{24} (-3',4',5'-(NHCH_3)_3)$: 
\[
DSC = 252 \\
EC_3 = 1.0379 \\
EC_4 = 1.0379
\]

$M_{25} (-3',4',5'-(SCH_3)_3)$: 
\[
DSC = 252 \\
EC_3 = 1.0073 \\
EC_4 = 1.0073
\]

$M_{26} (-4'\-NH_2\-3', 5'-(OCH_3)_2)$: 
\[
DSC = 218 \\
EC_3 = 1.1248 \\
EC_4 = 1.0644
\]

In comparison to the trimethoprime ($Y_{obs} = 8.87$) the proposed structures appear as promising compounds.

6.2. Activity of 2-Substituted 1, 2, 3-Thiadiazol-5-Sulphonamides

For validating the correlating ability of TI-MTD model, the activity of a set of 2-substituted 1, 2, 3-thiadiazol-5-sulphonamides is further investigated (Figure 12). They show an inhibitory activity upon the carbonic anhydrase (a metaloenzyme) higher than that shown by aromatic sulphonamides.
Note that this set was also tested by MTD, MVD and CoMFA approaches.53

![Figure 12. 2-Substituted 1, 2, 3-Thiadiazol-5-Sulphonamides](image)

Table 5 lists the set of 11 thiazol-5-sulfonamides along with the observed activity ($Y_{obs} = \log \text{II}_{50}$).

<table>
<thead>
<tr>
<th>No</th>
<th>X</th>
<th>log $\text{II}_{50}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NHSO$_2$C$_6$H$_4$-4-NHCOCH$_3$</td>
<td>3.38</td>
</tr>
<tr>
<td>2</td>
<td>NHSO$_2$C$_6$H$_4$-Cl</td>
<td>3.29</td>
</tr>
<tr>
<td>3</td>
<td>NHSO$_2$C$_6$H$_5$</td>
<td>3.16</td>
</tr>
<tr>
<td>4</td>
<td>NHCOC$_6$H$_5$</td>
<td>2.95</td>
</tr>
<tr>
<td>5</td>
<td>N(CH$_3$)COCH$_3$</td>
<td>2.66</td>
</tr>
<tr>
<td>6</td>
<td>NHC$_6$H$_5$</td>
<td>2062</td>
</tr>
<tr>
<td>7</td>
<td>NHCOCH$_3$</td>
<td>2.52</td>
</tr>
<tr>
<td>8</td>
<td>NHCHO</td>
<td>2.28</td>
</tr>
<tr>
<td>9</td>
<td>NHCH$_3$</td>
<td>1.96</td>
</tr>
<tr>
<td>10</td>
<td>H</td>
<td>1.90</td>
</tr>
<tr>
<td>11</td>
<td>NH$_2$</td>
<td>1.64</td>
</tr>
</tbody>
</table>

The hypermolecule for the set of 11 thiazol-5-sulfonamides is shown in Figure 13.

![Figure 13. The hypermolecule for the set of 11 thiazol-5-sulfonamides](image)
By using $DS$ as topological descriptor (see the Sect. 6.1) the following regression equation was obtained

$$Y_{\text{obs}} = 1.79877 + 0.00287DS_C$$

$$n = 11 \quad r = 0.988; \quad s = 0.0943; \quad F = 371$$

(44)

The plot $Y_{\text{obs}} = f(DS)$ is shown in Figure 14.

![Figure 14. The plot $Y_{\text{obs}} = f(DS)$ for the set of 11 tiadiazol-5-sulfonamides.](image)

The cross validation test (leave 5% out) performed on eq 44 led to eq 45

$$Y_{\text{obs}} = -0.01761 + 1.00387Y_{\text{cv}}$$

$$n = 11; \quad r_{\text{cv}} = 0.984; \quad s = 0.108; \quad F = 280$$

(45)

for which the plot $Y_{\text{obs}} = f(Y_{\text{cv}})$ is shown in Figure 15.
Figure 15. The plot $Y_{obs} = f(Y_{cv})$ for eq 45.

On the above set, the MTD and MVD approaches provided the following results:\textsuperscript{53}

\begin{align*}
Y_{obs} &= 4.745 - 0.350 \text { MTD} \\
n &= 11; r = 0.992; s = 0.078; F = 249 \quad (46)
\end{align*}

\begin{align*}
Y_{obs} &= 4.406 - 0.013 \text { MVD} \\
n &= 11; r = 0.993; s = 0.075; F = 272 \quad (47)
\end{align*}

In MVD the volumes are calculated by Monte Carlo method. It can be seen that the results of TI-MTD are quite close to those given by approaches using 3D geometries, at a much lower cpu cost.

7. Conclusions

TI-MTD is a 2D topological approach for ligand-receptor interaction description. It provides some grid-descriptors (steric and electrostatic) of such interaction and fine-tunes the model’s statistical quality by changing both the alignment modes and the receptor map until an optimal receptor map is obtained.

The method can be used in predicting novel bioactive structures. An advantage of this topological approach is the relatively reduced computational cost in comparison to the methods using molecular mechanic or quantum calculations. It can be used as a pre-test before investigation by more elaborated procedures. A further comparison of the TI-MTD results with those provided by other 3D-QSAR approaches could be relevant in validation of this method.
References


