Chemometric Study of the Antiproliferative Activity of Some New Hydantoin Derivatives: Assessment of Activity and Chromatographic Lipophilicity Data

Tatjana Lj. Djaković Sekulić, Adam Smoliński, Nemanja P. Trišović, Gordana S. Ušćumlić and Biljana Dj. Božić

*Department of Chemistry, Biochemistry and Environmental Protection, Faculty of Sciences, University of Novi Sad, 21000 Novi Sad, Republic of Serbia

Department of Energy Saving and Air Protection, Central Mining Institute, 40-166 Katowice, Poland

Faculty of Technology and Metallurgy and Faculty of Biology, University of Belgrade, 11000 Belgrade, Republic of Serbia

Cancer is the major health problem affecting the mankind of today. Most of the drugs used in traditional chemotherapy are very limited and the discovery of novel, more active, more selective and less toxic ones is still very intensive. A chemometric approach was applied in the study of antiproliferative activity against human colon cancer and breast cancer as well as in the study of lipophilicity of 3-(4-substituted benzyl)-5-ethyl-5-phenyl- and 3-(4-substituted benzyl)-5,5-diphenylhydantoins. Hierarchical clustering analysis (HCA) shows that the investigated hydantoins have higher antiproliferative activity against human breast cancer cells than against human colon cancer cells. However, some hydantoins at the highest applied concentration reverse antiproliferative effect, higher against the human colon cancer cells and lower against human breast cancer cells. Principal component analysis (PCA) gives better insight into the activity of hydantoins related to their structural changes. It distinguishes more active compounds from the less active ones according to various criteria. Generally, more lipophilic 5,5-diphenylhydantoins exhibit a higher antiproliferative activity comparing to less lipophilic 5-ethyl-5-phenylhydantoins. Also, a substituent attached to benzyl moieties affects the activity additionally. The activity is particularly pronounced for compounds with cyano, methyl, chloro and bromo group. Halogen substituent were superior in antiproliferative capacity particularly in the series of 5,5-diphenylhydantoins.

Keywords: 3-(4-substituted benzyl)-5-phenylhydantoins, antiproliferative activity, lipophilicity, hierarchical clustering analysis, principal component analysis

Introduction

The major health problem affecting the mankind of today is cancer. Several types of cancers have been identified that attack major organs such as lungs, brain, kidneys, colon, breasts and stomach. Regardless of major advances in many areas of modern medicine over the past 100 years, a successful treatment of cancer still remains a significant challenge. Most of the drugs used in traditional chemotherapy are very limited because the discovery of novel agents that can kill tumor cells selectively or inhibit their proliferation without general toxicity is very difficult. Hence, the search for new, more active, more selective and less toxic compounds is still very intense. One of the approaches is promotion of the structural diversity, which can be achieved by the attachment of pharmacophoric groups to the significant molecular scaffold. Such examples are di- and tri-substituted hydantoins, which have been widely used in biological screening resulting in numerous pharmaceutical applications.1-3 Until today, hydantoin derivatives (imidazolidine-2,4-dione) have been identified as anticonvulsants,4 antiarhythmics,5 antimuscarinics,6 antidiabetics,7 and antitumor agents.8-11

In this study, 15 newly synthesized hydantoins were investigated in order to better understand the structure-activity relationships of the new derivatives. Similarities and differences between the tested compounds were explored and visualized using two chemometric methods:
principal component analysis (PCA) and hierarchical clustering analysis (HCA). Besides antiproliferative activity against cell lines MDA-MB-231 (human breast adenocarcinoma) and HCT-116 (human colon cancer), the compounds were chromatographically characterized. The partition retention parameters were determined, i.e., lipophilicity of the compounds. Lipophilicity is closely correlated with adsorption, distribution, metabolism and excretion (ADME).

Experimental

Investigated compounds and their antiproliferative activity

The investigated hydantoin derivatives (Figure 1) were synthesized according to the procedure published earlier. The solutions of the investigated compounds were prepared by dissolving 0.02 g of each compound in 2 mL of ethanol.

Figure 1. General formula of investigated 3-(4-substituted benzyl)-5-phenylhydantoins. R
1
 denotes either ethyl (for compounds 1-7) or phenyl (for compounds 8-15) group; for R
2
, see Table 1.

The antiproliferative activity of the investigated compounds (Table 1) was taken from reference 12.

Chromatographic lipophilicity data

Chromatography was performed on 10 × 10 cm high-performance thin-layer chromatographic (HPTLC) plates of silica gel 60 RP-18 F254s (Merck, Darmstadt, Germany). Mobile phase used was acetonitrile-water (with increasing acetonitrile content from 55% to 75%, v/v); increments of 5% v/v. Acetonitrile (Acros organics, Geel, Belgium) of p.a. purity was used. Chromatograms were developed by the ascending technique at room temperature (22 ± 2 °C) without previous saturation of the chamber (Camag, Muttenz, Switzerland) with mobile phase vapor. The developing distance was 4.5 cm. After development, the plates were dried at room temperature, and individual chromatographic zones were detected under UV illumination at 254 nm.

The R
M
 values were calculated according to Bate-Smith and Westall equation: $R_M = \log \left( \frac{1}{R_p} \right) - 1$,

where R
p
 is the retardation factor defined as the distance travelled by the centre of the spot divided by the distance simultaneously travelled by the mobile phase. The R
p
 values were arithmetic mean of at least three measurements for each solute-stationary-mobile phase combination. The R
M
 value of each compound was examined as a function of mobile phase composition:

$$R_M = R_M^0 - S \phi$$  \hspace{1cm} (1)

In equation 1, $\phi$ is the volume fraction of organic solvent in the mobile phase, $R_M^0$ is the intercept, and S is the slope of equation 1. $R_M^0$ describes the overall effect of the intermolecular interaction of the compound with the stationary and mobile phases. It is often used as a measure of lipophilicity.

Values of R
M
 for the investigated compounds are given in Table 1.

Chemometric analysis

HCA and PCA were carried out using MATLAB software, version 6.1.0.450. For HCA, the studied data were organized in data matrix X (15 × 12), where rows correspond to 15 studied compounds (hydantoins), whereas columns represent antiproliferative activity evaluated against human breast cancer MDA-MB-231 (activity I-VI) and human colon cancer HCT116 (activity VII-XII) cell lines, respectively. A slightly modified data matrix X
1
 (15 × 13) was used for PCA. The additional column in data matrix is for chromatographic lipophilicity parameters (XIII).

A chemometric analysis of the studied compounds was performed for centred and standardized data Xc (15 × 12) and Xc
1
 (15 × 13).

Results and Discussion

The previous investigation shows that hydantoins 1-15 (see Figure 1 and Table 1) exhibit a considerable antiproliferative activity. The presence of different substituents on both hydantoin (substituent R
1
) and benzyl moieties (substituent R
2
) affects the antiproliferative activity. Activity is particularly pronounced for compounds 1, 10-12 and 14 (with R
1
 = cyano (CN), methyl (Me), chloro (Cl) and bromo (Br)).

Understanding the relationship between the activity, structure and the physicochemical properties of the examined compound enables identification of potentially bioactive compounds. Thorough classification and comparison of various hydantoins on the basis of their activity is objectively possible only by application of appropriate chemometric methods. Interpretation and discussion of results presented by tables or/and graphics are

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Compound & Activity & Lipophilicity & Conclusion \\
\hline
1 & High & High & Strong antioxidant activity \\
\hline
2 & Low & Low & Poor antioxidant activity \\
\hline
3 & Moderate & Moderate & Moderate antioxidant activity \\
\hline
4 & High & High & Strong antioxidant activity \\
\hline
5 & Low & Low & Poor antioxidant activity \\
\hline
6 & Moderate & Moderate & Moderate antioxidant activity \\
\hline
7 & High & High & Strong antioxidant activity \\
\hline
8 & Low & Low & Poor antioxidant activity \\
\hline
9 & Moderate & Moderate & Moderate antioxidant activity \\
\hline
10 & High & High & Strong antioxidant activity \\
\hline
11 & Low & Low & Poor antioxidant activity \\
\hline
12 & Moderate & Moderate & Moderate antioxidant activity \\
\hline
13 & High & High & Strong antioxidant activity \\
\hline
14 & Low & Low & Poor antioxidant activity \\
\hline
15 & Moderate & Moderate & Moderate antioxidant activity \\
\hline
\end{tabular}
\end{table}
Table 1. Antiproliferative effects of the investigated compounds on two cell lines and chromatographic retention data

<table>
<thead>
<tr>
<th>Substituent</th>
<th>Inhibition of proliferation of MDA-MB-231 cells / %</th>
<th>Inhibition of proliferation of HCT-116 cells / %</th>
<th>R² ※</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compound</td>
<td>100 μmol</td>
<td>50 μmol</td>
<td>10 μmol</td>
</tr>
<tr>
<td>R₁, R₂</td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>1</td>
<td>Et</td>
<td>H</td>
<td>20.88</td>
</tr>
<tr>
<td>2</td>
<td>Et</td>
<td>MeO</td>
<td>1.72</td>
</tr>
<tr>
<td>3</td>
<td>Et</td>
<td>CN</td>
<td>5.10</td>
</tr>
<tr>
<td>4</td>
<td>Et</td>
<td>Br</td>
<td>2.85</td>
</tr>
<tr>
<td>5</td>
<td>Et</td>
<td>Me</td>
<td>10.20</td>
</tr>
<tr>
<td>6</td>
<td>Et</td>
<td>Cl</td>
<td>4.09</td>
</tr>
<tr>
<td>7</td>
<td>Et</td>
<td>NO₂</td>
<td>7.47</td>
</tr>
<tr>
<td>8</td>
<td>Ph</td>
<td>H</td>
<td>6.58</td>
</tr>
<tr>
<td>9</td>
<td>Ph</td>
<td>MeO</td>
<td>6.47</td>
</tr>
<tr>
<td>10</td>
<td>Ph</td>
<td>CN</td>
<td>44.54</td>
</tr>
<tr>
<td>11</td>
<td>Ph</td>
<td>Me</td>
<td>15.72</td>
</tr>
<tr>
<td>12</td>
<td>Ph</td>
<td>Cl</td>
<td>30.20</td>
</tr>
<tr>
<td>13</td>
<td>Ph</td>
<td>t-Bu</td>
<td>19.04</td>
</tr>
<tr>
<td>14</td>
<td>Ph</td>
<td>Br</td>
<td>38.61</td>
</tr>
<tr>
<td>15</td>
<td>Ph</td>
<td>NO₂</td>
<td>7.18</td>
</tr>
</tbody>
</table>

Acronyms for functional groups: Me for methyl, Et for ethyl, t-Bu for tert-butyl, Ph for phenyl, MeO for metoxy, CN for cyano, NO₂ for nitro, Cl for chloro, and Br for bromo.

Hierarchical clustering analysis

Hierarchical clustering analysis (HCA) is a convenient method for studies of similarities (or dissimilarities) between objects in the variables space, or similarities (dissimilarities) between variables in the objects space.18–21 HCA is characterized by the similarity measure applied, and the way the resulting sub-clusters are merged (linked). Different clustering methods are available. Their choice depends on the data studied and the particular purpose of the method applied. In the present study, the Ward linkage method based on the Euclidean distance was used.22,23 The results of HCA presented as dendrograms, reveal data structure, i.e., the sub-groups of objects in the variables space or the sub-groups of variables in the objects space.

The dendrogram presented in Figure 2a groups the hydantoins into four main clusters: (i) cluster A contains compounds 1, 10 and 14; (ii) cluster B is constituted by hydantoins 2 and 15; (iii) cluster C is composed of hydantoins 3, 9 and 11-13; (iv) cluster D contains hydantoins 4-8.

The most active compounds are collected in cluster A, while the least active, mostly 5-ethyl-5-phenylhydantoins, in cluster D.

In the main clusters, C and D, the following sub-structures can be distinguished: cluster C contains two subgroups: C1 (compounds 3, 11 and 13) and C2 (compounds 9 and 12) whereas cluster D is divided into two subgroups: D1 (compounds 4, 7 and 8) and D2 (compounds 5 and 6).

The dendrogram constructed for the antiproliferative activity in the samples space (see Figure 2b) reveals three main classes: (i) Class A, containing the variables VIII-XII (antiproliferative activity against the human colon cancer cells HCT-116 at the concentration levels of 50, 10, 1, 0.1 and 0.01 μmol dm⁻³); (ii) Class B, constituted by variables I and VII (inhibition of proliferation at the concentration level of 100 μmol dm⁻³) and (iii) Class C, containing variables II-VI (inhibition of proliferation of human breast cancer cells MDA-MB-231 at the concentration levels of 50, 10, 1, 0.1 and 0.01 μmol dm⁻³).

The classes in Figure 2b show that antiproliferative activity of the investigated compounds against MDA-MB-231 (human breast cancer) is superior to that against HCT-116 cell lines (human colon cancer). The highest concentration level of 100 μmol dm⁻³ is specific and differs from lower concentrations. This is due to the opposite activity effect of compounds 3-9 and 12, which exhibit higher antiproliferative activity against the human colon cancer cells HCT-116 compared to the inhibition of cells in MDA-MB-231. Figure 2a does not give any obvious classification of the investigated compounds. The

extremely subjective. In order to avoid subjective analysis, we used the chemometric methods, PCA and HCA. Finally, since all active hydantoins are lipophilic, a further objective of our work was to determine their lipophilicity.
Chemometric Study of the Antiproliferative Activity of Some New Hydantoin Derivatives

A color map is more informative, though. It complements the dendrogram of the studied antiproliferative activity in the hydantoin compounds space of the studied data sorted according to the Ward linkage method (see Figure 3).

Hydantoins 1, 10 and 14 belonging to cluster A exhibit the highest antiproliferative activity against human breast cancer MDA-MB-231 at the concentration of 100 μmol dm⁻³ (antiproliferative activity I).

Hydantoins 2 and 15, collected in cluster B, stand out from the rest of the tested compounds due to the relatively high inhibition of proliferation of human breast cancer cells MDA-MB-231 at the concentration levels of 50 and 1 μmol dm⁻³ (antiproliferative activities II and IV). Moreover, compound 2 is characterized by the highest antiproliferative activities II and IV whereas compound 15 is additionally characterized by the highest antiproliferative activity V.

Cluster C, containing hydantoins 3, 9 and 11-13, is unique due to high antiproliferative activities I-IV and VII. Additionally, compound 11 is characterized by the lowest antiproliferative activity XII among all the studied compounds. Cluster D, which contains hydantoins 4-8, is characterized by the highest antiproliferative activity VII (the highest concentration level of 100 μmol dm⁻³ applied to human colon cancer cell lines).

To sum up, the dendrograms in Figure 2 reveal the data structure but do not enable interpretation of the relationships between the studied compounds and activity data. These limitations of the HCA were overcome by adding a color map of the studied data (see Figure 3) sorted according to the order of compounds and the order of antiproliferative activity observed on their respective dendrograms (as previously shown in Figure 2). Parallel interpretation of the dendrogram of antiproliferative activity and the color map of the data leads to conclusion that the highest antiproliferative activity of the investigated compounds is obtained at the concentration of 100 μmol dm⁻³ for both cell lines. At the concentrations below 100 μmol dm⁻³, antiproliferative activity against the human colon cancer

Figure 2. Dendrogram of the studied (a) hydantoins in the space of antiproliferative activity evaluated against human breast cancer MDA-MB-231 and human colon cancer HCT116 cell lines and (b) antiproliferative activity in the samples space by the Ward linkage method using Euclidean distance as the similarity measure.

Figure 3. Dendrogram of studied antiproliferative activity in hydantoin compounds space with the color map of the studied data sorted according to the Ward linkage method.
cells is lower comparing to inhibition of proliferation of human breast cancer cells. The previous investigations show that the investigated compounds are non-toxic to normal rat peritoneal cells. Therefore, the pronounced antiproliferative activity against MDA-MB-231 indicates their significant anticancer potential in the treatment of breast cancer.

**Principal component analysis**

The principal components analysis, PCA, is one of the most popular methods used in reduction of data dimensionality. If the reduction of data dimensionality is effective, it enables data visualization and partial reduction of experimental error. PCA results in a decomposition of a data matrix \(X (m \times n)\), into matrices \(S (m \times k)\) and \(D (k \times n)\), where \(k\) denotes the number of significant factors:

\[
X (m \times n) = S (m \times k) \times D (k \times n) + E (m \times n) \tag{2}
\]

Matrix \(S\) is called a score matrix and, in the case of the data set concerned, it includes information on the studied compounds. Matrix \(D\), also called a loading matrix, describes antiproliferative activity and chromatographic data of the investigated compounds. Columns of matrix \(S\) and rows of matrix \(D\) are called principal components (PCs) or eigenvectors and are constructed to maximize the description of the data variance. The way of construction of consecutive PCs is enforced by their orthogonality and their arrangement according to decreasing data variance they describe. Since the PCs provide information not only on the data structure but also on the experimental error, the optimal number of the principal components (k) needs to be selected, enabling effective data modeling and simultaneously, partial elimination of experimental error. Matrix \(E\) includes the part of data which are not modelled by the constructed model with \(k\) PCs. To determine the number of significant components (PCs) of the studied data organized in matrix \(X_{c1}\) (15 \(\times\) 13), the percent of modelled variance was used (Table 2).

The PCA model with six significant PCs, describes 93.11\% of data variance (Table 2). Score plots and loading plots obtained as a result of this analysis are presented in Figure 4.

PC1 (describing 46.88\% of the total data variance) reflects the difference between compound 12 and all the remaining compounds. The largest difference is observed between compound 12 and compounds 2 and 3 (see score plots, Figure 4a). Compound 12 differs from compounds 2 and 3 not only in activity but also in lipophilicity. Moreover, three groups of compounds can be distinguished along PC1. The first one collects the most active compounds 10, 11 and 14, the second contains samples 7-9, 13 and 15, and the third contains compounds 1-6 (all with the same \(R_1\) (Et)). Based on the loading plots (see Figure 4b), a conclusion may be drawn that compounds 7-9, 13 and 15, compounds 10, 11 and 14 and, in particular, compound 12 differ from the remaining samples by their relatively high antiproliferative activities.

PC2, describing 17.00\% of the total variance, shows the difference between sample 12 and samples 10 and 15. Based on the loading plots, these differences were attributed to relatively high antiproliferative activity of compounds 10 and 15 and low antiproliferative activity of compound 12 observed for the concentration of 0.01 \(\mu\)mol dm\(^{-3}\) for HCT-116 cells (parameter XII). Compound 12 has \(R_2\) of lower polarity (Me) than 10 and 15 (CN and NO\(_2\)).

PC3 describes 13.70\% of the total variance and reflects the differences between compounds 7 and 11 and compound 10. The latter is unique mainly because of relatively high antiproliferative activity at the concentration of 100 \(\mu\)mol dm\(^{-3}\) for MDA-MB-231 cells (parameter I). Compounds 7 and 11 are characterized by relatively high antiproliferative activity obtained at concentrations of 1, 0.1 and 0.01 \(\mu\)mol dm\(^{-3}\) for MDA-MB-231 cells (parameters IV, V and VI, respectively). PC4 (describing 6.80\% of data variance) displays the highest difference between compounds 13 and 14 and all the remaining samples. Compound 14 and in particular, compound 13 are unique due to a high value of \(R_m^0\) (parameter XIII). PC5, describing 4.61\% of the total variance, reveals the singularity of compounds 7 and 9. PC6, describing 4.12\% of the total

<table>
<thead>
<tr>
<th>Principal component</th>
<th>Total variance / %</th>
<th>Cumulative / %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46.8761</td>
<td>46.8761</td>
</tr>
<tr>
<td>2</td>
<td>17.0023</td>
<td>63.8784</td>
</tr>
<tr>
<td>3</td>
<td>13.7018</td>
<td>77.5802</td>
</tr>
<tr>
<td>4</td>
<td>6.8014</td>
<td>84.3816</td>
</tr>
<tr>
<td>5</td>
<td>4.6113</td>
<td>88.9929</td>
</tr>
<tr>
<td>6</td>
<td>4.1213</td>
<td>93.1142</td>
</tr>
<tr>
<td>7</td>
<td>2.5618</td>
<td>95.6760</td>
</tr>
<tr>
<td>8</td>
<td>1.9072</td>
<td>97.5802</td>
</tr>
<tr>
<td>9</td>
<td>1.3278</td>
<td>98.9110</td>
</tr>
<tr>
<td>10</td>
<td>0.6628</td>
<td>99.5738</td>
</tr>
<tr>
<td>11</td>
<td>0.3755</td>
<td>99.9493</td>
</tr>
<tr>
<td>12</td>
<td>0.0443</td>
<td>99.9936</td>
</tr>
<tr>
<td>13</td>
<td>0.0064</td>
<td>100.0000</td>
</tr>
</tbody>
</table>
Figure 4. (a) Score plots and (b) loading plots of PCA for centered and standardized $X_c$ ($15 \times 13$) data set.
variance, shows the difference between compound 14 and all the remaining compounds. Compounds 7 and 9 are characterized by high antiproliferative activity observed at the concentration of 100 μmol dm$^{-3}$ for HCT-116 cells (parameter VII). The differences between compound 14 and all the remaining compounds were attributed to high antiproliferative activity reported at the concentration of 100 μmol dm$^{-3}$ for MDA-MB 231 cells (parameter I).

Concerning the compounds with an alkyl group at the benzyl substituent, compound 11 with the methyl group shows better inhibition of MDAMB-231 cell proliferation at the lowest concentration compared to compound 13 with the tert-butyl group. Compounds with a halogen substituent as $R_e$ (12 and 14) have a superior antiproliferative capacity to the other compounds in series of 5,5-diphenylhydantoins.

The antiproliferative activity of 5,5-diphenylhydantoins 8 ($R_1 = \text{OH}$), 9 ($R_2 = \text{OCH}_3$), and 12 ($R_e = \text{Cl}$) has previously been reported against the renal cancer cell lines A498 and UO-31 at the concentration of 10 μmol dm$^{-3}$.[27] A significant inhibition was observed; for example, maximal inhibition of 58.66% and 35.82% for A498 and UO-31, respectively, was achieved for compound 12. The replacement of the benzyl moiety with 2-oxo-2-(4-piperidin-1-yl)phenyl group results in the increase of antitumor activity to 68.14% and 53.43% against A498 and UO-31 cell, respectively. It is assumed that this derivative of hydantoin is located well in the biological target. Molecular docking studies in the same paper show the structural features essential for the antitumor activity of hydantoins. Those features are: the presence of two aromatic rings separated by an average distance, a hydrophobic moiety, and an H-bonding acceptor. All those features are present in the structure of the hydantoins examined in our study.

Conclusions

Most of the investigated compounds show superior antiproliferative activity against MDA-MB-231 to that against HCT-116 cell line. HCA differentiates the activity of hydantoins at high and low concentrations of the investigated compounds. Only in the highest investigated concentration (100 μmol dm$^{-3}$) antiproliferative activity is reverse for compounds 3-9 and 12.

PCA makes it possible to distinguish active compounds at different concentration levels. PC4 is an exception because it differentiates two most lipophilic compounds: 13 and 14 (with $R_e = \text{IBu and Br}$) from all the remaining samples. In general, compounds from of 5,5-diphenylhydantoins (8-15) exhibit higher antiproliferative activity compared to 5-ethyl-5-phenylhydantoins (1-7). This might be attributed to the introduction of the additional phenyl group in position 5, which contributes to the formation of a strong lipophilic shield around the hydantoin moiety. Hence, the lipophilicity of the investigated compounds may be an indication for their activity. A substituent attached to benzyl moieties (substituent $R_e$) has an additional impact on the activity. Hydantoins with the strong electron-donating and electron-withdrawing $R_e$ groups, such as NO$_2$, CN or OCH$_3$, decrease lipophilicity. Conversely, the introduction of an alkyl or halogen substituent yields the highest values of lipophilicity. As expected, the replacement of one of the phenyl groups in position 5 with an ethyl group results in a decreased lipophilicity of 5-ethyl-5-phenylhydantoins (1-7) compared to the corresponding derivatives of 5,5-diphenylhydantoins (8-15). It might be concluded that the higher the lipophilicity of the compounds (11, 12, and 14), the higher the inhibition.

Supplementary Information

Supplementary data is available free of charge at http://jbc.org.br as PDF file.

Acknowledgements

The presented results are part of the research project 172013 supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References


