Available online at www.sciencedirect.com
SCIENCE DIRECT
Polymer 46 (2005) 2783–2790
polymer
www.elsevier.com/locate/polymer

A topological sub-structural approach to the mutagenic activity in dental monomers. 3. Heterogeneous set of compounds

Maykel Pérez González[a,b,*], Maria del Carmen Teran Moldes[e], Yagamare Fall[e], Luiz Carlos Dias[c], Aliuska Morales Helguera[b,d]

[a]*Service Unit, Department of Drug Design, Experimental Sugar Cane Station 'Villa Clara-Cienfuegos', Ranchuelo, Villa Clara, Cuba*
[b]*Department of Drug Design, Chemical Bioactives Center, Central University of Las Villas, Santa Clara, Villa Clara. Cuba*
[c]*Universidade Estadual de Campinas, Instituto de Química 13084-971, Cx.Postal 6154 Campinas, SP, Brazil*
[d]*Department of Chemistry, Faculty of Chemistry and Pharmacy, Central University of Las Villas, Santa Clara, Villa Clara, Cuba*
[e]*Departamento de Química Orgánica, Facultad de Química, Universidad de Vigo, 36200 Vigo, Spain*

Received 8 October 2004; received in revised form 19 January 2005; accepted 22 January 2005

Abstract

The topological sub-structural molecular design (TOPS-MODE) approach has been applied to the study of mutagenic properties in a heterogeneous set of dental monomers. A model able to describe close to 90% of the experimental variance in the values for mutagenic activity of 41 dental monomers through genetic algorithm was developed with the use of the mentioned approach. Also, a study for the determination of the optimal number of variables in the equation and potential outliers was carried out. Finally, the TOPS-MODE approach was used to derive the contribution of different fragments to the mutagenic activity.

Keywords: Epoxides; Methacrylates; QSAR

1. Introduction

The aim of toxicological testing is to predict possible adverse effect in humans when exposed to, e.g. chemicals whether used as industrial chemicals, pharmaceuticals or polymer or their monomers. Animal models are predominantly used in identifying potential chemical hazards. The use of laboratory animals raises ethical concern. For that reason, the theoretical models represent an alternative to the assaying of chemical compounds for determining their toxicological properties on living organisms in the laboratory. The quantitative structure–toxicity relationships (QSTR) are predictive tools for a preliminary evaluation of the hazard of chemical compounds by using computer-aided models [1–7]. The constructions of these models consist of data sets of compounds, their computational analysis, hypothesis generation and toxicity prediction, made and stored on a computer. By this means, the expensive, time-consuming and in many cases animal aggressive bioassays are made only after exploring the initial concepts with computational models. Thus, QSTR models have great potential for designing new dental resins that will possess favorable biocompatibility profiles. The method is somewhat new in dental materials research, but QSAR have been widely applied to rational drug design and successfully used to predict the structures of novel compounds and protein properties [8–12].

The idea of employing computational approaches for predicting the mutagenicity of novel chemicals such as polymer and their respective monomers [13] was developed by us in recent publications [14–16]. In these works we demonstrated the value of the TOPS MODE approach in this area. However, in order to demonstrate the generality of this approach it is absolutely necessary to carry out this study in dental monomers using a heterogeneous set of compounds.

* Corresponding author. Address: Service Unit, Department of Drug Design, Experimental Sugar Cane Station 'Villa Clara-Cienfuegos', Ranchuelo, Villa Clara, Cuba. Tel.: +53 422 81473; fax: +53 422 81230.
E-mail address: mpgonzalez76@yahoo.es (M.P. González).

0032-3861/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.polymer.2005.01.064

2. Materials and methods

2.1. The TOPS-MODE approach

The TOPS-MODE approach is based on the calculation of the spectral moments of the so-called bond matrix [17,18], whose theoretical basis has been described in previous reports [19–21]. Nevertheless, an overview of this approach will be given below.

The bond matrix is defined as a square and symmetric matrix whose entries are ones or zeros if the corresponding bonds are adjacent or not. The order of this matrix (m) is the number of bonds in the molecular graph, being two bonds adjacent if they are incident to a common atom. The spectral moments of the edge adjacency matrix are defined as the traces, that is, the sum of the main diagonal, of the different powers of such a matrix.

To apply the present approach to the structure–toxicity relationship, the following steps should be followed. First, to select an adequate training set with certain structural diversity. Second, to draw the hydrogen depleted molecular graphs for each molecule of the training set. The third step is to differentiate the molecular bonds with appropriate weights. The fourth, to compute the spectral moments of the bond matrix for each molecule of the data set. Fifth, to find a qualitative structure–toxicity relationship by using a regression analysis:

$$P = a_0\mu_0 + a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + \ldots a_k\mu_k + b \quad (1)$$

where P is the studied property, in our case, the slopes of revertants vs. nanomoles of test chemical in the *Salmonella* test strain TA100 with the natural logarithm of the slopes (ln TA100), μ_k is the kth spectral moment, and the a_ks are the coefficients obtained by the linear regression. Sixth, to test the predictive capacity of the regression model by cross-validation procedures and an external prediction set. Finally, to compute the contribution of the different substructures to determine their quantitative contribution to the mutagenicity of the molecules studied.

2.2. Selection of bond weights and calculation of molecular descriptors

Taking into consideration the QSAR's for structure–mutagenicity correlation in dental monomers described previously by Yourtee et al. and our recent works [13–16], several bond weights such as hydrophobicity [22], polar surface area [23], molar refraction [24], Gasteiger-Marsilli atomic charge [25], polarizability [26], and van der Waals atomic radii [27] were used for computing the spectral moments of the bond matrix. Due to the fact that most of the approaches for computing physicochemical properties from fragment are based on atom-additive methods, several transform from atomic to bond contributions were carried out. The way in which these atomic contributions were transformed into bond contributions have been described by Estrada et al. [28]:

$$w(i,j) = \frac{w_i}{\delta_j} + \frac{w_j}{\delta_i} \quad (2)$$

where w_i and δ_i are the atomic weight and vertex degree of the atom i. The calculation of the descriptors was carried out with the computer software MODESLAB 1.0 [29]. The input of the software consists of SMILES codes for each compound [30]. We calculated the first 15 spectral moments ($\mu_1-\mu_{15}$) for each bond weight and the number of bonds in the molecules (μ_0). In this sense, 81 molecular descriptors (variables) were included in the analysis.

2.3. Computation of substructure contributions

The quantitative contribution (negative or positive) of a given substructure to the mutagenic activity of dental monomers was calculated. The general methodology used in this computational approach is as follows. In the first step all the substructure whose contribution we would like to determine were selected. The spectral moments for each substructure were calculated and their contributions to the mutagenic activity were obtained by substitution of their spectral moments in the regression model.

In this study, 18 fragments were selected to compute their contribution to the mutagenicity. Considering those fragments with high or poor contribution to the mutagenic activity, new hypothetical polymers with or without mutagenic effect of their monomers could be designed.

2.4. Data set and computational strategies

The series of 53 dental monomers given in Table 1 for which mutagenicity data was reported in the literature [14–16] was used in the current work. The mutagenic parameter studied here is the slopes of revertants vs. nanomoles of test chemical in the Salmonella test strain TA100 with the natural logarithm of the slopes (ln(TAs100)) used in the QSARs models.

All statistical analysis and data exploration was carrying out using the Statistic 6.0 [31]. The most significant parameters were identified from the data set using genetic algorithm (GA) method [32]. The GA is a technique developed for model building. It begins with an initial population of QSAR models using randomly selected features. Least squares regression is used to generate the coefficients. The population is evolved by building new models based on variables of two better-scored models. The worst models in the populations are replaced by new models. The average fitness of the models increases as evolution proceeds. The equation term type was set to linear polynomial and the mutation probability was specified as 50%. The length of the equation was set to six terms and a constant. The

Table 1
Data set of dental monomers used in the current study

No.	Compound	ln TA100 Exp
1	4-methylphenyl glycidyl ether	0.860
2	4-t-butylphenyl glycidyl ether	−0.362
3	3,4-dimethoxyphenylpropylene oxide	−0.930
4	p-benzylphenylpropylene oxide	−1.080
5	p-biphenylpropylene oxide	0.620
6	R-glycidyl alcohol	−0.514
7	phenylpropylene oxide	−0.536
8	p-methoxyphenylpropylene oxide	−0.896
9	p-methylphenylpropylene oxide	−0.111
10	phenoxypropylene oxide	0.172
11	R-naphthyl glycidyl ether	2.230
12	S-naphthyl glycidyl ether	2.100
13	1.3-cyclohexadiene oxide	−2.160
14	3-pyranoylidene oxide	−3.320
15	4.5-cyclohexadiene oxide	−3.150
16	*cis*-3.6-dibromocyclohexene oxide	−0.530
17	*cis*-1.3-cyclohexatriene dioxide	−2.350
18	cyclohexene oxide	−3.190
19	cyclopentene oxide	−3.950
20	*trans*-3.6-dibromocyclohexene oxide	−0.450
21	*trans*-1.3-cyclohexadiene dioxide	−4.830
22	*trans*-1.3-cyclohexatriene dioxide	−2.550
23	*trans*-4.5–dibromocyclohexene dioxide	−4.020
24	urethane dimethacrylate	−3.470
25	glicidyl methacrylate	−1.920
26	bisphenol A dimethacrylate	−4.510
27	glicidyl acrylate	−0.750
28	2-chloro-3-dichloromethyl-4-methoxyfur-2-enone?	8.650
29	2-chloro-3-dibromomethyl-4-hydroxyfur-2-enone	8.610
30	2-bromo-3-dibromomethyl-4-hydroxyfur-2-enone	7.970
31	2-chloro-3-chloromethyl-4-hydroxyfur-2-enone	6.360
32	2-chloro-3-dibromomethylfur-2-enone	5.200
33	2,3-dichloro-4-hydroxyfur-2-enone	4.090
34	2-chloro-3-chloromethylfur-2-enone	1.590
35	2-chloro-3-bromomethylfur-2-enone	1.370
36	2,3-dichloro-4-methoxyfur-2-enone	0.990
37	2-chloro-3-methyl-4-ethoxyfur-2-enone	0.740
38	2-chloro-3-methyl-4-hydroxyfur-2-enone	0.410
39	2-bromo-3-methyl-4-hydroxyfur-2-enone	0.410
40	3-chloro-4-ethoxyfur-2-enone	−0.220
41	2-chloro-4-hydroxyfur-2-enone	−1.600
42	3-methyl-4-hydroxyfur-2-enone	−3.510
43	4-methoxyphenyl glycidyl ether	0.115
44	o-methoxyphenylpropylene oxide	−0.576
45	4-hydroxy-3-methoxyphenylpropylene oxide	−1.060
46	S-glycidyl alcohol	−1.040
47	4-pyranoylidene oxide	−3.960
48	norbornene oxide	−4.370
49	vinyl cyclohexene dioxide	−3.060
50	2-chloro-3-dichloromethyl-4-hydroxyfur-2-enone	8.750
51	2-bromo-3-bromomethyl-4-hydroxyfur-2-enone	6.040
52	2-bromo-3-chloromethylfur-2-enone	1.370
53	2,3-dichlorofur-2-enone	0.110

population size was established as 300. All equations were sorted by a statistical term, the correlation coefficients (R^2). The best equations were saved for subsequent studies for the examination of the regression coefficient, the standard deviation and the significance of the model.

2.5. Model validation

The last and most important part of QSAR model development is the validation. Most of the QSAR modeling methods implement the leave-one-out (LOO) cross-validation procedure. The outcome from the cross-validation procedure is cross-validate R^2 (q^2), which is used as a criterion of both robustness and predictive ability of the model [33].

Although, the low value of q^2 for the training set can indeed serve as an indicator of a low predictive ability of a model, the opposite is not necessarily true. Indeed, the high q^2 does not imply automatically a high predictive ability of the model. In order to both develop the model and validate it, one needs to split the whole available data set into the training and test set. Several methods can be used for this purpose. They include random selection, selection by groups of compounds where the whole group is included into training or a test set, selection of training set compounds with major features varying in a systematic way, etc. [34,35]. Here, we used a K-means cluster analysis for the mentioned selection of 11 compounds of the whole data set to perform the test set. These compounds were never used to develop the prediction function.

2.6. K-means cluster analysis

The k-MCA has been used in training and predicting series design [36–38]. The idea consists of carrying out a partition of the whole data set of compounds in several statistically representative classes of chemicals. Thence, one may select from the member of all these classes of training and predicting series. This procedure ensures that any chemical classes (as determined by the clusters derived from k-MCA) will be represented in both compounds series (training and predicting). It permits the designing of both training and predicting series, which are representative of the entire 'experimental universe'. Fig. 1 graphically illustrates the above-described procedure where a cluster analysis was carried out to select a representative sample for the prediction and training sets. A k-MCA splits dental monomers in four clusters with 10, 22, 8, 13 members and standard deviations of 0.62, 0.73, 0.53, and 0.81, respectively. Selection of the training and prediction set was carried out by taking, according to the Euclidean distance, compounds belonging to each cluster. To ensure a statistically acceptable data partition into several clusters, we took into account the number of members in each cluster and the standard deviation of the variables in the cluster (as low as possible). We also made an inspection of the standard deviation between and within clusters, the respective Fisher ratio and their p-level of significance considered to be lower than 0.05 [38]. All spectral moments (from μ_0 to μ_{15}) were used in both analysis; all variables show p-levels <0.05 for the Fisher test, and the results are depicted in Table 2.

The main conclusion should be achieved from k-MCA:

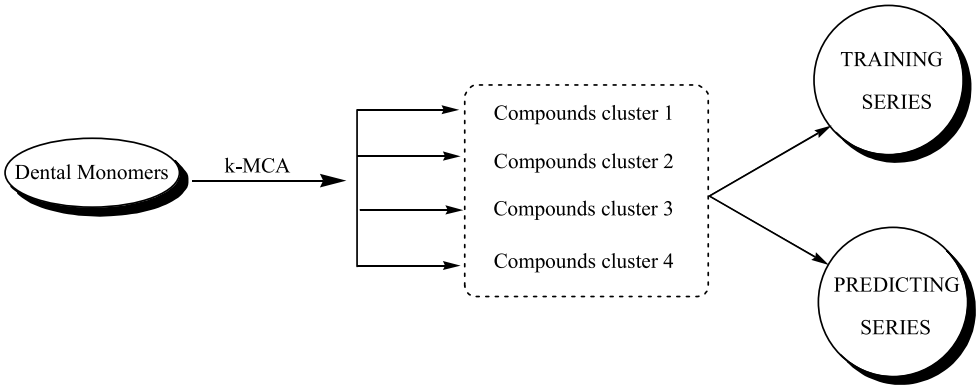

Fig. 1. Training and predicting series design throughout k-MCA.

the structural diversity of several up-to-date known dental monomers (as codified by TOPS-MODE descriptors) may be described at least by four statistically homogeneous clusters of chemicals. Anyhow, further conclusions about the mechanistic and molecular signification of these clusters seem to be speculative. Mainly, if it is considered that k-MCAs based partitions of data which consider not only 4 but also 3 or 5 clusters are statistically significant too (results not reported). However, the use of the k-MCA analysis here points to a structurally representative distribution of chemicals into training and predicting series [36].

3. Results and discussion

Once we perform a representative selection of training series it could be used to fit the model. The observed, predicted, residual and deleted residual of mutagenic activity values of dental monomers used in the training set appear in Table 3.

The best predictive model obtained for the mutagenic activity of the dental monomers is given below, together with the statistical parameters of the regression.

$$\ln TA100 = -0.189 + 4.06 \times 10^{-3} \mu_6^{\mathrm{Dip}} - 8.11 \times 10^{-9} \times \mu_{14}^{\mathrm{Dip}} - 0.004 \times \mu_3^{\mathrm{PS}} + 1.16 \times 10^{-5} \times \mu_5^{\mathrm{PS}} - 0.015 \times \mu_5^{\mathrm{GM}} \quad (3)$$

$$N = 42;\ S = 1.183;\ R^2 = 0.892;$$

$$F = 59.722;\ p < 0.0001;\ q^2 = 0.848;\ S_{\mathrm{cv}} = 1.412$$

where, N is the number of compounds used in the training set, R^2 is the correlation coefficient, S is the standard deviation of the regression, S_{cv} is the standard deviation of the leave-one-out cross-validation, q^2 is the correlation coefficient of the leave-one-out cross-validation, p is the significance of the variables in the model and F is the Fisher ratio at the 95% confidence level. The variables included in the model are designated as follows: the sub-index represents the order of the spectral moment and the super-index the type of bond weight used, i.e. *Dip* for dipole moment, *PS* for polar surface and *GM* for Gasteiger Marsili charges.

The model selection was subjected to the principle of parsimony. This principle calls for using models and

Table 2
Main results of the k-means cluster analysis for the dental monomers

Variance analysis				
Spectral moments	Between SS[a]	Within SS[b]	Fisher ratio (F)	p-level <[c]
Statistics for Active compound clusters (k-MCA 1)				
μ_6^{Dip}	58.60	23.39	19.97	0.0001
μ_{14}^{Dip}	62.99	19.00	28.34	0.0001
μ_3^{PS}	65.00	16.99	33.65	0.0001
μ_5^{PS}	65.15	16.84	34.07	0.0001
μ_5^{GM}	62.28	19.71	26.74	0.0001
μ_{15}^{GM}	64.83	17.16	33.13	0.0001

[a] Variability between groups.
[b] Variability within groups.
[c] Level of significance.

Table 3
The observed, predicted, residual and deleted residual of mutagenic activity values of dental monomers used in the training set

No.	Obs (ln (TA100))	Pred. (ln (TA100))	Residual	Deleted residual
1	0.860	0.140	0.720	0.763
2	−0.362	−0.372	0.010	0.013
3	−0.930	0.668	−1.598	−1.751
4	−1.080	−0.097	−0.983	−1.143
5	0.620	−0.020	0.640	0.699
6	−0.514	0.711	−1.225	−1.501
7	−0.536	−0.523	−0.013	−0.014
8	−0.896	0.072	−0.968	−1.029
9	−0.111	−0.566	0.455	0.491
10	0.172	0.183	−0.011	−0.012
11	2.230	0.602	1.628	1.727
12	2.100	0.602	1.498	1.589
13	−2.160	−1.834	−0.326	−0.349
14	−3.320	−2.348	−0.972	−1.023
15	−3.150	−1.350	−1.800	−1.975
16	−0.530	−0.304	−0.226	−0.263
17	−2.350	−1.956	−0.394	−0.424
18	−3.190	−4.698	1.508	1.655
19	−3.950	−5.669	1.719	1.994
20	−0.450	−0.304	−0.146	−0.170
21	−4.830	−6.336	1.506	1.873
22	−2.550	−1.956	−0.594	−0.640
23	−4.020	−2.515	−1.505	−2.426
24	−3.470	−3.330	−0.140	−0.251
25	−1.920	−1.793	−0.127	−0.135
26	−4.510	−3.794	−0.716	−1.338
27	−0.750	−1.588	0.838	0.915
28	8.650	8.070	0.580	0.756
29	8.610	8.070	0.540	0.704
30	7.970	8.132	−0.162	−0.235
31	6.360	3.754	2.606	3.111
32	5.200	6.494	−1.294	−1.554
33	4.090	1.938	2.152	2.538
34	1.590	1.964	−0.374	−0.422
35	1.370	1.787	−0.417	−0.471
36	0.990	1.349	−0.359	−0.412
37	0.740	−0.158	0.898	0.955
38	0.410	−0.158	0.568	0.604
39	0.410	−0.034	0.444	0.526
40	−0.220	−0.096	−0.124	−0.135
41	−1.600	0.400	−2.000	−2.408
42	−3.510	−1.676	−1.834	−2.267

procedures that contain all that is necessary for the modeling but nothing more. For example, as follows we show a model with six variables, one more than the previous model.

$$\ln TA100 = -1.94 + 4.64\times10^{-3}\mu_6^{\mathrm{Dip}} - 1.18\times10^{-8} \times\mu_{14}^{\mathrm{Dip}} - 0.005\times\mu_3^{\mathrm{PS}} + 1.23\times10^{-5} \times\mu_5^{\mathrm{PS}} - 0.017\times\mu_5^{\mathrm{GM}} + 1.25\times10^{-9} \times\mu_{15}^{\mathrm{GM}} \quad (4)$$

$N = 42;\ S = 1.154;\ R^2 = 0.901;$

$F = 53.251;\ p < 0.0001;\ q^2 = 0.847;\ S_{\mathrm{cv}} = 1.452$

As can be seen in this new equation in spite of having a relation among variables and cases superior to 5 [14,15], the variance of the mutagenic activity of the dental monomers explained for the variables in the Eq. (4) only increases in 1%. Also, the standard deviation shows a poor decrease (2.51%) reason why it is not significant from a statistical point of view to accept the inclusion of a new variable in the model. In addition, the regression coefficient of the cross-validation leave-one-out (q^2) remains practically constant. In this sense, the q^2 values can be considered a measure of the predictive power of a regression equation: whereas R^2 can always be increased artificially by adding more parameters (descriptors), q^2 decreases if a model is over-parameterized, and is therefore a more meaningful summary statistic for QSAR models. For that reasons, the five-dimensional models are characterized by the best

compromise between predictive power and model complexity. The addition of another variable does not lead to such an increase in predictive power such that the complexity increase is counterbalanced.

On the other hand, an analysis of potentially outlier compounds was carrying out. Consideration of the outliers removed from a QSAR is essential. An outlier to a QSAR is identified normally by having a large standard residual and can indicate the limits of applicability of a QSAR models. Although it is acceptable to remove a small number of outliers from QSAR it is noted that it is not acceptable to remove the outlier repeatedly from a QSAR analysis simply to improve a correlation. In the current work, the compound 31 present a large residual and should be consider as an outlier. At removal of this compound from the training set the following equation is obtained:

$$\ln TA100 = -1.621 + 3.96 \times 10^{-3} \mu_6^{\mathrm{Dip}} - 7.90 \times 10^{-9} \times \mu_{14}^{\mathrm{Dip}} - 0.005 \times \mu_3^{\mathrm{PS}} + 1.11 \times 10^{-5} \times \mu_5^{\mathrm{PS}} - 0.014 \times \mu_5^{\mathrm{GM}} \quad (5)$$

$$N = 41\ S = 1.103\ R^2 = 0.903\ F = 63.707\ p < 0.0001$$

$$q^2 = 0.852\ S_{\mathrm{cv}} = 1.324$$

Removal of the outlier did not improve the explanation of experimental variance significantly of Eq. (3) when compared to Eq. (5). It is not appropriate to remove compounds from a data set simply to improve a correlation, and indeed much important information may be gleaned from the analysis of outliers omitted from a QSAR. For this reason, here any compound was considered as a potential outlier.

Finally, using Eq. (3), the efficacy of our model for evaluating on external prediction set of 11 compounds was corroborated. The result of this analysis is shown in the following Table 4, in which we can see that the correlation coefficient for the observed vs predicted mutagenic activity of these dental monomers is 0.834, which is an excellent value. In addition, the standard error of estimation was of 1.782.

Table 4
Experimental, predicted and residual values of mutagenic property of dental monomers in the test set

No.	Obs (ln (TA100))	Pred. (ln (TA100))	Residual
43	0.115	0.778	−0.663
44	−0.576	−0.536	−0.040
45	−1.060	0.555	−1.615
46	−1.040	−0.817	−0.223
47	−3.960	−0.205	−3.755
48	−4.370	−8.013	3.643
49	−3.060	−4.331	1.271
50	8.750	10.542	−1.792
51	6.040	5.387	0.653
52	1.370	3.768	−2.398
53	0.110	1.983	−1.873

3.1. Study of fragments contribution to mutagenic property

In these years, individual QSARs have been developed for mutagenic endpoint. For instance, in the case of dental monomers we demonstrate the potentialities of the TOPS-MODE approach for the prediction of the mutagenic to *Salmonella typhimurium* TA100 in two series of epoxides compounds [14,39]. In this instance we find that for the case of the aromatic epoxides the best correlation with the mutagenicity is obtained when the molar refractivity is used as bond weights. This shows an interesting behavior if taken into account that the molar refractivity is the molecular volume corrected with the refraction index of the molecule. As we previously explained, the TOPS-MODE approach is able to compute the contribution of any structural fragment (real or hypothetical) to the biological property or activity studied. In the present case, we can find the positive and negative contributions of such fragments to the development of the mutagenicity activity. In Table 5 and Fig. 2 we show the fragments and their contributions to the mutagenicity calculated from Eq. (3). The analysis of the fragments F_9 and F_{10} point out to their positive contribution to molecular mutagenic. The mutagenic character of these cyclic ethers epoxides obeys to their vast reactivity so that the high torsion spanning of the three-member ring leads steady to the ring cleavage. The internal bond angles of the ring around 60° are far away from the 109.58 expected for a tetrahedral arrangement at carbon atom or to the divalent oxygen bonded to the carbon atoms in acyclic ethers. Since the atoms are not close enough in order to allow the maximal overlapping of the orbitals, thence the bonds are not so strong like normal ether and it is more reactive. The arrangement of the three atoms is normally accepted to look like a banana shape bond. Essentially, epoxides are electrophilic, reactive chemicals that may form DNA–protein cross-links and induce mutagenesis. However, the

Table 5
The contribution of different fragments to the mutagenic activity of the dental monomers under study

Studied Fragments	Fragment contribution	Studied Fragments	Fragment contribution
F_1	0.109	F_{10}	1.210
F_2	0.368	F_{11}	−0.821
F_3	0.799	F_{12}	−1.170
F_4	−0.227	F_{13}	−1.924
F_5	0.091	F_{14}	−0.625
F_6	−1.180	F_{15}	−0.654
F_7	−0.820	F_{16}	0.192
F_8	0.831	F_{17}	0.165
F_9	1.110	F_{18}	0.317

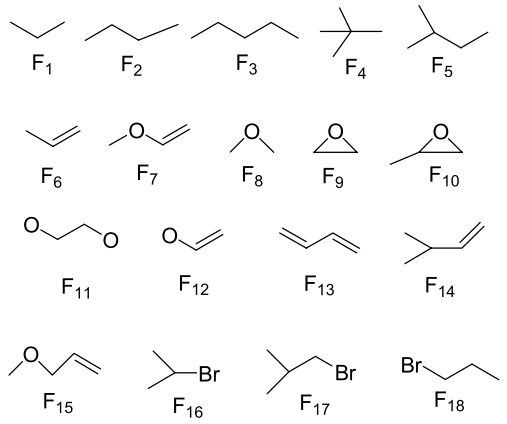

Fig. 2. Structures of selected fragments for which their contributions to the mutagenic activities were calculated.

chemical properties of the epoxides convert them in potential precursors of dental resins, when these monomers are photo excited, a polymerization process that leads to the building of resins takes place.

On the other hand, this research shows that an increase in the carbon lineal chain leads to an increase in the mutagenic activity. This affirmation is based on the analysis of the set of fragment from F_1 to F_3 where an increase in methyl group in each fragment increases the contribution to the property from 0.109 to 0.799. Nevertheless, when the branching of the fragments is increased, the contribution to the mutagenic property is minor as was observed in the fragments F_4 and F_5. This type of contribution associated with the ramification or branching of the carbon chain is in relation with the target point of each dental monomer in special [13]. We can observe that ramification in the groups of the linear aliphatic chain causes sudden increase in the activity. The fragments F_3, F_4, F_5 have the same number of carbons (almost the same hydrophobicity), but very different contributions. If we compare aliphatic groups with different ramifications and different number of methyl groups we would be able to appreciate higher changes in their contributions. In this case, an increase in the hydrophobicity leads to the amplification of the analyzed property. We got to this conclusion in a previous work [39], but the interpretation of this behavior should be confirmed in a heterogeneous set of compounds.

Finally, when a bromine atom appears in some fragments, their mutagenic activity increases (i.e. F_{16}, F_{17}). This behavior has been previously observed [1–3]. Accordingly all new dental monomers used for forming dental resins should not possess halogens in their structure.

4. Concluding remarks

We have shown that TOPS-MODE approach is able to describe mutagenicity for heterogeneous set of dental monomers with an appropriate degree of correlation and robustness. In fact we have developed a model for predicting mutagenicity of a data set of 41 dental monomers. This model explains more than 89% of the variance in the experimental mutagenicity with appropriate predictive power.

On the other hand, along this series we demonstrate that the spectral moments should be very useful tools in the prediction of the mutagenic property in non-congeneric and congeneric set of compounds. Therefore, this study could be the beginning of important investigations in the field of the mutagenesis of diverse series of compounds such as dental monomers, pesticides, drugs and organic compounds.

References

[1] Benigni R. J Exp Clin Cancer Res 2004;23(1):5–8.
[2] Mekenyan OG, Dimitrov SD, Pavlov TS, Veith GD. Curr Pharm Des 2004;10(11):1273–93.
[3] Mekenyan O, Dimitrov S, Schmieder P, Veith G. SAR QSAR Environ Res 2003;14(5–6):361–71.
[4] Benigni R, Zito R. Mutat Res 2004;566(1):49–63.
[5] Benigni R, Passerini L, Rodomonte A. Environ Mol Mutagen 2003; 42(3):136–43.
[6] Gramatica P, Consonni V, Pavan M. SAR QSAR Environ Res 2003; 14(4):237–50.
[7] González MP, González-Díaz H, Cabrera MA, Molina RR. Bioorg Med Chem 2004;12:735–44.
[8] González-Díaz H, Marrero Y, Hernández I, Bastida I, Tenorio I, Nasco O, et al. Chem Res Toxicol 2003;16:1318–24.
[9] González-Díaz H, Ramos de AR, Molina RR. Bioinformatics 2003; 19:2079–87.
[10] González-Díaz H, Molina RR, Uriarte E. Polymer 2004;45(4): 3845–53.
[11] González-Díaz H, Molina RR, Uriarte E. Bioorg Med Chem Lett 2004;14(6):4691–5.
[12] Ramos de Armas R, González-Díaz H, Molina RR, Uriarte E. Proteins Struct Func Bioinf 2004;23(56):715–24.
[13] Yourtee D, Holder A, Smith R, Morrill J, Kostoryz E, Brockmann W, et al. J Biomater Sci, Polym Ed 2001;1:89–105.
[14] González MP, Morales AH, Molina R. Polymer 2004;45:2773–83.
[15] González MP, Morales AH, González-Díaz H. Polymer 2004;45: 2073–83.
[16] Morales AH, González MP, Rieumont JB. Polymer 2004;45:2045–54.
[17] Estrada E. J Chem Inf Comput Sci 1995;35:31–3.
[18] Estrada E. J Chem Inf Comput Sci 1996;36:844–9.
[19] Estrada E. J Chem Inf Comput Sci 1997;37:320–8.
[20] Estrada E. J Chem Inf Comput Sci 1998;38:23–7.
[21] Cabrera MA, Bermejo M, González MP, Ramos R. J Pharm Sci 2004; 7:1701–17.
[22] Wang R, Gao Y, Lai L. Perspect Drug Discov Des 2000;19:47–66.
[23] Ertl P, Rohde B, Selzer P. J Med Chem 2000;43:3714–7.
[24] Ghose AK, Crippen GM. J Chem Inf Comput Sci 1987;27:21–35.
[25] Gasteiger J, Marsilli M. Tetrahedron Lett 1978;34:3181–4.
[26] Miller KJ. J Am Chem Soc 1990;112:8533–42.
[27] Bondi A. J Phys Chem 1964;68:441–51.
[28] Estrada E, Uriarte E, Gutierrez Y, González-Díaz H. SAR QSAR Environ Res 2003;14:145–63.
[29] Gutiérrez, Y., Estrada, E. 2002. MODESLAB 1.0 for Windows.
[30] Weininger D. J Chem Inf Comput Sci 1988;28:31–6.
[31] STATISTICA version. 6.0. Statsoft, Inc., 2002.
[32] Rogers D, Hopfinger AJ. J Chem Inf Comput Sci 1994;34:854–62.
[33] Golbraikh A, Tropsha A. J Mol Graph Model 2002;20:269–76.

[34] Golbraikh A, Shen M, Xiao Y, Lee K, Tropsha A. J Comput Aided Mol Des 2003;17:241–53.
[35] Golbraikh A, Tropsha A. J Comput Aided Mol Des 2002;16:357–69.
[36] González MP, González-Díaz H, Molina RR, Cabrera MA, Ramos RA. J Chem Inf Comput Sci 2003;43:1192–9.
[37] Molina E, González-Díaz H, González MP, Rodríguez E, Uriarte E. J Chem Inf Comput Sci 2004;44:515–21.
[38] González MP, Diaz LC, Morales AH, Morales YR, Gonzaga de Oliveira L, Torres LG, et al. Bioorg Med Chem 2004;12:4467–75.
[39] González MP, Dias LC, Morales AH. Polymer 2004;45:5353–9.