TOPOLOGICAL INDICES FOR DESCRIPTION OF THE PERTRACTION

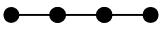
Piotr SZCZEPAŃSKI

Nicolaus Copernicus University, Faculty of Chemistry, Gagarina Str. 7,
87-100 Toruń, Poland; e-mail: piotrs@chem.uni.torun.pl

1. INTRODUCTION

Graph theory can be applied in chemistry to describe the structure of molecules [1]. According to this theory, chemical compounds are presented in the form of hydrogen-depleted structural graphs (molecular graphs) in which nodes (vertices) symbolize atoms and lines (edges) represent covalent bond. A structural graph can be used for creating an adjacency (A) or a distance matrix (D). Based on these matrices, various topological indices enabling the numerical interpretation of the molecule structure can be calculated.

The most important matrix representation of the graph is an adjacency matrix (A). This is a square matrix of n - dimensional, where n is the number of vertices in the hydrogen-depleted graph. In this matrix, each element $A_{i,j}$ equals 1 if there is an edge (bond) between adjacent vertices (atoms), and $A_{i,j}$ = 0 otherwise. For example, a graph of butane and the corresponding adjacency matrix is shown in Fig.1.



$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Fig.1. Molecular graph of butane and the corresponding adjacency matrix (A).

Another method of graph presentation is a distance matrix D, i.e. a square matrix with the diagonal values equal to 0. Each other element $d_{i,j}$ corresponds to the number of edges in the shortest path between vertices i and j (the number of bonds). For instance, the distance matrix for butane is presented in Fig.2.

$$D = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$$

Fig.2. The distance matrix (*D*) for butane.

It is evident that the distance matrix *D* is a better source of information when compared with the adjacency matrix *A* which can be regarded as a distance matrix containing the distances equal to unity only.

At the beginning of the twentieth century, Wiener [2] proposed a topological index calculated as half the sum of all the elements from the distance matrix:

$$W = \frac{1}{2} \sum_{i,j} d_{i,j} \tag{1}$$

For example, the Wiener index (*W*) for butane is equal to:

$$W = \frac{1}{2} \cdot 20 = 10 \tag{2}$$

Henceforward, over three thousand new topological indices were developed.

The chemical behavior of a compound may be related to the structure; therefore, topological indices are useful in the formulation of quantitative structure-activity relationships (QSAR) or quantitative structure-property relationships (QSPR).

In this presentation, the application of topological indices for a description of the benzoic acid pertraction in an agitated bulk liquid membrane system is discussed. For this purpose, some topological indices were linearly correlated with the experimental fluxes of benzoic acid. Moreover, the influence of the feed concentration and the liquid membrane volume on the benzoic acid fluxes was evaluated by chemometric approach.

2. EXPERIMENTAL

Experimental studies were performed in the agitated bulk liquid membrane system (ABLM) (Fig. 3) The feed phase with a volume of 190 cm^3 was placed in the left chamber of the membrane system. As the feed phase, the solution of benzoic acid (pure p.a., P.O.Ch., Gliwice) was used. Distilled water of the same volume was applied as the receiving solution (right chamber). As the liquid membrane phase, hexane, octane or decane (pure, Chempur, Sigma-Aldrich, Fluka) were used. Concentrations of benzoic acid in the aqueous solutions were determined conductometrically (ELMETRON CX-721 Multimeter).

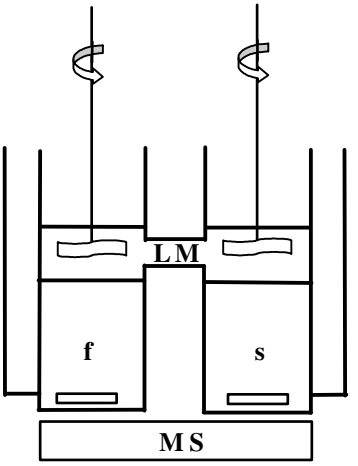


Fig.3. The agitated bulk liquid membrane system; f – feed phase, LM – liquid membrane phase, s – stripping phase, MS – magnetic stirrer.

3. RESULTS

Fig. 4 presents the typical results of the benzoic acid concentration dependences vs. time in the system ABLM with hexane as the organic phase.

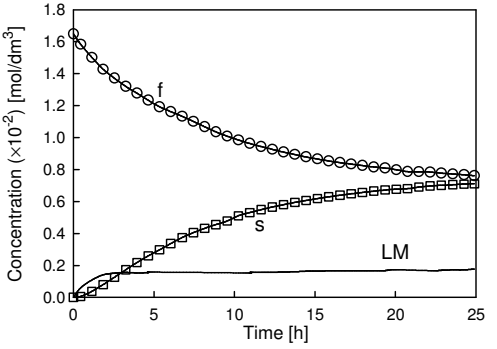


Fig.4. Typical experimental results of benzoic acid pertraction in the agitated bulk liquid membrane system with hexane as the liquid membrane. The concentration of benzoic acid in the LM was calculated from the difference between the concentration in the feed and receiving solution.

The experiments were performed according to an appropriate full factorial two-level design (2^1 – 2 pertraction experiments). To estimate the experimental error, each experiment was repeated 3 times (3×2^1 design, 6 pertraction experiments). The results of experiments indicate that the

transport in the system with limited volumes of aqueous phases occurs in non-stationary conditions. In such a case, time-dependent fluxes with a characteristic maximum (J_{max}) [2] are observed.

In order to determine the effect of the solvent on the pertraction efficiency, the selected molecular descriptors were calculated [3], and thereafter linearly correlated with the experimentally determined benzoic acid flux (J_{max}). All the linear equations were analyzed in terms of significance and quality of the fit. As an example, the following relationship for connectivity index ($^4\chi$) was found:

$$J_{max} = 2.53(\pm0.05)\times10^{-9} - 7.33(\pm0.47)\times10^{-10}\times{}^4\chi \quad (3)$$
$$r^2=0.9836,\ r^2_{Adj}=0.9796,\ S.D. = 4.1\times10^{-11},\ F = 241,\ N = 6$$

Another parameter affecting the pertraction efficiency is the feed phase concentration (c_f [g/dm^3]). In the case of benzoic acid pertraction in the ABLM system (the solution-diffusion mechanism), the $J_{max} = f(c_f)$ dependence is linear. Therefore, the full factorial design (3×2^2) was created and the additional six pertraction experiments were performed. The following regression model for J_{max}, connectivity index ($^4\chi$) and the feed concentration c_f was obtained:

$$J_{max} = 1.40(\pm1.06)\times10^{-10} - 4.51(\pm0.97)\times10^{-10}\cdot{}^4\chi + 1.07(\pm0.05)\times10^{-9}\times c_f \quad (4)$$
$$r^2=0.9846,\ r^2_{Adj}=0.9812,\ S.D. = 1.2\times10^{-10},\ F = 288,\ N = 12$$

According to the chemomertic approach, the model (relationship) should be as simple as possible; therefore, in first approximation, the best model can be represented by the first-degree polynomial (linear). To improve its prediction ability, the model can be extended by adding the interactions terms (between the independent variables). For example, the interaction between variables $^4\chi$ and c_f, leads to the following relationship:

$$J_{max} = -2.61(\pm0.43)\times10^{-10} + 1.9(\pm4.6)\times10^{-11}\times{}^4\chi + $$
$$+1.39(\pm0.03)\times10^{-9}\times c_f - 3.76(\pm0.32)\times10^{-10}\times{}^4\chi\times c_f \quad (5)$$
$$r^2=0.9992,\ r^2_{Adj}=0.9989,\ S.D. = 2.9\times10^{-11},\ F = 3213,\ N = 12$$

In this model, one of the coefficients ($^4\chi$ parameter) is of no statistical importance; therefore, the backward variable selection method (STATISTICA 6.0, StatSoft Inc.) was applied to reduce space of variables. The reduced model is expressed by:

$$J_{max} = -2.45(\pm0.16)\times10^{-10} + 1.38(\pm0.02)\times10^{-9}\times c_f - 3.65(\pm0.16)\times10^{-10}\times{}^4\chi\times c_f \quad (6)$$
$$r^2=0.9991,\ r^2_{Adj}=0.9989,\ S.D. = 2.8\times10^{-11},\ F = 5305,\ N = 12$$

The values of the correlation coefficient indicate that this relationship satisfactorily explains the variance of J_{max} vs. c_f and $^4\chi$.

In the next step of modelling, the liquid membrane volume (V_{LM}) was used as an additional independent variable.. In this case, the full factorial design (2^3) with center points was created, and appropriate parameter values were used in planned pertraction experiments (3×2^3 +3 = 27 pertraction experiments). This procedure results in the following regression dependence:

$$J_{max} = 4.5(\pm1.3)\times10^{-10} - 3.06(\pm0.68)\times10^{-12}\times V_{LM} + 8.94(\pm0.45)\times10^{-10}\times c_f - \\ - 3.70(\pm0.95)\times10^{-10}\times{}^4\chi \qquad (7)$$
$$r^2=0.9497,\ r^2_{Adj}=0.9432,\ S.D. = 1.6\times10^{-10},\ F = 145,\ N = 27$$

In order to improve the quality of the correlation (r^2=0.9497), Eq. (7) was extended by adding the interaction terms between all the independent variables. Further, the elimination procedure was carried out to identify the parameters without statistical importance (backward variable selection). After this data analysis, the final satisfactory model of good correlation quality takes the following form:

$$J_{max} = -1.56(\pm0.45)\times10^{-10} + 1.491(\pm0.037)\times10^{-9}\times c_f - \\ - 1.82(\pm0.57)\times10^{-10}\times{}^4\chi - 3.57(\pm0.20)\times10^{-12}\times V_{LM}\times c_f - \\ - 2.81(\pm0.33)\times10^{-10}\times c_f\times{}^4\chi + 1.63(\pm0.31)\times10^{-12}\cdot V_{LM}\cdot{}^4\chi \qquad (8)$$
$$r^2=0.9969,\ r^2_{Adj}=0.9962,\ S.D. = 4.3\times10^{-11},\ F = 1350,\ N = 27$$

4. CONCLUSIONS

The results of this investigation indicate that the influence of the feed concentration, liquid membrane volume and the liquid membrane solvent on the flux of the benzoic acid can be described by a linear model with the interactions between the independent variables. The chemometric methods of the experiment design were demonstrated as useful for estimating the optimal model coefficients.

REFERENCES

[1] P.J. Hansen, P.C. Jurs, *J. Chem. Edu.*, 65(7) (1988) 574.
[2] P. Szczepański, *Sep. Purif. Technol.*, 71 (2010) 121.
[3] VCCLAB, Virtual Computational Chemistry Laboratory, http://www.vcclab.org, 2005.