Influence of chemical structure on the lipophilicity of isolated free carotenoids from paprika oleoresin was studied by a quantitative structure-retention relationship (QSRR) approach. The chromatographic behavior of these compounds was investigated by reversed phase high-pressure liquid chromatography (RP HPLC). The retention mechanism was determined using acetone-water as the mobile phase on a reversed-phase column (SB-C18). A variety of lipophilicity parameters \( (\log P) \) were calculated using different software products. Based on the correlations, nonlinear structure–activity models were derived between the retention constants, \( t_r \) (retention time of investigation compounds) and \( \log P \) values. Five high quality QSRR models were found to have a good predictive ability and close agreement between experimental and predicted values. The study showed that the retention constants can be used as a measure of lipophilicity of investigated compounds at a high significant level.

**Keywords:** carotenoids, RP HPLC, Lipophilicity, QSRR.

Available online at the Journal website: http://www.ache.org.rs/HI/

**Abstract**

Influence of chemical structure on the lipophilicity of isolated free carotenoids from paprika oleoresin was studied by a quantitative structure-retention relationship (QSRR) approach. The chromatographic behavior of these compounds was investigated by reversed phase high-pressure liquid chromatography (RP HPLC). The retention mechanism was determined using acetone-water as the mobile phase on a reversed-phase column (SB-C18). A variety of lipophilicity parameters \( (\log P) \) were calculated using different software products. Based on the correlations, nonlinear structure–activity models were derived between the retention constants, \( t_r \) (retention time of investigation compounds) and \( \log P \) values. Five high quality QSRR models were found to have a good predictive ability and close agreement between experimental and predicted values. The study showed that the retention constants can be used as a measure of lipophilicity of investigated compounds at a high significant level.

**Keywords:** carotenoids, RP HPLC, Lipophilicity, QSRR.

Available online at the Journal website: http://www.ache.org.rs/HI/

Red pepper \( (Capsicum annuum \text{ L.}) \) is one of the most important vegetable cultures in the province of Vojvodina in Serbia. There are a number of foods containing paprika or its compounds. It is especially important because of its high nutritive and biological value \([1,2]\). In recent years, carotenoids have been a subject of research interest as potential antioxidants, based on studies that reported that a higher consumption of carotenoids leads to lower risk of cancer and cardiovascular diseases \([3,4]\). The yellow, orange and red colors of paprika fruit originate from carotenoids and the pigments are synthesized during ripening. They are the most widespread group of pigments – the number of naturally occurring carotenoids continues to rise and has reached about 750 \([5]\). More than 25 different pigments have been identified in the fruits of paprika: green chlorophylls, yellow-orange lutein, zeaxanthin, violaxanthin, antheraxanthin, \( \beta \)-cryptoxanthin, \( \beta \)-carotene, etc. The red pigments capsanthin, capsorubin and cryptoxanthin are unique to the Capsicum species. The colour of paprika is determined by the proportion of red to yellow pigments \([6,7]\), whereas \( \beta \), \( \alpha \), \( \gamma \)-carotene and \( \beta \)-cryptoxanthin, as provitamins, contribute to its nutritive value. The quality of paprika and its oleoresins is determined primarily by their color \([8-11]\). Capsanthin and capsorubin, which comprise 65–80%, contribute the red colour to paprika \([12-14]\). In the fruits of paprika during ripening, free carotenoids esterification with fatty acids occurs \([15]\).

Saturated fatty acids are known to be more stable and solid at room temperature. With an increase in the number of double bonds, the fluidity rises. On the other hand, unsaturated fatty acids are prone to autoxidation, which leads to the generation of free radicals and increasing risk of cancer. However, substituting saturated with unsaturated fats in diet lowers the level of total cholesterol and LDL cholesterol in blood. Also, linoleic \((18:2)\) and \( \alpha \)-linolenic \((18:3)\) are essential fatty acids, which play an important role in human nutrition.

Because of their potential biological activity, for initial chemical screening of activity of investigated compounds, it is first recommended to determine their lipophilicity. The most widely accepted measure of lipophilicity is the octanol-water partition coefficient, which is expressed in its logarithmic form as \( \log P \) \([16]\).

Literature is rich in research articles investigating similarities/dissimilarities between \( \log P \) and chromatographic retention. Determination of partition coefficient using the classical “shake-flask” technique has a series of disadvantages and has been successfully replaced by alternative chromatographic methods, since the partition coefficient (lipophilicity) of a compound between aqueous and organic phase determines both its permeation through biological membranes and retention in RP LC \([17-24]\).

Quantitative structure (chromatographic) retention relationships (QSRR) have been considered a model approach to establish strategy and methods of property predictions. QSRR analysis appears especially attractive from the general chemometric point of view.
because it provides the best testing of the applicability of individual structural parameters for property description.

Currently, QSRR studies can be used to: identify the most useful structural descriptors, predict retention for a new analyte and to identify unknown analytes; gain insight into molecular mechanism of separation operating in a given chromatographic system; quantitatively compare separation properties of individual types of chromatographic columns; evaluate properties, other than chromatographic physicochemical properties of analytes, such as lipophilicity; estimate relative bioactivities within sets of drugs and other xeno-biotics [25]. In QSRR studies, relations between molecular descriptors and retention have been explored [26]. The aim of this methodology is to derive a model to describe the chromatographic retention on a given chromatographic system, which then can be used for future retention prediction of new solutes. Thus, when a meaningful and statistically significant model is found, no additional experiments are needed to predict the retention for new solutes.

In QSRR studies, molecular descriptors are either determined from experiments or computed by molecular mechanics or even semi-empirical quantum chemical techniques. Chromatographic retention is a physical phenomenon that is primarily dependent on the interactions between the solute and the stationary phase. The compatibility of experimental and theoretical approaches for the determination of organic compound lipophilicity remains also a focus of scientific interest.

Considering the practical importance of isolated free carotenoids from paprika oleoresin, the main objective of this study was to examine the retention behavior of seven carotenoids in reversed-phase chromatographic systems (HPLC) of mobile phase (acetone:water). Obtained chromatographic data were correlated to selected lipophilicity parameters. This method includes data collection, molecular descriptor selection, correlation model development, and finally model evaluation. QSRR studies have prediction abilities.

**EXPERIMENTAL**

**Material**

Commercial ground paprika of the “Aleva N.K.” variety, harvested in 2005, was obtained from the Aleva a.d. company from Novi Kneževac, the most important producer of ground pepper in Serbia. The mean diameter of the particles was 0.224 mm. Ground pepper was placed into the thimble in the middle portion of the Soxhlet apparatus, the solvent was then poured in, and the process was continued until complete discoloration of sample was achieved. Soxhlet (Sx) extract of paprika was obtained using technical grade hexane. The solvent was evaporated from extract under vacuum.

**HPLC Analysis of carotenoid content**

Qualitative and quantitative analysis of samples-free carotenoids (Figure 1) was performed according to a previously described method [27]. HPLC was performed using a Hewlett-Packard liquid chromatograph HP 1090 equipped with Diode Array Detector (DAD). A reversed-phase column (Zorbax SB-C18, 5 µm, 3.0×250 mm² i.d.), protected by a guard column (Zorbax SB-C18, 5 µm, 4.6×12 mm² i.d., Agilent, USA) was used throughout this research. The mixture of mobile phase acetone:water (75:25; v/v) was used and the HPLC separations were performed by the following linear gradient: 0–25% of acetone in 10 min, then until 100% of acetone by 35 min, 100% of acetone by 45 min, 0% of acetone by 65 min, post time 15 min. The flow rate of the mobile phase was set at 0.500 mL/min. The oven temperature was set at 50°C.

**Figure 1. Chemical structures of the carotenoids from paprika studied.**
was operated at room temperature (25°C). The sample injection volume was 10 μL, and the injection was performed manually. The chromatograms were acquired in the range 460+4 nm by DAD detector; the spectra were recorded in the range of 350–550 nm. Carotenoid standards capsanthin, capsorubin, antheraxanthin, zeaxanthin, violaxanthin, β-carotene were obtained from “Carotenature”, Switzerland; 80-β-apo-carotenal was obtained from Fluka, Germany. All reagents used were HPLC grade.

**Molecular modelling and calculations of lipophilicity parameters**

Molecular modelling studies were performed using CS Chem-Office Software version 7.0 (Cambridge software) running on a P-III processor [28]. All molecules were constructed by using Chem Draw Ultra 7.0 and saved as the template structures. For every compound, the template structure was suitably changed considering its structural features, copied to Chem 3D 7.0, and subjected to energy minimization using molecular mechanics (MM3). The minimization was executed until the root mean square (RMS) gradient value reached a value smaller than 0.1 kcal/mol-A. The Austin Model-1 (AM-1) method was used for re-optimization until the RMS gradient attained a value smaller than 0.0001 kcal/mol-A using MOPAC. Partition coefficients and aqueous solubility values were calculated with different theoretical bases (atomic based prediction, fragment based prediction): Alog P, AClog P, Alog P, Mlog P, log P_kowin, Xlog P2, and Xlog P3 by applying different theoretical procedures (Table 1) [29,30].

**Statistical methods**

The complete regression analysis was carried out by PASS 2005, GESS 2006, NCSS Statistical Softwares [31].

**RESULTS AND DISCUSSION**

**Lipophilicity determination**

Lipophilicity is one of the important physico-chemical parameters that determine the activity. Hence, the estimation of the lipophilic character of new, potentially biologically active compounds is regarded as one of the first parameters to be determined at the earliest possible opportunity. It has been recognized that the retention of a compound in reversed-phase liquid chromatography (RP-LC) is governed by its lipophilicity and thus chromatography may be successfully used for the determination of lipophilicity. Although HPLC is a relatively new LC technique, it offers several practical advantages compared to the traditional shake-flask method, including a rapid and simple way of lipophilicity determination, reproducibility, broader dynamic range, insensitivity to impurities or degradation products, and a wide choice of adsorbents and solvents.

The main purpose of this study was to use chromatographic data – retention time (t_r) as a descriptor of the lipophilic character for carotenoids studied.

**Correlation of retention time, t_r, and log P**

To investigate the quantitative effects of the structural parameters of free carotenoids from paprika oleoresin retention time, QSRR analysis with seven different partition coefficients (log P) was employed. First, the correlation of each one of the log P values with each other was calculated. The resulting correlation matrix is shown in Table 2. The correlation matrix was constructed to find the interrelationship among the parameters, which shows that all the lipophilicity descriptors selected in the study are highly mutually correlated (r > 0.85). Therefore, any combination of these descriptors in multiple regression analysis may result with a model suffering from multi-colinearity.

Usually, lipophilicity parameters are linearly related to retention time, but in the more general case this relationship is not linear. Therefore, a complete regression analysis was performed, using linear, quadratic and cubic relationships. It is apparent from the data presented in Table 3 that the fit quality improved with higher (second or third) order polynomials.

Among the presented lipophilicity parameters, log P values that had the highest correlation with t_r were retained and used to find a nonlinear monoparametric models. The resulting models were:

**Table 1. Partition coefficients calculated by different theoretical methods**

<table>
<thead>
<tr>
<th>Compd.</th>
<th>Alog P_1</th>
<th>AClog P</th>
<th>Alog P</th>
<th>Mlog P</th>
<th>P_kowin</th>
<th>Xlog P2</th>
<th>Xlog P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.37</td>
<td>9.91</td>
<td>8.26</td>
<td>6.30</td>
<td>13.46</td>
<td>7.05</td>
<td>10.34</td>
</tr>
<tr>
<td>3</td>
<td>9.08</td>
<td>11.94</td>
<td>10.76</td>
<td>7.94</td>
<td>16.08</td>
<td>8.63</td>
<td>12.25</td>
</tr>
<tr>
<td>4</td>
<td>8.20</td>
<td>9.95</td>
<td>8.81</td>
<td>6.22</td>
<td>13.18</td>
<td>6.93</td>
<td>10.60</td>
</tr>
<tr>
<td>5</td>
<td>7.43</td>
<td>8.78</td>
<td>8.10</td>
<td>5.42</td>
<td>11.41</td>
<td>6.68</td>
<td>10.29</td>
</tr>
<tr>
<td>6</td>
<td>8.77</td>
<td>9.27</td>
<td>7.70</td>
<td>6.37</td>
<td>13.27</td>
<td>7.87</td>
<td>11.08</td>
</tr>
<tr>
<td>7</td>
<td>8.30</td>
<td>11.11</td>
<td>9.52</td>
<td>7.06</td>
<td>14.95</td>
<td>7.18</td>
<td>10.91</td>
</tr>
</tbody>
</table>
To confirm the predictive conclusion that the...

\[ r_t = -0.457 \text{AClog} P^3 + 15.689 \text{AClog} P^2 - 177.703 \text{AClog} P + 652.919 \]
\[ r^2 = 0.9328 \ , \ s = 17.7960 \]  

\[ r_t = -0.561 \text{Alog} P^3 + 18.276 \text{Alog} P^2 - 186.718 \text{Alog} P + 617.785 \]
\[ r^2 = 0.9600 \ , \ s = 10.5843 \]  

\[ r_t = -1.128 \text{Mlog} P^3 + 26.012 \text{Mlog} P^2 - 187.286 \text{Mlog} P + 436.234 \]
\[ r^2 = 0.9555 \ , \ s = 11.7775 \]  

\[ r_t = -0.159 \text{log} P_{\text{Kow}}^3 + 7.669 \text{log} P_{\text{Kow}}^2 - 115.682 \text{log} P_{\text{Kow}} + 564.124 \]
\[ r^2 = 0.9345 \ , \ s = 17.3226 \]  

\[ r_t = -1.423 \text{Xlog} P^3 + 55.120 \text{Xlog} P^2 - 689.682 \text{Xlog} P + 2818.871 \]
\[ r^2 = 0.9594 \ , \ s = 10.7601 \]  

(5)

where \( r^2 \) is the square of the correlation coefficient and s is the standard deviation.

It is obvious that the obtained regression equations have high statistical quality and these models can be used to predict the retention time for free carotenoids from paprika oleoresin. To confirm the predictive power of the models, the retention times were calculated by the theoretical models 1–5. From the data presented in Table 4, it can be concluded that the observed and the estimated \( t_r \) are very close to each other. This indicates the good predictability of the established models.

Figure 2 shows the linear plots of predicted vs. experimental values of the retention constants. To investigate the existence of a systemic error in developed model, the residuals of predicted values of \( t_r \) were plotted against the experimental values in Figure 3. The propagation of the residuals on both sides of zero indicates that no systematic error exists in the deve-
Figure 2. Plot of predicted versus experimentally observed values retention parameters of investigated carotenoids obtained using HPLC.
development of regression models, as suggested by Jalali-Heravi et al. [32].
These results illustrated that the chemometric analysis combined with a successful variable selection procedure is adequate to generate an efficient QSRR model for predicting the retention time of free carotenoids from paprika oleoresin. All the present results suggest a dependence of the lipophilicity parameters on retention constant of compounds investigated. By knowing the exact values for these parameters and correlations, we can accurately predict the retention behaviour of free carotenoids from paprika oleoresin.

CONCLUSION
It is well known that mechanisms of chromatographic separation are very complex and depend of many factors such as type of chromatographic system, physicochemical characteristics of analytes, experimental conditions, etc. Therefore, in order to under-
stand chromatographic processes, it is very useful to establish mathematical models that can predict the retention behavior of analytes based on their structural characteristics in the applied chromatographic system. Determination of the correlations between molecular structure and retention behavior of molecules in different chromatographic systems is the main task of quantitative structure–retention relationships (QSRR) chemometric method. Chemometric processing of chromatographic data can reveal systematic information both about the analytes (retention, physicochemical properties, etc.) and about the stationary phases studied (the molecular mechanism of separation). In QSRR models, the retention (e.g., the retention parameter, $t_r$) of solutes in specific chromatographic system is presented as a function of molecular descriptors of the analytes. The main parameters used in QSRR studies are physicochemical parameters. QSRR analysis was performed to find the quantitative effects of the molecular structure of isolated free carotenoids from paprika oleoresin on their retention time. Different lipophilicity parameters were calculated for each molecule using different software products. The obtained partition coefficients had good correlation among each other. A complete regression analysis was performed, using linear, quadratic and cubic relationships between the retention constants and lipophilicity parameters. The fit quality improved with higher (second or third) order polynomials. Five high quality nonlinear structure–activity models were derived between the retention constants and lipophilicity parameters of isolated free carotenoids from paprika oleoresin. The best QSRR mathematical models were used to predict retention time and close agreement between experimental and predicted values was obtained, indicating that these models can be successfully applied to predict the retention constants for this class of molecules.

Acknowledgement

These results are part of the projects No. 114-451-2707/2012-01, financially supported by the Provincial Secretariat for Science and Technological Development of Vojvodina and projects No. 172012, 172014 and 31055 supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, 2011–2014.

REFERENCES


Hem. ind. 67 (6) 933–940 (2013)


IZVOD

RETENCIONI PARAMETRI RP HPLC U KORELACIJI SA MOLEKULSKIM DESKRIPTORIMA LIPOFILNOSTI KAROTENOIDA

Sanja O. Podunavac-Kuzmanović, Lidija R. Jevrvić, Aleksandra N. Tepić, Zdravko Šumić

Univerzitet u Novom Sadu, Tehnološki fakultet, Novi Sad, Srbija

(Naunir rad)

Paprika (Capsicum annuum L.) je jedna od najznačajnijih povrtarskih kultura u svetu i kod nas. Osnovno merilo kvaliteta mlevene začinske paprike je njena ekstrahovana boja, površinska boja i kvalitativni i kvantitativni sastav karotenoida. U tom kontekstu, u ovom radu je upotrebom HPLC ekstrahovana boja, površinska boja i kvalitativni i kvantitativni sastav karotenoida paprike primenom QSRR analize. Kao pokretna faza je bila primenjena smeša vode i acetona sa različitim zapreminskim udaljenostima. U tom kontekstu, u ovom radu je upotrebom HPLC-a na obrnutim fazama ispitani uticaj hemijske strukture na lipofilnost izolovanih slobodnih karotenoida paprike, te analizirana korelacija između retencionalnih konstanti i odabranih parametara lipofilnosti.

Ključne reči: Karotenoidi • RP HPLC • Lipofilnost • QSRR

940