

Scientific paper

# Application of Solvatic Model for Prediction of Retention at Gradient Elution in Reversed-Phase Liquid Chromatography

Tina Snoj Ekmečič<sup>1,2</sup> and Irena Kralj Cigić<sup>1,\*</sup><sup>1</sup> Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia.<sup>2</sup> Krka d.d., R&D, Novo Mesto, Slovenia.

\* Corresponding author: E-mail: irena.kralj-cigic@fkkt.uni-lj.si

Received: 04-11-2019

## Abstract

There are several different approaches for LC method development; beside traditional, different software programs for method development and optimization are available. The solvatic retention model of reversed-phase LC was applied for prediction of retention in the gradient elution mode for aripiprazole and its related substances described in European Pharmacopoeia. As some of these compounds have very similar and others quite different chemical structure, their separation is challenge. Prediction was suitable on examined stationary phases (C18, C8 and phenyl-hexyl) with 0.1% phosphoric acid as aqueous mobile phase and acetonitrile or methanol as organic modifier. Predicted retention times take into account structural formulae of compounds and properties of stationary and mobile phases result in average difference of 14–17% compared to experimental ones on phenyl-hexyl stationary phase, where the highest matching was obtained. After utilisation of the retention models with data from one experimental run, the average difference decrease to maximal 7% and after contribution of data from two experimental runs, to maximal 2%. For majority of studied compounds difference between predicted and experimental values on all examined stationary phases is lower than 3%.

**Keywords:** High-performance liquid chromatography; quantitative structure-retention relationships (QSRRs); solvatic retention model; RP stationary phases; gradient elution; aripiprazole

## 1. Introduction

As development and optimization of HPLC methods can be very time-consuming,<sup>1,2</sup> for rapid method development a systematic, automated approach is needed.<sup>3</sup> This approach includes identification of most suitable initial conditions (column and mobile phase), analysis of multi-component mixture (e. g. active pharmaceutical compound with typical impurities) and evaluation of the results.<sup>4</sup> Presumably methods developed using computer-assisted procedures are expected to be more robust than those, developed by the traditional »trial and error« approach.<sup>5–7</sup> Three main factors affect analyte partition between the stationary and the mobile phase. Those are the chemical structure of analyte, characteristics of the stationary phase and physico-chemical properties of the mobile phase at the constant temperature.<sup>8</sup> Several software programs for method development and method optimization are available – ACD/LC simulator<sup>®</sup>,

ChromSmart<sup>®</sup>, ChromSword<sup>®</sup>, DryLab<sup>®</sup>, Osiris<sup>®</sup>, Pre-opt-W<sup>®</sup> and others. ChromSword<sup>®</sup> and ACD/LC simulator<sup>®</sup> contain algorithms for prediction of initial conditions from structural formulae of compounds. In ACD/LC simulator<sup>®</sup> correlation between logarithm of partition (log P) or logarithm of distribution (log D) and retention in reversed phase liquid chromatography is used. In ChromSword<sup>®</sup> the solvophobic (solvatic) model or reversed-phase liquid chromatography is applied for predictions. Different kind of quantitative structure-retention relationships (QSRRs) have also been described in literature for prediction of retention.<sup>9–15</sup> Optimization of gradient elution is more complex compared to isocratic one and requires more experiments.<sup>5,16–18</sup> In the literature successful prediction of isocratic retention parameters is reported.<sup>16</sup> In 2003, Baczek and Kalisz applied quantitative structure retention relationships (QSRRs) to predict retention in reversed-phase HPLC with linear gradient.<sup>19</sup> Well known is the model based on LSER (linear solvation

energy relationships) developed by Abraham and coworkers.<sup>20</sup> The most known QSRR approach and generally accepted is the solvophobic theory of reversed-phase liquid chromatography. The solvophobic model of interaction of an analyte with surrounding liquid was proposed by Sinanoglu and Haliloglu and applied by Horvath and coworkers<sup>21–23</sup> for a retention description in reversed-phase HPLC. Galushko used the solvophobic theory to calculate retention and in his approach a two-layer continuum model of reversed-phase liquid chromatography was applied.<sup>24–25</sup> In this approach a stationary phase is considered as a quasi-liquid layer that has its own characteristics, which vary with mobile and stationary phase composition. An analyte interacts with the surface layer and retention is determined by the difference in molecule solvation energies in the mobile phase and stationary phase:  $\ln k_x = aV_x^{2/3} + b\Delta G_{e.s.x.H_2O} + c$ , where  $V_x^{2/3}$  is the partial molecular volume of the analyte in water, which determines the value of energy to create a cavity in the mobile and stationary phases,  $\Delta G_{e.s.x.H_2O}$  is the energy of electrostatic interaction of the analyte with water and  $a$ ,  $b$ ,  $c$  are the parameters, determined by the properties of a reversed-phase column and a mobile phase. Thus, in the Galushko's model the molecular interactions of analytes with the stationary and mobile phase are assumed to be accounted by the partial molar surface,  $S$  ( $S = V^{2/3}$ ) that determines the energy to create a cavity in the phases.<sup>5</sup> The partial molar volume parameter ( $V$ ) appears to be a reasonably reliable parameter of structurally nonspecific determination of analyte retention. Another molecular parameter  $\Delta G$ , energy of interaction, reflects differences in the so-called electrostatic intermolecular interactions of analytes with a surrounding mobile phase.<sup>5</sup>

The main limitation for retention prediction of chemical compounds from their structure is the inadequacy of the translation of structural formulas into sets of numerical descriptors.<sup>19</sup> Kaliszan's approach needs three molecular descriptors and four regression coefficients, Abraham's approach needs five molecular descriptors, constant  $\log k_0$ , and five regression coefficients and Galushko's approach needs only two molecular parameters and three regression coefficients.<sup>2,20,24–25</sup> It should be noted that Kaliszan's approach can be used only for prediction on selected stationary phase and only for the same gradient that was used for deriving the model,<sup>16</sup> while Galushko's approach enables prediction of retention also for other stationary phases and multistep gradients.<sup>5</sup> Other practical important point to be considered is that in QSRR modeling retention behaviour is usually performed in one software, and different software is used to calculate the molecular descriptors.<sup>26</sup> ChromSword® enables all operations – structures drawing, molecular parameters calculation, chromatogram simulation for different columns and gradients in one software platform. In off-line mode ChromSword® software utilise defined relationships between separation, retention, and chromatographic condi-

tions for the prediction. Chemical structures of analytes are entered and the software model chromatographic retention behaviour consider organic modifier in mobile phase. Additionally, the software uses chromatographic data obtained from at least two initial experiments to predict optimum separation condition for different type of retention models.<sup>13</sup>

In the present work, the solvatic retention model in reversed-phase HPLC for retention prediction in gradient elution mode, using different types of stationary phases (C18, C8 and phenyl-hexyl) was applied for aripiprazole and its related substances described in European Pharmacopoeia. Also, two different organic modifiers (acetonitrile and methanol) were utilised. As some of the compounds have basic properties, effects of organic modifiers are important as methanol is concerned mainly in proton acceptor interactions, while acetonitrile mainly to dipole-dipole interactions.<sup>2</sup>

## 2. Experimental

### 2.1. Chemicals and Reagents

Acetonitrile and methanol were ultra-gradient HPLC grade and were obtained from J.T.Baker (Avantor Performance Materials, USA). Ultrapure water was obtained with Milli-Q water system (Millipore Merck, Germany). Acetic acid and phosphoric acid (49–51%) were obtained from Sigma-Aldrich, USA.

### 2.2. Standards

Standards of aripiprazole, impurity B (European Pharmacopoeia), impurity E (Eur. Ph.) and impurity F (Eur. Ph.) were obtained from USP (USA). Standards of impurity A (Eur.Ph.), impurity C (Eur.Ph.) and impurity D (Eur.Ph.) were obtained from Toronto research Chemicals, Canada. Impurity G (Eur.Ph.) was obtained from Molcan, Canada.

### 2.3. Mobile Phases and Dilution Solvent

Mobile phase A was prepared with addition of 2.0 mL of 49–51% phosphoric acid to 1000 mL of Milli Q water. Mobile phase B was either acetonitrile or methanol. Dilution solvent was mixture of acetic acid: methanol: acetonitrile: water = 1:10:30:60 (V/V/V/V).

### 2.4. Equipment

HPLC system Agilent 1260 series (Agilent Technologies, USA) equipped with quaternary pump, vacuum degasser, autosampler, temperature-controlled column compartment for five columns and diode array detector was used. The mobile phases were A aqueous solution of phosphoric acid, B acetonitrile and C methanol.

## 2. 5. Software

ChromSwordAuto<sup>®</sup> method development chromatography data system (ChromSword, Germany), version 5.0.234.300, was used for HPLC system control, acquisition of chromatographic data and the rapid automatic optimization of gradient methods. ChromSword<sup>®</sup> software for computer-assisted method development was used for retention prediction from structural formulae, stationary and mobile phase with simulation of gradient elution.

## 2. 6. HPLC Columns and Chromatographic Conditions

Five HPLC columns with the same dimensions were utilised (Table 1): three Halo columns with identical particle and pore size, the same surface area, but different type of reversed phase (C18, C8 and phenyl-hexyl) and two columns from different vendors, slightly different particle and pore size and the same phases as two Halo columns. Software input parameters concerning stationary phase were: column name, type of stationary phase, column and particle dimensions. Columns temperature was maintained at 30 °C. Flow rate of mobile phase was 1.0 mL/min and injection volume 5 µL. Detection was at 220 nm. Concentration of each compound in standard solution was 0.1 mg/mL. Gradient profiles for all five columns and both organic modifiers are shown in Table 2. Minimum and maximum percent of organic modifier in gradient was preselected and then software delivers initial gradient conditions: the same for all selected columns (1<sup>st</sup> run). Gradient was determined independent from chemical structures of analytes. Then software modelled two different gradient conditions (2<sup>nd</sup> and 3<sup>rd</sup> run).

## 3. Results and Discussion

### 3. 1. Investigated Compounds

Drug aripiprazole and its related substances were studied. All compounds are described and marked with

letters in European Pharmacopoeia (Figure 1). Aripiprazole is chemically 7-[4-[4-(2,3-dichlorophenyl)piperazin-1-yl]butoxy]-3,4-dihydroquinolin-2(1H)-one. Chemical structures of aripiprazole and impurities C, D, E, F and G are very similar as 7-hydroxy-3,4-dihydroquinolin-2(1H)-one is part of the structure. Aripiprazole and impurities B, E, F and G have 1-(2,3-dichlorophenyl)piperazine, but impurities C and D have only 1-(2-chlorophenyl)piperazine. These structural parts include basic N atoms, therefore analytes are bases. Piperazine part and dihydroquinolin-2(1H)-one part are linked via butyl in aripiprazole, impurity C, D, E, F and G. Impurity G is chemically the most different from aripiprazole, while impurities C and D have the most similar chemical structures (they are positional isomers).

Separation of analytes, which have very similar chemical structures, and separation of far different chemical structures in one analysis represents a challenge and can require a large number of experiments for separation optimisation, therefore this group of compounds was selected for automatic method optimisation.

### 3. 2. Method Development

Software ChromSword<sup>®</sup> was used for automatic method optimisation using the rapid method development algorithm, which usually performs 3 gradient runs (1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup>) to achieve separation of all analytes.<sup>13</sup> The optimisation of the method was performed on five different reversed-phase HPLC columns (Table 1) with acetonitrile or methanol as organic modifier.

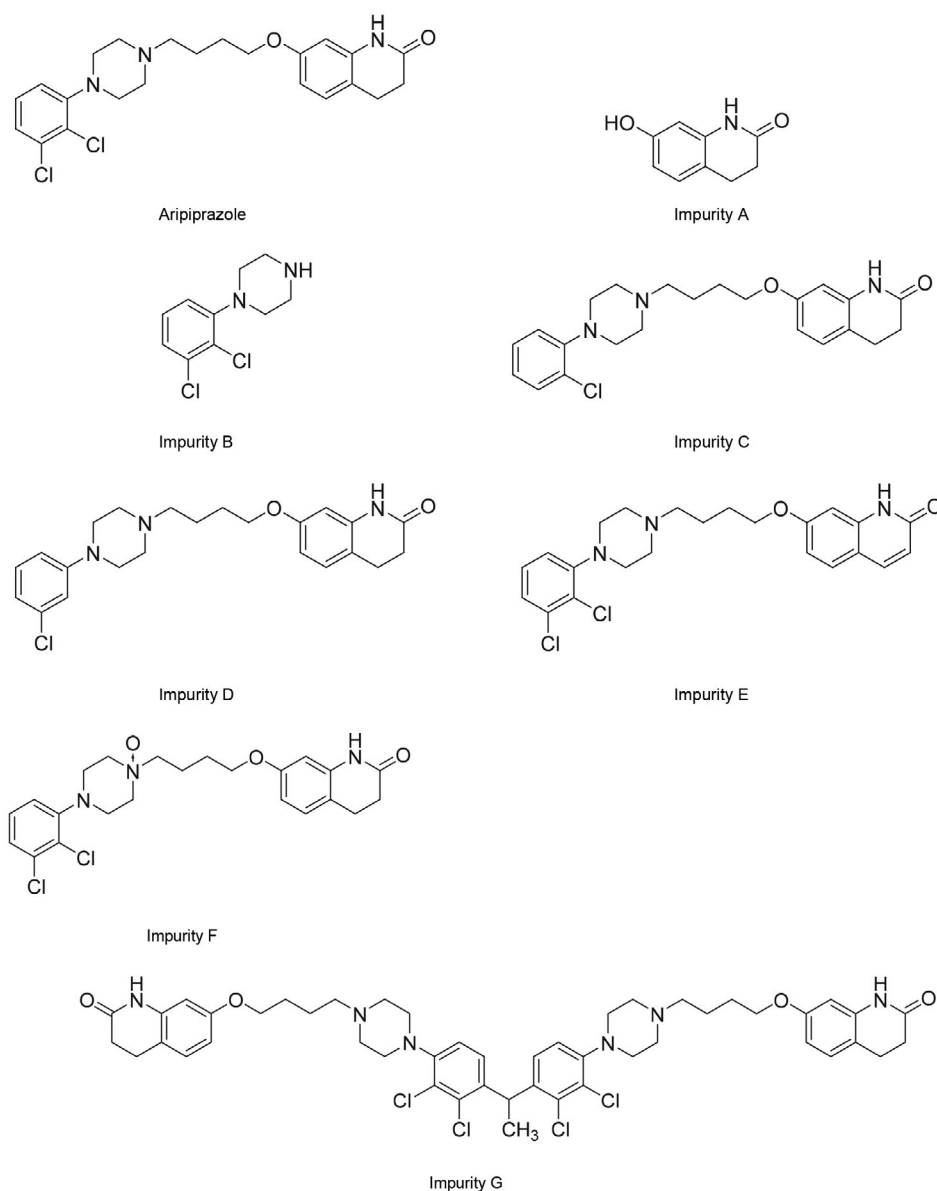
Retention prediction in the reversed-phase HPLC from chemical structure of analytes (Figure 1) and stationary/mobile phase characteristics can be considered as the zero approximation level and can be used for selection of initial conditions (the “first guess method”). Parameters of the impurity F contains the N-oxide fragment, which is charged and cannot be calculated with the ChromSword<sup>®</sup> software. Therefore, simulated chromatograms have seven peaks and the experimental one eight peaks. The main

Table 1. Utilised reversed-phase HPLC columns.

Column	Stationary phase	Column zero time, min	Length, mm	Internal diameter, mm	Particle size, µm	Pore size, nm	Surface area, m <sup>2</sup> /g	Carbon load, %	Vendor
Triart C18	Silica/C18	1.5	150	4.6	3.0	12	360	20.0	YMC, USA
Halo C18	Silica/C18	1.5	150	4.6	2.7	9	135	7.7	Advanced materials technology, USA
Halo Phenyl-Hexyl	Silica/Phenyl-Hexyl	1.5	150	4.6	2.7	9	135	7.1	Advanced materials technology, USA
Halo C8	Silica/C8	1.5	150	4.6	2.7	9	135	5.4	Advanced materials technology, USA
Symmetry C8	Silica/C8	1.5	150	4.6	3.5	10	335	11.7	Waters, USA

**Table 2.** Gradient profiles for selected columns with acetonitrile (B) as organic modifier or with methanol (C) as organic modifier.

1 <sup>st</sup> run	Column	2 <sup>nd</sup> run	3 <sup>rd</sup> run
0.0 min = 0% B, 20.0 min = 20% B or C 45.0 min = 70% B or C 55.0–60.0 min = 100% B or C	YMC Triart C18	0.0 min = 2% B, 32.6–49.6 min = 50% B	0.0 min = 0% B or C 5.8 min = 19% B, 12.8 min = 28% B, 15.3 min = 29% B, 23.6 min = 36% B, 30.7 min = 40% B 32.9–52.7 min = 47% B
		0.0 min = 1% C, 32.6–50.8 min = 100% C	0.0 min = 2% C, 16.3 min = 56% C, 23.6 min = 74% C, 31.5–64.5 min = 100% C
	Halo C18	0.0 min = 0% B, 32.8–53.4 min = 80% B	0.0 min = 0% B, 8.2 min = 27% B, 14.0 min = 34% B, 19.9 min = 39% B, 32.9–51.5 min = 82% B
		0.0 min = 0% C, 32.8–56.3 min = 97% C	0.0 min = 0% C, 16.9 min = 56% C, 21.0 min = 62% C, 23.2 min = 64% C, 32.9–58.5 min = 96% C
	Halo Phenyl-Hexyl	0.0 min = 0% B, 32.8–50.4 min = 82% B	0.0 min = 0% B, 5.5 min = 18% B, 14.7 min = 28% B, 17.5 min = 37% B, 22.6 min = 38% B, 32.4–49.8 min = 70% B
		0.0 min = 0% C, 32.8–54.8 min = 97% C	0.0 min = 0% C, 13.9 min = 46% C, 15.2 min = 49% C, 21.1 min = 68% C, 24.2 min = 69% C, 32.9–55.0 min = 97% C
	Halo C8	0.0 min = 0% B, 32.8–52.6 min = 98% B	0.0 min = 0% B, 2.3 min = 1% B, 10.7 min = 28% B, 13.2 min = 31% B, 32.9–51.8 min = 96% B
		0.0 min = 0% C, 32.5–51.6 min = 100% C	min = 0% C, 0.6 min = 1% C, 30.6–61.7 min = 100% C
	Symmetry C8	0.0 min = 0% B, 32.8–52.7 min = 96% B	0.0 min = 1% B, 7.0 min = 24% B, 12.0 min = 31% B, 32.9–51.9 min = 100% B
		0.0 min = 2% C, 30.9–49.5 min = 100% C	0.0 min = 3% C, 19.7 min = 67% C, 21.8 min = 68% C, 31.6–54.0 min = 100% C



**Figure 1.** Structure of aripiprazole and its impurities described in European Pharmacopoeia.

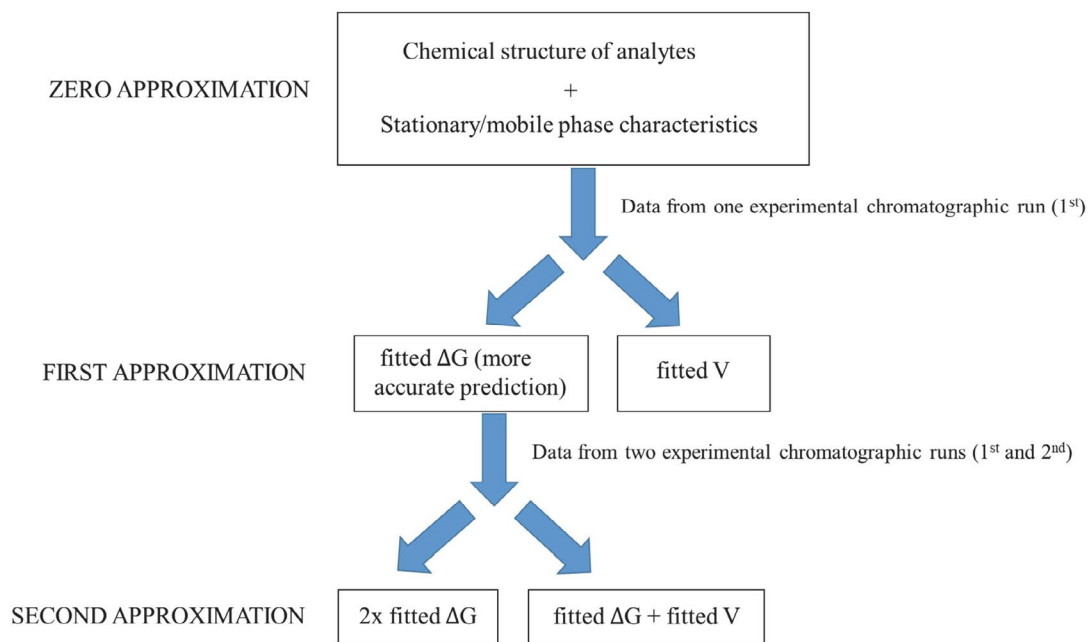
goal of these experiments was to predict conditions that provide as short as possible and practically acceptable retention times. Values for  $\Delta G$  and  $V$  for aripiprazole and its impurities are presented in Table 3. The software ChromSword<sup>®</sup> does not distinguish between positional isomers for the zero approximation level and the estimated  $\Delta G$  and  $V$  values are therefore the same for impurity C and impurity D.

Input and output data for individual approximation are presented in Scheme 1. For the first approximation the results of one experimental run (1<sup>st</sup>) are used to correct retention models for analytes which were derived from structural formulae and characteristics of the reversed-phase column and the mobile phase. The second approximation procedure provides fine tuning of the re-

**Table 3.** Estimated  $\Delta G$  and  $V$  for aripiprazole and its impurities.

Compound	$\Delta G$	$V$
Impurity A	-101	120
Impurity B	-93	184
Impurity C	-191	332
Impurity D	-191	332
Impurity E	-184	345
ARIPIPRAZOLE	-189	350
Impurity G	-380	717

tention models using results of two experimental runs (1<sup>st</sup> and 2<sup>nd</sup>). The first and the second approximations provide retention prediction much more precisely than the zero



**Scheme 1.** Input and output data for zero, first and second approximation.

**Table 4.** Differences of retention times at zero approximation for 1<sup>st</sup> run of rapid method development (acetonitrile as organic modifier) on all five tested columns.

Column Compound	Difference (%)				
	YMC Triart C18	Halo C18	Halo Phenyl-Hexyl	Halo C8	Symmetry C8
Impurity A	-26.3	-30.1	-16.5	-33.7	-10.9
Impurity B	-55.6	-72.7	-32.7	-68.6	-82.1
Impurity C	-28.6	-38.5	-17.4	-35.2	-43.5
Impurity D	-27.5	-38.1	-16.3	-34.3	-42.8
Impurity E	-33.7	-43.7	-18.7	-38.4	-45.7
ARIPIRAZOLE	-27.8	-37.8	-15.0	-33.2	-40.8
Impurity G	-11.4	-19.6	-2.9	-17.1	-23.4
Average	-30.1	-40.1	-17.1	-37.2	-41.3

**Table 5.** Differences of retention times at zero approximation for 1<sup>st</sup> run of rapid method development (methanol as organic modifier) on all five tested columns.

Column Compound	Difference (%)				
	YMC Triart C18	Halo C18	Halo Phenyl-Hexyl	Halo C8	Symmetry C8
Impurity A	-5.4	-21.6	-10.5	-26.1	-10.1
Impurity B	-41.6	-56.5	-31.8	-59.7	-58.8
Impurity C	-29.3	-34.8	-15.7	-39.4	-39.8
Impurity D	-28.1	-36.9	-16.4	-41.2	-41.8
Impurity E	-25.2	-31.0	-13.8	-35.2	-34.4
ARIPIRAZOLE	-23.7	-29.1	-10.9	-33.2	-32.6
Impurity G	-12.5	-14.3	-0.6	-18.5	-19.1
Average	-23.7	-32.0	-14.3	-36.2	-33.8

approximation. Using retention data from experimental run, either the energy of electrostatic interaction of the analyte with water ( $\Delta G_{e.s.x.H_2O}$ ) or the partial molecular volume of the analyte in water ( $V_x^{2/3}$ ) can be corrected. The results for zero approximation are presented in Table 4 and

Table 5 and for the first and second approximation in Table 6 and Table 7 respectively.

Results of the zero approximation step enable prediction of approximate elution order and approximate retention times. Table 4 and Table 5 (and additional tables in

Supplementary Material) represent results for zero approximation for all five tested columns. On phenyl-hexyl stationary phase the highest matching for zero approxima-

tion step was obtained with both organic modifiers. The lowest matching for zero approximation step was obtained on C8 stationary phase. Therefore further optimization

**Table 6.** Difference between experimental and predicted retention times (in%) for the first and the second approximation with acetonitrile as organic modifier on all selected columns.

YMC TRIART C18	First approximation for 2 <sup>nd</sup> consecutive run		Second approximation for 3 <sup>rd</sup> consecutive run	
	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-2.2	-3.4	-2.4	-2.5
Impurity B	no difference	-2.2	-0.7	-0.7
Impurity C	-0.1	0.4	-2.7	-2.7
Impurity D	-0.1	0.5	-2.8	-2.8
Impurity E	0.1	0.8	-1.9	-1.9
ARIPRAZOLE	0.1	0.7	-2.2	-2.2
Impurity G	0.7	0.9	-0.2	-0.2
Average	-0.2	-0.3	-1.8	-1.9
HALO C18	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-3.7	-5.6	-1.3	-1.4
Impurity B	-1.4	-9.1	-1.3	-1.3
Impurity C	-0.3	-1.5	no difference	no difference
Impurity D	-0.1	-1.3	0.2	0.2
Impurity E	-0.4	-1.5	0.8	0.8
ARIPRAZOLE	-0.4	-1.3	1.2	1.2
Impurity G	-0.3	-0.5	4.7	4.7
Average	-0.9	-3.0	0.6	0.6
HALO Phenyl-Hexyl	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-6.3	-7.5	-1.9	-2.0
Impurity B	-3.4	-9.4	0.8	0.9
Impurity C	-0.7	-1.5	-0.7	-0.7
Impurity D	-0.7	-1.4	-0.7	-0.7
Impurity E	-0.9	-1.7	-0.4	-0.3
ARIPRAZOLE	-1.1	-1.6	-0.5	-0.5
Impurity G	-2.5	-2.6	-0.8	-0.6
Average	-2.2	-3.7	-0.6	-0.6
Symmetry C8	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-8.5	-9.3	-1.2	-1.4
Impurity B	-5.1	-17.1	-0.7	-0.7
Impurity C	-1.3	-4.8	0.2	0.2
Impurity D	-0.9	-4.4	0.2	0.2
Impurity E	-1.3	-4.6	1.3	1.4
ARIPRAZOLE	-1.1	-3.8	1.8	1.8
Impurity G	-0.8	-1.3	0.9	0.9
Average	-2.7	-6.5	0.3	0.3
Halo C8	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-2.5	-5.3	0.6	0.7
Impurity B	-1.0	-9.8	-0.4	-0.5
Impurity C	-0.2	-2.2	0.1	0.1
Impurity D	0.2	-1.8	0.1	0.1
Impurity E	-0.7	-2.6	0.8	0.8
ARIPRAZOLE	-0.8	-2.2	0.7	0.8
Impurity G	-1.1	-1.4	0.1	0.1
Average	-0.9	-3.6	0.3	0.3

$\Delta G$  fitted = fitted energy of electrostatic interaction of the analyte with water; V fitted = fitted partial molecular volume of the analyte in water

was done on column Symmetry C8 (first and second approximation) and it is graphically represented in Figure 2 and Figure 3. In practice, the first guess method is used

mainly for retention optimization – retention time should be between 5 and 30 of the column zero time value. Obtained results are reasonable and practically useful to pre-

**Table 7.** Difference between experimental and predicted retention times (in %) for the first and the second approximation with methanol as organic modifier on all selected columns.

YMC TRIART C18	First approximation for 2 <sup>nd</sup> consecutive run		Second approximation for 3 <sup>rd</sup> consecutive run	
	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-10.3	-10.9	-0.1	-0.2
Impurity B	0.4	-3.4	0.2	0.2
Impurity C	0.2	-0.9	0.5	0.5
Impurity D	-0.1	-1.1	0.4	0.4
Impurity E	no difference	-0.9	0.2	0.2
ARIPRAZOLE	no difference	-0.8	0.2	0.2
Impurity G	-0.5	-0.8	1.0	1.0
Average	-1.5	-2.7	0.3	0.3
<hr/>				
HALO C18	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-1.2	-3.3	-0.2	-0.2
Impurity B	1.5	-3.9	-0.1	-0.1
Impurity C	-0.9	-2.0	-0.3	-0.3
Impurity D	-0.6	-1.8	-0.2	-0.2
Impurity E	-1.1	-2.1	-0.3	-0.3
ARIPRAZOLE	-1.2	-2.1	-0.2	-0.2
Impurity G	-1.3	-1.6	2.8	2.8
Average	-0.7	-2.4	0.2	0.2
<hr/>				
HALO Phenyl-Hexyl	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-1.6	-2.6	-0.1	-0.1
Impurity B	-0.5	-2.8	-0.3	-0.4
Impurity C	-1.6	-2.1	-0.3	-0.3
Impurity D	no difference	-0.6	-0.4	-0.3
Impurity E	-0.7	-1.2	-0.3	-0.3
ARIPRAZOLE	-0.7	-1.1	-0.5	-0.5
Impurity G	-1.3	-1.3	-0.8	-0.8
Average	-0.9	-1.7	-0.4	-0.4
<hr/>				
Symmetry C8	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	-5.1	-6.3	-0.1	-0.2
Impurity B	-0.1	-9.4	no difference	no difference
Impurity C	-0.5	-3.0	0.1	0.1
Impurity D	-0.2	-2.9	0.2	0.2
Impurity E	-0.7	-2.7	no difference	no difference
ARIPRAZOLE	-0.7	-2.5	-0.1	-0.1
Impurity G	-1.0	-1.6	-0.1	-0.1
Average	-1.2	-4.0	no difference	no difference
<hr/>				
Halo C8	$\Delta G$ fitted	V fitted	2x $\Delta G$ fitted	$\Delta G$ fitted + V fitted
Impurity A	0.2	-2.3	no difference	no difference
Impurity B	1.7	-4.5	-0.2	-0.1
Impurity C	-0.9	-2.4	-0.2	-0.2
Impurity D	-0.6	-2.1	-0.2	-0.2
Impurity E	-1.3	-2.5	-0.3	-0.3
ARIPRAZOLE	-1.3	-2.4	-0.3	-0.3
Impurity G	-1.5	-1.9	-0.3	-0.3
Average	-0.5	-2.6	-0.2	-0.2

$\Delta G$  fitted = fitted energy of electrostatic interaction of the analyte with water; V fitted = fitted partial molecular volume of the analyte in water



dict the first guess gradient just from structural formulae and stationary/mobile phase characteristics. Software ChromSword® utilise for zero approximation combination of eluents water-acetonitrile or water-methanol and stationary phase. Previously published predicted elution order was the same as the experimental one as buffer with pH 6.4 was used, which is similar as pH of water.<sup>2</sup> In our experiments acidic pH (0.1% H<sub>3</sub>PO<sub>4</sub>, pH = 2.4) was used. Analytes are bases and have charged structures at acidic conditions to have retention as low as possible. Additionally, predicted retention times are longer than experimental ones. This can be explained as all analytes except impurity A contain N atoms, which are strong bases and are protonated at acidic pH conditions (pH = 2.4), therefore their retention is lower than in the neutral eluent (that was obtained in 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> run).

Results in Table 6 and Table 7 demonstrate that retention prediction after the first approximation is much more precise: maximal average difference from experimental values is 6.5% with acetonitrile and 4.0% with methanol. At this step a correction of interaction energy of the analyte,  $\Delta G$  with water is fitted. Difference between predicted and experimental retention is lower if  $\Delta G$  is fitted compared to partial molecular volume,  $V$ . We consider that the reason for this is that partial molecular volume can be calculated from structure more precisely than energy of interaction.

After the second approximation (Table 6 and Table 7) even more precise prediction of experimental retention times on all columns with both organic modifiers was achieved. The average difference between predicted and experimental values is maximal 1.9% with acetonitrile and 0.4% with methanol. For all tested compounds on all five columns, exception is Impurity G, difference between predicted and experimental values is lower than 3% (maximal difference is 2.8% for Impurity D on column YMC Triart C18). Difference between predicted and experimental values is the highest (4.7%) for impurity G on column Halo

C18. Impurity G is dimeric impurity and its chemical structure differ from other compounds and this could be the reason.

For the second approximation step either the interaction energy can be fitted twice (option one) or both the interaction energy and the partial molecular volume (option two) can be fitted. According to the results in Table 6 and Table 7 there is no significant difference between these two options. Results show that second approximation gives much more precise results for predicted retention for 3<sup>rd</sup> consecutive run. However, both the first and second approximations give satisfactory prediction of elution order and retention time for all selected stationary phases (C18, C8 and phenyl-hexyl).

High correlation between experimental and predicted retention from the structure of examined compounds and data from two experiments (second approximation) on all selected columns in both cases – after use of twice  $\Delta G$  fitted (Table 8) or after use  $\Delta G$  fitted +  $V$  fitted (Table 9) was obtained. The highest correlation coefficients are obtained with phenyl-hexyl stationary phase and the lowest with C18 stationary phase, where  $R^2$  is 0.9973 and can be also considered as a very good result.

### 3. 4. Comparison of Experimental and Predicted Chromatograms

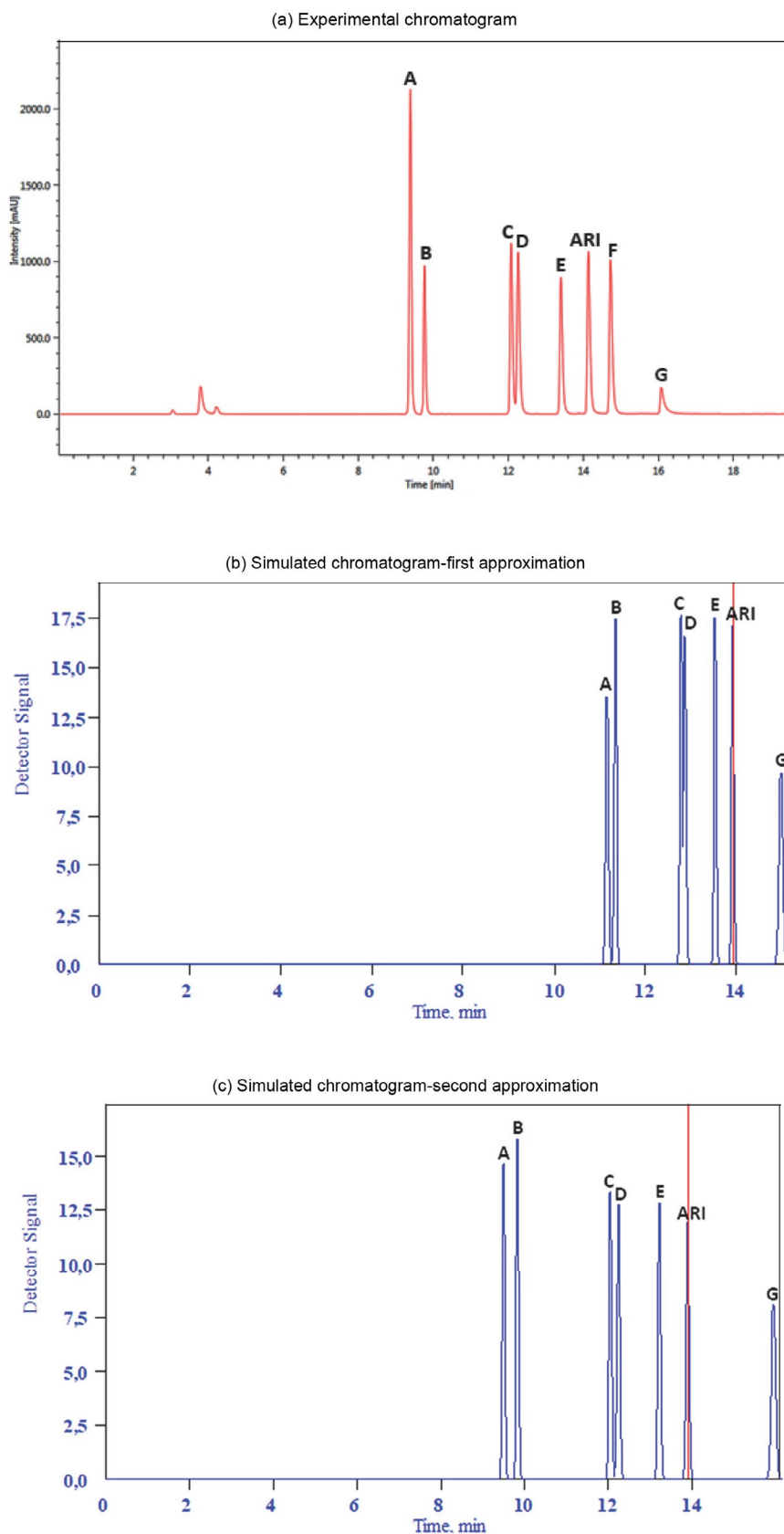
Figure 2 and Figure 3 represent chromatograms of aripiprazole (ARI) and its impurities (A to G as are marked in Pharmacopoeia) on column Symmetry C8. On experimental chromatogram presented in Figure 2a peaks for all eight analytes appear. Peaks for impurity C and impurity D are not baseline separated. That is expected, because their chemical structure is very similar (positional isomers) and therefore their separation is challenging. Figure 2b and Figure 2c represent predicted chromatograms – for the first and the second approximations, where only 7 chromatographic peaks are presented (peak for impurity F is

**Table 8.** Correlation between predicted and experimental retention times from the structure and data of two experiments (second approximation) on all selected columns after use of twice  $\Delta G$  fitted.

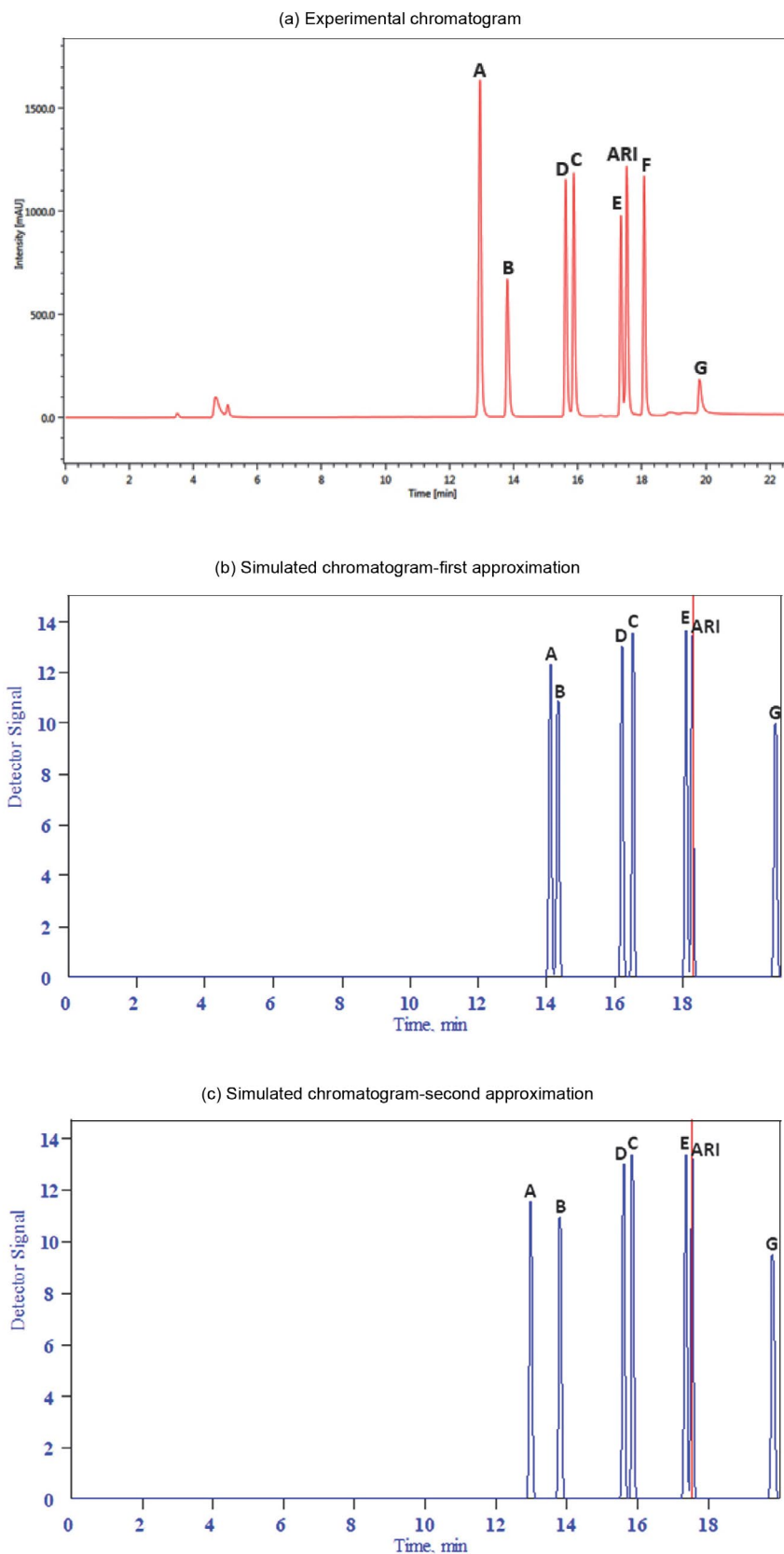
Column Organic modifier	Correlation coefficient ( $R^2$ )				
	YMC Triart C18	Halo C18	Halo Phenyl-Hexyl	Halo C8	Symmetry C8
Acetonitrile	0.9990	0.9982	0.9998	0.9994	0.9991
Methanol	0.9998	0.9973	1.0000	1.0000	1.0000

**Table 9.** Correlation between predicted and experimental retention times from the structure and data of two experiments (second approximation) on all selected columns after use of  $\Delta G$  fitted +  $V$  fitted.

Column Organic modifier	Correlation coefficient ( $R^2$ )				
	YMC Triart C18	Halo C18	Halo Phenyl-Hexyl	Halo C8	Symmetry C8
Acetonitrile	0.9991	0.9982	0.9997	0.9993	0.9990
Methanol	0.9998	0.9973	1.0000	1.0000	0.9999



**Figure 2.** (a) Experimental chromatogram of 3<sup>rd</sup> run of rapid method development (b) predicted chromatogram – first approximation and (c) predicted chromatogram – second approximation on column Symmetry C8, with acetonitrile as organic modifier. Chromatographic conditions are shown in Table 2.



**Figure 3.** (a) Experimental chromatogram of 3<sup>rd</sup> run of rapid method development (b) predicted chromatogram – first approximation and (c) predicted chromatogram – second approximation on column Symmetry C8, with methanol as organic modifier. Chromatographic conditions are shown in Table 2.

missing as explained in paragraph 3.2.). After the first approximation there is a low resolution between peaks for impurities A and B, but after the second approximation step separation practically corresponds to the experimental one.

On another experimental chromatogram presented in Figure 3a obtained with methanol as organic modifier on column Symmetry C8 again eight peaks for all analytes are separated. Separation of peaks for impurity C and impurity D (positional isomers) is better with methanol than with acetonitrile as organic modifier but elution order is reversed. Figure 3b and Figure 3c represent predicted chromatograms for the first and the second approximation correspondently. For these conditions (Figure 3b) there is better resolution between impurities A and B than with acetonitrile as an organic modifier (Figure 2b). After the second approximation (Figure 2c and Figure 3c) these two analytes are baseline separated in both cases (with acetonitrile or methanol as organic modifier).

## 4. Conclusions

The solvatic retention model of reversed-phase high-performance liquid chromatography was applied for prediction of aripiprazole and its impurities retention in multi-step gradient elution mode on C18, C8 and phenyl-hexyl stationary phases with acetonitrile or methanol and 0.1% phosphoric acid mobile phases. In the zero approximation step – only from structural formulae and column/mobile phase characteristics reasonable prediction of retention times was obtained. The average difference between experimental retention and predicted retention is 17.1% with acetonitrile and 14.3% with methanol as the organic modifier on phenyl-hexyl stationary phase, where highest matching was obtained. Retention data from one experimental run (1<sup>st</sup>) were used to correct the retention models. The first approximation step enables prediction of retention time with maximal average difference from experimental values 6.5% with acetonitrile and 4.0% with methanol. After the second approximation step retention data from two runs (1<sup>st</sup> and 2<sup>nd</sup>) were used for further fine-tuning of the retention models. The average difference between predicted and experimental values is maximal 1.9% with acetonitrile and 0.4% with methanol. For all tested compounds on all five columns difference between predicted and experimental values is lower than 3%. Exception is dimeric Impurity G, where on column Halo C18 difference is 4.7%.

Some of investigated compounds, aripiprazole and its related substances described in European Pharmacopoeia, have very similar chemical structure; two of them are even positional isomers. On the other hand, some of the compounds have very different chemical structure. Therefore, appropriate separation of all these compounds in one analysis represent complex problem and the opti-

misation can be time consuming. Using solvatic retention model optimisation was successful; exception is the compound with charged structure, as there is no option for prediction of retention time of such compounds.

In the present research solvatic retention model was investigated for prediction on different reversed stationary phases (C18, C8 and phenyl-hexyl), while in literature generally for reversed phase investigations C18 is used as model phase. In our study best matching between experimental and predicted retention was observed on phenyl-hexyl stationary phase, which was not described so far as model reversed phase and also comparison with C18 and C8 stationary phases was made. Additionally, there are several published investigations using isocratic elution, while few investigations describe results with gradient elution mode. In our research we perform experiments with multi-step gradient and obtained quality results (average difference between predicted and experimental values below 2%) compared to the literature data.

Results of this investigation, where solvatic model was utilised, can serve as support for further usage of this approach for fast development and optimisation of robust analytical methods.

## 5. Acknowledgment

The authors acknowledge the financial support from the Slovenian Research Agency (research core funding No. P1-0153) and Krka d.d., Novo Mesto.

## 6. References

1. A. Wang, P. W. Carr, *J. Chromatogr. A* **2002**, *965*, 3–23. DOI:10.1016/S0021-9673(01)01472-8
2. S. Vorslova, J. Golushko, S. Galushko, A. Viksna, *Proc. Est. Acad. Sci.* **2016**, *65*, 37–49. DOI:10.3176/proc.2016.1.03
3. K. P. Xiao, Y. Xiong, F. Z. Liu, A. M. Rustum, *J. Chromatogr. A* **2007**, *1163*, 145–156. DOI:10.1016/j.chroma.2007.06.027
4. E. F. Hewitt, P. Lukulay, S. Galushko, *J. Chromatogr. A* **2006**, *1107*, 79–87. DOI:10.1016/j.chroma.2005.12.042
5. S. Vorslova, J. Golushko, S. Galushko, A. Viksna, *Chromatographia* **2015**, *78*, 899–908. DOI:10.1007/s10337-014-2816-4
6. A. A. D'Archivio, M. A. Maggi, P. Mazzeo, F. Ruggieri, *Anal. Chim. Acta* **2008**, *628*, 162–172. DOI:10.1016/j.aca.2008.09.018
7. A. A. D'Archivio, M. A. Maggi, F. Ruggieri, *Anal. Chim. Acta* **2011**, *690*, 35–46. DOI:10.1016/j.aca.2011.01.056
8. S. Vorslova, J. Golushko, S. Galushko, A. Viksna, *Latvian J. Chem.* **2014**, *52*, 61–70. DOI:10.2478/ljc-2013-0007
9. www.chromsword.com (accessed: March 3, 2019)
10. www.molnar-institute.com (accessed: March 3, 2019)
11. www.acdlabs.com (accessed: March 3, 2019)
12. S. Fekete, S. Rudaz, J. Fekete, D. Guillarme, *J. Pharm. Biomed. Anal.* **2012**, *70*, 158–168. DOI:10.1016/j.jpba.2012.06.021

13. S. Kromidas, HPLC made to measure: A Practical Handbook for Optimization, first ed., Wiley-VCH, Weinheim, Germany, 2008. DOI:10.1002/9783527611973
14. A. A. D'Archivio, M. A. Maggi, F. Ruggieri, *J. Sep. Sci.* **2010**, 33, 155–166. DOI:10.1002/jssc.200900537
15. A. A. D'Archivio, M. A. Maggi, F. Ruggieri, *Anal. Bioanal. Chem.* **2015**, 407, 1181–1190. DOI:10.1007/s00216-014-8317-3
16. T. Baczek, R. Kaliszan, *J. Chromatogr. A* **2002**, 962, 41–55. DOI:10.1016/S0021-9673(02)00557-5
17. M. Szultka-Mlynska, B. Buszewski, *J. Chromatogr. A* **2016**, 1478, 50–59. DOI:10.1016/j.chroma.2016.11.057
18. A. A. D'Archivio, M. A. Maggi, C. Marinelli, F. Ruggieri, F. Stecca, *J. Chromatogr. A* **2015**, 1423, 149–157. DOI:10.1016/j.chroma.2015.10.082
19. T. Baczek, R. Kaliszan, *J. Chromatogr. A* **2003**, 987, 29–37. DOI:10.1016/S0021-9673(02)01701-6
20. T. Baczek, R. Kaliszan, K. Novotna, P. Jandera, *J. Chromatogr. A* **2005**, 1075, 109–115. DOI:10.1016/j.chroma.2005.03.117
21. C. Horvath, W. Melander, I. Molnar, *Anal. Chem.* **1977**, 49, 142–154. DOI:10.1021/ac50009a044
22. C. Horvath, W. Melander, I. Molnar, *J. Chromatogr. A* **1976**, 125, 129–156. DOI:10.1016/S0021-9673(00)93816-0
23. I. Molnar, *Chromatographia* **2005**, 62, 7–17. DOI:10.1365/s10337-005-0645-1
24. S. V. Galushko, A. A. Kamenchuk, G. L. Pit, *J. Chromatogr. A* **1994**, 660, 47–59. DOI:10.1016/0021-9673(94)85098-4
25. S. V. Galushko, *Chromatographia* **1993**, 36, 39–42. DOI:10.1007/BF02263833
26. H. Du, J. Wang, X. Yao, Z. Hu, *J. Chromatogr. Sci.* **2009**, 47, 396–404. DOI:10.1093/chromsci/47.5.396

## Povzetek

Za razvoj kromatografskih metod lahko uporabimo različne pristope; tako poleg tradicionalnega pristopa lahko optimiziramo metode z različno programsko opremo. Za napoved retencijskih časov pri gradientnem načinu za aripiprazol in njegove nečistote, opisane v Evropski farmakopeji, je bil uporabljen solvatni retencijski model za reverznofazno tekočinsko kromatografijo. Nekatere preučevane spojine imajo zelo podobno kemijsko strukturo, druge pa zelo različno, zato njihova separacija predstavlja izziv. Napoved retencije je bila zadovoljiva na vseh preiskovanih stacionarnih fazah (C8, C18 in fenil-heksil) z 0,1 % fosforno kislino kot vodno mobilno fazo in acetonitrilom ali metanolom kot organskim modifikatorjem. Pri napovedanih retencijskih časih, kjer se je upoštevala kemijska struktura spojin ter lastnosti stacionarne in mobilne faze, je bila povprečna razlika med napovedanimi in eksperimentalnimi retencijskimi časi 14–17 % v primeru stacionarne faze fenil-heksil, pri kateri je bilo ujemanje največje. Pri uporabi retencijskega modela skupaj s podatki enega eksperimenta se je povprečna razlika zmanjšala na največ 7 %, pri uporabi podatkov dveh eksperimentov pa se je povprečna razlika zmanjšala na največ 2 %. Za večino preiskovanih spojin je bila razlika med napovedanimi in eksperimentalno dobljenimi retencijskimi časi manjša od 3 % na vseh preiskovanih stacionarnih fazah.



Except when otherwise noted, articles in this journal are published under the terms and conditions of the Creative Commons Attribution 4.0 International License