

UNIVERZITET U BEOGRADU

HEMIJSKI FAKULTET

Jelena Đ. Trifković

**ODREĐIVANJE KVANTITATIVNOG  
ODNOSA STRUKTURE I RETENCIJE  
ARILPIPERAZINA PRIMENOM TEČNE  
HROMATOGRAFIJE I  
MULTIVARIJANTNIH HEMOMETRIJSKIH  
METODA**

doktorska disertacija

Beograd, 2013

UNIVERSITY OF BELGRADE

FACULTY OF CHEMISTRY

Jelena Đ. Trifković

**QUANTITATIVE STRUCTURE-  
RETENTION RELATIONSHIP STUDY OF  
ARYLPIPERAZINES BY LIQUID  
CHROMATOGRAPHY AND  
MULTIVARIATE CHEMOMETRIC  
METHODS**

Doctoral Dissertation

Belgrade, 2013

Mentor:

**dr Dušanka Milojković-Opsenica**

vanredni profesor Hemijskog fakulteta  
Univerziteta u Beogradu

Članovi komisije:

**dr Živoslav Tešić**

redovni profesor Hemijskog fakulteta  
Univerziteta u Beogradu

**dr Deana Andrić**

vanredni profesor Hemijskog fakulteta  
Univerziteta u Beogradu

**dr Danica Agbaba**

redovni profesor Farmaceutskog fakulteta  
Univerziteta u Beogradu

Datum odbrane:

*Ova doktorska disertacija je urađena na Katedri za analitičku hemiju Hemijskog fakulteta Univerziteta u Beogradu pod mentorstvom dr Dušanke Milojković-Opsenice, vanrednog profesora Hemijskog fakulteta u Beogradu.*

*Autor koristi priliku da se zahvali dr Dušanki Milojković-Opsenici na pomoći i podršci u toku izrade i pisanja rada, na korisnim sugestijama i razumevanju.*

*Zahvaljujem se dr Živoslavu Tešiću na ukazanom poverenju tokom niza godina zajedničkog rada, dr Deani Andrić, koja je sintetisala supstance ispitivane u okviru ove disertacije, na korisnim sugestijama pri završnoj izradi rada, kao i dr Danici Agbabi na stručnoj pomoći.*

*I hvala mojim roditeljima i Leni, Jani i Ivanu. Ovaj rad posvećujem njima.*

# **Određivanje kvantitativnog odnosa strukture i retencije arilpiperazina primenom tečne hromatografije i multivarijantnih hemometrijskih metoda**

Cilj ove disertacije bilo je sistematsko proučavanje hromatografskog ponašanja serije novosintetisanih arilpiperazina u uslovima planarne hromatografije, procena retencionih podataka na osnovu multivarijantnih statističkih metoda i utvrđivanje veze između retencionih karakteristika i fizičko-hemijskih parametara ispitivanih molekula.

Različiti reverzno-fazni hromatografski sistemi uspešno su primenjeni za razdvajanje ispitivanih arilpiperazina. Definisane su međumolekulske interakcije koje upravljaju retencionim ponašanjem ispitivanih supstanci. To su pre svega hidrofobne, elektrostatičke i interakcije koje uključuju građenje vodoničnih veza. Ponašanje arilpiperazina u primenjenim hromatografskim uslovima u saglasnosti je sa njihovim strukturnim svojstvima, kao i sa interakcijama za koje je utvrđeno da se uspostavljaju kada se ligand koji pokazuje afinitet prema dopaminskim D2 i serotoninским 5-HT<sub>1A</sub> receptorima približi aktivnom centru receptora.

Lipofilnost ispitivanih arilpiperazina određena je pomoću četiri hromatografska deskriptora,  $R_M^0$ ,  $b$ ,  $C_0$  i  $PC1$ . Na osnovu statističkih parametara linearne zavisnosti pomenutih deskriptora i izračunatih  $\log P$  vrednosti, utvrđeno je da varijabla koja direktno opisuje raspodelu supstanci između stacionarne i mobilne faze, kao što je parameter  $R_M^0$ , pouzdanije određuje lipofilnost ispitivanih arilpiperazina od parametra koji je proporcionalan hidrofobnoj površini molekula ( $b$ ) ili interpolisane veličine koja obuhvata sve hromatografske podatke jednog sistema ( $PC1$ ).

Analiza glavnih komponenata i klasteraska analiza kao i višestruka linearna regresija, regresija metodom delimičnih-najmanjih kvadrata i regresija glavnih komponenata primenjene su za identifikaciju molekulskih deskriptora koji na najbolji način opisuju ponašanje ispitivanih jedinjenja u hromatografskom sistemu, kao i za kvantifikaciju njihovog uticaja, a u cilju dobijanja matematičkih modela koji bi mogli da predvide hromatografsko ponašanje, odnosno lipofilnost novih srodnih molekula. Statistički parametri potvrđuju da su predloženi modeli statistički značajni i uporedivi. Molekulski deskriptori uključeni u konačne modele su slične prirode i njihov uticaj je približno jednak. Potvrđeno je da je hromatografsko ponašanje molekula određeno

dvema komponentama, veličinom molekula i sposobnošću građenja vodoničnih veza. Usled većeg uticaja parametara rastvorljivosti na  $R_M^0$  vrednosti, pretpostavljeno je da je u posmatranom hromatografskom sistemu particija dominantni mehanizam razdvajanja u odnosu na adsorpciju. Prediktivna svojstva dobijenih modela potvrđuju mogućnost njihovog korišćenja za predviđanje lipofilnosti novih, strukturno sličnih jedinjenja, kao i za razumevanje njihovog hromatografskog ponašanja.

***Ključne reči:*** Arilpiperazini, reverzno-fazna tankoslojna hromatografija, lipofilnost, analiza glavnih komponentata, klusterska analiza, višestruka linearna regresija, regresija glavnih komponentata, regresija metodom delimičnih-najmanjih kvadrata.

Naučna oblast: Hemija

Uža naučna oblast: Analitička hemija

UDK broj: 543

# **Quantitative structure-retention relationship study of arylpiperazines by liquid chromatography and multivariate chemometric methods**

The main goal of this PhD thesis was a systematic study of the retention of the newly synthesized arylpiperazines in planar chromatographic conditions, evaluation of the retention data by multivariate statistical methods and determination of the possible relationship between the retention characteristics and the physicochemical parameters of the investigated arylpiperazines.

Different reversed-phase chromatographic systems are successfully applied for the separation of the investigated arylpiperazines. Hydrophobic, electrostatic and hydrogen bond interactions were defined as the most important molecular interactions govern the retention process. The results obtained are in accordance with the structural properties of the investigated compounds and moreover with the interactions established for their binding affinity at the dopamine D2 and serotonin 5-HT1A receptors.

Four chromatographic descriptors,  $R_M^0$ ,  $b$ ,  $C_0$  i  $PC1$  are used for assessing the lipophilicity of unknown solutes. These parameters were compared with calculated  $\log P$  values, and statistical quality of these dependences shows that variables describing directly solute partitioning between stationary and mobile phase, such as  $R_M^0$ , are more suitable for lipophilicity estimation than parameters proportional to the molecular hydrophobic surface area ( $b$ ), or interpolated quantity which combines all chromatographic data in one single feature ( $PC1$ ).

Principal component analysis and cluster analysis followed by multiple linear regression, partial least square regression and principal component regression, were performed to identify the molecular descriptors that best describe the chromatographic behavior of the investigated compounds, and to quantify their influences, in order to create a model that could be able to predict chromatographic behavior or lipophilicity of a newly synthesized molecules. The statistical parameters revealed that the models are statistically significant and their statistical results are comparable. The descriptors included in the final models of these three methods are of similar nature and significance. All models confirm the importance of the size of the molecule and hydrogen bonding interactions in assessment of their chromatographic behavior.

According to larger influence of the solubility parameters on  $R_M^0$  values, it was concluded that partition was the dominant separation mechanism in studied chromatographic systems. Predictive ability of the obtained models and equations based on the physically meaningful parameters allows an estimation of lipophilicity for similar compounds and understanding of their chromatographic behavior.

**Keywords:** Arylpiperazines, reversed-phase thin-layer chromatography, lipophilicity, principal component analysis, cluster analysis, multiple linear regression, principal component regression, partial least square regression.

Scientific field: Chemistry

Field of Academic Expertise: Analytical Chemistry

UDC Number: 543



# Sadržaj

	Str.		Str.
<b>1. Uvod</b>	1		
<b>2. Opšti deo</b>	4	2.1 Tečna hromatografija	5
		2.1.1 Reverzno-fazna tankoslojna hromatografija	7
		2.2 Procena lipofilnosti jedinjenja	9
		2.3 Ispitivanje kvantitativnog odnosa strukture i retencije	14
		2.3.1 Deskriptori molekulske strukture	15
		2.3.2 Analiza glavnih komponenata	18
		2.3.3 Komplementarne metode za opisnu analizu podataka	21
		2.3.4 Kalibracija	24
		2.3.4.1 Regresija metodom najmanjih kvadrata	31
		2.3.4.2 Višestruki linearni regresioni model	33
		2.3.4.3 Regresija glavnih komponenata	37
		2.3.4.4 Regresija metodom delimičnih-najmanjih kvadrata	39
		2.4 Arilpiperazini: struktura i biološka aktivnost	41
<b>3. Eksperimentalni deo</b>	45	3.1 Reagensi	46
		3.2 Tankoslojna hromatografija	46
		3.3 Geometrijska optimizacija i izračunavanje vrednosti molekulskih deskriptora	47
		3.4 Multivarijantna statistička analiza i modelovanje	48
<b>4. Naši radovi</b>	50	4.1 Retenciono ponašanje ispitivanih arilpiperazina u uslovima planarne hromatografije	51
		4.2 Određivanje parametara lipofilnosti ispitivanih arilpiperazina	58
		4.3 Multivarijantna analiza podataka	65
		4.3.1 Uvodna multivarijantna analiza	65
		4.3.2 Analiza glavnih komponenata	65
		4.3.2.1 PCA primenjena na retencione podatke	66
		4.3.2.2 PCA primenjena na molekulske deskriptore	68
		4.3.3 Klasterska analiza	70
		4.3.4 QSRR analiza	72

	4.3.4.1	Višestruki linearni regresioni model	73
	4.3.4.2	Regresija glavnih komponenata	77
	4.3.4.3	Regresija metodom delimičnih-najmanjih kvadrata	81
	4.3.4.4	Uticaj pojedinih molekulskih deskriptora na hromatografsko ponašanje ispitivanih arilpiperazina	86
<b>5. Zaključak</b>			90
<b>6. Literatura</b>			93
<b>7. Prilog</b>			110

1

**Uvod**

## 1. Uvod

Veza između strukture supstanci i njihovog hromatografskog ponašanja je jedna od najvažnijih teorijskih oblasti izučavanja u hromatografiji poznata pod nazivom Kvantitativni odnosi strukture i retencije (Quantitative Structure-Retention Relationships, QSRR). Glavni ciljevi ovih istraživanja su predviđanje retencionog ponašanja novog jedinjenja na osnovu njegovih strukturnih karakteristika; pronalaženje relevantnih strukturnih parametara koji sadrže najviše korisnih informacija; karakterisanje stacionarnih faza; objašnjenje mehanizama odvajanja u hromatografskom sistemu; optimizacija hromatografskog sistema u određenim uslovima razdvajanja; procena lipofilnosti, odnosno relativne biološke aktivnosti jedinjenja.

Hromatografija je danas jedna od osnovnih metoda za analizu različitih uzoraka, prečišćavanje složenih smeša i razdvajanje i izolovanje pojedinačnih komponenata. Tečna hromatografija predstavlja metodu kod koje se raspodela analita vrši između čvrste stacionarne faze (tankog sloja ili kolone) i tečne mobilne faze. Na konstantnoj temperaturi, ova raspodela je određena hemijskom strukturom analita i fizičko-hemijskim svojstvima mobilne i stacionarne faze. Utvrđivanje i optimizacija matematičke zavisnosti između veličina koje opisuju strukturu jedinjenja i retencije u posmatranim hromatografskim uslovima, primenom različitih statističkih tehnika za multivarijantnu analizu podataka, čini osnov QSRR studija. Hemometrijske metode koje se pri tome najčešće koriste su: analiza glavnih komponenata, klasterka analiza, višestruka linearna regresija, regresija glavnih komponenata i regresija metodom delimičnih-najmanjih kvadrata.

U okviru ovog rada planirano je proučavanje uticaja strukture serije novosintetisanih arilpiperazina na njihovu retenciju u uslovima reverzno-fazne tečne, odnosno tankoslojne hromatografije. Jedinjenja su nastala uvođenjem atoma halogena u benzimidazolsku i benzimidazol-2-tionsku dopaminergičku farmakoforu 1-aralkil-4-arilpiperazina ispitanih u našim prethodnim radovima<sup>1-4</sup>. Analizirani arilpiperazini pripadaju grupi dopaminskih

liganada, koji su veoma značajni zbog svoje uloge u funkcionisanju velikog broja fizioloških procesa, kao što su koordinacija psihomotoričkih funkcija, organizacija procesa prepoznavanja i regulacija reproduktivnih ciklusa kod ljudi. Reverzno-faznom planarnom hromatografijom biće određena lipofilnost arilpiperazina, kao jedan od važnih faktora koji utiču na biološku aktivnost supstanci. Planirana je, takođe, analiza dobijenih retencionih podataka multivarijantnim statističkim metodama (kvalitativna analiza) i pronalaženje mogućeg odnosa između retencionih karakteristika i parametara molekulske strukture (kvantitativna analiza). Cilj navedenog istraživanja bio bi razumevanje mehanizama razdvajanja u primenjenim hromatografskim sistemima, pronalaženje strukturnih parametara koji sadrže najviše korisnih informacija, kao i dobijanje statistički značajnog modela koji bi poslužio za predviđanje retencije i lipofilnosti novih strukturno sličnih jedinjenja.

**2**

**Opšti deo**

## 2.1 Tečna hromatografija

U savremenoj analitičkoj hemiji, koja je zasnovana na primeni novih, naprednih metoda i izuzetno sofisticirane opreme, hromatografija zauzima značajno mesto. Hromatografske metode su, i danas, gotovo nezamenljive prilikom analize različitih uzoraka, prečišćavanja složenih smeša i razdvajanja i izolovanja pojedinačnih komponenata.

Tečna hromatografija predstavlja tehniku kod koje mobilna faza (eluent) prolazi kroz stacionarnu fazu (tanak sloj ili kolonu) noseći sa sobom rastvor uzorka. Pri tome dolazi do procesa adsorpcije i desorpcije komponenata uzorka i njihove raspodele između tečne mobilne faze i površine čvrstog sorbenta kao stacionarne faze. Ceo postupak razdvajanja se može posmatrati kao sekvencijalni proces u kojem se migracija analita (odvajane supstance) ponavlja korak po korak kroz čitavu stacionarnu fazu sve dok uzorak ne napusti sistem ili rastvarač ne dostigne određeno mesto na ploči<sup>5</sup>. Različiti fizički i hemijski faktori određuju mehanizam sorpcije, koji uključuje adsorpciju, particiju, jonsku izmenu i ekskluziju.

Planarna hromatografija predstavlja podvrstu tečne hromatografije u kojoj se kao stacionarna faza koristi tanak sloj sorbenta nanet na neku ravnu površinu (staklenu, metalnu ili plastičnu ploču). Zadržavanje, odnosno retencija, jedinjenja u ovom sistemu izražava se pomoću retardacionog faktora, odnosno  $R_F$  vrednosti, koja predstavlja odnos puta koji je prešla analizirana supstanca i puta rastvarača:

$$R_F = \frac{x}{f} \tag{1}$$

gde je  $x$  pređeni put supstance od polazne mrlje do težišta hromatografske zone, dok je  $f$  pređeni put rastvarača od polazne mrlje do fronta rastvarača.

Nagli razvoj tankoslojne hromatografije sedamdesetih godina XX veka bio je rezultat optimizacije uslova izvođenja, tj. razvoja visokoefikasnih ploča i mogućnosti višestrukog eluiranja. Na taj način je klasična tankoslojna hromatografija postala visoko efikasna kvantitativna metoda koja zbog svoje jednostavnosti, brzine i niske cene danas

predstavlja jednu od osnovnih metoda odvajanja i analize organskih i neorganskih jedinjenja u mnogim laboratorijama<sup>6</sup>. Visoko-efikasna tankoslojna hromatografija (High-Performance Thin-Layer Chromatography, HPTLC) je kompetitivna sa drugim hromatografskim tehnikama u pogledu tačnosti, preciznosti i reproduktivnosti. U odnosu na visoko-efikasnu tečnu hromatografiju (High Performance Liquid Chromatography, HPLC) HPTLC karakterišu izvesne prednosti, kao što su: a) mala količina uzorka potrebna za analizu, b) veća preciznost i tačnost usled mogućnosti istovremene analize većeg broja uzoraka pod istim hromatografskim uslovima, c) mala osetljivost na nečistoće, d) prethodne analize ne mogu da utiču na ishod eksperimenta jer se za svaku analizu koriste nove stacionarne faze, e) mali utrošak mobilne faze po uzorku, f) veliki izbor adsorbenasa i rastvarača za razvijanje, h) mogućnost primene na nepolarnija jedinjenja<sup>7,8</sup>.

Jedan od razloga široke primene planarne hromatografije je i mogućnost izbora i kombinacija velikog broja stacionarnih i mobilnih faza. Stacionarna faza, kao jedna od tri komponente hromatografskog sistema (stacionarna faza, mobilna faza i odvajana supstanca ili sorbend), igra značajnu ulogu u retencionom ponašanju različitih jedinjenja. Poznavanje svojstava stacionarnih faza omogućava njihov pravilan izbor prilikom rešavanja različitih problema vezanih za proces hromatografskog razdvajanja<sup>9</sup>.

Postoji više različitih načina klasifikacije sorbenata koji se koriste u TLC. Najpraktičnija je podela prema hemijskoj prirodi sorbenta na sorbente sa uniformnim hemijskim karakteristikama (silika-gel, aluminijum oksid, celuloza, poliamid) i sorbente sa kombinovanim hemijskim karakteristikama (modifikovani silika-gel).

Najčešće korišćeni sorbent sa uniformnim karakteristikama je silika-gel. Silika-gel je silicijum dioksid kod kojeg je svaki silicijumov atom okružen sa četiri atoma kiseonika gradeći tetraedar. Na površini silika-gela atomi kiseonika povezani su ili sa vodonikom (-Si-OH, silanolna grupa) ili sa drugim atomom silicijuma (-Si-O-Si-, siloksanska grupa). Silanolni aktivni centri se mogu u izvesnoj meri razlikovati, u zavisnosti od toga da li se javljaju kao izolovani ili vezani (tj. slobodni, vicinalni ili geminalni). Sve silanolne grupe ponašaju se kao proton-donori, dok siloksanske grupe imaju proton-akceptorska svojstva. U uslovima normalno-fazne (NP) hromatografije



supstanca se raspodeljuje između izrazito polarne stacionarne faze i manje polarne mobilne faze.

Sorbenti sa kombinovanim hemijskim karakteristikama dobijaju se reakcijom površinskih silanolnih grupa silika-gela sa različitim organskim i organo-silikonskim jedinjenjima. U zavisnosti od primenjenog organskog liganda dobijaju se hidrofobni (nepolarni) ili hidrofilni (polarni) sorbenti. Najznačajniji komercijalno dostupni hidrofobni sorbenti su metil- (RP-2), oktil- (RP-8) i oktadecil- (RP-18) modifikovani silika-gelovi, koji se, kombinovani sa polarnim mobilnim fazama, primenjuju kao sorbenti u reverzno-faznoj (RP) hromatografiji. Hidrofobnost pojedinih sorbenata zavisi ne samo od tipa liganda, već i od gustine prekrivenosti silika-gela alifatičnim ligandima<sup>10,11</sup>. Oktadecil-modifikovani silika-gel je najčešće korišćena stacionarna faza u RPTLC. Uz značajne prednosti, osnovni nedostatak ovog sorbenta je izražena heterogenost njegove površine, odnosno postojanje nemodifikovanih silanolnih grupa na površini koje učestvuju u građenju vodoničnih veza ili jonskim interakcijama sa molekulima odvajanih supstanci<sup>12</sup>.

Među hidrofilno modifikovanim silika-gelovima najveću primenu u hromatografiji na tankom sloju imaju cijano-propil- (CN), amino-propil- (NH<sub>2</sub>) i diol-modifikovani silika-gelovi. Polarne funkcionalne grupe, kod svih ovih sorbenata, vezane su za matriks silika-gela pomoću kratkog nepolarnog dela. Preko polarnih grupa ovi sorbenti ostvaruju specifične interakcije sa sorbendima (normalno-fazna hromatografija), a prisustvo nepolarnih ugljovodoničnih delova na njihovoj površini, omogućava primenu i u reverzno-faznoj hromatografiji<sup>9,13</sup>.

### **2.1.1 Reverzno-fazna tankoslojna hromatografija**

Retencija u reverzno-faznoj tečnoj hromatografiji predstavlja veoma složen proces određen nizom specifičnih i nespecifičnih interakcija koje uključuju vodonično vezivanje, van der Waals-ove i elektrostatičke interakcije<sup>14</sup>. Do interakcija dolazi između odvajane supstance i rastvarača, između molekula sorbenda i u mobilnoj i u stacionarnoj fazi, između molekula rastvarača u mobilnoj i eventualno stacionarnoj fazi,

a značajne su i intramolekulske interakcije alkil-nizova stacionarne faze i međumolekulske interakcije alkil-lanca sa molekulima odvajane supstance i rastvarača. Takođe, alkil modifikovani silika-gel sadrži 30-50 % nemodifikovanih silanolnih grupa sa kojima molekuli analita i rastvarača mogu da interaguju<sup>15</sup>. Jačina silanofilnih interakcija zavisi od broja efektivnih silanolnih grupa, strukture ispitivanih supstanci, kao i od elucione moći upotrebljene mobilne faze.

Postoji više pristupa kojima se objašnjava kompleksan mehanizam razdvajanja u RPTLC sistemu, od kojih su solvofobna teorija i particioni model najšire prihvaćeni. Prema solvofobnoj teoriji Horváth-a i saradnika<sup>16,17</sup>, stacionarna faza ima pasivnu ulogu i zadržavanje je u najvećoj meri određeno favorizovanim promenama slobodne energije mobilne faze, odnosno mobilna faza potiskuje analit prema stacionarnoj fazi. Solvofobna teorija je kritikovana od strane mnogih autora, jer isključuje mogućnost ulaska molekula sorbenda u šupljine stacionarne faze i interakcije u njima, već kao moguće navodi samo interakcije u šupljinama mobilne faze i asocijaciju dva molekula analita<sup>18,19</sup>. Prema Carr-u i saradnicima<sup>20,21</sup>, stacionarna faza ima značajnu ulogu u retenciji, odnosno ukupna slobodna energija retencije nepolarnih jedinjenja je, pre svega, određena doprinosom slobodne energije koja potiče od stacionarne faze. Takođe, prema ovim autorima RP retencioni mehanizam predstavlja više particioni proces nego adsorpciju. Pomenute pristupe, kao i druge koji u ovom radu nisu navedeni, ne treba uzimati *a priori* jer se jasna razlika između particije i adsorpcije u RP hromatografiji nepolarnih jedinjenja ne može uočiti na osnovu termodinamičke analize<sup>22</sup>.

Za uspešno hromatografsko razdvajanje primenom RPTLC važna komponenta je i mobilna faza. U tom smislu, optimizacija sastava mobilne faze ispitivana je od strane mnogih autora, a rezultati tih istraživanja sumirani u seriji radova<sup>23-25</sup>.

U reverzno-faznoj hromatografiji voda je komponenta koja reguliše elucionu moć mobilne faze, dok su komponente koje se koriste za postizanje određene selektivnosti ograničene na one koje su mešljive sa vodom. Eluciona moć rastvarača u velikoj meri zavisi od upotrebljenog sorbenta i predstavlja meru njegovog kapaciteta da stupi u intermolekulske interakcije, dok selektivnost rastvarača predstavlja meru njegove sposobnosti da ostvaruje specifične intermolekulske interakcije, kao što su Keesom-ove (dipol-dipol) interakcije i vodonične veze. Parametar selektivnosti se karakteriše sa dva

različita aspekta, polarnosti i parametra lokalizacije. Polarnost predstavlja meru solvatacione snage rastvarača i određuju ga parametri baznosti, kiselosti i dipolarnosti. Lokalizacioni parametar predstavlja meru sposobnosti rastvarača da ostvari interakcije sa sorbentom, tj. da interaguje na tačno definisan način<sup>26</sup>.

Postoji nekoliko metoda za rangiranje rastvarača prema selektivnosti od kojih je najviše korišćen Snyder-ov trougao selektivnosti<sup>27</sup>. Klasifikacija rastvarača prema pomenutoj metodi je izvršena na osnovu njihove ukupne polarnosti. Rastvarači koji pripadaju istoj grupi pokazuju slično ponašanje. Solvatohromna skala selektivnosti rastvarača, zasnovana na studijama Kamleta i Taft-a<sup>28-30</sup>, o uticaju rastvarača na pomeranje apsorpcionog maksimuma solvatohromnog indikatora je, takođe, često korišćena metoda odabira rastvarača. Solvatohromni parametri su parametar  $\pi^*$ , koji predstavlja meru dipolarnosti i polarizabilnosti rastvarača, kao i parametri  $\alpha_1$  i  $\beta_1$ , koji opisuju sposobnosti vodoničnog vezivanja kao proton-donori, odnosno proton-akceptori, redom. Rastvarači se na osnovu pomenutih parametara klasifikuju u određene grupe slične selektivnosti prema polarnim interakcijama. Najnovija, a istovremeno relativno široko prihvaćena, skala selektivnosti zasnovana je na Abraham-ovim parametrima<sup>31</sup>. Svojstva rastvarača data su veličinama E, S, A, B i V i odnose se redom na: polarizabilnost u višku (u odnosu na *n*-alkan iste molarne zapremine), dipolarnost, sposobnost građenja vodoničnih veza, pri čemu se supstanca može ponašati kao proton-donor, odnosno proton-akceptor. Date veličine korelišu se sa svojstvom rastvarača koje se odnosi na raspodelu istog između dve faze, za različite grupe supstanci, pri čemu se dobijaju odgovarajući regresioni koeficijenti<sup>32</sup>.

## 2.2 Procena lipofilnosti jedinjenja

Lipofilnost predstavlja fizičko-hemijsku karakteristiku koja opisuje sposobnost supstance da se više rastvara u lipidima nego u vodi. Ovaj koncept je od velikog značaja u oblasti medicinskih nauka i životne sredine. Lipofilnost utiče na rastvorljivost jedinjenja<sup>33</sup> i određuje pasivan transport kroz biološke membrane<sup>34</sup>. Pored drugih brojnih strukturnih zahteva utiče i na uspostavljanje interakcija u aktivnom mestu

receptora<sup>35</sup>. Lipofilnost je, takođe, povezana sa biodostupnošću i biokoncentrovanjem hemikalija značajnih za životnu sredinu u zemljištu, biljkama i životinjama<sup>36</sup>.

Lipofilnost nekog molekula je rezultat polarnih i nepolarnih interakcija do kojih dolazi u posmatranoj sredini. Napolarne interakcije (Londonove disperzione sile) mogu da se izjednače sa hidrofobnošću i izraze preko sternih parametara kao što su zapremina molekula, polarizabilnost i molarna refraktivnost. Hidrofobne interakcije pozitivno utiču na lipofilnost. Polarne interakcije, koje obuhvataju jon-jon, jon-dipol, dipol-dipol interakcije i vodonično vezivanje, prikazuju se pomoću elektronskih konstanti, dipolnog momenta ili parametara vodoničnog vezivanja. Na lipofilnost molekula u značajnoj meri utiču i intramolekulske interakcije određene izvesnim strukturnim karakteristikama molekula, kao što su supstitucija na aromatičnom ili heterocikličnom delu molekula, prisustvo  $\pi$  sistema i intramolekulske vodonične veze<sup>37</sup>. Takođe, konformacione promene u molekulu, uslovljene prirodom sredine u kojoj se molekul nalazi (polarna ili nepolarna), u zavisnosti od prirode grupa „zarobljenih” unutar molekula ili izloženih rastvaraču, mogu da učine molekul hidrofilnijim ili lipofilnijim nego što bi se to očekivalo<sup>38,39</sup>. Jonizacija molekula, takođe, značajno menja vrstu međumolekulskih interakcija u koje stupa molekul<sup>40</sup>.

Za izražavanje lipofilnosti jedinjenja obično se koristi particioni koeficijent između vode i organskog rastvarača nemešljivog sa vodom. Opsežno ispitivanje lipofilnosti sredinom dvadesetog veka, od strane Hansch-a i Fujita-e<sup>41</sup> i Leo-a i saradnika<sup>42</sup>, rezultovalo je uvođenjem oktanol-voda sistema, kao referentnog sistema za određivanje ovog svojstva. *n*-Oktanol poseduje izvesne karakteristike zahvaljujući kojima je pogodan za simuliranje biološke membrane. Prisustvo dugačkog alkilnog (hidrofobnog) lanca sa funkcionalnom grupom sposobnom da se ponaša i kao donor i kao akceptor vodonične veze, kao i delimično zasićenje *n*-oktanola vodom, omogućavaju relativno lakše i pouzdanije eksperimentalno određivanje lipofilnosti<sup>42</sup>. Lipofilnost izražena kao oktanol-voda particioni koeficijent, u logaritamskom obliku,  $\log P_{O/W}$ , postala je jedna od najčešće ispitivanih fizičko-hemijskih karakteristika lekova, pesticida i drugih hemikalija<sup>36,43,44</sup>.

Najčešće korišćena metoda za određivanje lipofilnosti u tečno-tečnom sistemu je metoda mućkanja („shake-flask”) u kojoj se vrši raspodela supstance između različitih

zapremina 1-oktanola i vode, u levku za odvajanje. Međutim, značajni nedostaci ove metode su slaba reproduktivnost, dužina trajanja eksperimenta, potreba za većom količinom čistih supstanci, kao i nemogućnost primene na veoma hidrofilne, a naročito veoma lipofilne supstance ( $\log P_{O/W} > 5$ ), zbog nepouzdanosti dobijenih rezultata uslovljene malom količinom oktanola u vodi<sup>45,46</sup>. Lipofilnost jedinjenja se može izraziti i pomoću koeficijenta raspodele supstance u čvrsto-tečnom sistemu, odnosno retencije u reverzno-faznoj visoko-efikasnoj tečnoj ili tankoslojnoj hromatografiji. Osnovne prednosti ovih metoda u odnosu na tradicionalnu proceduru su dužina trajanja eksperimenta, reproduktivnost, neosetljivost na prisustvo nečistoća i degradacionih proizvoda koje se tokom hromatografskog procesa izdvajaju, širi dinamički opseg, on-line detekcija, manje količine uzoraka potrebne za analizu, kao i lakša priprema uzorka<sup>47</sup>. Revije Hartman-a i Schmitt-a<sup>48</sup> i Giaginis-a i Tsantili-Kakoulidou-a<sup>49</sup> daju pregled radova koji opisuju hromatografske metode za određivanje lipofilnosti, stacionarne i mobilne faze koje su u tu svrhu korišćene, poteškoće koje se mogu javiti tokom particionog razdvajanja i načina za njihovo rešavanje, kao i pregled naprednih hromatografskih tehnika.

Reverzno-fazna TLC metoda predstavlja najčešće korišćenu indirektnu metodu za određivanje lipofilnosti jedinjenja. Retenciona konstanta,  $R_M$ , predložena od strane Bate-Smith-a i Westall-a<sup>50</sup>, je najčešće korišćeni retencioni parametar u tankoslojnoj hromatografiji za izražavanje lipofilnosti jedinjenja. Ova vrednost predstavlja logaritamsku funkciju retardacionog faktora:

$$R_M = \log\left(\frac{1}{R_F} - 1\right) \quad (2)$$

Između  $R_M$  vrednosti i koncentracije organske komponente u mobilnoj fazi najčešće postoji linearna zavisnost<sup>51</sup>:

$$R_M = R_M^0 + bC \quad (3)$$

U zavisnosti od primenjene organske komponente u mobilnoj fazi, u slučajevima kada se posmatra širok opseg sastava mobilne faze, mogu se uočiti manje ili veće devijacije

od linearne zavisnosti<sup>52,53</sup>. U tim slučajevima, kvadratna jednačina bolje opisuje retenciju supstance u zavisnosti od sastava mobilne faze<sup>54</sup>:

$$R_M = R_M^0 + b_1 C + b_2 C^2 \quad (4)$$

Odsečak  $R_M^0$  predstavlja  $R_M$  vrednost ekstrapolisiranu na 0 Vol% organske komponente u mobilnoj fazi, odnosno na čistu vodu. Parametar  $R_M^0$  predstavlja meru raspodele supstance između nepolarne stacionarne faze i polarne (vodene) mobilne faze, opisujući na taj način lipofilnost jedinjenja. Vrednost  $R_M^0$  se zbog toga uglavnom koristi kao kvantitativni retencioni deskriptor u TLC umesto  $R_M$  vrednosti koja u velikoj meri zavisi od primenjenih hromatografskih uslova<sup>55,56</sup>.

Parametar  $b$ , odnosno nagib u jednačini 3, ukazuje na zavisnost rastvorljivosti supstance, a samim tim i retencije, od sastava mobilne faze. Nagib se uglavnom posmatra kao karakteristika specifične hidrofobne površine odvajane supstance. Dve teorije opisuju vezu između parametra  $b$  i specifične hidrofobne površine. Prva teorija povezuje nagib sa brojem molekula mobilne faze koji se nalaze u solvacionoj sferi supstance, a koji su oslobođeni prilikom formiranja kompleksa stacionarna faza-sorbend<sup>57</sup>. U slučaju reverzno-fazne hromatografije ova zavisnost je određena nepolarnom (hidrofobnom) površinom molekula. Prema drugoj teoriji, površinski napon mobilne faze se menja sa njenim sastavom, uslovljavajući na taj način i promenu energije formiranja kompleksa stacionarna faza-odvajana supstanca, na mestu koje je upražnjeno oslobađanjem molekula mobilne faze<sup>58</sup>.

Zavisnost između odsečka,  $R_M^0$  i nagiba,  $b$ , za seriju strukturno sličnih jedinjenja, je uglavnom linearna i može se izraziti sledećom jednačinom:

$$R_M^0 = a_0 + a_1 b \quad (5)$$

Prisustvo jonogenih grupa koje modifikuju interakcije jedinjenja sa nepolarnom stacionarnom fazom ili polarnom mobilnom fazom, mogu da naruše pomenutu zavisnost.

Parametar hidrofobnosti  $C_0$ , koji predstavlja odnos odsečka ( $R_M^0$ ) i nagiba ( $b$ ), uveden je od strane Bieganowska-e i saradnika<sup>59</sup>:

$$C_0 = -\frac{R_M^0}{b} \quad (6)$$

Parametar  $C_0$  se može shvatiti kao koncentracija organske komponente u mobilnoj fazi za koju je  $R_M = 0$ , ili kao hidrofobnost po jedinici specifične hidrofobne površine.

Sârbu i saradnici su izneli pretpostavku da visoka linearna zavisnost između odsečka ( $R_M^0$ ) i nagiba ( $b$ ) ne može biti objektivan kriterijum za pronalaženje homologe serije jedinjenja u okviru serije strukturno sličnih molekula, jer su korišćenjem metode rasplinite regresije (fuzzy regression) pokazali da su pomenuti parametri uvek visoko korelisani<sup>60</sup>. Uzimajući u obzir ove rezultate, kao i druge nedostatke regresionog modela, kao što su logaritamska skala i ekstrapolacija, predložili su korišćenje skorova dobijenih primenom metode analize glavnih komponentata (Principal Component Analysis, PCA) na retencionim podacima, kao parametara za opisivanje lipofilnosti<sup>61-63</sup>. U hromatografiji su dve glavne komponente ( $PC1$  i  $PC2$ ) uglavnom dovoljne za opisivanje najvećeg dela varijabiliteta među retencionim podacima. Jedna od komponentata uglavnom pokazuje visoku korelaciju sa lipofilnošću, veličinom molekula ili sternim faktorima, dok je druga komponenta više određena dipol-dipol interakcijama i elektronskim faktorima.

Sva četiri navedena hromatografska deskriptora,  $R_M^0$ ,  $b$ ,  $C_0$  i  $PC1$ , se u istoj meri koriste za određivanje lipofilnosti supstanci i podjednako su zastupljeni u hromatografskoj literaturi<sup>64,65</sup>.

Lipofilnost je bila prva fizičko-hemijska karakteristika za koju su razvijene metode za predviđanje. Danas postoji dosta softverskih paketa za izračunavanje lipofilnosti. Računske metode se zasnivaju ili na podstrukturama u molekulu (Substructure-Based Methods)<sup>66</sup> ili na odgovarajućim svojstvima molekula (Property-Based Methods)<sup>67</sup>. Osnovni princip izračunavanja zasnovanog na podstrukturama je podela molekula na fragmente (tzv. fragmentaciona metoda) ili svođenje na nivo jednog atoma (tzv. atomska metoda) i naknadnog sabiranja dosprinosu pojedinih podstrukturna konačnoj  $\log P$  vrednosti. Metode koja se zasnivaju na svojstvima molekula podrazumevaju korišćenje deskriptora koji opisuju ceo molekul (trodimenzionalni i topološki deskriptori) za izračunavanje  $\log P$  vrednosti. Na osnovu nekoliko

fragmentacionih metoda, kao što su Fujita-Hansch-ova metoda<sup>68</sup>, Rekker-ova metoda<sup>69</sup> i Hancsh-Leo-ova metoda<sup>70</sup>, kao i njihovih kombinacija, razvijeni su različiti softverski paketi. Neki od njih su: KLOGP, K<sub>OW</sub>WIN, CLOGP, ACD/LogP. Softverski paketi koji se zasnivaju na izračunavanjima na osnovu atoma su MOLCAD, TSAR, PrologP, ALOGP98 i XLOGP. Neki od programa za izračunavanje logP vrednosti kod kojih se uzimaju u obzir određene karakteristike molekula su: ABSOLV, ChemProp, VolSurflogP, MlogP, AlogPS, CSlogP.

Uprkos velikoj raznolikosti razvijenih metoda, svaka ukazuje na značaj veličine molekula i sposobnosti građenja vodoničnih veza za tačno predviđanje lipofilnosti.

## 2.3 Ispitivanje kvantitativnog odnosa strukture i retencije

Od 1977. godine u literaturi počinju da se pojavljuju radovi koji govore o nečemu što se danas označava kao ispitivanje kvantitativnog odnosa strukture i retencije (*Quantitative Structure-Retention Relationships, QSRR*)<sup>71,72</sup>. QSRR nastaje kao rezultat primene metodologije koja se koristi u ispitivanjima kvantitativnog odnosa strukture i biološke aktivnosti (*Quantitative Structure-Activity Relationships, QSAR*) na analizu hromatografskih podataka. Ovakve studije se u velikoj meri primenjuju na retencionim podacima dobijenim HPLC, GC i TLC metodama. Osnovni ciljevi, problemi i zablude koje prate ovakva istraživanja, kao i validacije modela i sugestije za dalji rad, kroz opsežan pregled publikacija, dati su u nekoliko revijalnih radova<sup>73-76</sup>.

Dva tipa ulaznih podataka su potrebna za QSRR analize: set kvantitativno uporedivih retencionih podataka za dovoljno veliku grupu jedinjenja i set koji objedinjuje različite parametre koji odražavaju strukturne karakteristike tih jedinjenja<sup>77</sup>. Za relevantan QSRR model neophodan je pouzdan ulazni set podataka, odnosno pravilan odabir molekula i njihovih deskriptora, kao i odgovarajuća statistička analiza. Poželjno je da ispitivana jedinjenja pripadaju homologoj seriji, jer je u suprotnom teško dobiti linearnu zavisnost između retencionih parametara i deskriptora<sup>78</sup>. Ukoliko se QSRR analiza primenjuje na set strukturno različitih jedinjenja, najbolji rezultati se postižu ukoliko se početni set podeli na grupe koje sadrže strukturno slične fragmente<sup>79</sup>.



Dobijanje QSRR modela sastoji se u utvrđivanju i optimizaciji matematičke zavisnosti između deskriptora i hromatografskih parametara primenom različitih statističkih tehnika za multivarijantnu analizu podataka. Najčešće hemometrijske metode koje se koriste u QSRR studijama su: analiza glavnih komponentata (Principal Component Analysis, PCA), klasterka analiza (Cluster analysis), višestruka linearna regresija (Multiple Linear Regression, MLR), regresija glavnih komponentata (Principal Component Regression, PCR), regresija metodom delimičnih-najmanjih kvadrata (Partial Least Square, PLS), veštačke neuronske mreže (Artificial Neural Networks, ANN)<sup>80</sup>. Kao prvi korak u analizi multivarijantnih podataka preporučuje se primena kategorijski nedefinisanih metoda (unsupervised methods) u cilju sagledavanja strukture podataka (detekcije klastera ili spoljašnjih vrednosti) što može biti značajno za dalje razvijanje klasifikacionih ili kalibracionih modela.

Ukoliko se dobije statistički značajan i razumljiv QSRR model, on može da se iskoristi za predviđanje retencije novih, srodnih jedinjenja<sup>81</sup>; za identifikaciju strukturnih deskriptora koji najviše utiču na retenciono ponašanje u posmatranim uslovima; za objašnjenje mehanizma razdvajanja u primenjenom hromatografskom sistemu<sup>82</sup>; za određivanje složenih fizičko-hemijskih svojstava jedinjenja (različitih od hromatografskih)<sup>83,84</sup>; kao i za predviđanje relativne biološke aktivnosti saglasno činjenici da je ista određena lipofilnošću jedinjenja i interakcijama sa aktivnim centrom receptora<sup>85</sup>.

### **2.3.1 Deskriptori molekulske strukture**

Molekulski deskriptori predstavljaju rezultat „prevođenja molekula u brojeve” čime je omogućena matematička obrada hemijskih informacija sadržanih u molekulu<sup>78</sup>. Za određenu hemijsku strukturu pomenuti parametri se dobijaju na osnovu hemijske građe, topologije, geometrije, talasne funkcije i površinske potencijalne energije molekula. U literaturi su opisane hiljade deskriptora, od jednostavnih koji daju broj veza ili atoma u molekulu, do složenih koji uzimaju u obzir trodimenzionalnu strukturu i konformacione promene<sup>86</sup>.

Prema definiciji Todeschini-ja i Consonni-ja molekularni deskriptor je konačan rezultat logičkog i matematičkog postupka kojim se transformiše hemijska informacija, kodirana u simboličnom predstavljanju molekula, u koristan broj ili rezultat nekog standardizovanog eksperimenta<sup>87</sup>.

Prema ovoj definiciji, molekularni deskriptori su podeljeni u dve kategorije: *eksperimentalne vrednosti*, odnosno *empirijski deskriptori*, kao što su  $\log P$ , rastvorljivost u vodi ( $\log S$ ), kiselinske konstante i *teorijski (neempirijski) molekularni deskriptori*, koji proističu iz simboličke prezentacije molekula, kao što su molarna refraktivnost, polarizabilnost, kvantno-hemijski indeksi izračunati semi-empirijskim metodama, strukturni deskriptori dobijeni iz molekularne formule (topološki deskriptori), itd. Eksperimentalno određeni parametri su veoma korisni u ispitivanjima odnosa strukture i različitih svojstava jedinjenja jer na najbolji način odražavaju hemijsku strukturu molekula. Ovakvi deskriptori su, međutim, dostupni za mali broj jedinjenja, zbog čega je u pomenutim analizama neophodno koristiti izračunate deskriptore. Upoređujući dve vrste molekularnih deskriptora, dobijenih za seriju odabranih primarnih amina, primenom PCA i PLS metoda, Andersson i saradnici su potvrdili da izračunati deskriptori sadrže slične informacije o hemijskoj strukturi molekula kao i eksperimentalno određeni parametri<sup>88</sup>.

Za izračunavanje teorijskih molekularnih deskriptora primenjuje se više različitih kompjuterskih programa zasnovanih na matematičkim modelima kojima se opisuje položaj i ponašanje atoma. Za proučavanje molekularnih sistema primenjuju se klasično-mehaničke i kvantno-mehaničke teorije. Klasično-mehanički pristup predstavlja metodu za lociranje optimalne geometrije molekula. Molekularno-mehanička MM+ metoda je najčešće primenjivana i zasniva se na pretpostavci da potencijalna energija rasporeda atomskih jezgara u molekulu zavisi samo od međusobnog položaja jezgara, a ne i od položaja elektrona. Kvantno-mehaničke metode se zasnivaju na rešavanju Šredingerove jednačine semi-empirijskim i *ab initio* pristupom. Semi-empirijske metode koriste eksperimentalne podatke za pojednostavljenje Šredingerove jednačine, dok *ab initio* metode koriste samo matematičke aproksimacije<sup>89</sup>. Najčešće korišćena semi-empirijska metoda je Austin model 1 (AM1)<sup>90</sup>. Ova metoda se često koristi kod izračunavanja kvantno-mehaničkih deskriptora glomaznih molekula za koje *ab initio* proračuni jako

dugo traju i prilično su skupi. Takođe, naelektrisanja, dipoli i dužine veza izračunati AM1 metodom pouzdaniji su od nekih *ab initio* metoda<sup>91</sup>. Metoda AM1 pouzdano reprodukuje elektronske osobine velikog broja supstituenata<sup>92</sup>.

Za izračunavanja molekulskih deskriptora u okviru ove doktorske disertacije korišćeni su programi *HyperChem* i *Molecular Modeling Pro*. *HyperChem* je veoma pogodan za optimizaciju, manipulaciju i vizualizaciju molekula, kao i za molekulsko mehanička, semi-empirijska i *ab initio* kvantno-mehanička izračunavanja koja se izvode simultano. Izračunavanja različitih energija minimizacije (geometrijska optimizacija) i molekulske dinamike daju uvid u molekulsku strukturu i fleksibilnost molekula. *Molecular Modeling Pro* program je specijalizovan za brza i efikasna određivanja odnosa struktura-svojstvo<sup>78</sup> i u velikoj meri ga primenjuju hemičari koji se bave ispitivanjima odnosa između različitih svojstava molekula (QSAR, QSPR, QSRR).

Molekulski deskriptori koji se najčešće koriste u QSRR analizi, klasifikovani na osnovu povezanosti sa određenom karakteristikom molekula, navedeni su u Tabeli 1.

Tabela 1. Najčešće korišćeni deskriptori molekulске strukture

Fizičko-hemijski empirijski i semi-empirijski parametri	Hammett-ova konstanta, hidrofobne konstante, Taft-ove sterne konstante, rastvorljivost, solvatohromni parametri, particioni koeficijenti, temperatura ključanja
Deskriptori povezani sa geometrijom molekula	Dužina molekula, širina molekula, dubina molekula
Particioni koeficijenti, parametri za opisivanje hidrofobnosti molekula	Parametar lipofilnosti - $\log P$ , hidrofilno-lipofilni balans, hidrofilna površina, polarna površina
Deskriptori povezani sa veličinom molekula (Konstitucioni deskriptori)	Molekulska masa, refraktivnost, polarizabilnost, van der Walls-ova zapremina i površina
Deskriptori povezani sa polarnošću molekula	Dipolni momenat, ukupni dipol, energija molekulskih orbitala, akceptor i donor vodonične veze
Deskriptori povezani sa molekulskom grafikom (Topološki parametri)	Indeksi konektivnosti, Kappa indeksi, Valencioni indeksi
Deskriptori povezani sa osobinama polimera i surfaktanata	Površinski napon, rastvorljivost

### 2.3.2 Analiza glavnih komponentata

Analiza glavnih komponentata je multivarijantna metoda za određivanje linearnih latentnih varijabli (komponentata). Može se shvatiti kao metoda pomoću koje se formira novi koordinatni sistem sačinjen od međusobno normalnih latentnih varijabli usmerenih na pravac najvećeg varijabiliteta među podacima<sup>93</sup>. PCA uzima u obzir sve varijable i može se primeniti na bilo koju matricu podataka  $X$ , pri čemu se  $y$ -podaci ne uzimaju u obzir.

Smanjenje dimenzija podataka koristi se u cilju:

- Vizualizacije multivarijantnih podataka pomoću grafika rasutosti (scatter plot).
- Prevođenja visoko-korelisanih  $x$ -varijabli u manji set nekorelisanih latentnih varijabli koje mogu da se koriste u daljoj hemometrijskoj analizi.
- Razdvajanja značajnih informacija sadržanih u podacima (opisanih pomoću nekoliko latentnih varijabli) od onih koje su manje bitne.

PCA primenjena za analizu retencionih podataka dobijenih RPTLC daje značajne informacije o ispitivanim jedinjenjima. Vrednosti skorova (linearna kombinacija retencionih podataka) koji odgovaraju prvoj glavnoj komponenti mogu se smatrati parametrima lipofilnosti, kako je naglašeno u ranijem tekstu. Grafik zavisnosti skorova prve od druge glavne komponente razdvaja jedinjenja po lipofilnosti ukazujući na postojanje odgovarajućih kongenernih serija, dok grafik koeficijenata latentnih varijabli daje uvid u mehanizam razdvajanja i ponašanje jedinjenja u primenjenim hromatografskim sistemima<sup>63,94-96</sup>. PCA razdvaja značajne informacije sadržane u retencionim podacima od informacija koje ne doprinose određivanju lipofilnosti jedinjenja, čime se eliminišu eksperimentalne greške uočene u regresionoj proceduri, kao i precenjene vrednosti dobijene postupkom ekstrapolacije.

U QSRR studijama PCA se izvodi kao uvodna multivarijantna analiza podataka dobijenih hromatografskom analizom i izračunatih primenom različitih kompjuterskih programa. Primena PCA na retencione podatke ukazuje na sličnosti koje postoje između ispitivanih supstanci, a koje su određene njihovim strukturnim karakteristikama, kao i specifičnim interakcijama do kojih dolazi u različitim hromatografskim sistemima<sup>13,55,97</sup>. PCA može da ukaže i na različit uticaj organskih modifikatora i

stacionarnih faza na retenciju, postojanje razlika između organskih rastvarača koji se koriste kao komponente mobilnih faza<sup>98,99</sup>, kao i između primenjenih stacionarnih faza<sup>100</sup>. Takođe, PCA primenjena na vrednosti molekulskih deskriptora grupiše jedinjenja na osnovu sličnosti u strukturi<sup>13,55,97,98</sup>. PCA je osetljiva na prisustvo rezultata koji su posledica grubih grešaka. Ovakve vrednosti u velikoj meri povećavaju varijansu, a pošto su glavne komponente u PCA usmerene na pravac najvećeg varijabiliteta, one bi značajno uticale na ishod analize. U tom smislu, ukoliko se posmatra homologa serija jedinjenja, PCA može da ukaže na prisustvo spoljašnjih vrednosti koje bi trebalo ukloniti pre konačnog modelovanja<sup>13,59,97,101</sup>.

PCA se sastoji u pronalaženju pravca među varijablama koji na najbolji način odražava relativno rastojanje između objekata i opisuje maksimum varijanse skorova. Skorovi predstavljaju projekcije vrednosti podataka na latentnu varijablu, odnosno pravac. Ovaj pravac se naziva prva glavna komponenta (*PC1*) i definiše se pomoću vektora koeficijenata latentnih varijabli (loading vector):

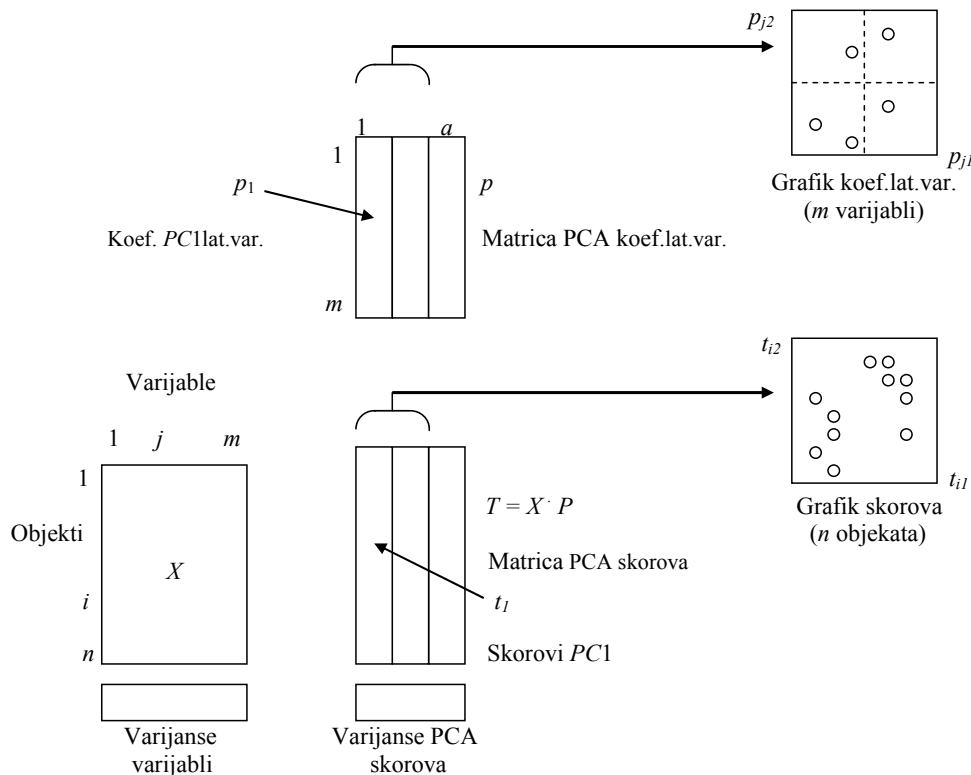
$$\mathbf{p}_1 = (p_1, p_2, \dots, p_m) \quad (7)$$

gde je  $m$  broj varijabli. Odgovarajući skorovi (označeni sa  $t$ ) predstavljaju linearnu kombinaciju koeficijenata i varijabli. Za  $n$  objekata sortiranih u redove u matrici  $\mathbf{X}$ , skor vektor  $t_1$  prve glavne komponente određen je sledećom zavisnošću:

$$\mathbf{t}_1 = \mathbf{X} \cdot \mathbf{p}_1 \quad (8)$$

Druga glavna komponenta (*PC2*) predstavlja pravac normalan na *PC1* kojim se opisuje maksimum preostalog varijabiliteta među podacima. Sve dalje glavne komponente su normalne na prethodne i njihov pravac pokriva varijanse među podacima projektovane na posmatrani pravac. Ukupan broj glavnih komponenata jednak je broju varijabli<sup>93</sup>.

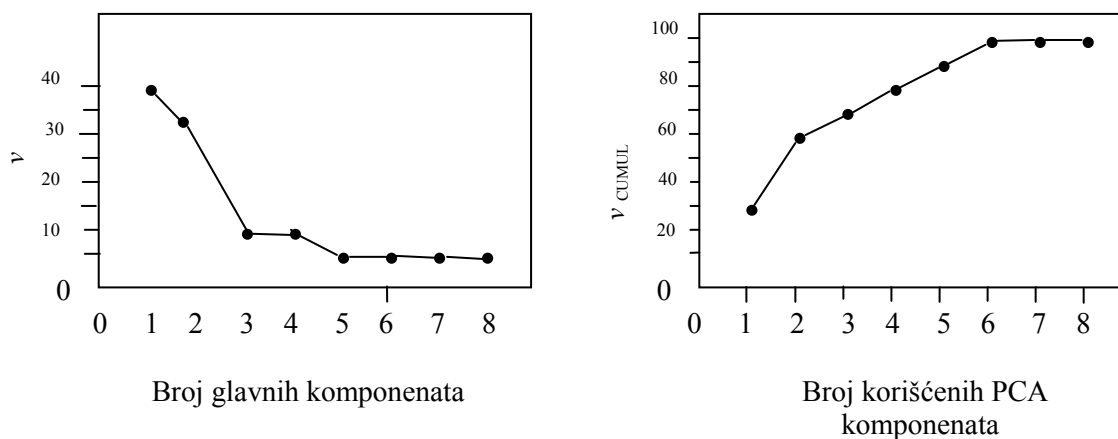
Na Slici 1 dat je šematski prikaz PCA; svi vektori koeficijenata latentnih varijabli predstavljeni su kao kolone u matrici koeficijenata,  $\mathbf{P}$ , dok su svi vektori skorova predstavljeni kao kolone u matrici skorova,  $\mathbf{T}$  ( $\mathbf{T} = \mathbf{X} \cdot \mathbf{P}$ ).



Slika 1. Šematski prikaz PCA metode

Broj glavnih komponentata se u najvećem broju slučajeva određuje pomoću grafika zavisnosti broja komponentata od varijanse PCA skorova (Slika 2). Saglasno definiciji, *PC1* poseduje najveću varijansu dok svaka naredna glavna komponenta opisuje sve manji udeo varijabiliteta. Na pomenutom grafiku kriva naglo opada posle nekoliko prvih komponentata, jer je najveći deo varijanse objašnjen pomoću njih. Preostale glavne komponente sa niskim vrednostima varijansi pružaju informacije koje su od malog značaja za posmatranu analizu. Tačka u kojoj kriva odjednom postaje ravna predstavlja broj glavnih komponentata koje bi trebalo uzeti u obzir za dalju analizu. Pojedini programi za obradu podataka koriste svojstvene vrednosti na pomenutim graficima zavisnosti, koje su identične varijansama skorova. Kumulativna varijansa,  $v_{\text{cumul}}$ , PCA skorova pokazuje koliki deo ukupne varijanse je obuhvaćen određenim brojem glavnih komponentata (Slika 2b). Po pravilu, broj glavnih komponentata koji se

uzima za dalju obradu bi trebalo da objašnjava najmanje 80% ukupne varijanse podataka<sup>80</sup>.



Slika 2. Grafik zavisnosti broja glavnih komponenata od a) varijanse PCA skorova (udeo u ukupnoj varijansi), b) kumulativne varijanse PCA skorova.

Početni set podataka bi trebalo centrirati u odnosu na srednju vrednost ili autoskalirati na nulu srednje vrednosti i jedinicu varijanse pre početka PCA. Na taj način svaka varijabla dobija istu „težinu” za PCA. Međutim, u izvesnim slučajevima skaliranje podataka ima neželjene efekte. Naime, varijable za koje se zna da sadrže malu količinu informacija, posle centriranja, odnosno autoskaliranja postaće značajne koliko i varijable koje sadrže značajnije informacije. PCA ne može da napravi razliku između značajnih i informacija od malog značaja, već izražava što je moguće više varijabiliteta među podacima. U slučajevima kada je značaj pojedinih varijabli poznat, a među varijablama postoje one koje su manje značajne, podatke ne bi trebalo skalirati pre početka analize<sup>80</sup>.

### 2.3.3 Komplementarne metode za opisnu analizu podataka

PCA je najčešće korišćena metoda za opisnu analizu (exploratory analysis) multivarijantnih podataka. Razvijene su, međutim, i nelinearne metode,

komplementarne PCA metodi, kao što su klasterska analiza, faktorska analiza, Kohonen-ovo mapiranje i Sammon-ovo nelinearno mapiranje. Navedene metode daju bolji prikaz strukture podataka od PCA u slučajevima kada linearna projekcija varijabli ne daje dovoljno informacija zbog kompleksne strukture podataka<sup>80</sup>. Prednosti PCA se ogledaju u jasnoj definiciji metode i mogućnosti interpretacije rezultata preko grafika skorova i grafika koeficijenata latentnih varijabli. Takođe, nedostatak nelinearnih metoda u odnosu na PCA je postojanje parametara koji se mogu podešavati, kao što su metode za određivanje udaljenosti objekata i metode za povezivanje istih kod klusterske analize, redukcija podataka kod faktorske analize, itd.

Klasterska analiza predstavlja nelinearnu metodu kojom se u setu multivarijantnih podataka pronalaze grupe sličnih objekata (u prostoru varijabli), ili varijable (u prostoru objekata), ili varijable i objekti istovremeno. Ova metoda pretpostavlja da su informacije o strukturi grupa sadržane u podacima, ali i da određeni objekat, odnosno varijabla, ne pripada samo jednoj grupi, već može biti deo dve ili više grupa<sup>80,93</sup>. U tom smislu, klasterska analiza često daje rezultate koje je teško objasniti, u smislu sličnosti objekata ili varijabli, zbog čega su u literaturi predloženi mnogi algoritmi koji se zasnivaju na različitim principima. Najznačajnije metode klusterske analize su sledeće:

- Particiona metoda, u kojoj je svaki objekat, odnosno varijabla, pripisana samo jednoj grupi.
- Hijerarhijska metoda, u kojoj su objekti, odnosno varijable i particije organizovane hijerarhijski. Odgovarajuća grafička prezentacija rezultata ima izgled dendrograma. Na dobijenom grafiku je moguće odrediti optimalan broj klastera i videti hijerarhijski odnos između različitih grupa objekata, odnosno varijabli.
- Metoda rasplnutog klasterovanja (Fuzzy clustering methods), u kojoj je svaki objekat, odnosno varijabla, pripisana određenom klasteru na osnovu vrednosti koeficijenta pripadnosti toj grupi. Ovaj koeficijent je normiran i nalazi se u intervalu  $[0, 1]$ , a može da se interpretira kao verovatnoća da je neki objekat, odnosno varijabla, pripisana određenom klasteru.



- Metoda grupisanja po određenom modelu, u kojoj različiti klasteri prate određeni model, kao što je multivarijantna normalna distribucija sa određenom srednjom vrednošću i kovarijansom.

Hijerarhijska metoda je najčešće korišćena metoda za klastersku analizu. U okviru nje postoje dve procedure za formiranje klastera:

- Aglomerativna metoda, pri kojoj na prvom hijerarhijskom nivou svaki od objekata, odnosno varijabli, formira zasebne klaster. U sledećem nivou dva najbliža klastera se spajaju, pri čemu se postupak nastavlja sve dok svi objekti ne budu u jednom klasteru.
- Metoda razdvajanja (separaciona metoda), pri kojoj su na prvom hijerarhijskom nivou svi objekti, odnosno varijable, u jednom klasteru. Na sledećem nivou, prvobitni klaster se razdvaja na dva manja klastera, od kojih se svaki u narednom nivou razdvaja na još dva i postupak ponavlja sve dok svaki od objekata, odnosno varijabli, ne formira zaseban klaster.

Razdvajanje ili spajanje klastera u pomenutim hijerarhijskim metodama zasniva se na sličnostima i udaljenosti između klastera. Udaljenost između dva klastera određuje se različitim metodama od kojih su najzastupljenije: potpuno povezivanje (complete linkage), pojedinačno povezivanje (single linkage), prosečno povezivanje (average linkage), metoda centroida (centroid method) i Ward-ova metoda. Sve pomenute metode predstavljaju funkciju rastojanja između objekata, odnosno varijabli, kao što su Euclidean-ova razdaljina, Manhattan-ova razdaljina, Minkowski razdaljina ili Mahalanobis-ova razdaljina. Metoda potpunog povezivanja uzima u obzir maksimum svih međusobnih rastojanja između objekata koji pripadaju klasteru  $j$  i onih koji pripadaju klasteru  $l$ . Posledica ovakvog određivanja rastojanja između klastera je da, ukoliko su samo dva objekta iz dva klastera daleko, rastojanje između klastera će biti veliko. Suprotno ovoj metodi, kod pojedinačnog povezivanja se koristi minimalno rastojanje između dva objekta iz susednih klastera, zbog čega će ono biti malo ukoliko postoje dva objekta koja su dovoljno blizu, nezavisno od ostalih objekata iz posmatranih klastera. Metoda prosečnog povezivanja izračunava srednju vrednost rastojanja između svih parova objekata u dva klastera, odnosno izračunava sumu rastojanja svih

međusobnih klastera i deli je sa brojem parova. Metoda centroida određuje rastojanje između centroida dva klastera. Pomenute metode, međutim, ne dovode do povećanja rastojanja unutar klastera zbog čega je vizualizacija rezultata otežana. Ovaj problem je prevaziđen korišćenjem korekcionog faktora koji se primenjuje u Ward-ovoj metodi<sup>80</sup>.

U QSRR studijama klastera analiza se izvodi kao dodatak PCA metodi u cilju grupisanja sličnih objekata (supstanci) ili varijabli (molekulskih deskriptora) na jednostavniji način i kao pomoć u razumevanju rezultata dobijenih PCA metodom<sup>13</sup>.

### 2.3.4 Kalibracija

Jedan od osnovnih zadataka u nauci i tehnologiji je modelovanje određene osobine  $y$  sa jednom ili više varijabli  $x$ . Modeli opšteg oblika  $y = f(x)$  ili  $y = f(x_1, x_2, \dots, x_m)$  mogu biti linearni ili nelinearni, mogu biti izraženi jednostavnim jednačinama zavisnosti ili manje uočljivim algoritamskim korelacijama. Ovaj rad je fokusiran na linearne modele opšteg oblika:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_jx_j + \dots + b_mx_m + e \quad (9)$$

gde je  $b_0$  odsečak,  $b_1 - b_m$  su regresioni koeficijenti,  $m$  je broj varijabli i  $e$  je odstupanje (rezidual). Regresija može biti izvedena korelisanjem zavisno promenljive sa vrednostima varijabli, kao što je to slučaj kod višestruke linearne regresije, ili pomoću malog seta intermedijalnih linearnih latentnih varijabli, odnosno komponenata, kod metoda kao što su regresija metodom delimičnih-najmanjih kvadrata i regresija glavnih komponenata. Osnovni cilj pomenutih metoda je što tačnije određivanje vrednosti  $b_0, b_1 - b_m$ , kao i greške predikcije.

Svi modeli za predviđanje imaju smisla samo ako je odgovarajući kriterijum za određivanje performansi modela definisan i primenjen. Za analize zasnovane na regresiji, reziduali, tj. greške predviđanja,  $e_i$ , su osnova za određivanje karakteristika modela:

$$e_i = y_i - \hat{y}_i \quad (10)$$

sa  $y_i$  vrednostima koje odgovaraju eksperimentalnim podacima i  $\hat{y}_i$  predviđenim, tj. modelovanim vrednostima, određenog objekta  $i$ . Parametri modela koji samo pokazuju koliko se dobro model uklapa u dobijene podatke nisu prihvatljivi; neophodno je realno određivanje performansi primenom modela na novim podacima. U zavisnosti od veličine seta podataka, odnosno broja raspoloživih objekata, postoji nekoliko načina modelovanja<sup>80</sup>:

1. Ukoliko su na raspolaganju podaci za veliki broj objekata, najbolje je podatke podeliti na tri seta: set za obuku (training set), koji sadrži približno 50% objekata i služi za kreiranje modela, set za validaciju (validation set), sa oko 25% objekata, za optimizaciju modela u cilju povećanja njegove efikasnosti i set za predikciju (prediction, test set), za testiranje krajnjeg modela na novim objektima u cilju dobijanja realne slike o mogućnosti predikcije. Sva tri seta se obrađuju posebno. Primena opisane tehnike u hemiji je retka zbog toga što je često nemoguće obezbediti dovoljan broj objekata za ovakva modelovanja<sup>93</sup>.
2. U realnim uslovima, podatke je najčešće moguće podeliti na set za kalibraciju, koji se koristi za kreiranje i optimizaciju modela i set za predikciju. Kalibracioni set je podeljen na set za obuku i validacioni set pomoću unakrsne validacije (cross validation, CV). Pomenutom tehnikom najpre se određuje optimalna kompleksnost modela, kao optimalan broj PLS komponenata ili varijabli za MLR analizu, a potom se konstruiše model na čitavom kalibracionom setu i dobijena zavisnost primenjuje na test setu.
3. U slučaju manjeg broja objekata, CV metodom je moguće podeliti početne podatke na kalibracioni i test set. Kalibracioni set se koristi za dobijanje optimalnog modela koji se potom primenjuje na setu za predikciju. Svi objekti ulaze i u set za obuku, i u set za validaciju i u set za predikciju, ali se jedan određeni objekat nikada ne koristi istovremeno i za kreiranje i za testiranje modela.
4. Kada se obrađuje jako mali broj objekata, svi podaci mogu da se koriste kao kalibracioni set, dok se CV metoda koristi za određivanje kompleksnosti modela. Zbog nedostatka test seta, predikciona moć dobijene zavisnosti je u najvećem broju slučajeva previše optimistična.

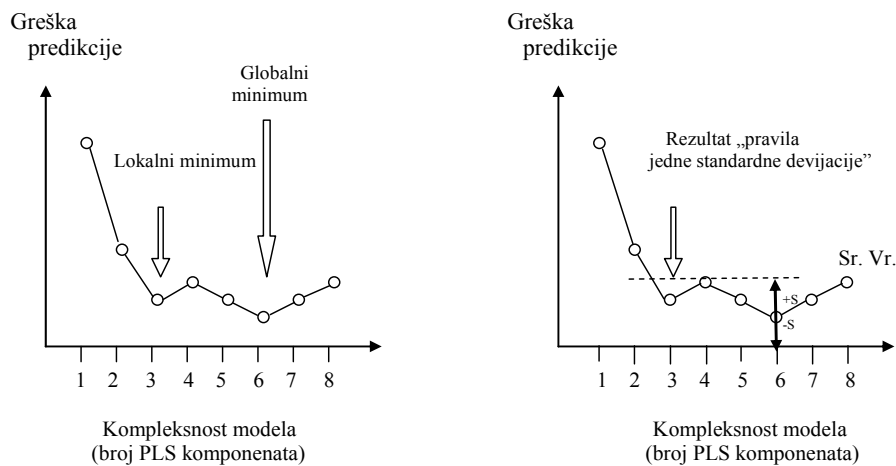
5. Model može da se kreira i iz svih podataka bez bilo kakve metode za optimizaciju i da se potom primeni na istim podacima. U ovom slučaju rezultujuća greška predikcije je uglavnom previše optimistična i ovaj postupak bi trebalo izbegavati.

Podela objekata na set za obuku, validaciju i predikciju se, u najvećem broju slučajeva, izvodi jednostavnom metodom slučajnog uzorka. Međutim, postoje i sofisticiranije metode zasnovane na eksperimentalnom dizajnu i teoriji uzorkovanja. U hemometriji se često koristi Kennard-Stone-ov algoritam po kojem se kalibracioni set formira tako da bude reprezentativan za čitavu populaciju, da uniformno pokriva  $x$ -prostor i da obezbedi odgovarajuću težinu objektima udaljenim od centra<sup>102</sup>. Ovo se ostvaruje odabirom objekata na maksimalnoj razdaljini u  $x$ -prostoru (npr. Euclidean-ova razdaljina).

Sa povećanjem broja nezavisno promenljivih u modelu, odnosno sa povećanjem kompleksnosti modela, greška predikcije za kalibracioni set se smanjuje. To bi značilo da se dovoljno kompleksan model može uklopiti u bilo koje podatke sa zanemarljivom devijacijom (rezidualima) između eksperimentalnih (pravih) i modelovanih (predviđenih)  $y$ -vrednosti. Ovakav model je, međutim, najverovatnije precenjen (overfitted) i često nije koristan za nove objekte. Generalno, greške predikcije za test set su minimalne u „srednje” kompleksnim modelima. U hemometriji, kompleksnost modela se kontroliše brojem PLS i PCA komponenata, kao i broja varijabli, a optimalna kompleksnost modela se određuje pomoću unakrsne validacije<sup>103</sup>. CV omogućava određivanje greške predviđanja za svaki objekat u kalibracionom setu, na svakom nivou kompleksnosti modela. Slika 3 prikazuje tri moguća načina određivanja optimalne kompleksnosti regresionog modela<sup>80</sup>.

1. Određivanje tzv. globalnog minimuma, koji predstavlja kompleksnost modela koji pokazuje minimalnu vrednost greške predviđanja za validacioni set. Ovakav model je obično precenjen.
2. Određivanje tzv. lokalnog minimuma za dobijanje jednostavnijih modela koji nisu precenjeni.
3. Prema “pravilu jedne standardne devijacije” pretpostavlja se da za model određene kompleksnosti postoji više vrednosti grešaka predikcije, što se postiže

primenom unakrsne validacije. Srednja vrednost i standardna greška se izračunavaju za modele različite kompleksnosti i bira se najjednostavniji model čija je srednja vrednost greške predikcije jednu standardnu devijaciju iznad minimalne srednje greške predikcije. Slika 3b prikazuje ovaj postupak, pri čemu tačke na grafiku predstavljaju srednje vrednosti grešaka predikcije, dok je horizontalna linija povučena kroz tačku koja odgovara vrednosti globalnog minimuma plus jedna standardna devijacija. Najjednostavniji model se bira tako da srednja greška predikcije bude ispod pomenute linije.



Slika 3. Načini određivanja optimalne kompleksnosti regresionog modela.

Kvalitet modela određen je greškom predikcije, odnosno rezidualima,  $y_i - \hat{y}_i$ , dobijenim za nezavisni test set, ili unakrsnom validacijom. Postoji veći broj različitih parametara koji se koriste za opisivanje raspodele grešaka, a izvode se iz reziduala. Često korišćena veličina je tzv. *standardna greška predikcije* (standard error of prediction, *SEP*) koja se definiše na sledeći način:

$$SEP = \sqrt{\frac{1}{z-1} \sum_{i=1}^z (y_i - \hat{y}_i - bias)^2}$$

$$bias = \frac{1}{z} \sum_{i=1}^z (y_i - \hat{y}_i)$$
(11)

gde je  $y_i$  eksperimentalna vrednost,  $\hat{y}_i$  predviđena (modelovana) vrednost,  $z$  broj predikcija (koji u slučaju unakrsne validacije može biti veći od broja objekata), *bias* pristrasnost, odnosno greška predubedenja. Greška predubedenja predstavlja aritmetičku sredinu grešaka predikcije i ukoliko ne postoje sistematske greške, trebalo bi da bude bliska nuli. Primenom prethodne jednačine na kalibracioni set dobija se veličina koja se označava kao *standardna greška kalibracije* (standard error of calibration, *SEC*).

*Srednje kvadratno odstupanje ili srednja kvadratna greška* (mean squared error, *MSE*) predstavlja srednju vrednost kvadrata grešaka:

$$MSE = \frac{1}{z} \sum_{i=1}^z (y_i - \hat{y}_i)^2 \quad (12)$$

*MSEC* se odnosi na rezultate iz kalibracionog seta, *MSECV* na rezultate dobijene unakrsnom validacijom i *MSEP* na rezultate iz test seta. *MSE* i *SEP* se nalaze u sledećem odnosu:

$$SEP^2 = MSE - bias^2 \quad (13)$$

Pojmovi koji se koriste u multivarijantnoj analizi podataka su i *koren srednjeg kvadratnog odstupanja* (root mean squared error, *RMSE*) koji može biti određen za kalibraciju (*RMSEC*), unakrsnu validaciju (*RMSECV*) i predikciju (*RMSEP*). U slučaju zanemarljive vrednosti sistematske greške, *RMSEP* i *SEP* su gotovo identični, kao i *MSEP* i *SEP*<sup>2</sup>.

$$RMSE = \sqrt{\frac{1}{z} \sum_{i=1}^z (y_i - \hat{y}_i)^2} \quad (14)$$

U unakrsnoj validaciji često se koristi veličina označena kao *suma kvadrata odstupanja od predviđenih vrednosti* (predicted residual error sum of squares, *PRESS*), koja predstavlja sumu kvadrata grešaka:

$$PRESS = \sum_{i=1}^z (y_i - \hat{y}_i)^2 = z \cdot MSE \quad (15)$$

Postupak odabira manjeg seta ( $m$ ) od velikog broja varijabli ( $n$ ) naziva se selekcija varijabli i koristi se prilikom konstruisanja višestrukog linearnog regresionog modela. Parametar koji se koristi kao kriterijum za odabir najboljeg MLR modela u procesu selekcije varijabli je tzv. *korigovani (podešeni) koeficijent determinacije* (adjusted coefficient of determination,  $R^2_{adj}$ ) koji se može izraziti na jedan od sledeća dva načina:

$$R^2_{adj} = 1 - \frac{n-1}{n-m-1} (1 - R^2)$$

(16)

$$R^2_{adj} = 1 - \frac{RSS/(n-m-1)}{TSS/(n-1)}$$

(17)

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$TSS = \sum (y_i - \bar{y})^2$$

gde je  $R^2$  *koeficijent determinacije*, tj. deo varijanse koji je objašnjen modelom,  $RSS$  (residual sum of squares) je *suma kvadrata odstupanja*,  $TSS$  (total sum of squares) je *ukupna suma kvadrata odstupanja*, tj. suma kvadrata razlika aritmetičke sredine eksperimentalnih vrednosti  $y$  i svake vrednosti  $y$  pojedinačno.

Unakrsna validacija predstavlja korisnu alatku koja rešava problem dve kritične funkcije u hemometriji:

- omogućava određivanje optimalne kompleksnosti modela, odnosno broja glavnih komponentata u PCA ili PCR modelu, ili broja latentnih varijabli u PLS modelima, i
- omogućava određivanje performansi modela kada se primeni na nepoznatim podacima.

Unakrsna validacija se zasniva na nizu analiza koje se označavaju kao *sub-validacioni* eksperimenti. Svaki eksperiment se izvodi tako što se za odabrani set podataka konstruiše model koji se potom primeni na preostale podatke. Na ovaj način, svaki sub-validacioni eksperiment uključuje testiranje modela na objektima koji nisu korišćeni za njegovo dobijanje. Unakrsna validacija uglavnom uključuje više od jednog

sub-validacionog eksperimenta, od kojih svaki odabira različite podsetove uzoraka za građenje modela i njegovo testiranje. U statistici postoji nekoliko različitih metoda za unakrsnu validaciju koje se razlikuju prema načinu odabira podsetova za sub-validacioni eksperiment<sup>80</sup>.

Svojstva pet različitih metoda za unakrsnu validaciju dostupnih u okviru statističkog paketa PLS\_Toolbox za Matlab, koji je korišćen u ovoj disertaciji za statističku obradu rezultata, biće u ovom delu rada prodiskutovane. Pri narednim opisivanjima pojedinih metoda unakrsne validacije,  $n$  predstavlja ukupan broj objekata u setu podataka, dok je  $s$  broj podela podataka specifičan za određenu proceduru, koji mora biti manji od  $n/2$ . Pet pomenutih metoda su: metoda preklapanja (venetian blinds), metoda susednih blokova (contiguous blocks), metoda slučajnog podseta (random subset), metoda izostavljanja jednog objekta (leave-one-out) i uobičajena metoda (custom).

- Metoda preklapanja predstavlja metodu kod koje se test set formira tako što se među podacima jedan objekat označi sa  $s$ - i svi objekti počevši od onog numerisanog kao prvi do  $s$ -tog objekta čine test set<sup>104,105</sup>. Ova metoda je jednostavna i laka za primenu ukoliko postoji relativno veliki broj objekata kod kojih nije utvrđena određena pravilnost u redosledu.
- Kod metode susednih blokova svaki test set je određen odabirom susednih blokova koji sadrže  $n/s$  objekata iz seta podataka, počevši od prvog objekta u nizu<sup>105</sup>. Slično prethodnoj metodi, ovaj način unakrsne validacije je jednostavan i primenljiv na set većeg broja objekata između kojih ne postoji određena pravilnost.
- Kod metode slučajnog podseta određuje se  $s$  različitih test setova slučajnim odabirom  $n/s$  objekata iz seta podataka, tako da jedan određeni objekat bude samo u jednom test setu<sup>105</sup>. Metoda slučajnog podseta obično daje dobre rezultate, naročito kada se selekcija podsetova ponavlja više puta. Ovom metodom se prevazilaze problemi koji se javljaju u unakrsnoj validaciji, kao što je mogućnost da objekti iz test seta i seta za modelovanje imaju različit fizički smisao ili da se isti objekat pojavljuje u oba seta. Povećanjem broja podela može se pravazići problem različitog



fizičkog smisla objekata iz kalibracionog i test seta, dok se smanjenjem broja podela prevazilazi problem pojavljivanja objekata u oba seta.

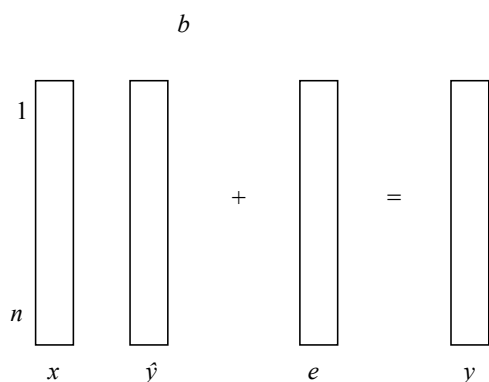
- Kod metode izostavljanja jednog objekta svaki pojedinačni objekat u setu podataka se koristi kao test set<sup>105</sup>. Ova metoda se uglavnom primenjuje kod manjeg broja objekata u setu ( $n < 20$ ).
- Kod uobičajene metode korisnik bira svaki test set. Na ovaj način određeni objekti se mogu naći u svakom test setu, mogu biti izostavljeni iz test setova ili mogu biti izostavljeni iz postupka unakrsne validacije. Uobičajena metoda zahteva dodatno vreme za određivanje kalibracionog i test seta za svaki sub-validacioni eksperiment, ali njena fleksibilnost može biti presudna u slučaju malog broja objekata ili kada su podaci generisani iz eksperimentalnog dizajna koji može prouzrokovati da test set i model set imaju različit fizički smisao.

#### 2.3.4.1 Regresija metodom najmanjih kvadrata

Linearni model kojim se povezuje jedna zavisna promenljiva sa jednom nezavisnom varijablom široko se primenjuje u hemiji. Predstavlja se sledećom relacijom:

$$y = a + bx + e \tag{18}$$

gde su  $a$  i  $b$  regresioni koeficijenti,  $a$  je odsečak,  $b$  je nagib, dok vektor  $e$  obuhvata rezidualne (odstupanja) usled odstupanja podataka od idealnog linearnog odnosa<sup>106</sup>. Šematski prikaz ovog modela dat je na Slici 4.



Slika 4. Šematski prikaz regresije metodom najmanjih kvadrata

Osnovni zadatak prilikom konstruisanja linearnog modela je da se regresioni koeficijenti odaberu tako da greške modela budu svedene na minimum. Predviđene, tj. modelovane vrednosti  $\hat{y}_i$ , za uzorak  $i$  i greška predikcije posmatrane zavisnosti  $e_i$  dati su sledećim relacijama:

$$\begin{aligned} \hat{y}_i &= a + bx_i \\ e_i &= y_i - \hat{y}_i \end{aligned} \quad (19)$$

Najčešće korišćena metoda za određivanje prave koja se na najbolji način uklapa u eksperimentalne rezultate je metoda najmanjih kvadrata. Pri tome se traži onaj pravac (vrednost nagiba i odsečka) koji daje najmanju vrednost sume kvadrata reziduala,  $\sum e_i^2$ . Ovaj uslov je ispunjen kada je vrednost nagiba i odsečka data sledećim relacijama:

$$\begin{aligned} b &= \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} \\ a &= \bar{y} - b\bar{x} \end{aligned} \quad (20)$$

gde je  $\bar{x}$  srednja vrednost svih vrednosti  $x_i$ , a  $\bar{y}$  srednja vrednost svih  $y_i$  vrednosti. Greške nagiba i odsečka, kao i odgovarajući intervali pouzdanosti, proporcionalni su standardnoj grešci modela  $s_{y/x}$  i dati su sledećim relacijama:

$$s_{y/x} = \sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{n-2}}$$

(21)

$$s_a = s_{y/x} \sqrt{\frac{\sum_i x_i^2}{n \sum_i (x_i - \bar{x})^2}}$$

(22)

$$s_b = \frac{s_{y/x}}{\sqrt{\sum_i (x_i - \bar{x})^2}}$$

(23)

$$L_{1,2a} = a \pm t_{(n-2)} s_a$$

$$L_{1,2b} = b \pm t_{(n-2)} s_b$$

gde je  $t$  kritična vrednost Studentove  $t$ -raspodele za određeni nivo poverenja i broj stepeni slobode. Standardna greška modela predstavlja kvadratni koren sume kvadrata rezidualnih vrednosti podeljenih brojem stepeni slobode koji u slučaju linearne regresije iznosi  $n - 2$ , gde je  $n$  broj kalibracionih tačaka.

#### 2.3.4.2 Višestruki linearni regresioni model

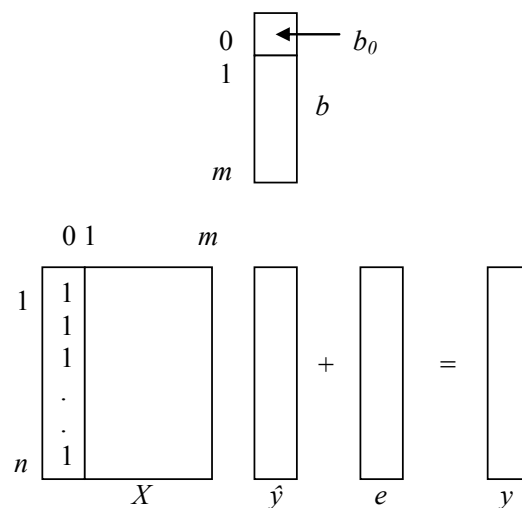
Linearni model kojim se povezuje više nezavisno promenljivih sa jednom zavisno promenljivom je višestruki linearni regresioni model i prikazuje se sledećom relacijom:

$$y_n = a + b_1 x_{n1} + b_2 x_{n2} + \dots + b_m x_{nm} + e_n \quad (24)$$

gde je  $n$  broj objekata, a  $m$  broj varijabli<sup>106</sup>. Mnogo pogodniji način prikazivanja prethodne zavisnosti je pomoću matrica. Ukoliko se nezavisno promenljive prikažu pomoću matrice  $\mathbf{X}$ , dimenzija  $n \times (m+1)$ , koja u prvoj koloni (koja određuje vrednost odsečka) sadrži  $n$  vrednosti jednakih 1, a preostale kolone vrednosti  $x_{ij}$  ( $i = 1, \dots, n; j = 1, \dots, m$ ), tada osnovna zavisnost dobija sledeći oblik:

$$y = \mathbf{X}b + e \quad (25)$$

sa regresionim koeficijentima  $\mathbf{b} = (a, b_1, \dots, b_m)^T$ . Opisani model šematski je prikazan na Slici 5. Za srednje centrirane vrednosti  $x$  (za koje je  $\bar{x} = 0, \bar{y} = 0$ , odnosno  $a = 0$ ) prva kolona u matrici  $\mathbf{X}$  može biti izostavljena.



Slika 5. Šematski prikaz višestrukog linearnog regresionog modela

Optimalne vrednosti regresionih koeficijenata se dobijaju za najmanju vrednost sume kvadrata reziduala,  $e^T e$ , što je postignuto kada vrednosti regresionih koeficijenata iznose:

$$b = (X^T X)^{-1} X^T y \quad (26)$$

a fitovane  $y$  vrednosti:

$$\hat{y} = Xb \quad (27)$$

Regresioni koeficijenti (uključujući i odsečak) daju vektor  $b$  koji se može smatrati vektorom koeficijenata latentnih varijabli u  $x$ -prostoru, odnosno to je latentna varijabla koja daje maksimalnu vrednost Pearson-ovog korelacionog koeficijenta zavisnosti između vrednosti  $y$  i  $\hat{y}$ . Neophodno je naglasiti da se za izračunavanje vrednosti regresionih koeficijenata koristi inverzna matrica  $X^T X$ , koja je povezana sa korelacionom matricom  $x$ -varijabli, što dovodi do problema kada se posmatraju visoko korelisane  $x$ -varijable. U slučaju približno kolinearnih varijabli, inverzna matrica bi postala nestabilna ili bi je bilo nemoguće izračunati. To bi se isto desilo i kada bi broj nezavisno promenljivih bio veći od broja objekata ( $m > n$ ), zbog čega ne bi bilo moguće predvideti  $y$ -vrednosti. Ovaj problem bi mogao biti rešen smanjenjem broja varijabli u

regresionom modelu ili korišćenjem multivarijantnih metoda kao što su PCR ili PLS<sup>80,103</sup>.

Korišćenje svih raspoloživih varijabli dovodi do boljeg uklapanja modela u podatke seta za kalibraciju jer reziduali postaju manji, a koeficijent determinacije veći. Međutim, osnovni cilj linearnog modelovanja nije povećanje uklapanja u podatke seta za kalibraciju, već povećanje prediktivnih svojstava test seta. Takođe, regresioni model sa velikim brojem varijabli je gotovo nemoguće interpretirati. Metoda MLR primenjena na set odabranih varijabli je, stoga, podjednako značajna i primenjivana kao i PCR i PLS metoda.

Da bi se smanjio broj nezavisno promenljivih potrebno je pronaći kriterijum za odbacivanje onih varijabli koje u najmanjoj meri doprinose  $y$ -vrednostima, odnosno zadržati varijable koje nose najveći deo informacija sadržanih u polaznom setu. Pri tome bi trebalo ispuniti uslov da regresioni model bude realan i da u njega budu uključene sve varijable koje imaju uticaj na njega, ali da pri tome njihov broj bude što manji<sup>107,108</sup>.

U okviru različitih statističkih programa dostupan je veliki broj metoda za selekciju varijabli. Najčešće korišćene metode, koje su istovremeno sastavni deo NCSS programa korišćenog u ovom radu su<sup>80</sup>:

- Selekcija unapred (Forward selection), tj. metoda kod koje je prvi model bez promenljivih, a prva varijabla koja se prihvata je ona koja daje model sa najvećom vrednošću parametra  $R^2$ . Svaka naredna varijabla doprinosi povećanju pomenute veličine, pri čemu se proces prekida uključivanjem promenljive koja ne utiče značajno na vrednost  $R^2$ . Varijabla koja se jednom nađe u modelu više ne može biti izostavljena iz njega.
- Selekcija unazad (Backward elimination) koja započinje modelom sa uključenim svim varijablama, nakon čega se iz modela postepeno eliminišu varijable sa najmanjim doprinosom vrednosti *PRESS*.
- Postepena regresija (Stepwise regression), tj. postupak koji kombinuje prethodne dve metode. Ona je u osnovi selekcija unapred, ali u svakom stupnju postoji mogućnost eliminacije promenljive iz modela. Potrebno je odabrati dva nivoa

značajnosti na kojima se radi, jedan niži, kao kriterijum za uključivanje novih varijabli, i drugi, za njihovo isključivanje. Broj varijabli koji se zadržava u modelu zasniva se na ta dva nivoa značajnosti.

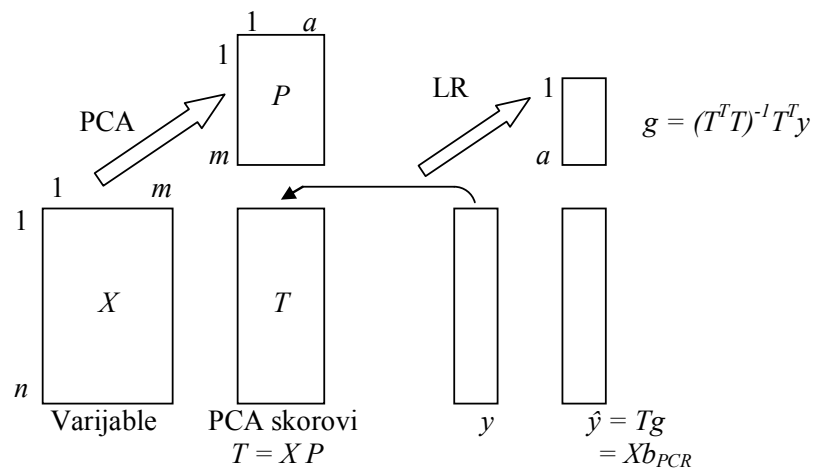
- Regresija svih promenljivih (All possible regression). To je metoda za selekciju varijabli kojom se proveravaju sve korelacije sa jednom, dve, ili više varijabli. Kriterijum za selekciju se beleži za svaku regresiju na osnovu čega je moguće dobiti najbolji deskriptor za jednoparametrijsku ili set deskriptora za višeparametrijsku jednačinu.
- Multivarijantna selekcija varijabli (Multivariate variable selection), tj. metoda kojom se u okviru datog seta varijabli traži manji set sa maksimalnom vrednošću parametra  $R^2$ . Nakon odabira prve varijable, program iz preostalog seta bira drugu, pri čemu istovremeno proverava da li bi neka druga promenljiva mogla da zameni prvoizabranu. Ukoliko takva promenljiva postoji, ona ulazi u model, a lošija varijabla biva uklonjena. Potom se pokreće novo pretraživanje među preostalim promenljivim sve dok zamena varijabli dovodi do povećanja vrednosti  $R^2$ .

Pre građenja višestrukog linearnog regresionog modela neophodno je obraditi polazni set nezavisno promenljivih. Deskriptori koji nisu dostupni za sve ispitivane objekte se odbacuju. Uklanjaju se, takođe, i varijable koje imaju konstantne vrednosti za sve objekte, kao i one koje su međusobno visoko korelisane. Nakon pomenutih provera može se primeniti neki od postupaka za selekciju varijabli. Nezavisno promenljive koje su rezultat prethodnih operacija koriste se za pravljenje MLR modela.

MLR je najčešće primenjivan način regresione analize u QSRR studijama<sup>85,109-111</sup>. Upoređujući modele dobijene MLR i PLS metodama primenjenim na istom setu podataka, ustanovljeno je da su statistički parametri koji opisuju kvalitet dobijenih modela uporedivi, kao i da modele određuju deskriptori slične prirode, sa približno istim uticajem na zavisno promenljivu<sup>85,112-114</sup>.

### 2.3.4.3 Regresija glavnih komponentata

Selekcija varijabli je jedan od načina za redukciju broja regresionih varijabli, kao i isključivanje varijabli koje su međusobne kolinearne. Na ovaj način dobija se regresioni model koji je lako objašnjiv, ali u slučaju velikog broja  $x$ -varijabli konstruisanje ovakvog modela zahteva mnogo vremena i truda. PCR predstavlja multivarijantnu metodu koja, takođe, rešava problem kolinearnosti i redukcije broja nezavisno promenljivih, ali regresione varijable ne predstavljaju više originalne  $x$ -vrednosti već njihovu linearnu kombinaciju, odnosno skorove glavnih komponentata  $x$ -varijabli. Na taj način, PCR je kombinacija PCA i MLR metode<sup>80,103</sup>. Metode PCR šematski je prikazana na Slici 6.



Slika 6. Šematski prikaz metode regresije glavnih komponentata

PCA razlaže centriranu matricu  $X$  na skorove  $T$  i koeficijente latentnih varijabli  $P$ . Za određeni broj  $a$  glavnih komponentata, koji je uglavnom manji od broja varijabli u originalnoj  $X$ -matrici, pomenuto razlaganje se može predstaviti sledećom zavisnošću:

$$X = TP^T + E \quad (28)$$

sa greškom izraženom kroz matricu  $E$ . Prema MLR modelu  $y = Xb + e$ , i zamenom matrice  $X$  matricom skorova  $T$  dobija se sledeći regresioni model:

$$y = Xb + e = (TP^T)b + e_T = Tg + e_T \quad (29)$$

sa novim regresionim koeficijentom  $g = P^T b$  i greškom  $e_T$ . Visoko korelisane  $x$ -varijable su na ovaj način svedene na nekoliko vektora skorova koji međusobno nisu korelisani. Kompleksnost PCR regresionog modela može biti optimizovana korišćenjem određenog broja glavnih komponentata. Regresioni koeficijenti, saglasno tekstu iz prethodnih poglavlja, dati su sledećim izrazom:

$$g = (T^T T)^{-1} T^T y \quad (30)$$

pa su konačni regresioni koeficijenti

$$b_{PCR} = Pg \quad (31)$$

PCR je alternativa PLS metodi. U najvećem broju slučajeva njihove performanse su veoma slične<sup>55,115</sup>. PCR uglavnom rezultuje modelom sa većim brojem komponentata u odnosu na PLS jer se za izračunavanje PCA skorova ne koriste informacije sadržane u  $y$ -vrednostima. Ovo, međutim, ne mora uvek biti nedostatak jer se na taj način koristi veći deo varijanse sadržane u matrici  $X$ , što može učiniti model stabilnijim.

Optimalan broj glavnih komponentata za PCR model bira se tako da one na najbolji način predviđaju  $y$ -varijablu, pri čemu procenat objašnjene varijanse u  $x$ -vrednostima nije od primarnog značaja. Drugim rečima, traži se optimalan podset određenog broja glavnih komponentata, a ne prvih nekoliko komponentata sa najvećom varijansom, jer je vrlo često neka dalja glavna komponenta pogodna za predviđanje  $y$ -varijable. Odabir dobrog seta PCA skorova vrši se na jedan od sledeća dva načina<sup>74</sup>:

- a) selekcija prvih nekoliko PCA komponentata koje obuhvataju određeni procenat ukupne varijanse podataka  $X$  (npr. 95%);
- b) selekcija PCA komponentata koje pokazuju maksimalnu korelaciju sa  $y$ .



Prediktivna sposobnost finalnog modela proverava se primenom neke od metoda unakrsne validacije.

#### 2.3.4.4 Regresija metodom delimičnih-najmanjih kvadrata

PLS je jedna od najčešće korišćenih metoda za multivarijantnu kalibraciju koja povezuje matricu  $X$  sa vektorom  $y$  ili matricom  $Y$ , tj. više zavisno promenljivih<sup>55,116-118</sup>.

Po svojoj strukturi PLS i PCR modeli su praktično isti; vrednosti nezavisno promenljivih ( $x$ -podaci), se prvo prevedu u set od nekoliko intermedijernih linearnih latentnih varijabli (komponentata), koje se potom koriste za regresiju (MLR) sa zavisno promenljivom  $y$ . PCR koristi skorove glavnih komponentata (dobijenih isključivo iz matrice  $X$ ) kao komponente, dok PLS koristi komponente koje su povezane sa  $y$ . Kriterijum koji se najčešće koristi u PLS za dobijanje intermedijarnih latentnih varijabli je maksimum kovarijanse između skorova i vrednosti  $y$ . Kovarijanse kombinuje veći deo varijanse u matrici  $X$  i velike vrednosti korelacionog koeficijenta sa posmatranom karakteristikom, zbog čega se PLS smatra kompromisom između PCR (maksimum varijanse u modelovanju  $X$ ) i MLR (maksimum korelacije u modelovanju  $y$ ) metoda. PLS i PCR su linearne metode, odnosno konačne latentne varijable koje predviđaju modelovanu karakteristiku,  $y$ , predstavljaju linearnu kombinaciju originalnih varijabli, kao i u MLR. Međutim, regresioni koeficijenti u konačnim MLR, PCR i PLS modelima su različiti, kao i njihova prediktivna svojstva.

PLS se najčešće koristi za regresiju  $x$ -varijabli sa jednom  $y$ -varijablom na sledeći način<sup>80,93,103</sup>:

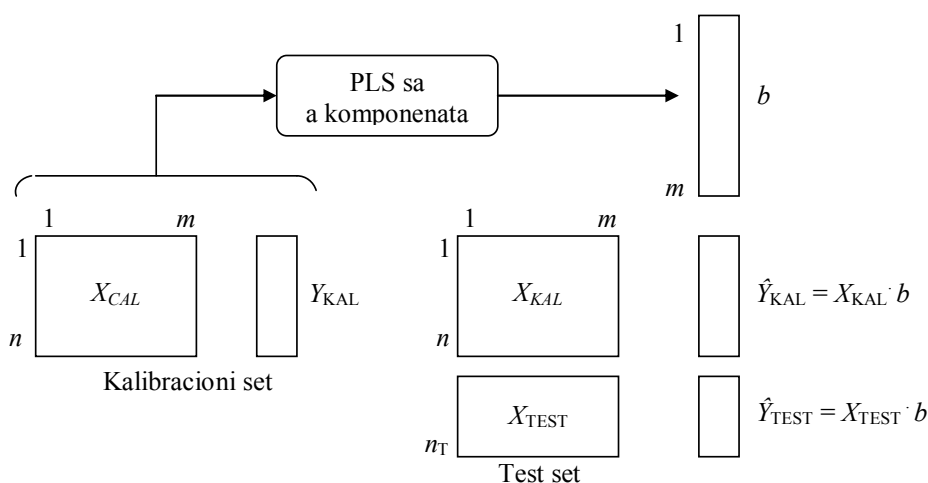
1. Prvo se izračunava PLS komponenta kao latentna varijabla koja pokazuje maksimum kovarijanse između skorova i modelovane karakteristike  $y$ . Ovde je potrebno naglasiti da je kriterijum „kovarijanse” zapravo kompromis između maksimalne vrednosti korelacionog koeficijenta (MLR) i maksimalne varijanse (PCA).
2. Informacije (varijansa) prve komponente se uklanjaju iz  $x$ -podataka. Pomenuti proces predstavlja projekciju  $x$ -prostora na ravan koja je

normalna na pravac prve komponente. Rezultujuća *matrica odstupanja* (residual matrix,  $X_{RES}$ ) ima isti broj varijabli kao i originalna matrica, ali je unutrašnja dimenzionalnost umanjena za jedan.

3. Iz matriksa odstupanja odabira se sledeća PLS komponenta, ponovo sa maksimalnom kovarijansom između skorova  $x$  i  $y$ .
4. Prethodno opisani postupak se ponavlja sve dok se uočavaju poboljšanja u modelovanju  $y$ . Broj PLS komponentata određuje kompleksnost modela. Takođe, optimalan broj komponentata određuje se pomoću unakrsne validacije.

Optimalan broj PLS komponentata je uglavnom manji od optimalnog broja PCA komponentata kod PCR metode, zbog činjenice da su PLS komponente razvijene kao latentne varijable visoko korelisane sa  $y$ . Sa povećanjem broja komponentata povećava se i sličnost između PCR i PLS modela.

Šematski prikaz PLS metode dat je na Slici 7.



Slika 7. Šematski prikaz regresije metodom delimičnih najmanjih kvadrata

Predikcija većeg broja različitih  $y$ -varijabli (matrica  $Y$ ) iz  $x$ -varijabli metodom delimičnih najmanjih kvadrata označava se kao PLS2. Višedimenzionalni  $x$ -prostor

projektuje se na manji broj PLS komponenata (sumirajući  $x$ -varijable), dok se višedimenzionalni  $y$ -prostor projektuje na manji broj PLS- $y$ -komponentata (sumirajući  $y$ -varijable). Komponente  $x$ - i  $y$ - su međusobno povezane matricom kovarijansi skorova<sup>80</sup>.

## 2.4 Arilpiperazini: struktura i biološka aktivnost

Nakon otkrića fenotiazina i zadovoljavajućih rezultata u terapijama psihotičkih oboljenja, razvijena je serija lekova sličnih svojstava, antagonista dopaminskih  $D_2$  receptora, koji su bili označeni kao „klasični antipsihotici”. Primena ovih lekova je, međutim, dovela do ozbiljnih propratnih efekata, a bili su i neefikasni protiv negativnih simptoma šizofrenije, zbog čega je razvijena druga generacija antipsihotika označena kao „netipični antipsihotici” ili „dopamin-serotonin sistem stabilizatori”. Ovi derivati poseduju kombinovani antagonistički efekat prema  $D_2$  receptorima i serotoninским  $5-HT_2$  receptorima, ili deluju kao antagonisti  $D_2$  receptora i interaguju sa  $5-HT_{1A}$  receptorima. Drugi farmakološki pristup se bazira na dizajnu i razvoju novih lekova koji po svojoj strukturi predstavljaju i  $D_2$  antagoniste i  $5-HT_{1A}$  agoniste<sup>119</sup>.

Dopaminski receptori su veoma značajni zbog svoje uloge u funkcionisanju velikog broja fizioloških procesa, kao što su koordinacija psihomotoričkih funkcija, organizacija procesa prepoznavanja i regulacija reproduktivnih ciklusa kod ljudi. Na molekularnom nivou, dopaminski receptori imaju značajnu ulogu u prenosu signala kroz ćelijsku membranu. Poremećaji na nivou dopaminergičkog sistema blisko su povezani sa nekim ozbiljnim neurološkim, psihičkim i endokrinološkim bolestima<sup>119</sup>.

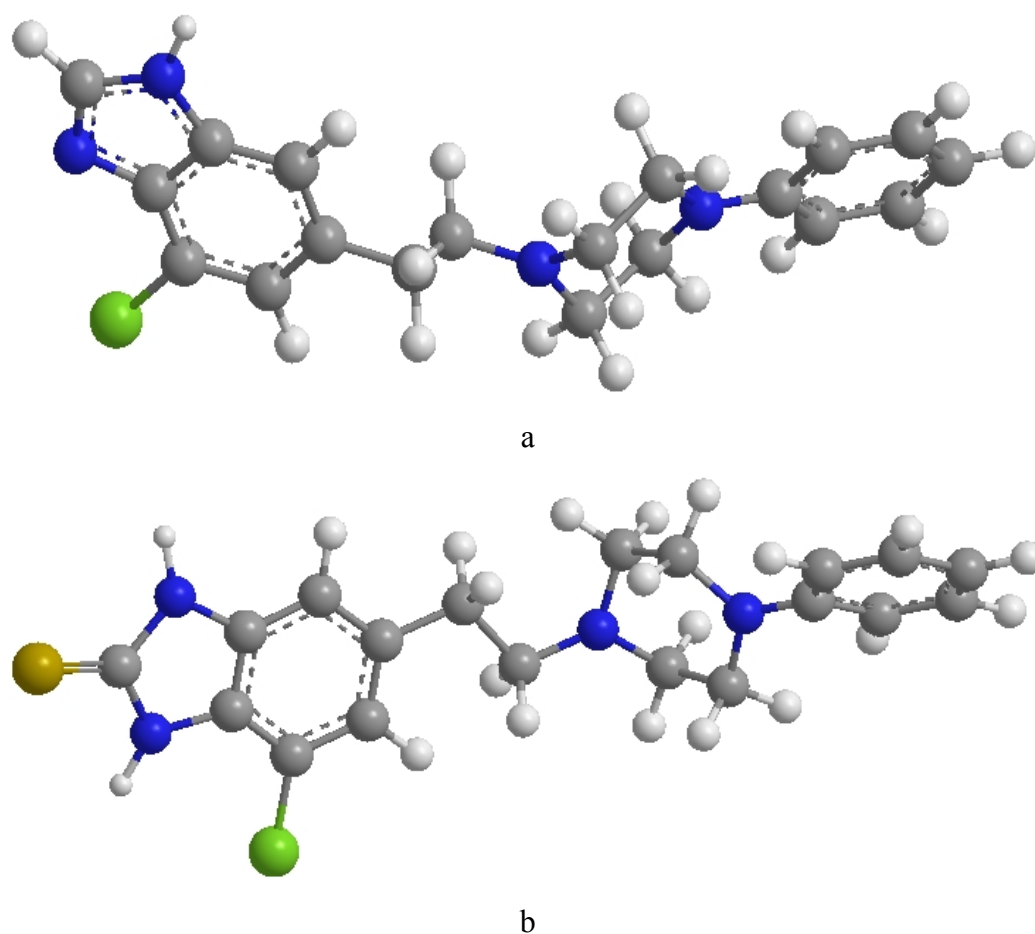
Na osnovu raspoloživih molekularnih modela dopaminskih receptora i objavljenih QSAR studija različitih dopaminergičkih liganada utvrđeno je da tri ključna farmakoforna strukturalna zahteva određuju afinitet dopaminskih agonista: aromatično jezgro, most koji obezbeđuje povoljnu udaljenost između aromatičnog dela i baznog azota i bazni azot<sup>120</sup>. U cilju dobijanja  $D_2$  dopaminergičkih liganada koji ujedno pokazuju i afinitet prema serotoninским receptorima, razvijena je serija arilpiperazinskih jedinjenja<sup>119</sup>. Pokazano je da se ligandi koji imaju visok afinitet prema  $D_2/5-HT_{1A}$  receptorima mogu dobiti povezivanjem arilpiperazina i 2-supstituisanog

benzimidazola preko fleksibilnog lanca<sup>120</sup>. Utvrđeno je da afinitet dobijenih liganada za vezivanje za D<sub>2</sub> receptore zavisi od strukture benzimidazolskog i arilpiperazinskog dela molekula, kao i od njihovog međusobnog rastojanja<sup>120</sup>. U tom smislu, sintetisana je serija 6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazola, kao i 1-{2-[5-(1*H*-benzimidazol-2-tion)]etil}-4-arilpiperazina i ispitivan uticaj različitih supstituenata na stabilnost ligand-D<sub>2</sub> kompleksa<sup>120,121</sup>. Interakcije dobijenih jedinjenja sa dopaminskim receptorom ostvaruju se preko jonskih interakcija između protonovanog N1 atoma piperazinskog prstena i Asp 86, kao i niza vodoničnih veza i elektronskih i sternih interakcija, koje u velikoj meri zavise od elektrostatičkog površinskog potencijala (raspodele elektronske gustine) u benzimidazolskom prstenu.

Oslanjajući se na strukturu prethodno navedenih jedinjenja kao polaznu tačku, kao i na činjenice da na raspodelu elektronske gustine utiče priroda supstituenta u heterocikličnom prstenu, uveden je atom halogena u benzimidazolsku farmakoforu, sa očekivanjem da će se glavna preraspodela elektronske gustine ostvariti posredstvom velikog elektron-privlačnog dejstva halogena<sup>122-124</sup>. Sintetisana je serija novih 4-halo-6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazola i 6-[2-(4-arilpiperazin-1-il)etil]-4-halo-1,3-dihidro-2*H*-benzimidazol-2-tion sa arilpiperazinskim delom molekula izabranim shodno poznatim zahtevima o odnosu strukture i reaktivnosti (Slika 8). Za sva novosintetisana jedinjenja određen je afinitet vezivanja za dopaminske i serotoninske receptore u *in vitro* eksperimentima kompeticije sa radio-ligandima. Kao izvor dopaminskih i serotoninskih receptora korišćene su sinaptozomalne membrane izolovane iz goveđeg nukleusa kaudatusa i hipokampusu.

Raspodela elektronske gustine u benzimidazolskom prstenu utiče na biološku aktivnost arilpiperazina. Efekat halogena na stabilnost kompleksa receptor-ligand može se objasniti transferom naelektrisanja ili vodoničnim vezivanjem halogena i receptora<sup>125,126</sup>. Za seriju 4-halo-6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazola utvrđeno je da halogenovanje povećava afinitet novih liganada prema oba tipa receptora u sledećem nizu: Br > Cl > H. Pokazano je da je najaktivniji ligand prema 5-HT<sub>1A</sub> receptoru 4-bromo-6-{2-[4-(2-hlorofenil)piperazin-1-il]etil}-1*H*-benzimidazol, čija je aktivnost povećana 1735 puta u odnosu na nehalogenovani analog. Sa druge strane, najaktivniji ligand prema D<sub>2</sub> receptoru bio je 4-bromo-6-{2-[4-(2-metoksifenil)piperazin-1-il]etil}-

1*H*-benzimidazol, kod kojeg je halogenovanje povećalo afinitet prema pomenutom receptoru 40 puta. Ovakvo povećanje afiniteta za vezivanje ukazuje na postojanje privlačnih interakcija između atoma halogena i aminokiselinskih ostataka aktivnog mesta vezivanja oba ispitivana receptora<sup>122</sup>. Doking analize na modelima D<sub>2</sub> receptora pokazale su da uvođenje halogena u benzimidazolski deo liganda dovodi do građenja vodoničnih veza i jonskih interakcija između atoma halogena i Ser 122. Slično tome, jonske interakcije su uočene i između halogenovanog liganda i Trp 358 kod 5-HT<sub>1A</sub> receptora.



Slika 8. Geometrijski optimizovane strukture: a) 4-hloro-6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazola, b) 6-[2-(4-arilpiperazin-1-il)etil]-4-halo-1,3-dihidro-2*H*-benzimidazol-2-tion

Za seriju 6-[2-(4-arilpiperazin-1-il)etil]-4-halo-1,3-dihidro-2*H*-benzimidazol-2-tion utvrđeno je, takođe, da raspodela elektronske gustine u benzimidazolskom prstenu prouzrokovana uvođenjem atoma halogena, kao i grupe koje se ponašaju kao akceptori vodonične veze (metoksi grupa) u orto položaju 1-arilpiperazinskog dela molekula, stabilizuju kompleks ligand-D<sub>2</sub> receptor. Slično, fenil, 2-hlorofenil i 3-(fluorometil)fenil derivati pokazuju veći afinitet prema 5-HT<sub>1A</sub> receptorima<sup>123,124</sup>.

Veći afinitet halogenih derivata arilpiperazina prema D<sub>2</sub> i 5-HT<sub>1A</sub> receptorima ukazuje na potrebu određivanja njihove lipofilnosti, odnosno razumevanja hromatografskog ponašanja, u cilju rasvetljavanja interakcija koji se odigravaju na ćelijskoj membrani, odnosno aktivnom mestu receptora.

# **3 Eksperimentalni deo**

### 3. Eksperimentalni deo

#### 3.1 Reagensi

Sinteza i karakterizacija ispitivanih arilpiperazina objavljena je u ranijim radovima<sup>122-124</sup>. Strukture proučavanih jedinjenja prikazane su u Tabeli 2. Metanol, dioksan i dimetilsulfoksid, koji su korišćeni kao organske komponente mobilne faze, bili su analitičkog stepena čistoće i nabavljeni od Merck-a (Darmstadt, Nemačka).

#### 3.2 Tankoslojna hromatografija

Za hromatografska ispitivanja primenjena je tehnika horizontalne tankoslojne hromatografije. Hromatografski sistemi sastojali su se od komercijalno dostupnog RP-18 F<sub>254S</sub> silika-gela (Merck, Nemačka) i mobilnih faza: (1) metanol:voda (sa 80 – 100 Vol% metanola), (2) dioksan:voda (sa 60 – 80 Vol% dioksana), (3) dimetilsulfoksid (DMSO):voda (sa 80 – 100 Vol% DMSO-a). U svim navedenim slučajevima, sadržaj organske komponente u mobilnoj fazi je povećavan u porcijama od 5 Vol%.

Ispitivane supstance rastvarane su u smeši DMSO:acetone (1:1, v/v). Na ploče, dimenzija 10×10 cm, nanošeno je oko 1μl sveže pripremljenog rastvora arilpiperazina približne koncentracije 0,01 mol/dm<sup>3</sup>.

Pre razvijanja, ploče sa nanetim uzorcima su sušene na vazduhu tokom 2 sata. Pređeni put rastvarača iznosio je oko 4,5 cm. Detekcija hromatografskih zona vršena je pod UV-lampom (254 nm). Sva ispitivanja su izvedena na sobnoj temperaturi (22 ± 2 °C).



### 3.3 Geometrijska optimizacija i izračunavanje vrednosti molekulskih deskriptora

Geometrijska optimizacija i modelovanje molekula urađeni su pomoću programa *Hyperchem* (verzija 7,0, Hybercube, Inc.). Strukture ispitivanih arilpiperazina nacrtane su u *Hyperchem Release 7* programu. Prva optimizacija izvedena je pomoću molekulsko-mehaničke MM+ metode, koja je zasnovana na MM2 algoritmu. U sledećem koraku, optimizacija je urađena korišćenjem *Polak-Ribier Conjugate gradient (PRCG)* algoritma. Strukture optimizovane na prethodno opisan način služile su kao početne za semi-empirijsku optimizaciju. Krajnja geometrijska optimizacija se zasniva na *Austin Model 1 (AMI)* semi-empirijskoj molekulsko-orbitanoj metodi<sup>90</sup>, a urađena je primenom algoritma *Steepest Descent* i *Conjugate Directions* algoritma.

Set molekulskih deskriptora izračunat pomoću *Hyperchem* programa sastojao se od serije kvantno-hemijskih deskriptora kao što su: polarizabilitet (Pol), energija najviše popunjene molekulske orbitale (HOMO) i energija najniže slobodne elektronske orbitale (LUMO), dipolni momenat (DM), ukupni dipol (TD) i QSAR parametara, površine (SA) i refraktivnosti. Program *Molecular Modeling Pro Plus (MMP Plus)*<sup>127</sup> korišćen je za izračunavanje seta dodatnih deskriptora za opisivanje veličine i račvanja molekula, kao što su dužina molekula (ML), širina molekula (MWe), dubina molekula (MD); topoloških parametara, Randić, Hall i Kier indeksa konektivnosti 0 – 4, Randić, Hall i Kier indeksa valentnosti 0 – 4; parametara koji opisuju osobine polimera i surfaktanata (Van Krevelen-ovi i Hoftzyer-ovi parametri rastvorljivosti), površinski napon (ST), molarna zapremina (MV), hidrofilno-lipofilni balans (HLB), površinski polaritet (SP), disperzija (D), polarnost (P), vodonično vezivanje (HB), udeo hidrofilne površine (HSA), akceptor vodonične veze (HBA), donor vodonične veze (HBD), lipofilnost ( $\log P$ ). Lipofilnost ispitivanih jedinjenja izračunata je primenom programa koji je dostupan na internet stranici *Virtual Computational Chemistry Laboratory*<sup>128,129</sup>.

### 3.4 Multivarijantna statistička analiza i modelovanje

Regresija metodom najmanjih kvadrata i statistički test za proveru značajnosti korelacionog koeficijenta urađeni su primenom statističkog paketa *Data Analysis* u okviru Microsoft Office Excel-a (verzija 7.0).

Metode PCA, hijerarhijska klsterska analiza (HCA), PCR i PLS izvedene su korišćenjem PLS\_Toolbox statističkog paketa (Eigenvectors Inc. v. 5.7) u okviru MATLAB verzije 7.4.0.287 (R2007a) (MathWorks INC, Natick, MA, USA).

Za PCA je korišćen *Singular Value Decomposition (SVD)* algoritam. Ovo je široko primenjivana tehnika za razlaganje originalne matrice podataka u nekoliko novih matrica koje otkrivaju mnoga korisna svojstva originalne matrice. Prema SVD metodi matrica  $X$  (dimenzija  $n \times m$ ) može da se razloži na tri nove matrice:  $X = T_0 \times S \times P^T$ . Za matricu  $X$  sa podacima koji su centrirani u odnosu na srednju vrednost, matrica  $T_0$  ima dimenzije  $n \times m$  i sadrži PCA skorove normirane na dužinu 1.  $S$  je dijagonalna matrica veličine  $m \times m$  koja sadrži tzv. jedinične vrednosti u svojoj dijagonali, a koje su jednake standardnoj devijaciji skorova.  $P^T$  je transponovana matrica koeficijenata PCA latentnih varijabli dimenzija  $m \times m$ . PCA skorovi,  $T$ , izračunavaju se iz odnosa:  $T = T_0 \times S$ <sup>80</sup>. Vrednosti  $Q$  i  $T^2$  za Hotelling-ovu raspodelu određeni su na 95% nivo značajnosti i korišćeni za utvrđivanje prisustva rezultata koji su posledica grubih grešaka.

HCA je izvedena primenom aglomerativne metode uz korišćenje Ward-ovog modela za određivanje rastojanja između dva klastera i Euclidean-ove razdaljine za merenje udaljenosti između uzoraka<sup>80</sup>.

Kod PLS regresije primenjen je SIMPLS algoritam. On izračunava PLS faktore kao linearnu kombinaciju originalnih varijabli tako da se kriterijum kovarijantnosti učini što značajnijim, poštujući pri tome određena ograničenja koja se odnose na ortogonalnost i normiranje. SIMPLS algoritam se razlikuje od ostalih PLS algoritama po načinu ortogonalizacije PLS komponentata. Naime, SIMPLS algoritam opisuje skorove kao

direktnu kombinaciju originalne  $X$  matrice a ne korišćenjem osiromašene  $X$  matrice (deflated  $X$  matrix) kao što je to slučaj kod drugih algoritama. Drugim rečima, analiza svojstvenih vrednosti se izvodi na originalnoj  $XY$  matrici projektovanoj na vektore koeficijenata latentnih varijabli iz prethodnih komponenata, a ne projektovanjem osiromašene matrice. Algoritam je predložen od strane Jong-a i označen kao „Straightforward Implementation of a Statistically Inspired Modification of the PLS algorithm”<sup>80,130</sup>. Validacija PLS modela izvedena je korišćenjem metode slučajnih uzoraka.

MLR modelovanje je izvedeno korišćenjem demo verzije NCSS statističkog paketa (Hintze, J. (2001), NCSS and PASS Number Cruncher Statistical Systems, Kaysville, Utah). Za selekciju varijabli korišćena je metoda postepene regresije.

# 4

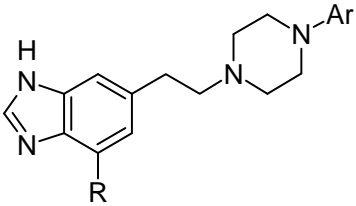
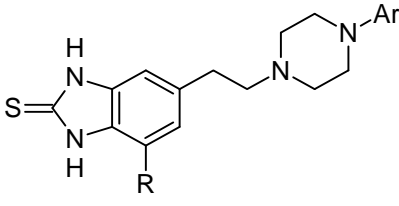
# Naši radovi

## 4.1 Retenciono ponašanje ispitivanih arilpiperazina u uslovima planarne hromatografije

U okviru ovog rada planirano je proučavanje hromatografskog ponašanja serije novosintetisanih arilpiperazina u uslovima planarne hromatografije, procena retencionih podataka na osnovu multivarijantnih statističkih metoda i utvrđivanje moguće veze između retencionih karakteristika i fizičko-hemijskih parametara ispitivanih molekula, u cilju dobijanja podataka značajnih za određivanje lipofilnosti jedinjenja, kao i razumevanje mehanizama razdvajanja u primenjenim hromatografskim sistemima. Ispitivanje hromatografskog ponašanja arilpiperazina složene strukture, uslovljene prisustvom hidrofilnih i hidrofobnih grupa (Tabela 2), trebalo bi da obezbedi korisne informacije o njihovim farmakološki značajnim fizičko-hemijskim parametrima.

Ispitivani molekuli poseduju tri zajednička strukturalna dela: benzimidazolski ili benzimidazol-2-tionski deo, arilpiperazinski deo i fleksibilni etilenski lanac koji čini vezu između prva dva dela. Svaki od navedenih delova učestvuje u nekoliko vrsta interakcija. Građenje vodonične veze na benzimidazolskom delu molekula može biti ostvareno preko aromatične –NH grupe koja se ponaša kao proton donorski centar čineći proton akceptorske osobine molekulskog okruženja još izraženijim, kao i preko –S atoma iz tio grupe benzimidazol-2-tionskog derivata, koji ima izražena proton akceptorska svojstva. Prisustvo jako polarizovanog  $\pi$  elektronskog sistema u arilpiperazinskom i benzimidazolskom delu molekula uslovljava dipolarne interakcije između ispitivanih molekula i stacionarne, odnosno mobilne faze. Etilenski most obezbeđuje odgovarajuću udaljenost arilpiperazinskog i heterocikličnog dela molekula, sprečavajući da sterne smetnje onemoguće prethodno navedene interakcije.

Tabela 2. Strukture ispitivanih arilpiperazina

					
Jed.	R	Ar	Jed.	R	Ar
1	Cl	Phe	13	Cl	Phe
2	Cl	2-MeOPhe	14	Cl	2-MeOPhe
3	Cl	2-ClPhe	15	Cl	2-ClPhe
4	Cl	pirimidin-2-il	16	Cl	3-CF <sub>3</sub> Phe
5	Cl	2-NO <sub>2</sub> Phe	17	Cl	pirimidin-2-il
6	Br	Phe	18	Cl	2-NO <sub>2</sub> Phe
7	Br	2-MeOPhe	19	Br	Phe
8	Br	2-ClPhe	20	Br	2-MeOPhe
9	Br	3-CF <sub>3</sub> Phe	21	Br	2-ClPhe
10	Br	pirimidin-2-il	22	Br	3-CF <sub>3</sub> Phe
11	Br	2-NO <sub>2</sub> Phe	23	Br	pirimidin-2-il
12	H	2-NO <sub>2</sub> Phe	24	Br	2-NO <sub>2</sub> Phe
			25	H	Phe
			26	H	2,3-diMePhe
			27	H	naftil
			28	H	2-MeOPhe
			29	H	4-MeOPhe
			30	H	4-ClPhe
			31	H	4-FPhe
			32	H	2-NO <sub>2</sub> Phe
			33	H	3-NO <sub>2</sub> Phe

Za proučavanje hromatografskog ponašanja ispitivanih molekula korišćena su tri organska rastvarača različite polarnosti, 1,4-dioksan (nepolaran rastvarač), dimetilsulfoksid (DMSO) (polaran neprotičan rastvarač) i metanol (polaran protični rastvarač). U Tabeli 3 prikazane su vrednosti dielektrične konstante, dipolnog momenta i Hansen-ovih parametara rastvorljivosti ( $\delta_D$  – energija disperzionih veza,  $\delta_P$  – energija dipolarnih međumolekulskih sila i  $\delta_H$  – energija vodoničnih veza između molekula), za korišćene organske rastvarače.

Polarne rastvarače karakterišu više vrednosti  $\delta P$  i dielektrične konstante, pri čemu polarni protični rastvarači imaju izraženije vrednosti  $\delta H$ , dok polarni neprotični rastvarači imaju više vrednosti dipolnih momenata. Metanol sa molekulima ispitivane supstance gradi vodonične veze, dok DMSO, zahvaljujući visokoj vrednosti dipolnog momenta (razdvajanje delimično pozitivnog i delimično negativnog naelektrisanja unutar istog molekula), interaguje sa pozitivno naelektrisanim česticama preko negativnog dela dipola. Dioksan se klasifikuje kao nepolaran rastvarač zbog jako niske vrednosti dielektrične konstante, ali zahvaljujući kiseoniku koji može da se ponaša kao Lewis-ova baza on je sposoban da solvatizuje pojedine molekule.

Tabela 3. Karakteristike primenjenih rastvarača

Rastvarač	Dielektrična konstanta	Dipolni moment (D)	$\delta D$	$\delta P$	$\delta H$
1,4-Dioksan	2,3	0,45	17,5	1,8	9,0
Dimetilsulfoksid	46,7	3,96	18,4	16,4	10,2
Metanol	33,0	1,70	14,7	12,3	22,3

$\delta D$  – energija disperzionih veza;  $\delta P$  – energija dipolarnih međumolekulskih sila;  $\delta H$  – energija vodoničnih veza između molekula

Rezultati hromatografskog određivanja dati su kao  $R_F$ , odnosno  $R_M$  -vrednosti, za svaki hromatografski sistem pojedinačno, u Prilogu (Tabele P1 – P3). U reverzno-faznim sistemima, u kojima je solvatacioni efekat veoma značajan, retencija je određena molekulskom strukturom jedinjenja. Uvođenje malih polarnih ili nepolarnih grupa u strukturu velikih molekula značajno utiče na njihovo hromatografsko ponašanje. Retencija, koja se ostvaruje nespecifičnim interakcijama ispitivane supstance i stacionarne faze (disperzione i indukcione sile, odnosno hidrofobne interakcije), kao i specifičnim interakcijama supstanci sa polarnom mobilnom fazom, jača je ukoliko su ispitivane supstance manje polarne.

Ispitivana jedinjenja podeljena su u dve grupe u zavisnosti od prisustva tio-grupe u benzimidazolskom delu molekula. Značajno različito hromatografsko ponašanje derivata benzimidazol-2-tiona, usled prisustva dodatnog –S atoma u molekulu, uočeno je samo u sistemu RP-18/metanol:voda. Slabija retencija derivata benzimidazol-2-tiona u navedenom hromatografskom sistemu verovatno je posledica građenja vodonične veze sa hidroksilnom grupom metanola. Razlike u hromatografskom ponašanju ispitivanih jedinjenja unutar svake od grupa potiču od prisustva atoma halogena (Cl, Br) vezanog za benzimidazolsku i benzimidazol-2-tionsku dopaminergičku farmakoforu. Uvođenje voluminoznog supstituenta bogatog elektronima pojačava retenciju ovih jedinjenja u odnosu na analoge koji u istom položaju sadrže vodonik. Takođe, jedinjenja sa bromom u heterocikličnom delu molekula se najjače zadržavaju na stacionarnoj fazi u uslovima reverzno-fazne hromatografije. Treći faktor koji utiče na različito ponašanje analiziranih molekula je priroda supstituenta vezanog za arilpiperazinski deo molekula (Phe, 2-MeOPhe, 4-MeOPhe, 2-ClPhe, 4-ClPhe, 4-FPhe, 3-CF<sub>3</sub>Phe, pirimidin-2-il, 2-NO<sub>2</sub>Phe, 3-NO<sub>2</sub>Phe, 2,3-diMePhe, naftil). Jedinjenja sa hidrofobnom grupom (2,3-diMePhe, naftil, 3-CF<sub>3</sub>Phe) pokazuju najjaču retenciju, dok se uvođenjem metoksi grupe u molekul ili zamenom benzenovog prstena pirimidinskim povećava rastvorljivost u mobilnoj fazi usled čega slabe interakcije sa stacionarnom fazom.

Svojstva organske komponente mobilne faze u značajnoj meri utiču na retenciju ispitivanih jedinjenja. Proton akceptorske osobine, izražene u sistemima DMSO:voda i dioksan:voda, ostvaruju se između –NH grupe benzimidazolskog i benzimidazol-2-tionskog prstena i parcijalno negativno naelektrisanog kiseonika iz DMSO-a, tj. slobodnog elektronskog para kiseonika iz dioksana. Proton donorska svojstva mobilne faze, izražene u sistemu metanol:voda, pokazuju značajan efekat samo u slučaju *o*- i *p*-metoksi derivata, zbog sposobnosti metoksi grupe da se ponaša kao jak proton akceptor. U svim ostalim slučajevima ovaj efekat je zanemarljiv jer u molekulu, pored azota u piperazinskom delu, koji je mnogo slabiji proton akceptor od metoksi grupe, ne postoji drugi dobar proton akceptorski centar. Stoga, retencija metoksi derivata je manja u odnosu na ostale analoge iz ispitivane serije. U slučaju derivata benzimidazol-2-tiona pomenute interakcije su manje



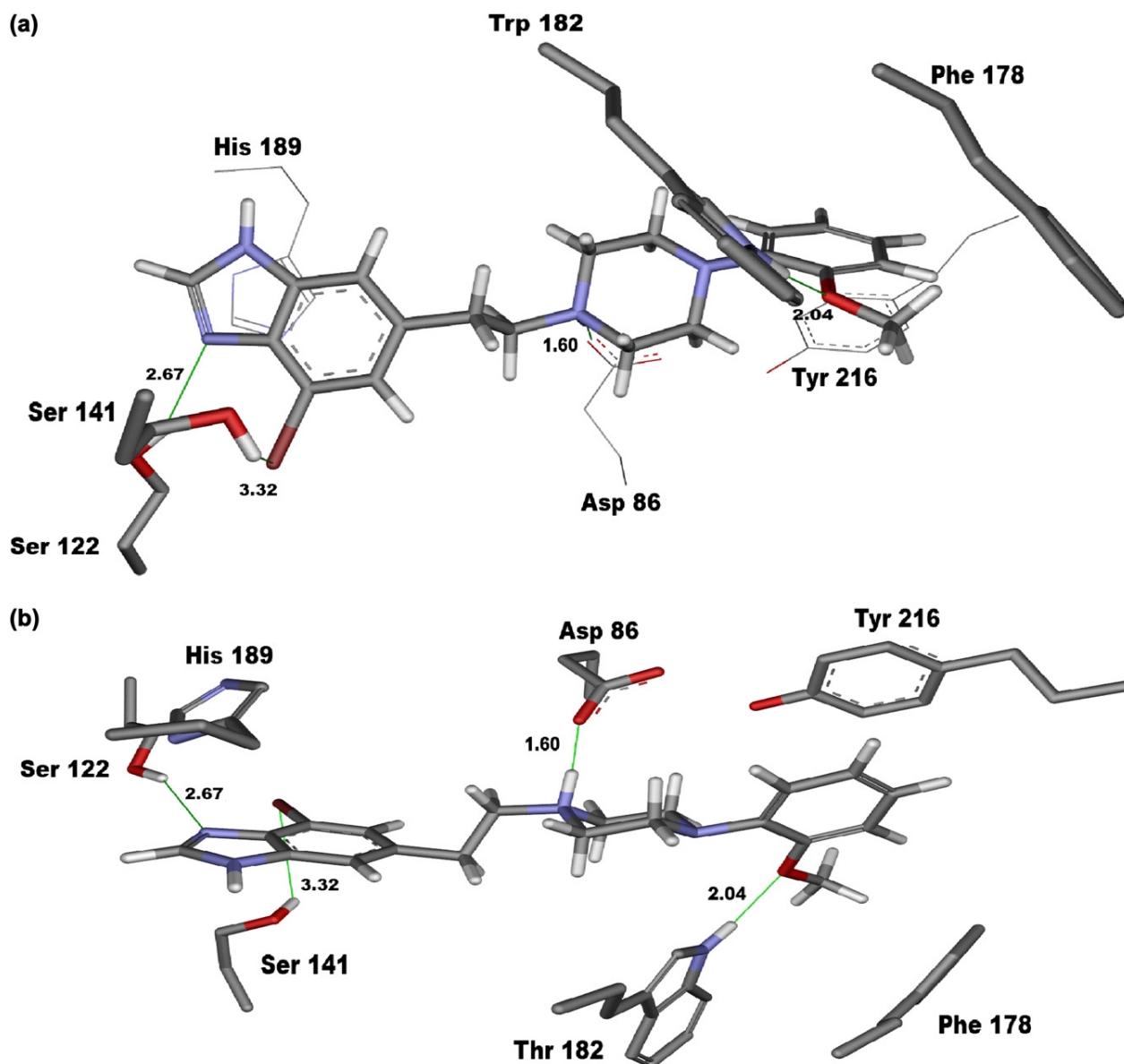
moгуće od interakcija koje vodonik iz hidroksilne grupe metanola ostvaruje sa sumporom iz tio-grupe, pa je pomenuta razlika u hromatografskom ponašanju analoga manje izražena.

Sposobnost metanola da jako interaguje sa molekulima vode značajno utiče na ponašanje jedinjenja u hromatografskim sistemima koji kao organsku komponentu mobilne faze sadrže pomenuti rastvarač<sup>131</sup>. Vodeni rastvor metanola sadrži sam metanol, vodu, kao i metanol vezan za vodu, zbog čega se ovaj rastvarač može smatrati pre tercijarnim sistemom nego binarnom mobilnom fazom. Metanol-voda asocijacija umanjuje efekat organskog modifikatora uslovljavajući na taj način slabljenje retencije. Sa druge strane, DMSO i dioksan su u manjoj meri asosovani sa molekulima vode. Veće vrednosti retencionih parametara dobijene za sisteme RP-18/dioksan:voda mogu se objasniti fizičko-hemijskim karakteristikama ovog cikličnog etra. Zbog svoje hidrofobne strukture, dioksan snažno interaguje sa C-18 lancima stacionarne faze stvarajući monosloj na njenoj površini sa polarnim delom molekula okrenutim ka analitima. Na ovaj način se u velikoj meri smanjuje hidrofobnost originalne stacionarne faze i kao posledica toga, retencija ispitivanih supstanci.

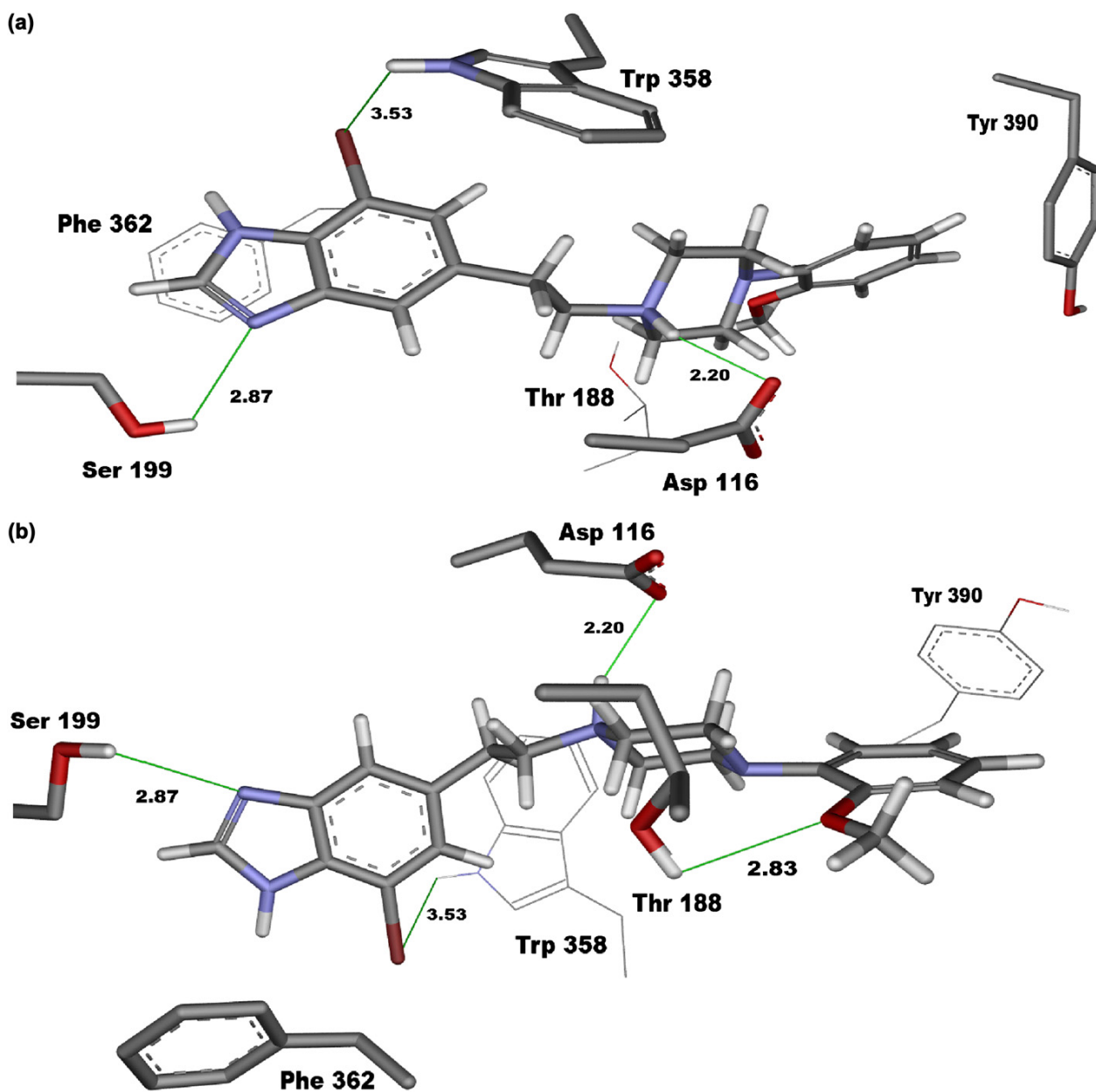
Dobijeni rezultati u saglasnosti su sa strukturnim svojstvima ispitivanih supstanci kao i sa interakcijama za koje je potvrđeno da se uspostavljaju kada se ligand koji pokazuje afinitet prema dopaminskim D2 i serotoninским 5-HT<sub>1A</sub> receptorima približi aktivnom centru receptora<sup>122</sup>. Osnovne interakcije između liganda i D2 receptora su sledeće: a) vodonično vezivanje između benzimidazolskog dela molekula i Ser 141 i Ser 122, koje se ostvaruje između –NH grupe liganda i halogena i amino kiselinskog ostatka; b) hidrofobnih interakcija između aromatičnog prstena arilpiperazinskog ili benzimidazolskog dela molekula i ostataka Phe 178, Tyr 216, Trp 182 i His 189; c) dodatne vodonične veze koje metoksi derivati (koji se ponašaju kao proton akceptori) mogu da grade sa Trp 182 (koji se ponaša kao proton donor); d) jonske interakcije između protonovanog N1 atoma piperazinskog prstena i Asp 86 (Slika 9).

Slično, osnovne interakcije između liganda i 5-HT<sub>1A</sub> receptora su: a) jonske interakcije između protonovanog N1 atoma piperazinskog prstena i negativno naelektrisanog Asp 116, b) vodonično vezivanje između supstituenata prisutnih na

benzimidazolskom delu molekula i Ser 199 i Trp 358, c) hidrofobnih interakcija između aromatičnog prstena arilpiperazinskog dela molekula i ostataka Phe 361 i Tyr 390 (Slika 10).



Slika 9. Prikaz osnovnih interakcija između jedinjenja 7 i D<sub>2</sub>-receptora posmatrano iz različitih uglova (Slika 9a i 9b).



Slika 10. Prikaz osnovnih interakcija između jedinjenja 7 i 5HT<sub>1A</sub>-receptora posmatrano iz različitih uglova (Slika 10a i 10b).

## 4.2 Određivanje parametara lipofilnosti ispitivanih arilpiperazina

Hromatografija je široko korišćena metoda za određivanje lipofilnosti, zbog odnosa koji postoji između odgovarajućeg retencionog parametra i particionog koeficijenta. Za uspešno hromatografsko merenje lipofilnosti mora biti utvrđena linearna zavisnost između retencionog parametra i koncentracije organske komponente mobilne faze.

Dobijeni rezultati (Tabele P1-P3) pokazuju da retencija ispitivanih jedinjenja opada sa povećanjem koncentracije organske komponente u mobilnoj fazi u svim primenjenim hromatografskim sistemima. Istovremeno, utvrđena je linearna zavisnost između  $R_M$  vrednosti, dobijenih na osnovu jednačine 2, i koncentracije organske komponente mobilne faze. Vrednosti  $R_M^0$  dobijene su linearnom ekstrapolacijom  $R_M$  vrednosti na čistu vodu. Vrednosti nagiba ( $b$ ) i odsečka (hromatografski određen parametar lipofilnosti,  $R_M^0$ ), kao i odgovarajući statistički parametri (Pearson-ov korelacioni koeficijent,  $r$ , standardna greška modela,  $s$  i nivo značajnosti,  $P$ ), utvrđene za prethodno pomenutu linearnu zavisnost navedeni su u Tabelama P4-P6 u Prilogu, za svaki primenjeni hromatografski sistem posebno.

Iz dobijenih rezultata se uočava da derivati benzimidazol-2-tiona pokazuju manju lipofilnost, odnosno niže  $R_M^0$  vrednosti, samo u slučaju metanola kao organskog modifikatora verovatno usled veće rastvorljivosti pomenutih jedinjenja u mobilnoj fazi. Uvođenje polarne metoksi grupe u molekul ima negativan uticaj na lipofilnost. Takođe, jedinjenja sa većim hidrofobnim delom u arilpiperazinskom delu molekula (2,3-diMePhe, naftil, 3-CF<sub>3</sub>Phe) su lipofilnija. Prisustvo atoma halogena u benzimidazolskom i benzimidazol-2-tionskom delu molekula rezultuje većim  $R_M^0$  vrednostima u poređenju sa njihovim nehalogenim analogima, usled povećanja elektronske gustine oko ovog dela molekula. Doking studije izvršene na ispitivanim arilpiperazinima pokazale su da 4-halo-derivati imaju veći afinitet prema D<sub>2</sub> i 5HT<sub>1A</sub> receptorima od njihovih nehalogenih analoga. U slučaju D<sub>2</sub> receptora između Cl, odnosno Br i Ser 122 dolazi do građenja vodoničnih veza i transfera naelektrisanja, dok se kod 5HT<sub>1A</sub> receptora transfer naelektrisanja vrši

između atoma halogena i Trp 358<sup>122</sup>. Na ovaj način je još jednom potvrđena mogućnost primene hromatografije za određivanje lipofilnosti ispitivanih arilpiperazina u cilju mapiranja procesa do kojih dolazi u biološkim sistemima.

Uticaj prirode organske komponente u mobilnoj fazi na  $R_M^0$  vrednosti analiziran je primenom neparametrijskog Kruskal-Wallis-ovog testa, kojim su upoređivane vrednosti medijana retencionih parametara analiziranih supstanci za tri primenjena hromatografska sistema. Kruskal-Wallis-ovim testom je potvrđeno da podaci ne pripadaju istoj populaciji, tj. da se medijane statistički značajno razlikuju ( $\chi^2 = 66,78$ ,  $\chi^2_{kr} = 5,99$ ). U cilju utvrđivanja seta podataka koji dovodi do odstupanja primenjena je Fisher-ova metoda za određivanje najmanje značajne razlike i utvrđeno da se  $R_M^0$  vrednosti sva tri seta statistički značajno razlikuju. Iako su Biagi i saradnici<sup>132,133</sup> ustanovili da su  $R_M^0$  vrednosti nezavisne od prirode organske komponente u mobilnoj fazi kada se u tu svrhu koriste aceton, acetonitril ili metanol, ovakav zaključak ne može biti prihvaćen generalno za svaki hromatografski proces<sup>13</sup>. Značajne razlike u  $R_M^0$  vrednostima posledica su jakih specifičnih interakcija između supstance i eluenta, tj. građenja vodoničnih veza koji utiču na rastvorljivost analita u mobilnoj fazi<sup>134</sup>. Zbog utvrđenih statistički značajnih razlika između  $R_M^0$  vrednosti analiziranih arilpiperazina, sledeći korak u analizi dobijenih rezultata sastojao se u odabiru hromatografskog sistema koji na najbolji način opisuje lipofilnost ispitivanih arilpiperazina.

Hromatografsko ponašanje jedinjenja određeno je nizom međumolekulskih interakcija između analita, stacionarne i mobilne faze. Linearna zavisnost između retencionog faktora,  $R_M^0$  vrednosti, i parametra lipofilnosti,  $\log P$ , može se očekivati saglasno činjenici da je retencija jedinjenja u reverzno-faznoj tečnoj hromatografiji određena hidrofobnim interakcijama. Određivanje linearne zavisnosti između parametra lipofilnosti dobijenog hromatografskim ispitivanjima i izračunatih  $\log P$  vrednosti je neizostavan korak u QSRR analizama. Statistički značajna korelacija između pomenutih veličina potvrđuje mogućnost primene RPTLC retencionog parametra za određivanje lipofilnosti. Za ispitivana jedinjenja izračunat je set  $\log P$  vrednosti i dobijeni rezultati prikazani u Tabeli P7 u Prilogu. Ekstrapolisane  $R_M^0$  vrednosti, za sva tri primenjena

hromatografska sistema, upoređene su sa izračunatim  $\log P$  vrednostima, a dobijeni statistički parametri ovih korelacija, prikazani u Tabeli 4.

Zbog nižih  $r$  vrednosti dobijenih za analizirane zavisnosti proverena je značajnost koeficijenta korelacije, imajući u vidu veliki broj tačaka korišćenih za kalibraciju. Statističkim testovima za proveru značajnosti linearne korelacije ( $t$ - i  $F$ -test) utvrđeno je da između  $R_M^0$  vrednosti, dobijenih za sistem sa dioksanom kao organskom komponentom mobilne faze, i predviđenih  $\log P$  vrednosti ne postoji linearna zavisnost. Vrednosti parametara  $t$  i  $F$ , zajedno sa odgovarajućim kritičnim vrednostima, navedene su u Tabeli 4. Stoga se  $R_M^0$  vrednosti, dobijene primenom dioksana, ne mogu koristiti kao mera lipofilnosti ispitivanih arilpiperazina.

Utvrđena je statistički značajna linearna zavisnost između  $R_M^0$  vrednosti, dobijenih za sisteme u kojima su korišćeni metanol i DMSO kao organski modifikatori, i izračunatih  $\log P$  vrednosti (Tabela 4). Na ovaj način, umesto zahtevnom i sporom metodom mućkanja, lipofilnost arilpiperazina je uspešno određena na osnovu retencionih konstanti, dobijenih primenom relativno brze i jednostavne tankoslojne hromatografije.

Prilikom tumačenja vrednosti  $t$  i  $F$  treba biti obazriv, jer se sa porastom broja merenja statistička značajnost može dokazati za relativno niske vrednosti Pearson-ovog koeficijenta iako za takve vrednosti bilo kakva kvantitativna zavisnost nema smisla. Zbog relativno velikog broja jedinjenja ( $n = 33$ ) u Tabeli 4 su date i vrednosti koeficijenta determinacije. Iako rezultati  $t$ - i  $F$ -testa pokazuju da se  $R_M^0$  vrednosti, dobijene primenom sistema metanol:voda, mogu upotrebiti kao parametar lipofilnosti ispitivanih jedinjenja, vrednosti koeficijenta determinacije i standardne greške, ukazuju da se pouzdaniji statistički parametri dobijaju za sistem sa DMSO-om. Na osnovu ovoga, kao i na osnovu najveće selektivnosti u razdvajanju ispitivanih arilpiperazina, kao najbolji sistem za određivanje lipofilnosti ispitivanih arilpiperazina odabran je onaj koji sadrži DMSO kao organski modifikator. Rezultati dobijeni primenom pomenutog hromatografskog sistema korišćeni su za dalja ispitivanja.

Mogućnost primene vrednosti nagiba regresione jednačine 3 kao parametra lipofilnosti, proverena je na osnovu korelacije između vrednosti odsečka,  $R_M^0$  i pomenutog parametra. Postojanje statistički značajne linearne zavisnosti potvrđeno je  $t$ - i  $F$ -testom:

$$b = -0,389(\pm 0,601) - 99,662(\pm 10,208)R_M^0$$

$$r = 0,8687 \quad s = 0,5414 \quad F = 95,31 \quad F_{kr(0,05;1,31)} = 4,16 \quad t = 9,76 \quad t_{kr(0,05;31)} = 2,04 \quad (32)$$

Dobijena statistički značajna zavisnost ukazuje da vrednost nagiba može biti posmatrana kao alternativa  $R_M^0$  vrednostima pri iskazivanju lipofilnosti ispitivanih jedinjenja. Vrednosti nagiba su dalje upoređivane sa izračunatim  $\log P$  vrednostima. Odgovarajući statistički parametri (koeficijent korelacije, standardna greška, koeficijent determinacije, 95% nivo značajnosti, vrednosti  $t$ - i  $F$ -testa) prikazani su u Tabeli 5. Na osnovu  $t$  i  $F$  vrednosti, ali i niskih vrednosti koeficijenta determinacije može se zaključiti da se nagib ne bi mogao koristiti kao parametar lipofilnosti.

Vrednosti parametra  $C_0$ , dobijenog na osnovu jednačine 6, za svaki od primenjenih hromatografskih sistema, navedene su u Tabelama P4-P6 u Prilogu. Između ovog parametra i vrednosti  $R_M^0$  nije utvrđena statistički značajna korelacija:

$$C_0 = 5,57(\pm 0,23) - 0,0079(\pm 0,0063)R_M^0$$

$$r = 0,2206 \quad s = 1,0660 \quad F = 1,58 \quad F_{kr(0,05;1,31)} = 4,16 \quad t = 1,26 \quad t_{kr(0,05;31)} = 2,04 \quad (33)$$

Razlog za ovako lošu korelaciju može biti činjenica da je sam nagib loša procena lipofilnosti, kako je navedeno u prethodnom izlaganju.

Vrednosti skorova prve glavne komponente, dobijenih primenom PCA metode na retencione podatke, ispitani su kao parametri za opisivanje lipofilnosti. Vrednosti  $PC1$ , za tri primenjena hromatografska sistema, date su u Tabelama P4-P6 u Prilogu. Statistički značajna zavisnost utvrđena je između vrednosti  $R_M^0$  i  $PC1$ :

$$PC1 = 5,40(\pm 0,11) - 0,393(\pm 0,052)R_M^0$$

$$r = 0,8066 \quad s = 0,6462 \quad F = 57,71 \quad F_{kr(0,05;1,31)} = 4,16 \quad t = 7,60 \quad t_{kr(0,05;31)} = 2,04 \quad (34)$$

Uzimajući u obzir visoku vrednost korelacionog koeficijenta i vrednosti  $t$  i  $F$  koje premašuju kritične vrednosti, parametar  $PCI$  korelisan je sa izračunatim  $\log P$  vrednostima, a dobijeni statistički parametri prikazani u Tabeli 6. Kao i u slučaju nagiba, dobijene su niske vrednosti koeficijenta determinacije, pa varijabla  $PC1$  nije dalje razmatrana kao mogući parametar lipofilnosti.

Primenom reverzno-fazne planarne hromatografije određena je lipofilnost ispitivanih arilpiperazina, jednog od važnih faktora koji utiču na biološku aktivnost supstanci. Određena su četiri hromatografska deskriptora,  $R_M^0$ ,  $b$ ,  $C_0$  i  $PC1$ , koji su u istoj meri zastupljeni u literaturi kao mera lipofilnosti. Utvrđeno je da varijabla koja direktno opisuje raspodelu supstanci između stacionarne i mobilne faze, kao što je parametar  $R_M^0$ , pouzdanije određuje lipofilnost ispitivanih arilpiperazina od parametra koji je proporcionalan hidrofobnoj površini molekula ( $b$ ) ili interpolisane veličine koja obuhvata sve hromatografske podatke jednog sistema ( $PC1$ ). Navedeni rezultati u saglasnosti su sa podacima dobijenim prilikom određivanja lipofilnosti nekih prirodnih i sintetisanih kumarina primenom RPTLC<sup>135</sup>, kao i određivanja particionog koeficijenta zemljište-voda ( $\log K_{OW}$ )<sup>64</sup>. U tom smislu, u QSRR analizama koje slede u daljem delu teksta, vrednosti  $R_M^0$ , dobijene za sistem RP-18/DMSO:voda, korišćene su kao parametar lipofilnosti koji je korelisan sa molekulskim deskriptorima.



Tabela 4. Statistički parametri linearne zavisnosti između eksperimentalno određene i izračunate vrednosti lipofilnosti za tri primenjena hromatografska sistema

$R_M^0 - f(\log P)$	Organska komponenta	Jednačina	$r$	$s$	$P$	$t^a$	$F^b$	$R^2$
$R_M^0 - \text{Alog } P_s$	Metanol	$R_M^0 = 1,469 (\pm 0,622) + 0,686 (\pm 0,187) \text{ Alog } P_s$ (35)	0,5496	0,6080	0,0009	3,66	13,42	0,3021
	Dioksan	$R_M^0 = 3,117 (\pm 0,764) + 0,207 (\pm 0,262) \text{ Alog } P_s$ (36)	0,1404	0,7206	0,4359	0,79	0,62	0,0197
	DMSO	$R_M^0 = 1,124 (\pm 0,499) + 0,480 (\pm 0,091) \text{ Alog } P_s$ (37)	0,6877	0,5284	<0,0001	5,27	27,82	0,4729
$R_M^0 - \text{AClog } P$	Metanol	$R_M^0 = 1,764 (\pm 0,573) + 0,581 (\pm 0,173) \text{ AClog } P$ (38)	0,5174	0,5602	0,0020	3,37	11,33	0,2677
	Dioksan	$R_M^0 = 1,958 (\pm 0,621) + 0,592 (\pm 0,213) \text{ AClog } P$ (39)	0,4475	0,5855	0,0090	2,79	7,76	0,2003
	DMSO	$R_M^0 = 1,359 (\pm 0,453) + 0,427 (\pm 0,082) \text{ AClog } P$ (40)	0,6810	0,4794	<0,0001	5,18	26,81	0,4638
$R_M^0 - \text{Alog } P$	Metanol	$R_M^0 = 1,761 (\pm 0,552) + 0,713 (\pm 0,166) \text{ Alog } P$ (41)	0,6096	0,5405	0,0002	4,28	18,33	0,3716
	Dioksan	$R_M^0 = 2,726 (\pm 0,679) + 0,474 (\pm 0,232) \text{ Alog } P$ (42)	0,3440	0,6400	0,0450	2,03	4,16	0,1183
	DMSO	$R_M^0 = 1,380 (\pm 0,411) + 0,503 (\pm 0,0751) \text{ Alog } P$ (43)	0,7691	0,4358	<0,0001	6,70	44,89	0,5915
$R_M^0 - \text{Mlog } P$	Metanol	$R_M^0 = 0,661 (\pm 0,309) + 0,730 (\pm 0,093) \text{ Mlog } P$ (44)	0,8150	0,3029	<0,0001	7,83	61,32	0,6642
	Dioksan	$R_M^0 = 2,405 (\pm 0,542) + 0,224 (\pm 0,186) \text{ Mlog } P$ (45)	0,2120	0,5108	0,2363	1,21	1,46	0,0449
	DMSO	$R_M^0 = 0,962 (\pm 0,314) + 0,387 (\pm 0,057) \text{ Mlog } P$ (46)	0,7725	0,3319	<0,0001	6,77	45,88	0,5968
$R_M^0 - \log P_{KOWWIN}$	Metanol	$R_M^0 = 1,377 (\pm 0,573) + 0,722 (\pm 0,173) \log P_{KOWWIN}$ (47)	0,6006	0,5609	0,0002	4,18	17,49	0,3607
	Dioksan	$R_M^0 = 2,544 (\pm 0,711) + 0,415 (\pm 0,244) \log P_{KOWWIN}$ (48)	0,2926	0,6708	0,0984	1,70	2,90	0,0856
	DMSO	$R_M^0 = 0,783 (\pm 0,384) + 0,548 (\pm 0,070) \log P_{KOWWIN}$ (49)	0,8151	0,4063	<0,0001	7,83	61,37	0,6644
$R_M^0 - \text{Xlog } P_2$	Metanol	$R_M^0 = -0,278 (\pm 0,550) + 1,0516 (\pm 0,166) \text{ Xlog } P_2$ (50)	0,7519	0,5377	<0,0001	6,35	40,32	0,5654
	Dioksan	$R_M^0 = 3,205 (\pm 0,865) - 0,015 (\pm 0,296) \text{ Xlog } P_2$ (51)	-0,0090	0,8156	0,9603	-0,05	0,00	0,0001
	DMSO	$R_M^0 = 0,139 (\pm 0,537) + 0,560 (\pm 0,097) \text{ Xlog } P_2$ (52)	0,7167	0,5687	<0,0001	5,72	32,74	0,5137
$R_M^0 - \text{Xlog } P_3$	Metanol	$R_M^0 = 0,778 (\pm 0,5108) + 0,881 (\pm 0,154) \text{ Xlog } P_3$ (53)	0,7170	0,4995	<0,0001	5,73	32,80	0,5141
	Dioksan	$R_M^0 = 3,297 (\pm 0,757) + 0,126 (\pm 0,259) \text{ Xlog } P_3$ (54)	0,0870	0,7140	0,6302	0,49	0,24	0,0076
	DMSO	$R_M^0 = 0,818 (\pm 0,435) + 0,527 (\pm 0,079) \text{ Xlog } P_3$ (55)	0,7669	0,4599	<0,0001	6,65	44,27	0,5881

$t_{kr}(\alpha=0,05; \nu=31) = 2,04$ ,  $F_{kr}(\alpha=0,05; \nu_1=1, \nu_2=31) = 4,16$

Tabela 5. Statistički parametri linearne zavisnosti između vrednosti nagiba i izračunatih  $\log P$  vrednosti za sistem RP-18/DMSO:voda

$b - f(\log P)$	Jednačina		$-r$	$s$	$R^2$	$P$	$t^a$	$F^b$
$b - \text{Alog } P_s$	$b = 1,318 (\pm 0,680) - 41,188 (\pm 11,556) \text{ Alog } P_s$	(56)	0,5392	0,6129	0,2907	0,0012	3,56	12,71
$b - \text{AClog } P$	$b = 1,486 (\pm 0,609) - 37,478 (\pm 10,345) \text{ AClog } P$	(57)	0,5454	0,5488	0,2975	0,0010	3,62	13,13
$b - \text{Alog } P$	$b = 1,417 (\pm 0,579) - 46,021 (\pm 9,843) \text{ Alog } P$	(58)	0,6431	0,5221	0,4136	<0,0001	4,68	21,86
$b - \text{Mlog } P$	$b = 0,974 (\pm 0,440) - 35,727 (\pm 7,478) \text{ Mlog } P$	(59)	0,6512	0,3967	0,4240	<0,0001	4,78	22,83
$b - \log P_{KOWWIN}$	$b = 0,947 (\pm 0,590) - 48,054 (\pm 10,020) \log P_{KOWWIN}$	(60)	0,6526	0,5315	0,4259	<0,0001	4,80	23,00
$b - \text{Xlog } P_2$	$b = 0,414 (\pm 0,754) - 47,288 (\pm 12,817) \text{ Xlog } P_2$	(61)	0,5523	0,6798	0,3050	0,0009	3,69	13,61
$b - \text{Xlog } P_3$	$b = 1,087 (\pm 0,643) - 44,278 (\pm 10,923) \text{ Xlog } P_3$	(62)	0,5886	0,5794	0,3464	0,0003	4,05	16,43

$t_{kr(\alpha=0,05; \nu=31)} = 2,04$ ;  $F_{kr(\alpha=0,05; \nu_1=1, \nu_2=31)} = 4,16$

Tabela 6. Statistički parametri linearne zavisnosti između vrednosti  $PCI$  i izračunatih  $\log P$  vrednosti za sistem RP-18/DMSO:voda

$PCI - f(\log P)$	Jednačina		$r$	$s$	$R^2$	$P$	$t^a$	$F^b$
$PCI - \text{Alog } P_s$	$PCI = 3,712 (\pm 0,114) + 0,142 (\pm 0,053) \text{ Alog } P_s$	(63)	0,4341	0,6557	0,1884	0,0116	2,68	7,20
$PCI - \text{AClog } P$	$PCI = 3,664 (\pm 0,100) + 0,141 (\pm 0,046) \text{ AClog } P$	(64)	0,4793	0,5746	0,2297	0,0048	3,04	9,25
$PCI - \text{Alog } P$	$PCI = 4,092 (\pm 0,097) + 0,178 (\pm 0,045) \text{ Alog } P$	(65)	0,5792	0,5558	0,3355	0,0004	3,96	15,65
$PCI - \text{Mlog } P$	$PCI = 3,050 (\pm 0,068) + 0,157 (\pm 0,031) \text{ Mlog } P$	(66)	0,6687	0,3886	0,4472	<0,0001	5,01	25,07
$PCI - \log P_{KOWWIN}$	$PCI = 3,740 (\pm 0,098) + 0,188 (\pm 0,045) \log P_{KOWWIN}$	(67)	0,5965	0,5630	0,3558	0,0002	4,14	17,12
$PCI - \text{Xlog } P_2$	$PCI = 3,162 (\pm 0,120) + 0,195 (\pm 0,056) \text{ Xlog } P_2$	(68)	0,5325	0,6903	0,2836	0,0014	3,50	12,27
$PCI - \text{Xlog } P_3$	$PCI = 3,660 (\pm 0,102) + 0,185 (\pm 0,047) \text{ Xlog } P_3$	(69)	0,5749	0,5864	0,3305	0,0005	3,91	15,30

$t_{kr(\alpha=0,05; \nu=31)} = 2,04$ ;  $F_{kr(\alpha=0,05; \nu_1=1, \nu_2=31)} = 4,16$

## 4.3 Multivarijantna analiza podataka

### 4.3.1 *Uvodna multivarijantna analiza*

U cilju boljeg razumevanja rezultata dobijenih prilikom hromatografskog razdvajanja ispitivanih arilpiperazina izvršena je uvodna multivarijantna analiza kojom je ukazano na eventualno prisustvo rezultata koji su posledica grubih grešaka, kao i na postojanje moguće međusobne zavisnosti između objekata (ispitivanih jedinjenja) i varijabli (retencionih podataka i molekulskih deskriptora). U tom smislu primenjene su analiza glavnih komponentata i klasteraska analiza.

Set molekulskih deskriptora odabran je tako da reflektuje geometrijska, elektronska i fizičko-hemijska svojstva molekula. Deskriptori korišćeni u daljim analizama, kao i njihove vrednosti navedeni su u Tabeli P8 u Prilogu. Vrednosti molekulskih deskriptora su centrirane u odnosu na srednju vrednost i normirane na jediničnu vrednost standardne devijacije. Nakon izvršenog autoskaliranja svi objekti su simetrično raspoređeni oko koordinatnog početka, varijable su istih dimenzija i promenjen je međusobni odnos objekata. Dimenzije seta podataka iznosile su  $N \times K$  ( $33 \times 30$ ), gde je  $N$  broj ispitivanih arilpiperazina, dok je  $K$  broj varijabli, odnosno molekulskih deskriptora.

### 4.3.2 *Analiza glavnih komponentata*

Analiza glavnih komponentata izvršena je na podacima koji opisuju hromatografsko ponašanje ispitivanih arilpiperazina, kao i na molekulskim deskriptorima istih. Primena PCA na retencione podatke može ukazati na određene sličnosti između ispitivanih jedinjenja koje su uslovljene kako njihovim strukturnim karakteristikama tako i specifičnim interakcijama do kojih dolazi u različitim hromatografskim sistemima. Odgovarajući grafik koeficijenata latentnih varijabli može da ukaže na hromatografski sistem koji u najvećoj meri određuje utvrđeno grupisanje. Sa druge strane, PCA primenjena na set izračunatih

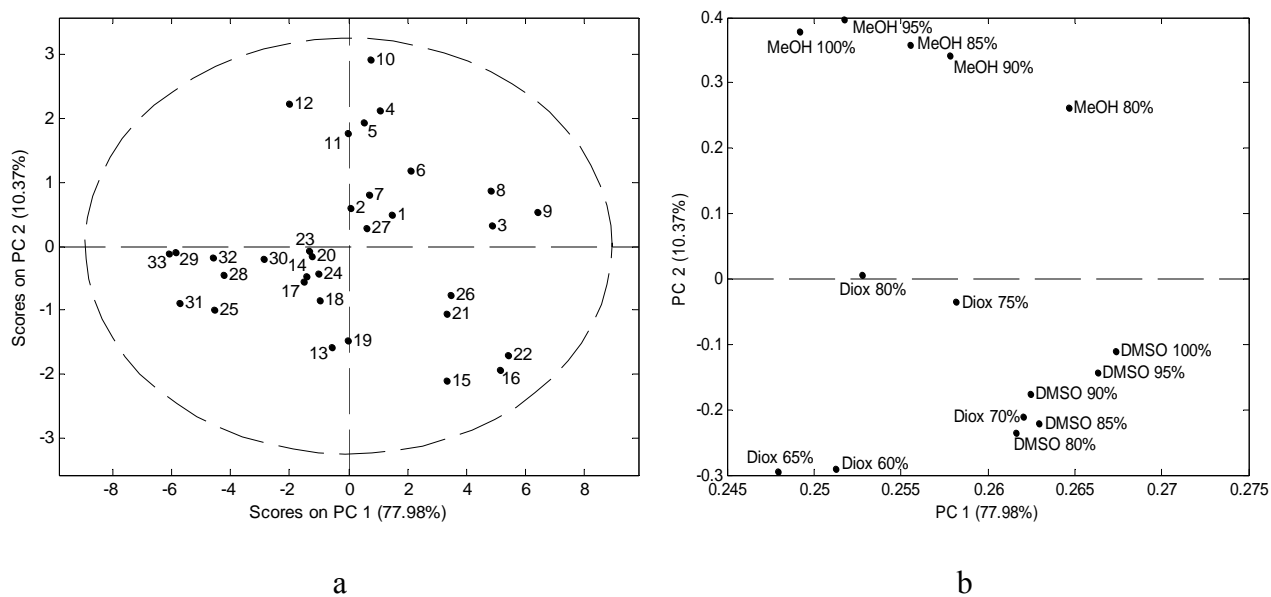
molekulskih deskriptora dovodi do objedinjavanja jedinjenja samo na osnovu sličnih strukturnih karakteristika.

#### 4.3.2.1 PCA primenjena na retencione podatke

PCA primenjena na hromatografske podatke ( $R_M$  vrednosti), dobijene za tri hromatografska sistema, rezultovala je trokomponentnim modelom kojim se objašnjava 95,14% ukupnog varijabiliteta među podacima. Grafik zavisnosti broja glavnih komponenata od svojstvenih vrednosti dat je u Prilogu (Slika P1). Prva glavna komponenta obuhvata 77,98% varijabiliteta, dok druga glavna komponenta ( $PC2$ ) pokriva 10,37%. Informacije sadržane u preostalim varijablama nisu bile značajne za analizu posmatranih podataka.

Vrednosti skorova, odnosno njihova uzajamna projekcija, za prve dve glavne komponente, prikazana je na Slici 11a. Prisustvo rezultata koji su posledica grubih grešaka nije utvrđeno, odnosno među jedinjenjima nema onih koji odstupaju od posmatranog seta, jer se svi podaci nalaze unutar Hotelling  $T^2$  elipse. Uzajamna projekcija koeficijenata latentnih varijabli vektora prikazana je Slici 11b. U odnosu na  $PC2$ -osu jedinjenja su podeljena u dva klastera, saglasno prisustvu tio-grupe u molekulu, odnosno 6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazoli su razdvojeni od 1-{2-[5-(1*H*-benzimidazol-2-tion)]etil}-4-arilpiperazina. Formiranje klastera na osnovu retencionih podataka ostvareno je isključivo duž  $PC2$  ose. Naime, opseg u kojem figurišu varijable u  $PC1$  je izuzetno uzak (0,245 – 0,275) i primenom ove metode nije utvrđen značajan uticaj sastava mobilne faze duž ovog pravca na razdvajanje jedinjenja. Najveći pozitivan uticaj na  $PC2$  ima sistem metanol:voda, odnosno jedinjenja sa vrednostima  $PC2$  skorova iznad nule (derivati 6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazola) pokazuju jaču retenciju u hromatografskom sistemu koji sadrži metanol. Najveći negativan uticaj na razdvajanje duž  $PC2$  pravca imaju sistemi DMSO:voda i dioksan:voda. Jedinjenja sa negativnim vrednostima  $PC2$  skorova (derivati 1-{2-[5-(1*H*-benzimidazol-2-tion)]etil}-4-arilpiperazina) pokazuju jaču retenciju u hromatografskim sistemima u kojima su kao organska komponenta mobilne faze

primenjeni DMSO i dioksan. Prema  $PC1$ , jedinjenja se grupišu na osnovu prisustva atoma halogena u benzimidazolskoj i benzimidazol-2-tionskoj farmakofori, pri čemu se u slučaju benzimidazol-2-tion jedinjenja javljaju izvesna odstupanja. Naime, derivati 1-{2-[5-(1*H*-benzimidazol-2-tion)]etil}-4-arilpiperazina koji u arilpiperazinskom delu molekula sadrže grupe sa proton-akceptorskim svojstvima (2-MeOPhe, pirimidin-2-il, 2-NO<sub>2</sub>Phe) nalaze se u istom delu grafika nezavisno od prisustva atoma halogena u benzimidazolskom delu. Takođe, jedinjenja koja sadrže hidrofobne grupe (2,3-diMePhe, naftil, 3-CF<sub>3</sub>Phe) i u benzimidazolskom delu molekula sadrže vodonik, nalaze se u grupi sa halogenim derivatima. Jedinjenja koja se nalaze unutar jedne grupe imaju sličnu lipofilnost. Saglasno rezultatima PCA primenjene na retencione podatke, hromatografsko ponašanje ove klase jedinjenja određeno je isključivo prirodom supstituenata u arilpiperazinskom delu molekula.



Slika 11. PCA primenjena na retencione podatke, a) grafik skorova, b) grafik koeficijenata latentnih varijabli

#### 4.3.2.2 PCA primenjena na molekulske deskriptore

PCA je izvedena na kompletnom setu podataka u cilju utvrđivanja spoljašnjih vrednosti, kao i međusobnih odnosa, odstupanja i grupisanja među objektima. Matrica podataka sastojala se od 33 objekta (ispitivanih arilpiperazina) i 31 varijable, od kojih su 3 zavisno promenljive ( $R_M^0_{\text{DMSO}}$ ,  $R_M^0_{\text{MeOH}}$ ,  $R_M^0_{\text{Diox}}$ ) dok su preostalih 28 molekulske deskriptori. Dobijeni PCA model sastoji se od sedam glavnih komponenata kojima se objašnjava 92,71% ukupnog varijabiliteta među podacima. Prva glavna komponenta obuhvata 42,24% varijabiliteta, dok kumulativna varijansa objašnjena pomoću prve dve komponente iznosi 63,53%. Grafik zavisnosti broja glavnih komponenata od svojstvenih vrednosti dat je u Prilogu (Slika P2).

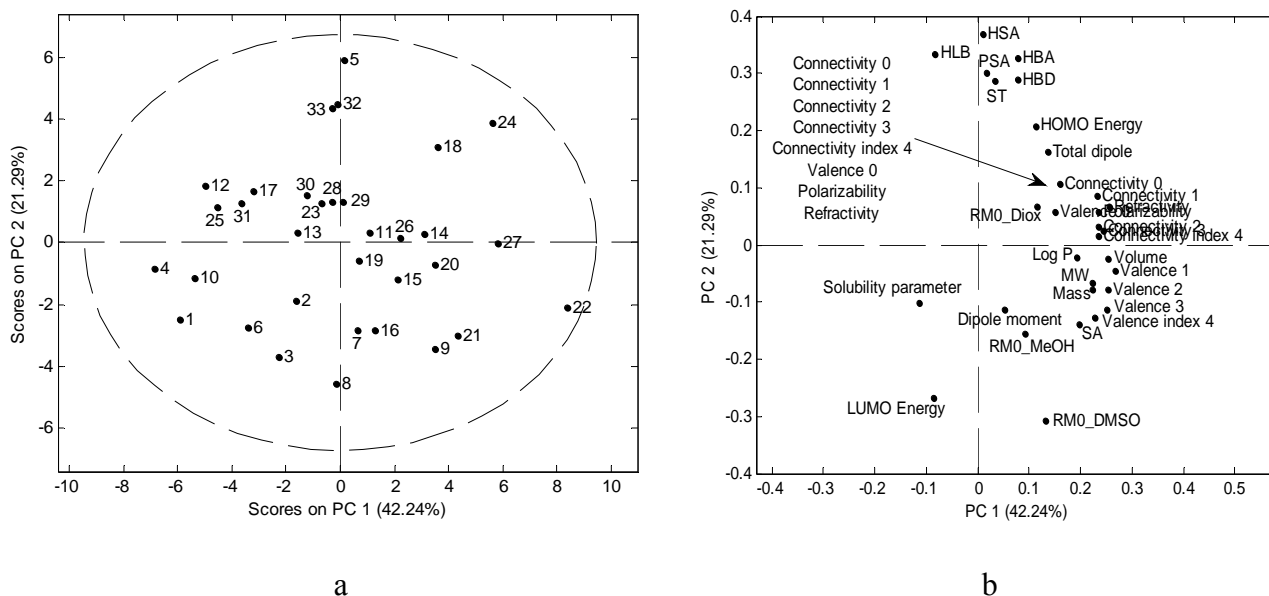
Uzajamne projekcije skorova i odgovarajućih koeficijenata latentnih varijabli, za prve dve glavne komponente, prikazane su na Slici 12. Prisustvo rezultata koji su posledica grubih grešaka nije utvrđeno. Na osnovu vrednosti  $PC2$  jedinjenja sa atomom halogena u benzimidazolskom i benzimidazol-2-tionskom prstenu razdvojena su od jedinjenja sa vodonikom u istom položaju. Izuzetak predstavljaju supstance koje u arilpiperazinskom delu molekula sadrže grupe sa proton-akceptorskim svojstvima (2-NO<sub>2</sub>Phe, pirimidin-2-il), a koja se i pored atoma halogena u dopaminergičkoj farmakofori nalaze u grupi sa nehalogenim derivatima. Svojstva pomenutih jedinjenja određena su, dakle, prisustvom proton-akceptorskih grupa u mnogo većoj meri nego prisustvom halogenog atoma ili tio-grupe.

Molekulske deskriptori koji određuju grupisanje halogenih derivata arilpiperazina opisuju geometriju i veličinu molekula, njihovu lipofilnost, molekulske grafiku ( $\log P$ , MW, masa, indeksi valentnosti 1-4, zapremina), kao i LUMO energiju. Grupisanje nehalogenih derivata i halogenih derivata koji u arilpiperazinskom delu sadrže proton-akceptorske grupe određeno je deskriptorima koji su povezani sa polarnošću molekula (HBA, HBD, ukupni dipol, HSA, HLB, PSA, HOMO energija). Najveći pozitivan i negativan uticaj na  $PC2$  imaju deskriptori koji opisuju polarnost molekula, zbog čega se ova komponenta može smatrati diskriminišućim faktorom koji analizirana jedinjenja razdvaja saglasno njihovim

polarnim karakteristikama. Najveći pozitivan uticaj na vrednosti  $PC1$  imaju topološki deskriptori, parametar lipofilnosti, veličina molekula i njihova geometrija. Značaj deskriptora koji na negativan način utiču na  $PC1$  je zanemarljiv. Ovim je potvrđena pretpostavka Sârbu-a i saradnika da su u hromatografiji dve glavne komponente uglavnom dovoljne za opisivanje najvećeg dela varijabiliteta među retencionim podacima, kao i da jedna od komponenata uglavnom pokazuje visoku korelaciju sa lipofilnošću, veličinom molekula ili sternim faktorima, dok je druga komponenta više određena dipol-dipol interakcijama i elektronskim faktorima<sup>61-63</sup>.

Na grafiku koeficijenata latentnih varijabli se uočava visoka korelacija između molekulskih deskriptora MW, mase, indeksa valentnosti 1-4, zapremine, površine i  $\log P$ , zbog čega bi se iste informacije o hromatografskom ponašanju arilpiperazina dobile korišćenjem bilo koje od ovih veličina. Pomenuti deskriptori su obrnuto proporcionalni parametrima HLB, HSA, PSA, HBA, HBD i ST. Visoka korelacija uočena je i između indeksa konektivnosti 0-4, polarizabiliteta i refraktivnosti.

Slika 12b ukazuje, takođe, na razlike u hromatografskom ponašanju arilpiperazina kada se koriste različiti modifikatori u mobilnim fazama, čime je još jednom potvrđeno da su  $R_M^0$  vrednosti zavisne od prirode organske komponente. Hromatografski određeni parametri lipofilnosti, dobijeni za sisteme DMSO:voda i metanol:voda, visoko su korelisani sa deskriptorima kao što su dipolni moment, masa, MW,  $\log P$ , kao i izvesnim topološkim parametrima, dok su obrnuto proporcionalni deskriptorima koji su povezani sa polarnošću molekula (HBA, HBD, ukupni dipol, HSA, HLB, PSA, HOMO energija). Nasuprot tome,  $R_M^0$  vrednosti dobijene za sistem dioksan:voda, su u obrnutim odnosima sa opisanim molekulskim deskriptorima.



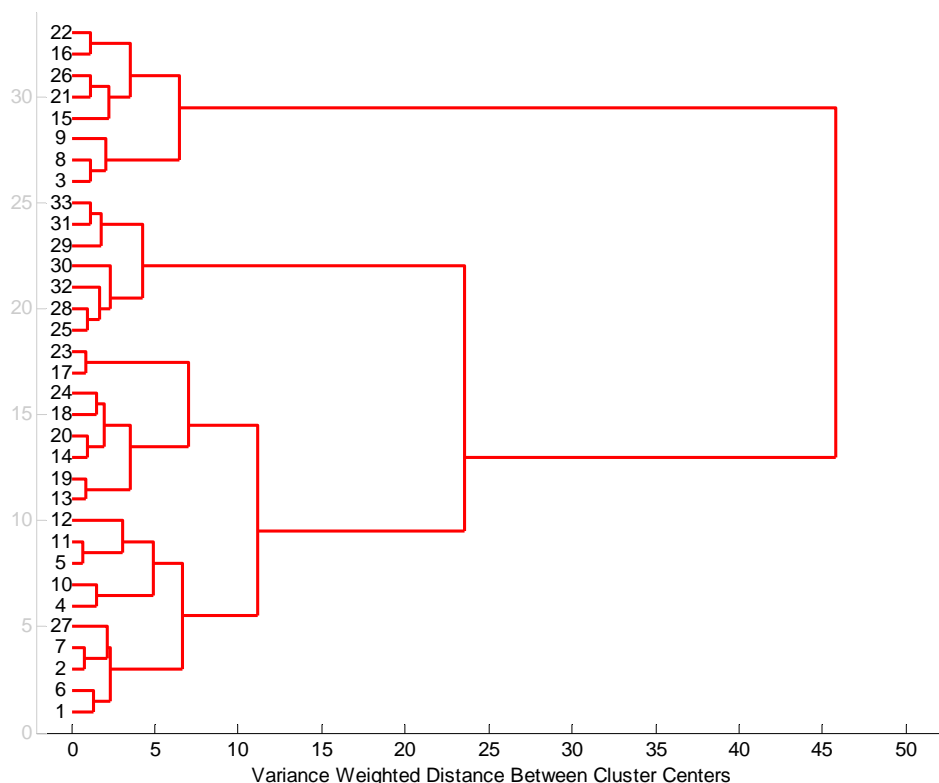
Slika 12. PCA primenjena na molekulske deskriptore, a) grafik skorova, b) grafik koeficijentata latentnih varijabli

### 4.3.3 Klusterska analiza

Klusterska analiza je izvedena kao uvodna analiza kojom je set multivarijantnih podataka preveden u klustere koji su olakšali razumevanje rezultata dobijenih primenom PCA, ali i potvrdili pretpostavljeno grupisanje. HCA je primenjena na podatke koji opisuju retenciju ispitivanih arilpiperazina u uslovima planarne hromatografije, kao i na molekulske deskriptore.

Na Slici 13 je prikazan dendrogram dobijen za retencione podatke za tri primenjena hromatografska sistema. Na razdaljini od 25 mogu se razlikovati dva klastera u kojima su jedinjenja razdvojena na osnovu polarnosti grupe u arilpiperazinskom delu molekula. U prvom klasteru se nalaze supstance (**3**, **8**, **9**, **15**, **16**, **21**, **22**, **26**) koje sadrže hidrofobne grupe (2,3-diMePhe, 3-CF<sub>3</sub>Phe, 2-ClPhe), dok su u drugom klasteru jedinjenja koja sadrže grupe sa proton-akceptorskim karakteristikama.

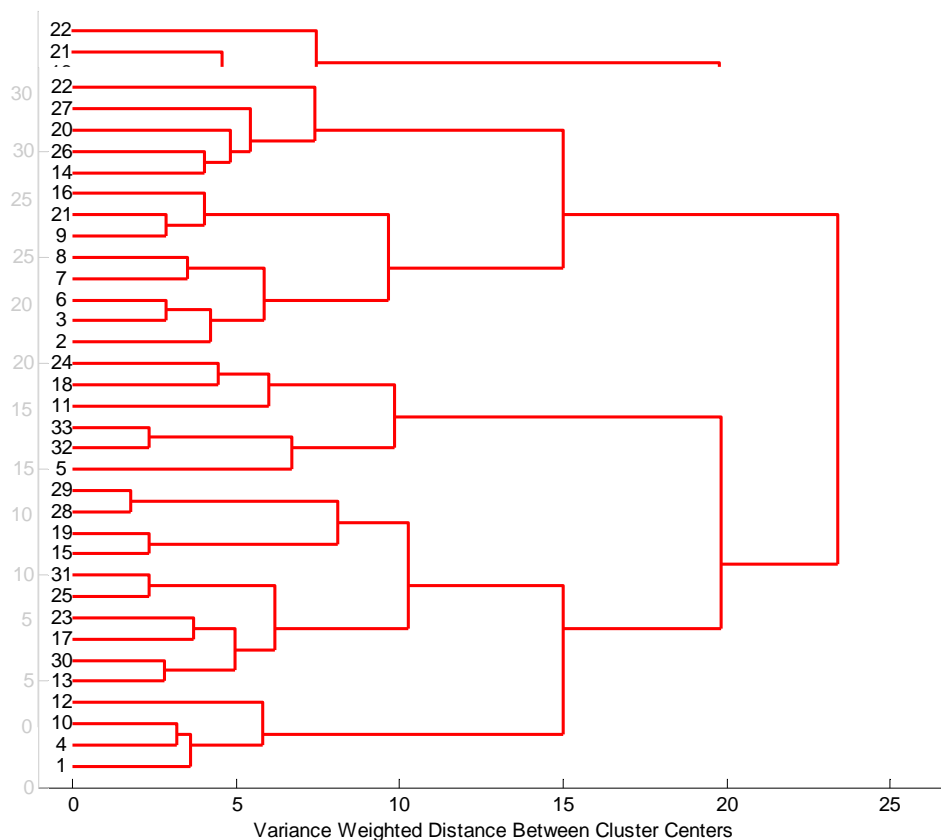


Slika 13. Dendrogram dobijen za  $R_M$  vrednosti

Na razdaljini od 8 mogu se uočiti četiri klastera. U prvom klasteru se nalaze derivati 6-[2-(4-arilpiperazin-1-il)etil]-1*H*-benzimidazola, pri čemu se supstance **3**, **6** i **9**, koje pripadaju pomenutoj klasi ali sadrže hidrofobne grupe, nalaze u četvrtom klasteru, zajedno sa istim jedinjenjima iz grupe derivata 1-{2-[5-(1*H*-benzimidazol-2-tion)]etil}-4-arilpiperazina. U drugom klasteru se nalaze derivati benzimidazol-2-tiona koji u arilpiperazinskom delu molekula sadrže polarne grupe (pirimidin-2-il, 2-NO<sub>2</sub>Phe, 2-MeOPhe), dok treći klaster čine jedinjenja iz benzimidazol-2-tionske grupe koja u benzimidazolskom delu molekula ne poseduju atom halogena.

Na dendrogramu dobijenom za molekulske deskriptore (Slika 14) uočavaju se dva klastera, na razdaljini od 17, pomoću kojih se ispitivana jedinjenja grupišu na osnovu

prisustva tio-grupe u benzimidazolskom delu molekula. Unutar grupe 1-{2-[5-(1*H*-benzimidazol-2-tion)]etil}-4-arilpiperazina nije moguće objasniti formirane podgrupe, što upućuje na zaključak da se na osnovu molekulskih deskriptora pomenuta jedinjenja malo razlikuju.



Slika 14. Dendrogram dobijen za molekulske deskriptore

#### 4.3.4 *QSRR analiza*

Hemometrijske metode višestruka linearna regresija, regresija metodom delimičnih-najmanjih kvadrata i regresija glavnih komponenta primenjene su za identifikaciju molekulskih deskriptora koji na najbolji način opisuju ponašanje ispitivanih jedinjenja u

hromatografskom sistemu, kao i za kvantifikaciju njihovog uticaja, a u cilju dobijanja matematičkih modela na osnovu kojih bi moglo da se predvidi hromatografsko ponašanje, odnosno lipofilnost novih srodnih molekula. U QSRR analizama koje slede zavisno promenljive su bili parametri lipofilnosti,  $R_M^0$ , dobijeni za hromatografski sistem u kojem je korišćen DMSO kao organska komponenta mobilne faze, dok su kao nezavisno promenljive korišćeni molekulski deskriptori.

Pre regresione analize podaci su podeljeni u dve grupe, na kalibracioni i test set. Molekuli koji su uključeni u određeni set odabrani su nasumice. Garkani-Nejad i Ahmadvand<sup>136,137</sup> su, proučavajući uticaj veličine test seta na kvalitet modela, utvrdili da ukoliko test set sadrži 15 – 40% jedinjenja iz osnovnog seta, model konstruisan sa kalibracionim setom može uspešno da predvidi test set. Ukoliko test set sadrži manje od 15% jedinjenja iz osnovnog seta korelacioni koeficijent test seta je nepouzdan, dok ukoliko je pomenuti procenat visok, model dobijen za kalibracioni set nije kompletan i ne može tačno da predvidi posmatranu promenljivu. U ovom radu test set se sastojao od približno 39% jedinjenja odabranih iz osnovnog seta. U kalibracionom setu se nalazilo 20 jedinjenja (**2, 3, 4, 5, 6, 7, 10, 11, 12, 15, 16, 19, 20, 22, 24, 25, 26, 28, 29, 30**) koja su korišćena za konstruisanje modela, dok je test set koji je obuhvatao 13 jedinjenja (**1, 8, 9, 13, 14, 17, 18, 21, 23, 27, 31, 32, 33**) korišćenih za procenu dobijenog matematičkog modela.

#### **4.3.4.1 Višestruki linearni regresioni model**

Višestruki linearni regresioni model se ne može primeniti za modelovanje kompleksnog seta podataka u kojem broj nezavisnih varijabli prevazilazi broj objekata. Stoga, korak koji prethodi pravljenju modela je redukcija broja nezavisno promenljivih varijabli. Molekulski deskriptori čije su vrednosti približno jednake za sva ispitivana jedinjenja, odnosno veličine kod kojih nije postojao varijabilitet među podacima, isključene su iz dalje analize jer nisu pružale informacije o razlikama između ispitivanih arilpiperazina (HOMO energija, zapremina, parametar rastvorljivosti i molarna refraktivnost). Deskriptori

kod kojih postoji visoka međusobna korelacija pružaju iste informacije o jedinjenjima i jednu od njih bi trebalo zanemariti. Dva deskriptora su međusobno nezavisna ukoliko je vrednost korelacionog koeficijenta manja od 0,8000<sup>39</sup>. Korelaciona matrica sa vrednostima Pearson-ovih koeficijenata ( $r_{ij}$ ), koji predstavljaju linearne regresione koeficijente dobijene korelisanjem  $i$ -te nezavisne promenljive sa  $j$ -tom nezavisno promenljivom, za set molekulskih deskriptora izračunatih za seriju arilpiperazina, dati su u Tabeli P9 u Prilogu. Visoke vrednosti korelacionog koeficijenta su dobijene za deskriptore hidrofilna površina i hidrofilno-lipofilni balans, hidrofilna površina i akceptor vodonične veze, kao i za površinski napon i polarnu površinu. Od prethodno navedenih deskriptora zadržani su oni koji su pokazivali bolju korelaciju sa zavisno promenljivom, odnosno molekulski deskriptori hidrofilna površina i polarna površina su odbačeni. Deskriptori koji su preostali nakon izvršene selekcije prikazani su u Tabeli P10 u Prilogu i korišćeni su kao ulazni podaci za MLR metodu.

Povećanje broja deskriptora u regresionoj jednačini rezultuje boljim prediktivnim karakteristikama modela. Sa druge strane, uvođenjem previše varijabli dobio bi se model koji je najverovatnije precenjen. U cilju odabira seta deskriptora koji će dati model sa najboljom prediktivnošću, izvršena je selekcija varijabli metodom postepene regresije. Rezultati primenjene metode dati su u Tabeli 7.

Tabela 7. Selekcija varijabli metodom postepene regresije

Broj ponavljanja		Varijabla	$R^2$	RMSE
1	Dodato	HLB	0,5951	0,666
2	Dodato	CI0	0,7438	0,545
3	Dodato	CI2	0,8825	0,381
4	Dodato	ST	0,9463	0,266
5	Dodato	TD	0,9623	0,230
6	Dodato	VI3	0,9762	0,190
7	Nepromenjeno		0,9762	0,190

Molekulski deskriptori odabrani kao varijable koje značajno utiču na retencione parametre ispitivanih arilpiperazina su: ukupni dipol, polarizabilnost, hidrofilno-lipofilni balans, površinski napon, indeks konektivnosti 0 i 2 i indeks valentnosti 3.

Modelovanjem  $R_M^0$  vrednosti jedinjenja iz kalibracionog seta sa vrednostima prethodno navedenih molekulskih deskriptora, dobijena je MLR jednačina koju karakteriše visoka vrednost korigovanog koeficijenta determinacije ( $R_{adj}^2 = 0,9652$ ) i  $F$ -testa za korelaciju ( $F = 88,729$ ), kao i niske vrednosti srednjeg kvadratnog odstupanja ( $MSE = 0,036$ ), nivoa verovatnoće ( $P < 0,001$ ) i sume kvadrata odstupanja od predviđenih vrednosti ( $PRESS = 1,13$ ). Vrednosti regresionih koeficijenata, njihove standardne devijacije i intervali pouzdanosti za 95% nivo značajnosti, dati su u Tabeli 8.

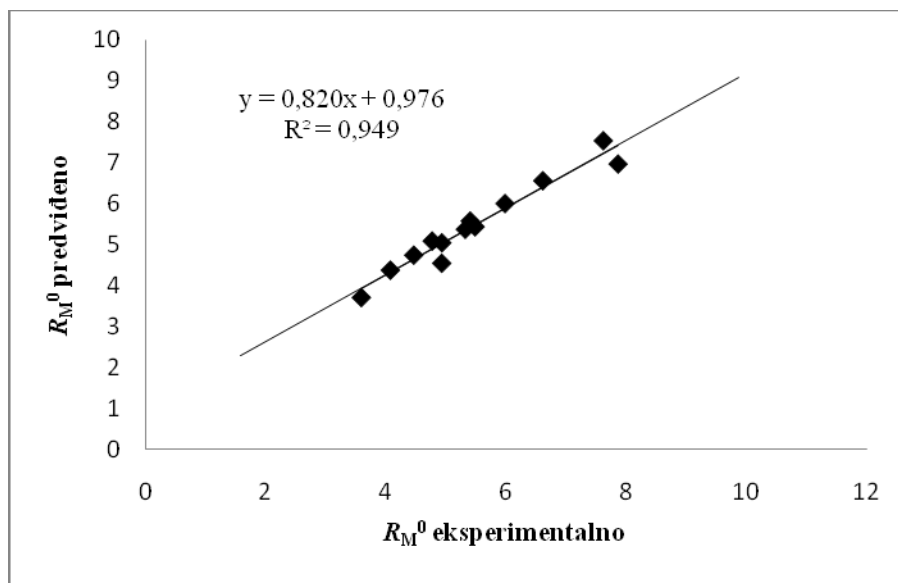
Tabela 8. Statistički parametri MLR modela

Nezavisna varijabla	Regresioni koeficijent	s	$L_1$	$L_2$	t-test $H_0: b_i=0$	P	$H_0: b_i \neq 0$
Odsečak	5,072	1,371	2,109	8,035	3,698	0,0027	Da
CI0	-0,189	0,021	-0,234	-0,144	-9,138	<0,0001	Da
CI2	0,816	0,081	0,642	0,991	10,096	<0,0001	Da
HLB	-0,269	0,055	-0,388	-0,149	-4,850	0,0003	Da
ST	-0,022	0,004	-0,029	-0,014	-6,134	<0,0001	Da
TD	-0,128	0,040	-0,215	-0,040	-3,159	0,0075	Da
VI3	0,613	0,224	0,130	1,096	2,743	0,0168	Da

U cilju utvrđivanja postojanja statistički značajne linearne zavisnosti između određene nezavisne varijable i zavisno promenljive, na vrednosti regresionih koeficijenata primenjen je  $t$ -test. Vrednost  $t$  se izračunava iz odnosa  $t = \frac{b}{s_b}$ , gde je  $b$  nagib u regresionom modelu, a  $s_b$  standardna devijacija nagiba, dok je broj stepeni slobode  $n-k-1$  ( $n$  – veličina uzorka,  $k$  – broj nezavisnih promenljivih). Prema nultoj hipotezi vrednost nagiba je jednaka nuli. Ukoliko je vrednost nagiba u regresionoj pravoj značajno različita od nule, između

posmatrane nezavisno promenljive i  $R_M^0$  vrednosti postoji statistički značajan odnos. Rezultati  $t$ -testa za regresione koeficijente dati su u Tabeli 8. Postojanje značajne zavisnosti utvrđeno je za sve molekulske deskriptore koji su uključeni u model.

Predložena MLR jednačina primenjena je na jedinjenja iz test seta i  $R_M^0$  vrednosti su upoređene sa eksperimentalno dobijenim. Visoke vrednosti  $R^2_{adj}$  (0,9652) i  $F$  (88,729), kao i niske vrednosti  $MSE$  (0,036) i  $P$  ( $< 0,001$ ) ukazuju na dobru prediktivnu moć modela. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti dat je na Slici 15.



Slika 15. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti dobijenih na osnovu MLR modela

#### 4.3.4.2 Regresija glavnih komponenata

Kod PCR metode originalne varijable zamenjuju se sa nekoliko glavnih komponenata, dobijenih nakon PCA, redukujući na taj način početni set podataka. PCR metoda primenjena je na jedinjenja iz kalibracionog seta, pri čemu su svi molekularni deskriptori korelisani sa zavisno promenljivom i kvalitet dobijenog modela procenjen na osnovu sledećih parametara:  $R^2_{\text{Cal}}$  (koeficijent determinacije koji se odnosi na rezultate iz kalibracionog seta),  $R^2_{\text{CV}}$  (koeficijent determinacije koji je rezultat unakrsne validacije),  $RMSEC$  (koren srednjeg kvadratnog odstupanja koji se odnosi na rezultate iz kalibracionog seta),  $RMSECV$  (koren srednjeg kvadratnog odstupanja koji je rezultat unakrsne validacije)<sup>80,103</sup>. Unakrsna validacija je izvršena metodom slučajnih uzoraka, uz 5 podela i 20 ponavljanja i korišćena je za proveru prediktivnih svojstava PCR modela.

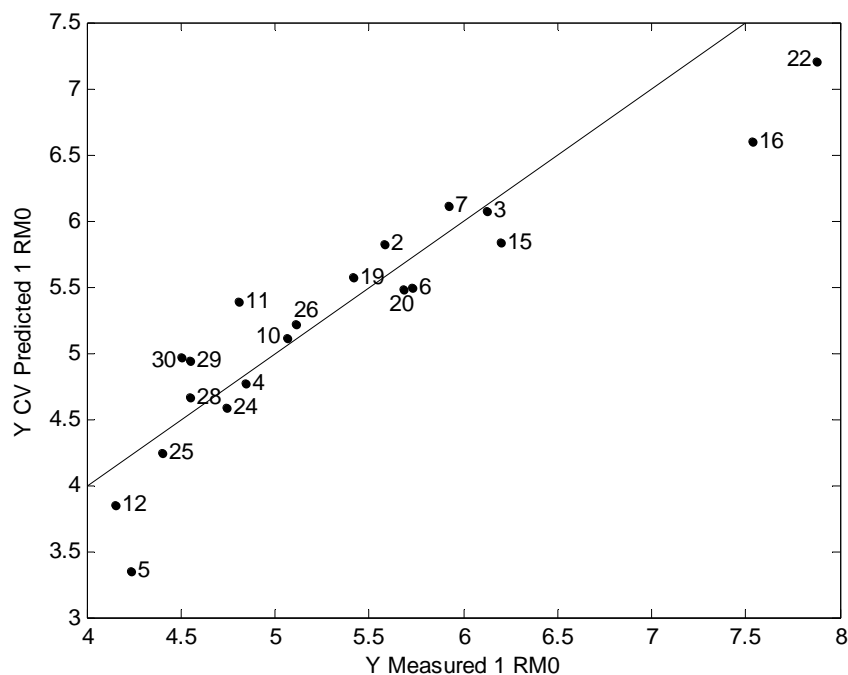
Prvih osam glavnih komponenata, kojima je objašnjeno 97,25% ukupnog varijabiliteta u matrici podataka  $X$ , korišćeno je za modelovanje. Grafik zavisnosti broja glavnih komponenata od odnosa  $RMSEC/RMSECV$  dat je na Slici P3 u Prilogu. Broj glavnih komponenata odabran je tako da odnos  $RMSEC/RMSECV$  bude što niži a da pri tome udeo objašnjenog varijabiliteta u modelu bude što viši.

Tabela 9. Udeo pojedinačnih komponenata u PCR modelu

Glavna komponenta	X - Blok		Y - Blok	
	Udeo pojedinačne komponente	Ukupan udeo	Udeo pojedinačne komponente	Ukupan udeo
1	46,01	46,01	17,66	17,66
2	19,49	65,49	48,75	66,40
3	13,62	79,12	6,99	73,39
4	7,91	87,03	0,22	73,61
5	4,42	91,45	12,51	86,12
6	2,64	94,09	9,93	96,05
7	1,92	96,01	0,36	96,41
8	1,24	97,25	1,44	97,85

U Tabeli 9 prikazani su udeli pojedinačnih komponenta u predloženom PCR modelu. Najveći uticaj na model ima  $PC2$  kojom se objašnjava 19,49% ukupne varijanse u matrici podataka  $X$  i sa  $y$ -vrednostima daje vrednost korelacionog koeficijenta u iznosu od 0,49. Značajno manji uticaj na model imaju  $PC1$  i  $PC5$  kod kojih vrednosti korelacionih koeficijenata iznose 0,18, odnosno 0,13, redom. Pri tome  $PC1$  opisuje 46,01% ukupnog varijabiliteta među podacima. Glavne komponente koje su pokazivale malu sposobnost predviđanja  $R_M^0$  vrednosti smatrane su manje značajnim.

Statistički parametri PCR modela dobijenog za kalibracioni set su:  $R^2_{\text{Cal}} = 0,9784$  i  $R^2_{\text{CV}} = 0,8810$ ,  $RMSEC = 0,1457$  i  $RMSECV = 0,3950$ .

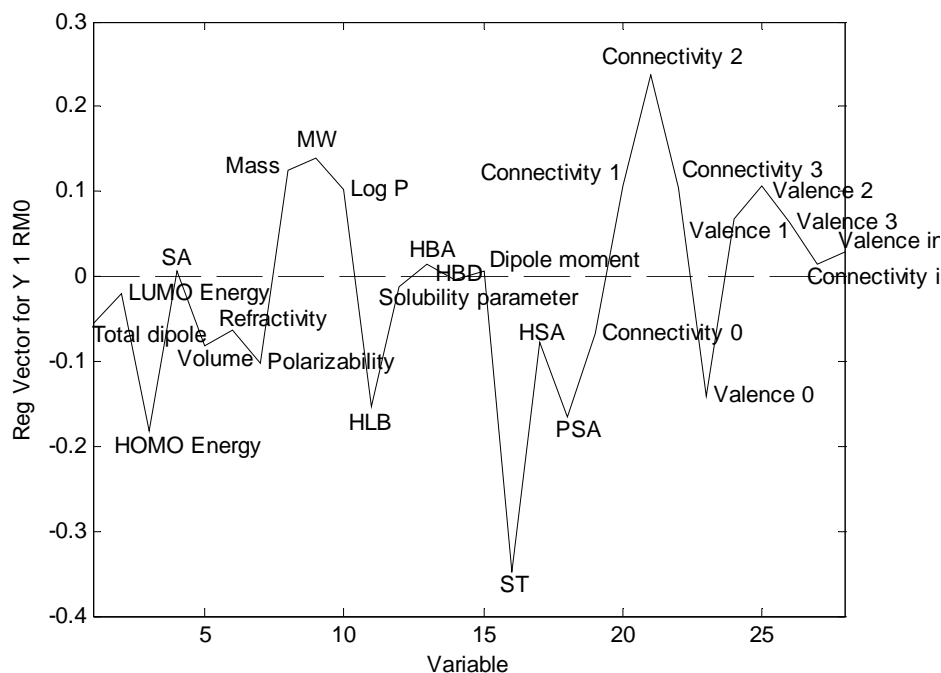


Slika 16. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti dobijenih na osnovu PCR modela



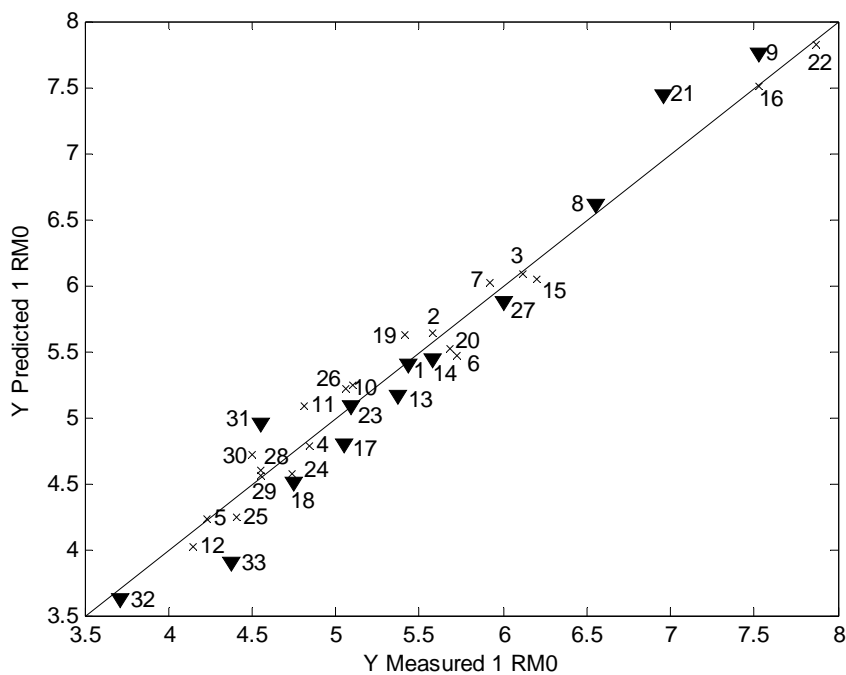
Grafik zavisnosti reziduala od  $Y_{\text{modela}}$  ( $R_M^0$  vrednosti predviđene modelom) dat je na Slici P4 u Prilogu. Ovakva vrsta grafika može biti veoma korisna u smislu procene uklapanja predloženog modela u set podataka. Ukoliko su vrednosti reziduala raspoređene nasumično na grafiku, odnosno loše korelisane sa predviđenim  $R_M^0$  vrednostima, smatra se da se model dobro uklapa u podatke. Slika P4 pokazuje slučajnu raspodelu reziduala sa niskom vrednošću korelacionog koeficijenta ( $r = -0,2985$ ) između  $Y_{\text{modela}}$  i vrednosti reziduala, ukazujući na činjenicu da se predloženi PCR model dobro uklapa u podatke. Ovakav zaključak potvrđen je i grafikom zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti (Slika 16).

Vrednosti koeficijenata molekulskih deskriptora u predloženom PCR modelu date su na Slici 17.



Slika 17. Grafik vrednosti koeficijenata molekulskih deskriptora u PCR modelu

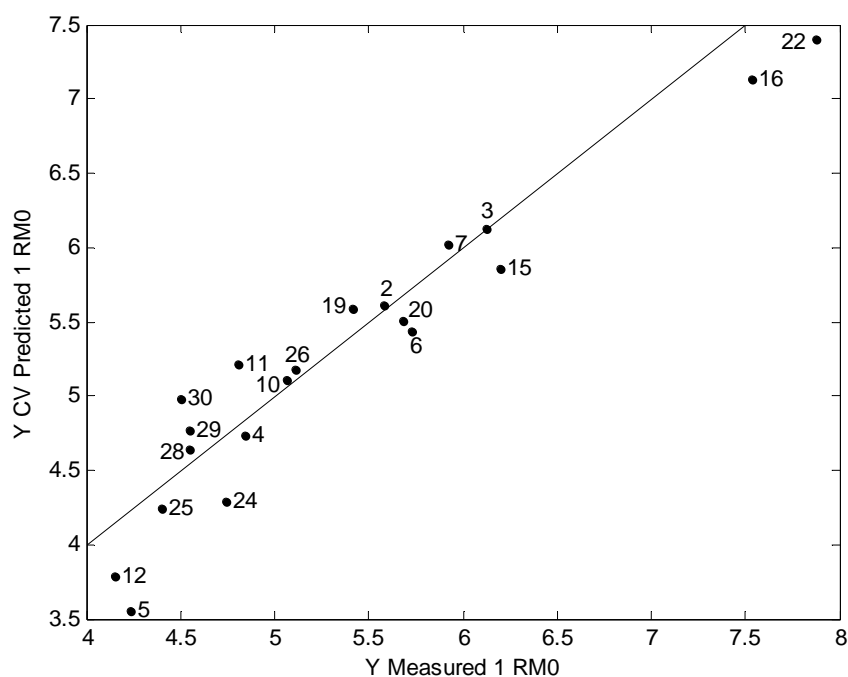
Prediktivna sposobnost PCR modela proverena je na podacima iz test seta pri čemu su dobijeni sledeći statistički parametri:  $R^2_{\text{Cal}} = 0,9785$ ,  $R^2_{\text{CV}} = 0,8385$  i  $R^2_{\text{Pred}} = 0,9639$ ;  $RMSEC = 0,1457$ ,  $RMSECV = 0,4717$  i  $RMSEP = 0,2628$ , gde je  $R^2_{\text{Pred}}$  i  $RMSEP$  koeficijent determinacije, odnosno koren srednjeg kvadratnog odstupanja koji se odnose na podatke iz test seta, redom. Na osnovu dobijenih statističkih parametara može se zaključiti da je predloženi PCR model dobar. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti za kalibracioni i test set prikazan je na Slici 18.



Slika 18. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti za kalibracioni (x) i test set (▼), dobijenih na osnovu PCR modela

#### 4.3.4.3 Regresija metodom delimičnih-najmanjih kvadrata

Regresija metodom delimičnih-najmanjih kvadrata je jedna od najčešće primenjenih metoda za regresionu analizu multivarijantnih podataka u hemometriji.  $R_M^0$  vrednosti jedinjenja iz kalibracionog seta korelisane su sa svim molekulskim deskriptorima uz primenu unakrsne validacije metodom slučajnih uzoraka, uz 5 podela i 20 ponavljanja. Broj latentnih varijabli je odabran na osnovu minimalne vrednosti  $RMSECV$ . Grafik zavisnosti broja glavnih komponenta od odnosa  $RMSEC/RMSECV$  dat je na Slici P5 u Prilogu. Minimalna vrednost  $RMSECV$  dobijena je za petokomponentni model.



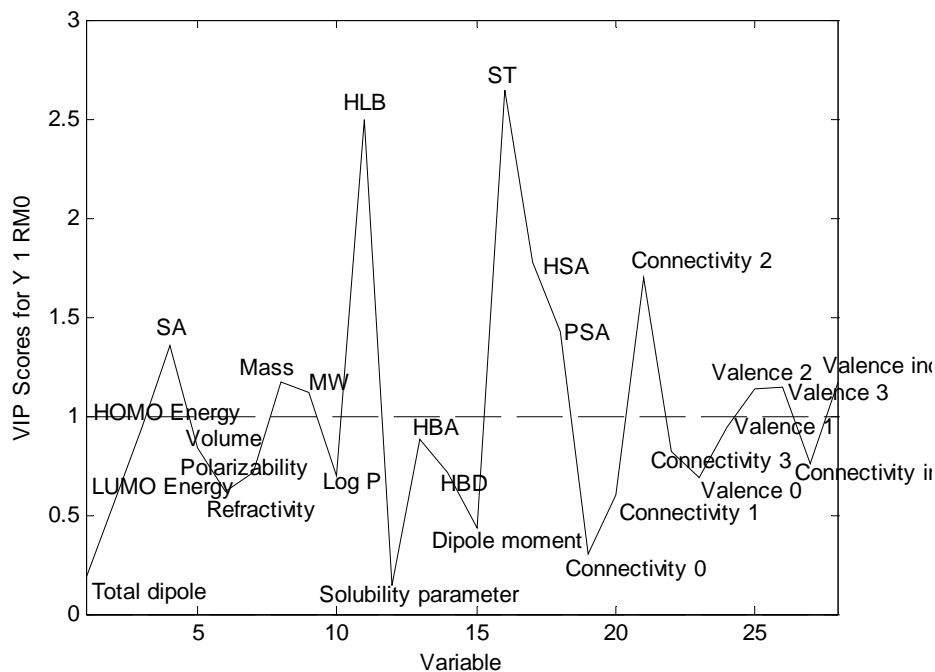
Slika 19. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti dobijenih na osnovu PLS modela

Statistički parametri PLS modela za kalibracioni set su sledeći:  $R^2_{\text{Cal}} = 0,9832$  i  $R^2_{\text{CV}} = 0,9095$ ,  $RMSEC = 0,1287$  i  $RMSECV = 0,3588$ . Grafik zavisnosti reziduala od  $Y_{\text{modela}}$  (Slika P6 u Prilogu), sa niskom vrednošću korelacionog koeficijenta ( $r = -0,2219$ ) između  $Y_{\text{modela}}$  i vrednosti reziduala, kao i grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti (Slika 19), potvrđuje dobro uklapanje predloženog modela u set podataka.

Kod PLS metode varijablu koja značajno utiče na vrednost zavisno promenljive karakteriše visoka vrednost regresionog koeficijenta. Varijabla, međutim, može biti značajna i za modelovanje matrice nezavisno promenljivih, što se iskazuje kroz visoku vrednost koeficijenta latentne varijable. Značaj određene varijable, u smislu njenog udela u objašnjenju zavisno promenljive i varijanse nezavisno promenljivih, izražava se kroz tzv. VIP vrednost (Variable Importance for the Projection – varijable značajne za projekciju). Ova vrednost se za varijablu  $j$  izračunava po formuli:

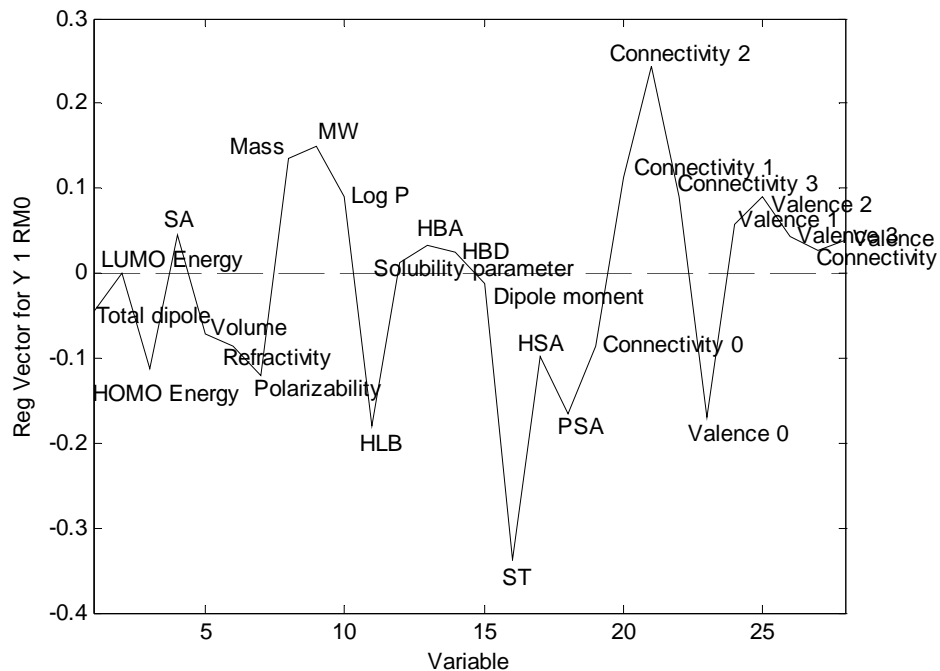
$$VIP_j = \sqrt{\frac{p}{\sum_{m=1}^M SS(b_m t_m)} \times \sum_{m=1}^M \omega^2_{mj} SS(b_m t_m)} \quad (70)$$

pri čemu je  $p$  broj varijabli,  $M$  je broj latentnih varijabli kojima se objašnjava PLS model,  $\omega_{ij}$  udeo  $j$ -te varijable u  $m$ -toj latentnoj varijabli,  $SS(b_m t_m)$  procenat podataka sadržanih u  $y$ -vrednostima objašnjen pomoću  $m$  latentnih varijabli. VIP vrednosti predstavljaju doprinos određene varijable modelu saglasno njihovom udelu u svakoj latentnoj komponenti, dok regresioni koeficijenti ukazuju na značaj određene varijable u modelu<sup>138,139</sup>. Na osnovu formule za izračunavanje VIP vrednosti srednja vrednost kvadrata VIP-skorova jednaka je 1, pa se pomenuta vrednost koristi kao kriterijum za selekciju varijabli<sup>139</sup>. Varijable sa vrednostima VIP-skorova većim od 1 smatraju se najznačajnijim za objašnjenje podataka zavisno promenljive. Grafik zavisnosti varijabli od vrednosti VIP-skorova dat je na Slici 20. Molekulski deskriptori koji imaju najveći uticaj na  $R_M^0$  vrednosti su ST, HLB, HSA, PSA, CI2, SA, masa, MW, VI2-4.



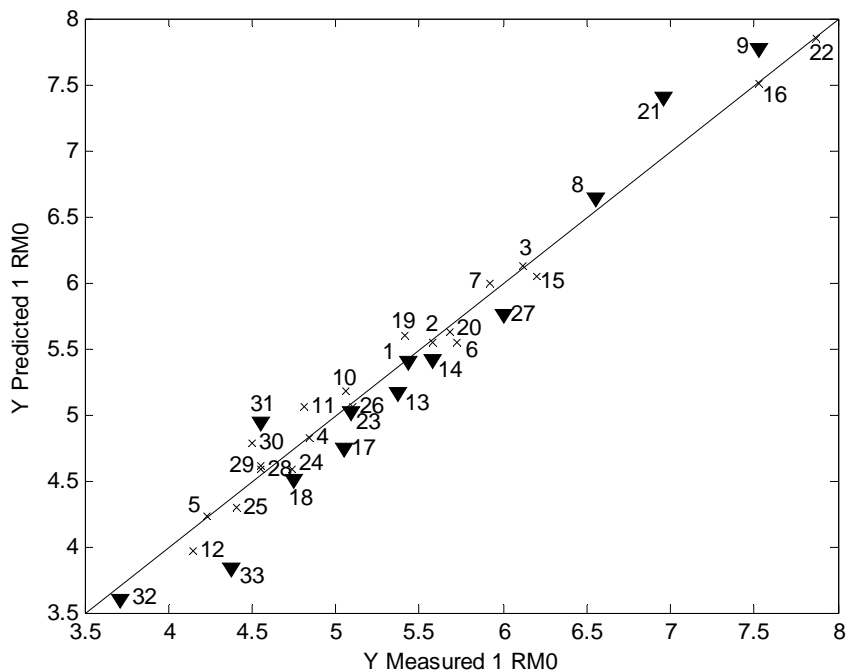
Slika 20. Grafik zavisnosti varijabli od vrednosti VIP-skorova za PLS model

Vrednosti korelacionih koeficijenata varijabli u predloženom PLS modelu date su na Slici 21. Najveći pozitivan uticaj na retencioni parametar imaju IC2, masa, MW i log  $P$ , dok najveći negativan uticaj imaju ST, HLB, PSA i VI0. Molekulski deskriptori koji opisuju geometriju i veličinu molekula, njihovu lipofilnost i molekulsku grafiku direktno su proporcionalni hromatografski određenom parametru lipofilnosti. Sa druge strane, prema dobijenom modelu jedinjenja sa većim vrednostima deskriptora koji su povezani sa polarnim karakteristikama molekula imaju niže  $R_M^0$  vrednosti, odnosno manju lipofilnost.



Slika 21. Vrednosti korelacionih koeficijenata varijabli u PLS modelu

Prediktivna sposobnost PLS modela proverena je na podacima iz test seta. Statistički parametri dobijenog modela,  $R^2_{\text{Cal}} = 0,9832$ ,  $R^2_{\text{CV}} = 0,8710$  i  $R^2_{\text{Pred}} = 0,9612$ ;  $\text{RMSEC} = 0,1287$ ,  $\text{RMSECV} = 0,4387$  i  $\text{RMSEP} = 0,2796$ , potvrđuju da je predloženi PLS model dobar. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti, za kalibracioni i test set, dat je na Slici 22.



Slika 22. Grafik zavisnosti eksperimentalno određenih od predviđenih  $R_M^0$  vrednosti za kalibracioni (x) i test set ( $\blacktriangledown$ ), dobijenih na osnovu PCR modela

Tri predložene multivarijantne metode (MLR, PCR i PLS) karakterišu dobri statistički parametri. MLR je najosetljivija na vrstu i broj odabranih varijabli. U slučajevima kada postoje suvišne varijable ili varijable koje imaju konstantne vrednosti i ne razlikuju uzorke međusobno, daje modele lošijeg kvaliteta, zbog čega su dobijeni rezultati neočekivani. PLS i PCR su izuzetno značajne u hemometriji jer se pomoću njih može modelovati veliki broj međusobno visoko korelisanih varijabli. PLS prilikom formiranja latentnih varijabli koristi dva kriterijuma, pravac najveće varijanse unutar seta podataka i najbolju korelaciju sa vrednostima zavisno promenljive, zbog čega bi ova metoda trebalo da rezultuje modelom sa najvećom prediktivnom moći. Latentne varijable kod PCR objašnjavaju varijabilitet unutar seta nezavisno promenljivih varijabli zbog čega je ova metoda ograničena u odnosu na PLS.

#### 4.3.4.4 Uticaj pojedinih molekulskih deskriptora na hromatografsko ponašanje ispitivanih arilpiperazina

Visoke vrednosti  $R_{\text{adj}}^2$  i  $R_{\text{pred}}^2$ , kao i niske vrednosti  $MSE$  i  $RMSEP$ , potvrđuju da su MLR, PCR i PLS modeli, dobijeni korelisanjem retencionih parametara sa molekulskim deskriptorima, validni. Korak koji sledi je pronalaženje fizičkog smisla u molekulskim deskriptorima koji su označeni kao važni za opisivanje i predviđanje hromatografskog ponašanja ispitivanih arilpiperazina. Kod MLR metode, varijable koje su pokazivale najveće vrednosti regresionog koeficijenta, i Student-ove  $t$ -vrednosti, kao i najniže vrednosti nivoa pouzdanosti  $P$ , odabrane su kao varijable koje imaju najveći uticaj na zavisno promenljivu. U slučaju PCR metode, deskriptori sa najvećim uticajem na  $PC2$  komponentu, koja je pokazivala najveći korelacioni koeficijent sa  $y$ -vrednostima, posmatrani su kao deskriptori koji u najvećoj meri određuju  $R_M^0$  vrednosti ispitivanih arilpiperazina. Kod PLS metode, doprinos varijabli  $y$ -vrednostima određen je korišćenjem vrednosti VIP skorova.

U predloženim modelima značajan uticaj na  $R_M^0$  vrednosti imaju molekulski deskriptori ST i HLB. Površinski napon je fizičko svojstvo određeno strukturom molekula. Najznačajnije karakteristike molekulske strukture koje utiču na površinski napon su površina molekula i proton-donorska, odnosno proton-akceptorska svojstva određenih atoma u molekulu<sup>140</sup>. Topološki deskriptori koji opisuju veličinu, oblik i račvanja u molekulu i nisu povezani sa geometrijom molekula, daju informaciju o površini molekula i direktno određuju površinski napon. Međutim, najveći pozitivan uticaj na pomenuto svojstvo imaju deskriptori koji opisuju sposobnost vodoničnog vezivanja. U modelu koji povezuje strukturu sa površinskim naponom nalaze se i izvesni elektronski deskriptori koji ukazuju na značaj polarnih interakcija pri određivanjima pomenutog fizičkog svojstva molekula<sup>140</sup>. U slučaju ispitivanih arilpiperazina elektrostatičke interakcije (jon-jon, jon-dipol, dipol-dipol interakcije), kao i vodonično vezivanje, određivaće vrednosti površinskog napona. Jedinjenja čiji atomi stupaju u ovu vrstu interakcija imaju više ST vrednosti, a u hromatografskom sistemu se slabije zadržavaju. Opisani odnos površinskog napona i



retencionog parametra potvrđen je i u svim predloženim modelima u kojima ova varijabla pokazuje negativan regresioni koeficijent. Za razliku od MLR modela u kojem ST pokazuje mali negativan uticaj na retenciju, u PCR i PLS modelu ST ima najveći koeficijent i u odnosu na sve druge varijable ima najveći negativan uticaj na  $R_M^0$  vrednosti.

Hidrofilno-lipofilni balans predstavlja odnos između hidrofilnih i hidrofobnih delova molekula i opisuje koji efekat dominira u molekulu. Prema Griffin-ovoj metodi za neutralne molekule, HLB je direktno proporcionalan molekulskoj masi hidrofilnog dela molekula i obrnuto proporcionalan masi celog molekula i ima vrednosti od 0 do 20. Vrednost HLB jednaka 0 odgovara potpuno hidrofobnom molekulu, dok vrednost 20 imaju jako hidrofilni molekuli<sup>141</sup>. QSRR i QSAR studije ukazuju na značajan uticaj ovog deskriptora na reticioni parametar<sup>142</sup>, odnosno biološku aktivnost<sup>143</sup>. Ispitivani arilpiperazini imaju HLB vrednosti u intervalu od 7 do 12, odnosno pokazuju i hidrofilna i lipofilna svojstva. U svim predloženim modelima pomenuti parametar karakteriše negativna vrednost regresionog koeficijenta, tj. sa povećanjem ovog balansa smanjuje se  $R_M^0$  vrednost. Hidrofobnija jedinjenja koja pokazuju jače nespecifične disperzione interakcije između nepolarnih delova molekula i nepolarne stacionarne faze, više se zadržavaju pod primenjenim hromatografskim uslovima.

U PLS i PCR modelima kao varijable koje značajno utiču na zavisno promenljivu pojavljuju se i parametri rastvorljivosti, PSA i HSA. Polarna površina se definiše kao deo površine molekula u kojem se nalaze izrazito elektronegativni atomi kao što su kiseonik, azot, sumpor, i vodonik vezan za prethodno navedene atome. Ovi polarni atomi učestvuju u vodoničnom vezivanju i drugim elektrostatičkim interakcijama. Jedinjenja sa većim brojem ovih atoma, koji nisu sterno zaklonjeni i dostupni su drugim polarnim atomima, rastvorljiviji su u polarnoj mobilnoj fazi hromatografskog sistema, zbog čega je regresioni koeficijent PSA u dobijenim modelima sa negativnim predznakom. PSA je deskriptor koji se pojavljuje u mnogim QSAR/QSRR modelima što je i očekivano sa obzirom na strukturne karakteristike ćelijske membrane i silika-gela<sup>117,144,145</sup>. Hidrofilna površina molekula predstavlja sumu svih hidrofilnih delova na površini molekula i mera je njegovih hidrofilnih svojstava<sup>146</sup>. Polarniji molekuli imaju veće vrednosti HLB parametra,

rastvorljiviji su u mobilnoj fazi i pokazuju slabiju retenciju u uslovima reverzno-fazne hromatografije. Ovo potvrđuje negativan koeficijent pomenutog parametra u dobijenim modelima.

U PLS modelu se kao varijabla koja značajno utiče na retencioni parametar pojavljuje i MW. Molekulska težina je najčešće korišćeni deskriptor za opisivanje veličine molekula jer za njegovo izračunavanje nisu neophodni optimizacija molekula i proračuni. Vrednosti ostalih parametara za opisivanje veličine i oblika molekula, kao što su zapremina, površina, refraktivnost, polarizabilnost, su različite u odnosu na MW, ali informacije koje pomenute varijable pružaju su praktično iste<sup>147</sup>. U slučaju ispitivanih arilpiperazina MW pozitivno utiče na  $R_M^0$  vrednosti, odnosno sa povećanjem težine molekula raste njegova lipofilnost. Kod PCR modela MW značajno utiče na PC1 varijablu, koja posle PC2 ima najveći korelacioni koeficijent sa  $y$ -vrednostima.

Lipofilnost i permeabilnost nekog jedinjenja u velikoj meri zavise od sposobnosti za vodonično vezivanje<sup>147,148</sup>. Fizičko-hemijski deskriptori koji određuju broj donora i akceptora vodoničnih veza u molekulu, imaju značajan uticaj na PC2. U dobijenom PCR modelu, međutim, HBA i HBD parametri imaju izrazito niske vrednosti regresionih koeficijenata (približno jednake nuli), zbog čega je njihov uticaj na model zanemarljiv. U PLS i MLR modele pomenuti deskriptori nisu uključeni. Ovi deskriptori, naime, ne opisuju jačinu vodoničnog vezivanja niti uzimaju u obzir postojanje intramolekulskog vodoničnog vezivanja. Ovakve informacije sadržane su u sofisticiranijim deskriptorima, kao što su PSA, HSA i ST, a koje imaju značajan uticaj na lipofilnost arilpiperazina<sup>149</sup>.

U dobijenom MLR modelu kao deskriptor koji u značajnoj meri utiče na lipofilnost arilpiperazina javlja se i ukupni dipol. Ovaj parametar opisuje sposobnost molekula da učestvuje u dipol-dipol i dipol-indukovani dipol interakcijama sa komponentama mobilne i stacionarne faze. Zbog negativnog regresionog koeficijent ovog parametra u modelu sa smanjenjem njegove vrednosti povećava  $R_M^0$  vrednost, zbog čega su pomenute interakcija javljaju između jedinjenja i mobilne faze, više nego između jedinjenja i hidrofobne

stacionarne faze. Značaj uticaj dipolnog momenta na retencioni parametar potvrđen je i u literaturi<sup>150,151</sup>.

Kao varijable koje značajno utiču na sva tri predložena modela javljaju se i topološki deskriptori (indeksi konektivnosti i indeksi valentnosti). Ovi parametri opisuju veličinu i oblik molekula, račvanja u strukturi, nezasićenja, prisustvo heteroatoma i prstenova, broj atoma i način na koji su oni povezani u molekulu. Struktura nekog jedinjenja se predstavlja pomoću molekulskog grafika u kojem su atomi prikazani kao tačke (rogljevi), a veze kao ivice. Veza između atoma može biti opisana različitim topološkim matricama koje matematičkim rešavanjem daju jednu vrednost, označenu kao topološki indeks. Valencioni indeksi dodatno koriste broj valentnih elektrona u atomu<sup>152</sup>. Topološki deskriptori koji uključuju valentne i nevalentne molekulске indekse konektivnosti igraju značajnu ulogu u QSRR analizama. Veliki broj radova potvrđuje primenu topoloških deskriptora za predviđanje retencionih parametara različitih klasa organskih jedinjenja<sup>153</sup>. U modelima koji utvrđuju zavisnost parametra lipofilnosti arilpiperazina sa njihovom strukturom, topološki parametri imaju pozitivan regresioni koeficijent, odnosno povećanje veličine molekula povećava površinu za interakciju sa stacionarnom fazom što vodi ka jačanju retencije.

Modeli koji opisuju zavisnost retencije serije arilpiperazina i molekulskih deskriptora, potvrđuju činjenicu da je lipofilnost molekula određena dvema komponentama, veličinom molekula i sposobnošću građenja vodoničnih veza<sup>147</sup>. Veličina molekula je opisana molekulskom težinom ili topološkim deskriptorima, dok su elektrostatičke interakcije (jon-jon, jon-dipol, dipol-dipol interakcije), kao i vodonično vezivanje prikazane pomoću deskriptora PSA, HSA, HLB i ST. Negativni regresioni koeficijenti parametara rastvorljivosti (PSA, HSA, HLB i ST) imaju veći uticaj na  $R_M^0$  vrednosti od parametara koji opisuju veličinu molekula. Na osnovu ovakvih rezultata može se pretpostaviti da je u posmatranom hromatografskom sistemu particija dominantni mehanizam razdvajanja u odnosu na adsorpciju. Statistički parametri koji opisuju prediktivna svojstva dobijenih modela potvrđuju mogućnost njihovog korišćenja za predviđanje lipofilnosti novih, strukturno sličnih jedinjenja.

# 5

# Zaključak

## 5. Zaključak

U radu je proučeno hromatografsko ponašanje serije novosintetisanih arilpiperazina u uslovima reverzno-fazne planarne hromatografije. Dobijeni retencioni podaci su procenjeni na osnovu multivarijantnih statističkih metoda i utvrđena je veza između retencionih karakteristika i fizičko-hemijskih parametara ispitivanih molekula.

Zahvaljujući amfifilnim svojstvima ispitivanih supstanci, različiti reverzno-fazni sistemi uspešno su primenjeni za njihovo razdvajanje. Utvrđeno je da su razlike u retenciji analiziranih jedinjenja posledica prisustva dodatnog atoma sumpora u molekulu derivata benzimidazol-2-tiona, prisustva atoma halogena (Cl, Br) vezanog za benzimidazolsku i benzimidazol-2-tionsku dopaminergičku farmakoforu i različite prirode supstituenta na benzenovom prstenu vezanom za piperazinski deo molekula. Ponašanje arilpiperazina u primenjenim hromatografskim uslovima u saglasnosti je sa njihovim strukturnim svojstvima, kao i sa interakcijama za koje je potvrđeno da se uspostavljaju kada se ligand koji pokazuje afinitet prema dopaminskim D2 i serotoninским 5-HT<sub>1A</sub> receptorima približi aktivnom centru receptora.

Lipofilnost ispitivanih arilpiperazina određena je pomoću četiri hromatografska deskriptora,  $R_M^0$ ,  $b$ ,  $C_0$  i  $PC1$ . Deskriptori su izračunati za sva tri primenjena hromatografska sistema. Takođe, utvrđena je linearna zavisnost između parametara lipofilnosti i izračunatih  $\log P$  vrednosti u cilju potvrde mogućnosti primene RPTLC retencionog parametra za određivanje lipofilnosti. Na osnovu vrednosti statističkih parametara pomenutih korelacija, kao i najveće selektivnosti u razdvajanju ispitivanih arilpiperazina, kao najbolji sistem za procenu lipofilnosti ispitivanih arilpiperazina odabran je sistem koji sadrži DMSO kao organsku komponentu mobilne faze. Rezultati dobijeni primenom pomenutog hromatografskog sistema korišćeni su za dalje analize. Takođe je utvrđeno da varijabla koja direktno opisuje raspodelu supstanci između stacionarne i mobilne faze, kao što je parameter  $R_M^0$ , pouzdanije određuje lipofilnost ispitivanih

arilpiperazina od parametra koji je proporcionalan hidrofobnoj površini molekula (*b*) ili interpolisane veličine koja obuhvata sve hromatografske podatke jednog sistema (*PC1*).

U cilju boljeg razumevanja utvrđenog retencionog ponašanja ispitivanih arilpiperazina urađena je uvodna multivarijantna analiza. Analiza glavnih komponenta i klasteraska analiza su primenjene radi utvrđivanja eventualnog prisustva rezultata koji su posledica grubih grešaka, kao i moguće međusobne zavisnosti između objekata (ispitivanih jedinjenja) i varijabli (retencionih podataka i molekulskih deskriptora). PCA potvrđuje da je ispitivana jedinjenja moguće klasifikovati saglasno njihovim strukturnim karakteristikama. Vrednosti prve dve glavne komponente pokazuju da se jedinjenja grupišu na osnovu prisustva tio-grupe u molekulu ili prisustva atoma halogena u benzimidazolskom, odnosno benzimidazol-2-tionskom delu molekula. Rezultati su potvrđeni klasteraskom analizom.

Hemometrijske metode MLR, PCR i PLS primenjene su za identifikaciju molekulskih deskriptora koji na najbolji način opisuju ponašanje ispitivanih jedinjenja u hromatografskom sistemu, kao i za kvantifikaciju njihovog uticaja, a u cilju dobijanja matematičkih modela koji bi mogli da predvide hromatografsko ponašanje, odnosno lipofilnost novih srodnih molekula. Statistički parametri potvrđuju da su predloženi modeli statistički značajni i uporedivi. Molekulski deskriptori koji su uključeni u konačne modele su slične prirode i njihov uticaj je približno jednak.

Pomoću modela koji opisuju zavisnost retencije serije arilpiperazina i molekulskih deskriptora utvrđeno je da je hromatografsko ponašanje molekula određeno dvema komponentama, veličinom molekula i sposobnošću građenja vodoničnih veza. Veličinu molekula opisuju parametri molekulska težina ili topološki deskriptori, dok elektrostatičke interakcije (jon-jon, jon-dipol, dipol-dipol interakcije) i vodonično vezivanje određuju molekulski deskriptori PSA, HSA, HLB i ST. Dobijeni modeli su pokazali da negativni regresioni koeficijenti parametara rastvorljivosti (PSA, HSA, HLB i ST) imaju veći uticaj na  $R_M^0$  vrednosti od parametara koji opisuju veličinu molekula. Na osnovu ovakvih rezultata pretpostavljeno je da je u posmatranom hromatografskom sistemu particija dominantni mehanizam razdvajanja u odnosu na adsorpciju. Na osnovu analize statističkih

parametara koji opisuju prediktivna svojstva dobijenih modela potvrđena je mogućnost njihovog korišćenja za predviđanje lipofilnosti novih, strukturno sličnih jedinjenja.

# 6

# Literatura



## 6. Literatura

1. J. Trifković, Uticaj supstituenata na aromatičnim jezgrima 1-aralkil-4-arilpiperazina na njihovu retenciju u uslovima planarne hromatografije, Magistarski rad, Hemijski fakultet, Beograd, 2006.
2. J. Veličković, D. Andrić, G. Roglić, Ž. Tešić, D. Milojković-Opsenica, Planar chromatography of some 1-arylpiperazines behaving as dopaminergic ligands, *Journal of Planar Chromatography – Modern TLC*, 17 (2004) 255–260.
3. J. Veličković, Ž. Tešić, D. Milojković-Opsenica, Evaluation of the lipophilicity of some 1-arylpiperazines by planar chromatography, *Journal of Planar Chromatography – Modern TLC*, 18 (2005) 358–363.
4. F.Lj. Andrić, J. Trifković, Ž.Lj. Tešić, D.M. Milojković-Opsenica, An approximate linear solvation energy relationships model based on Snyder's selectivity parameters. Chromatographic behavior of some 1-aralkyl-4-arylpiperazines, *Chromatographia*, 68 (2008) 453–458.
5. N. Grinberg, *Modern thin-layer chromatography*, Chromatography Science, Vol 52. Marcel Dekker, Inc. New York, USA, 1990.
6. J.M. Miller, *Chromatography – Concepts and contrasts*, John Wiley & Sons, Hoboken, New Jersey, 2005.
7. Sz. Nyiredy, *Planar chromatography, a retrospective view for the third millenium*, Research Institute for medicinal plants, Springer, 2001.
8. C.F. Poole, Thin-layer chromatography: challenges and opportunities, *Journal of Chromatography A*, 1000 (2003) 963–984.
9. D. Milojković-Opsenica, Sorbenti u hromatografiji na tankom sloju, *Hemijski pregled*, 41 (2000) 38-43.
10. J.J. Kirkland, Development of some stationary phases for reversed-phase high-performance liquid chromatography, *Journal of Chromatography A*, 1060 (2004) 9–21.

11. H.A. Claessens, M.A. Van Straten, Review on the chemical and thermal stability of stationary phases for reversed-phase liquid chromatography, *Journal of Chromatography A*, 1060 (2004) 23–41.
12. R.J.M. Vervoort, A.J.J. Debets, H.A. Claessens, C.A. Cramers, G.J. de Jong, Optimisation and characterisation of silica-based reversed-phase liquid chromatographic systems for the analysis of basic pharmaceuticals, *Journal of Chromatography A*, 897 (2000) 1–22.
13. S. Šegan, F. Andrić, A. Radoičić, D. Opsenica, B. Šolaja, M. Zlatović, D. Milojković-Opsenica, Correlation between structure, retention and activity of cholic acid derived *cis–trans* isomeric bis-steroidal tetraoxanes, *Journal of Separation Science*, 34 (2011) 2659–2667.
14. A. Vailaya, Fundamentals of Reversed Phase Chromatography: Thermodynamic and Exothermodynamic Treatment, *Journal of Liquid Chromatography & Related Technologies*, 28 (2005) 965–1054.
15. P. Nikitas, A. Pappa-Louisi, Retention models for isocratic and gradient elution in reversed-phase liquid chromatography, *Journal of Chromatography A*, 1216 (2009) 1737–1755.
16. C. Horváth, W. Melander, I. Molnar, Solvophobic interactions in liquid chromatography with nonpolar stationary phase, *Journal of Chromatography*, 125 (1976) 129-156.
17. A. Vailaya, C. Horvath, Solvophobic theory and normalized free energies of nonpolar substances in reversed phase chromatography, *Journal of Physical Chemistry B*, 101 (1997) 5875-5888.
18. J.G. Dorsey, K.A. Dill, The molecular mechanism of retention in reversed-phase liquid chromatography, *Chemical Reviews*, 89 (1989) 331-346.
19. L.C. Tan, P.W. Carr, Revisionist look at solvophobic driving forces in reversed-phase liquid chromatography II. Partitioning vs. adsorption mechanism in monomeric alkyl bonded phase supports, *Journal of Chromatography A*, 775 (1997) 1–12.

20. P.W. Carr, L.C. Tan, J.H.J. Park, Revisionist look at solvophobic driving forces in reversed-phase liquid chromatography III. Comparison of the behavior of nonpolar and polar solutes, *Journal of Chromatography A*, 724 (1996) 1-12.
21. J.H. Park, Y.K. Lee, Y.C. Weon, L.C. Tan, J. Li, L. Li, J.F. Evans, P.W. Carr, Revisionist look at solvophobic driving forces in reversed-phase liquid chromatography IV. Partitioning vs. adsorption mechanism on various types of polymeric bonded phases, *Journal of Chromatography A*, 767 (1997) 1-10.
22. A. Vailaya, C. Horvath, Retention in reversed-phase chromatography: partition or adsorption? *Journal of Chromatography A*, 829 (1998) 1–27.
23. C.F. Poole, N.C. Dias, Practitioner's guide to method development in thin-layer chromatography, *Journal of Chromatography A*, 892 (2000) 123–142.
24. A.A. Amiri, B. Hemmateenejad, A. Safavi, H. Sharghi, A.R.S. Bena, M. Shamsipur, Structure–retention and mobile phase–retention relationships for reversed-phase high-performance liquid chromatography of several hydroxythioxanthone derivatives in binary acetonitrile–water mixtures, *Analytica Chimica Acta*, 605 (2007) 11–19.
25. J.R. Torres-Lapasio, M.J. Ruiz-Angel, M.C. Garcia-Alvarez-Coque, Comparative study of solvation parameter models accounting the effects of mobile phase composition in reversed-phase liquid chromatography, *Journal of Chromatography A*, 1166 (2007) 85–96.
26. S. Gocan, Mobile phases in thin-layer chromatography. U: N. Grinberg, *Modern thin-layer chromatography*, Chromatography Science, Vol 52. Marcel Dekker, Inc. New York, USA, 1990, str. 139-174.
27. L.R. Snyder, Classification of the solvent properties of common liquids, *Journal of Chromatography A*, 92 (1974) 223–230.
28. M.J. Kamlet, R.W. Taft, The solvatochromic comparison method. I. The  $\beta$ -Scale of solvent hydrogen-bond acceptor (HBA) basicities, *Journal of the American Chemical Society*, 98 (1976) 377-383.

29. R.W. Taft, M.J. Kamlet, The solvatochromic comparison method. 2. The  $\alpha$ -Scale of solvent hydrogen-bond donor (HBD) acidities, *Journal of the American Chemical Society*, 98 (1976) 2886-2894.
30. M.J. Kamlet, J.L. Abboud, R.W. Taft, The solvatochromic comparison method. 6. The  $\pi^*$  scale of solvent polarities, *Journal of the American Chemical Society*, 99 (1977) 6027-6038.
31. M.H. Abraham, C.F. Poole, S.K. Poole, Classification of stationary phases and other materials by gas chromatography, *Journal of Chromatography A*, 842 (1999) 79–114.
32. F.Lj. Andrić, Mapiranje interakcija ostvarenih u uslovima reverzno-fazne tečne hromatografije primenom dizajna smeše i Snyder-ovog koncepta polarnosti i selektivnosti rastvarača, Magistarski rad, Hemijski fakultet, Beograd, 2009.
33. A.P. Hill, R.J. Young, Getting physical in drug discovery: a contemporary perspective on solubility and hydrophobicity, *Drug Discovery Today*, 15 (2010) 648-655.
34. Y. Henchoz, B. Bard, D. Guillarme, P.A. Carrupt, J.L. Veuthey, S. Martel, Analytical tools for the physicochemical profiling of drug candidates to predict absorption/distribution, *Analytical & Bioanalytical Chemistry*, 394 (2009) 707–729.
35. M.J. Waring, Defining optimum lipophilicity and molecular weight ranges for drug candidates-Molecular weight dependent lower logD limits based on permeability, *Bioorganic & Medicinal Chemistry Letters*, 19 (2009) 2844–2851.
36. S.K. Poole, C.F. Poole, Separation methods for estimating octanol–water partition coefficients, *Journal of Chromatography B*, 797 (2003) 3–19.
37. D. Vrakasa, A. Tsantili-Kakoulidou, D. Hadjipavlou-Litinab, Exploring the consistency of logP estimation for substituted coumarins, *QSAR & Combinatorial Science*, 22 (2003) 622-629.
38. P. Gaillard, P. Carrupt, B. Testa, The conformation-dependent lipophilicity of morphine glucuronides as calculated from their molecular lipophilicity potential, *Bioorganic & Medicinal Chemistry Letters*, 4 (1994) 737-742.

39. A. Tsantili-Kakoulidou, A. Varvaresou, T. Siatra-Papastaikoudi, O.A. Raevsky, A comprehensive investigation of the partitioning and hydrogen bonding behavior of indole containing derivatives of 1,3,4-thiadiazole and 1,2,4-triazole by means of experimental and calculative approaches, *Quantitative Structure-Activity Relationships*, 18 (1999) 482-489.
40. D. Dellis, C. Giaginis, A. Tsantili-Kakoulidou, Physicochemical profile of nimesulide exploring the interplay of lipophilicity, solubility and ionization, *Journal of Pharmaceutical and Biomedical Analysis*, 44 (2007) 57–62.
41. C. Hansch, T. Fujita,  $\rho$ - $\sigma$ - $\pi$  analysis. A method for the correlation of biological activity and chemical structure, *Journal of the American Chemical Society*, 86 (1964) 1616–1626.
42. A. Leo, C. Hansch, D. Elkins, Partition coefficients and their uses, *Chemical Reviews*, 71 (1971) 525–616.
43. C. Giaginis, A. Tsantili-Kakoulidou, Alternative measures of lipophilicity: from octanol–water partitioning to IAM retention, *Journal of Pharmaceutical Sciences*, 97 (2008) 2984–3004.
44. Ł. Komsta, R. Skibiński, A. Berecka, A. Gumieniczek, B. Radkiewicz, M. Radon, Revisiting thin-layer chromatography as a lipophilicity determination tool - A comparative study on several techniques with a model solute set. *Journal of Pharmaceutical and Biomedical Analysis*, 53 (2010) 911–918.
45. L-G. Danielsson, Y-H. Zhang, Methods for determining n-octanol-water partition constants, *Trends in Analytical Chemistry*, 15 (1996) 188-196.
46. A. Finizio, M. Vighi, D. Sandroni, Determination of n-octanol/water partition coefficient ( $K_{ow}$ ) of pesticide critical review and comparison of methods, *Chemosphere*, 34 (1997) 131–161.
47. H. Van de Waterbeemd, M. Kansy, B. Wagner, H. Fischer, Lipophilicity measurement by high performance liquid chromatography (RPHPLC). U: V. Pilska, B. Testa, H. van de Waterbeemd, Lipophilicity in drug action and toxicology. Weinheim, Germany, VCH Publishers, 1996, str. 73–87.

48. T. Hartmann, J. Schmitt, Lipophilicity – beyond octanol/water: a short comparison of modern technologies, *Drug Discovery Today: Technologies*, 1 (2004) 431-439.
49. C. Giaginis, A. Tsantili-Kakoulidou, Current state of the art in HPLC methodology for lipophilicity assessment of basic drugs. A review, *Journal of Liquid Chromatography & Related Technologies*, 31 (2008) 79–96.
50. E.C. Bate-Smith, R.G. Westall, Chromatographic behavior and chemical structure.1. Some naturally occurring phenolic substances, *Biochimica et Biophysica Acta*, 4 (1950) 427–440.
51. E. Soczewinski, G. Matysik, Two types of  $R_M$ -composition relationships in liquid-liquid partition chromatography, *Journal of Chromatography*, 32 (1968) 458–471.
52. J.R. Torres-Lapasió, M. Rosés, E. Bosch, M.C. Garcia-Alvarez-Coque, Interpretive optimisation strategy applied to the isocratic separation of phenols by reversed-phase liquid chromatography with acetonitrile–water and methanol–water mobile phases, *Journal of Chromatography A*, 886 (2000) 31–46.
53. S.B. Gaica, D.M. Opsenica, B.A. Šolaja, Ž.Lj. Tešić, D.M. Milojković-Opsenica, The retention behavior of some cholic acid derivatives on different adsorbents, *Journal of Planar Chromatography - Modern TLC*, 15 (2002) 299-305.
54. P.J. Shoenmakers, H.A.H. Billiet, L. De Galan, Influence of organic modifiers on the retention behaviour in reversed-phase liquid chromatography and its consequences for gradient elution, *Journal of Chromatography A*, 185 (1979) 179–195.
55. J.Đ. Trifković, F.Lj. Andrić, P. Ristivojević, D. Andrić, Ž.Lj. Tešić, D.M. Milojković-Opsenica, Structure-retention relationship study of arylpiperazines by linear multivariate modeling, *Journal of Separation Science*, 33 (2010) 2619-2628.
56. D. Dabić, M. Natić, Z. Džambaski, M. Stojanović, R. Marković, D. Milojković-Opsenica, Ž. Tešić, Estimation of lipophilicity of N-substituted 2-alkylidene-4-oxothiazolidines by means of reversed-phase thin-layer chromatography, *Journal of Liquid Chromatography & Related Technology*, 34 (2011) 791-804.
57. F. Murakami, Retention behavior of benzene derivatives on bonded reversed-phase columns, *Journal of Chromatography*, 178 (1979) 393–399.

58. C. Horváth, W. Melander, I. Molnar, Solvophobic interactions in liquid chromatography with nonpolar stationary phases, *Journal of Chromatography*, 125 (1976) 129–156.
59. M.L. Bieganowska, A. Doraczynska-Szopa, A. Petruczynik, The retention behavior of some sulfonamides on different TLC plates. 2. Comparison of the selectivity of the systems and quantitative determination of hydrophobicity parameters, *Journal of Planar Chromatography – Modern TLC*, 8 (1995) 122–128.
60. H.F. Pop, C. Sârbu, A new fuzzy regression algorithm, *Analytical Chemistry*, 68 (1996) 771–778.
61. C. Sârbu, K. Kuhajda, S. Kevrešan, Evaluation of the lipophilicity of bile acids and their derivatives by thin-layer chromatography and principal component analysis. *Journal of Chromatography A*, 917 (2001) 361–366.
62. C. Onisor, M. Poša, S. Kevrešan, K. Kuhajda, C. Sârbu, Estimation of chromatographic lipophilicity of bile acids and their derivatives by reversed-phase thin layer chromatography, *Journal of Separation Science*, 33 (2010) 3110–3118.
63. D. Casoni, A. Kot-Wasik, J. Namiesnik, C. Sârbu, Lipophilicity data for some preservatives estimated by reversed-phase liquid chromatography and different computation methods, *Journal of Chromatography A*, 1216 (2009) 2456–2465.
64. F.Lj. Andrić, J.Đ. Trifković, A.D. Radoičić, S.B. Šegan, Ž.Lj. Tešić, D.M. Milojković-Opsenica, Determination of the soil-water partition coefficients ( $\log K_{OC}$ ) of some mono- and poly-substituted phenols by reversed-phase thin-layer chromatography, *Chemosphere*, 81 (2010) 299–305.
65. J. Odović, K. Karljiković-Rajić, J. Trbojević-Stanković, B. Stojimirović, S. Vladimirov, The lipophilicity examination of some ACE inhibitors and hydrochlorothiazide on cellulose in RP thin-layer chromatography, *Iranian Journal of Pharmaceutical Research*, 11 (2012) 763–770.
66. R. Mannhold, C. Ostermann, Prediction of  $\log P$  with substructure - based methods. U: *Molecular drug properties measurement and prediction*, Editor: R. Mannhold. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2008, str. 357–379.

67. I.V. Tetko, G.I. Poda, Prediction of log P with property - based methods. U: Molecular drug properties measurement and prediction, Editor: R. Mannhold. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, 2008, str. 381-406.
68. T. Fujita, J. Iwasa, C. Hansch, A New Substituent Constant,  $\pi$ , Derived from partition coefficients, *Journal of American Chemical Society*, 86 (1964) 5175-5180.
69. R.F. Rekker, The hydrophobic fragmental constant, *Pharmacochemistry Library*, Vol. 1, Elsevier, New York, 1977.
70. C. Hansch, A.J. Leo, Substituent constants for correlation analysis in chemistry and biology, John Wiley, New York, 1979.
71. R. Kaliszan, H. Foks, The relationship between the  $R_M$  values and the connectivity indices for pyrazine carbothioamide derivatives, *Chromatographia*, 10 (1977) 346-349.
72. R. Kaliszan, Correlation between the retention indices and the connectivity indices of alcohols and methyl esters with complex cyclic structure, *Chromatographia*, 10 (1977) 529-531.
73. Q.S. Wang, L. Zhang, Review of research on quantitative structure-retention relationships in thin-layer chromatography, *Journal of Liquid Chromatography & Related Technologies*, 22 (1999)1-14.
74. K. Heberger, Quantitative structure-(chromatographic) retention relationships, *Journal of Chromatography A*, 1158 (2007) 273-305.
75. L. Zhang, Q-S. Wang, Quantitative structure-retention relationships: TLC analysis, U: *Encyclopedia of chromatography* (3rd Edition). CRC Press, Taylor & Francis Group, Boca Raton, Florida, USA, 2010, str. 1980-1982.
76. R. Put, Y. Vander Heyden, Review on modelling aspects in reversed-phase liquid chromatographic quantitative structure-retention relationships, *Analytica Chimica Acta*, 602 (2007) 164-172.
77. R. Kaliszan, Chromatography and capillary electrophoresis in modelling the basic processes of drug action, *Trends in Analytical Chemistry*, 18 (1999) 400-410.



78. Y. Polyakova, K.H. Row, Quantitative structure-retention relationships applied to reversed-phase high-performance liquid chromatography, *Medicinal Chemistry Research*, 14 (2005) 488-522.
79. Y. Polyakova, L.M. Jin, K.H. Row, QSPR models for chromatographic retention of some azoles with physicochemical properties, *Bulletin of Korean Chemical Society*, 27 (2006) 211-218.
80. K. Varmuza, P. Filzmoser, Introduction to multivariate statistical analysis in chemometrics. CRC Press, Taylor & Francis Group, Boca Raton, Florida, USA, 2009.
81. L-T. Qin, S-S. Liu, F. Chen, Q-F. Xiao, Q-S. Wu, Chemometric model for predicting retention indices of constituents of essential oils, *Chemosphere*, 90 (2013) 300–305.
82. T. Cserhati, Relationship between the physicochemical parameters of 3,5-dinitrobenzoic acid esters and their retention behaviour on  $\beta$ -cyclodextrin polymer support, *Analytica Chimica Acta*, 292 (1994) 17-22.
83. A.A. Qasimullah, S. Muzaffar, A. Andrabi, P.M. Qureshi, Solvent polarity as a function of  $R_F$  in thin-layer chromatography of selected nitro functions, *Journal of Chromatographic Science*, 34 (1996) 376-378.
84. A. Pyka, Topological indexes and  $R_M$  values for prediction of the pKa values of isomeric methylphenols, chlorophenols, and nitrophenols. Part X, *Journal of Planar Chromatography – Modern TLC*, 9 (1996) 52-55.
85. M. Koba, T. Baczek, M.P. Marszał, Importance of retention data from affinity and reverse-phase high-performance liquid chromatography on antitumor activity prediction of imidazoacridinones using QSAR strategy, *Journal of Pharmaceutical and Biomedical Analysis*, 64– 65 (2012) 87– 93.
86. R. Todeschini, V. Consonni, Molecular descriptors for chemoinformatics, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2009.
87. R. Todeschini, V. Consonni, Handbook of molecular descriptors, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2000.

88. P.M. Andersson, M. Sjoström, S. Wold, T. Lundstedt, Comparison between physicochemical and calculated molecular descriptors, *Journal of Chemometrics*, 14 (2000) 629–642.
89. M.M. Natić, Hemometrijski pristup proučavanju retencionog ponašanja diastereomernih 2-alkiliden-4-okskotiazolidina, Doktorska disertacija, Hemijski fakultet, Beograd, 2008.
90. M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, AM1: A new general purpose quantum mechanical molecular model, *Journal of American Chemical Society*, 107 (1985) 3902-3909.
91. K.M. Gough, K. Belohorcova, K.L.E. Kaiser, Quantitative structure-activity relationships (QSARs) of *Photobacterium phosphoreum* toxicity of nitrobenzene derivatives, *The Science of the Total Environment*, 142 (1994) 179-190.
92. T. Sotomatsu, Y. Murata, T. Fujita, Correlation analysis of substituent effects on the acidity of benzoic acids by the AM1 method, *Journal of Computational Chemistry*, 10 (1989) 94–98.
93. H. Van de Waterbeemd, *Chemometric methods in molecular design*, VCH Verlagsgesellschaft mbH, Weinheim, Federal Republic of Germany, 1995.
94. C. Sarbu, S. Todor, Determination of lipophilicity of some non-steroidal antiinflammatory agents and their relationships by using principal component analysis based on thin-layer chromatographic retention data, *Journal of Chromatography A*, 822 (1998) 263–269.
95. C. Sarbu, D. Casoni, M. Darabantu, C. Maieranu, Quantitative structure—retention and retention—activity relationships of some 1,3-oxazolidine systems by RP-HPTLC and PCA, *Journal of Pharmaceutical and Biomedical Analysis*, 35 (2004) 213–219.
96. R.D. Briciu, A. Kot-Wasik, A. Wasik, J. Namiesnik, C. Sârbu, The lipophilicity of artificial and natural sweeteners estimated by reversed-phase thin-layer chromatography and computed by various methods, *Journal of Chromatography A*, 1217 (2010) 3702–3706.

97. D. Dabić, M. Natić, Z. Džambaski, R. Marković, D. Milojković-Opsenica, Ž. Tešić, Quantitative structure–retention relationship of new N-substituted 2-alkylidene-4-oxothiazolidines, *Journal of Separation Science*, 34 (2011) 2397-2404.
98. T. Đaković-Sekulić, A. Smoliński, N. Trišović, G. Ušćumlić, Multivariate evaluation of the correlation between retention data and molecular descriptors of antiepileptic hydantoin analogs, *Journal of Chemometrics*, 26 (2012) 95–107.
99. F. Tache, R.D. Nascu-Briciu, C. Sârbu, F. Micale, A. Medvedovici, Estimation of the lipophilic character of flavonoids from the retention behavior in reversed phase liquid chromatography on different stationary phases: A comparative study, *Journal of Pharmaceutical and Biomedical Analysis*, 57 (2012) 82– 93.
100. K. Heberger, Evaluation of polarity indicators and stationary phases by principal component analysis in gas–liquid chromatography, *Chemometrics and Intelligent Laboratory Systems*, 47 (1999) 41–49.
101. H. Kemp, M. Kempe, QSRR analysis of  $\beta$ -lactam antibiotics on a penicillin G targeted MIP stationary phase, *Analytical and Bioanalytical Chemistry*, 398 (2010) 3087–3096.
102. R.W. Kennard, L.A. Stone, Computer aided design of experiments, *Technometrics*, 11 (1969) 137-148.
103. R. G. Brereton, *Chemometrics. Data analysis for the laboratory and chemical plant*, John Wiley & Sons Ltd, Chichester, England, 2003.
104. R.N. Feudale, H. Tan, S.D. Brown, Piecewise orthogonal signal correction, *Chemometrics and Intelligent Laboratory Systems*, 63 (2002) 129– 138.
105. V. Sikirzhyski, K. Virkler, I.K. Lednev, Discriminant analysis of raman spectra for body fluid identification for forensic purposes, *Sensors*, 10 (2010) 2869-2884.
106. J.N. Miller, J.C. Miller, *Statistics and chemometrics for analytical chemistry*, Sixth edition, Pearson Education Limited, Harlow, England, 2010.
107. E. Anderssen, K. Dyrstad, F. Westad, H. Martens, Reducing over-optimism in variable selection by cross-model validation, *Chemometrics and Intelligent Laboratory Systems*, 84 (2006) 69–74.

108. M. Forina, S. Lanteri, M.C. Cerrato Oliveros, C. Pizarro Millan, Selection of useful predictors in multivariate calibration, *Analytical and Bioanalytical Chemistry*, 380 (2004) 397–418.
109. N. Perišić-Janjić, R. Kaliszan, N. Milošević, G. Ušćumlić, N. Banjac, Chromatographic retention parameters in correlation analysis with in silico biological descriptors of a novel series of N-phenyl-3-methyl succinimide derivatives, *Journal of Pharmaceutical and Biomedical Analysis*, 72 (2013) 65–73.
110. J. Yana, D-S. Caoa, F-Q. Guo, L-X. Zhang, M. He, J-H. Huang, Q-S. Xu, Y-Z. Liang, Comparison of quantitative structure–retention relationship models on four stationary phases with different polarity for a diverse set of flavor compounds, *Journal of Chromatography A*, 1223 (2012) 118–125.
111. C.I. De Matteis, D.A. Simpson, M.R. Euerby, P.N. Shawb, D.A. Barrett, Chromatographic retention behaviour of monosubstituted benzene derivatives on porous graphitic carbon and octadecyl-bonded silica studied using molecular modelling and quantitative structure–retention relationships, *Journal of Chromatography A*, 1229 (2012) 95–106.
112. A.G. Fragkakia, A. Tsantili-Kakoulidou, Y.S. Angelis, M. Koupparis, C. Georgakopoulos, Gas chromatographic quantitative structure–retention relationships of trimethylsilylated anabolic androgenic steroids by multiple linear regression and partial least squares, *Journal of Chromatography A*, 1216 (2009) 8404–8420.
113. V.K. Gupta, H. Khani, B. Ahmadi-Roudi, S. Mirakhorli, E. Fereyduni, S. Agarwal, Prediction of capillary gas chromatographic retention times of fatty acid methyl esters in human blood using MLR, PLS and back-propagation artificial neural networks, *Talanta*, 83 (2011) 1014–1022.
114. K. Bodziocha, A. Durand, R. Kaliszan, T. Baczek, Y. Vander Heyden, Advanced QSRR modeling of peptides behavior in RPLC, *Talanta*, 81 (2010) 1711–1718.
115. L-T. Qin, S-S. Liu, H-L. Liu, J. Tong, Comparative multiple quantitative structure–retention relationships modeling of gas chromatographic retention time of essential oils using multiple linear regression, principal component regression, and

- partial least squares techniques, *Journal of Chromatography A*, 1216 (2009) 5302–5312.
116. E.H.M.A. Rabtti, M.M. Natić, D.M. Milojković-Opsenica, J.Đ. Trifković, I.M. Vučković, V.E. Vajs, Ž.Lj. Tešić, RP TLC-based lipophilicity assessment of some natural and synthetic coumarins, *Journal of Brazilian Chemical Society*, 23 (2012) 522-530.
117. S. Šegan, J. Trifković, T. Verbić, D. Opsenica, M. Zlatović, J. Burnett, B. Šolaja, D. Milojković-Opsenica, Correlation between structure, retention, property, and activity of biologically relevant 1,7-bis(aminoalkyl)diazachrysene derivatives, *Journal of Pharmaceutical and Biomedical Analysis*, 72 (2013) 231-239.
118. G. Ermondi, G. Caron, Molecular interaction fields based descriptors to interpret and compare chromatographic indexes, *Journal of Chromatography A*, 1252 (2012) 84– 89.
119. A. Zhang, J.L. Neumeyer. R.J. Baldessarini, Recent progress in development of dopamine receptor subtype-selective agents: potential therapeutics for neurological and psychiatric disorders, *Chemical Reviews*, 107 (2007) 274-302.
120. V. Šoškić, J. Joksimović, Bioisosteric approach in the design of the new dopaminergic/serotonergic ligands, *Current Medicinal Chemistry*, 5 (1998) 493-512.
121. V. Šukalovic, M. Zlatović, D. Andrić, G. Roglić, S. Kostić-Rajačić, V. Šoškić, Modeling of the D<sub>2</sub> dopamine receptor arylpiperazine binding site for 1-{2-[5-(1H-benzimidazole-2-thione)]ethyl}-4-arylpiperazines, *Archiv der Pharmacie*, 337 (2004) 502-512.
122. D. Andrić, G. Roglić, V. Šukalović, V. Šoškić, S. Kostić-Rajačić, Synthesis, binding properties and receptor docking of 4-halo-6-[2-(4-arylpiperazin-1-yl)ethyl]-1H-benzimidazoles, mixed ligands of D<sub>2</sub> and 5-HT<sub>1A</sub> receptors, *European Journal of Medicinal Chemistry*, 43 (2008) 1696-1705.
123. D. Andrić, G. Tovilović, G. Roglić, V. Šoškić, M. Tomić, S. Kostić-Rajačić, 6-[2-(4-Arylpiperazin-1-y)]-4-halo-1,3-dihydro-2H-benzimidazole-2-thiones: synthesis

- and pharmacological evaluation, *Journal of Serbian Chemical Society*, 72 (2007) 747-755.
124. D. Andrić, G. Tovilović, G. Roglič, Đ. Vasković, V. Šoškić, M. Tomić, S. Kostić-Rajačić, Synthesis and pharmacological evaluation of several N-(2-nitrophenyl)piperazine derivatives, *Journal of Serbian Chemical Society*, 72 (2007) 429-435.
125. P. Metrangolo, H. Neukirch, T. Pilati, G. Resnati, Halogen bonding based recognition processes: a world parallel to hydrogen bonding, *Accounts of Chemical Research*, 38 (2005) 386-395.
126. P. Politzer, P. Lane, M.C. Concha, Y. Ma, J.S. Murray, An overview of halogen bonding, *Journal of Molecular Modeling*, 13 (2007) 305-311.
127. <http://www.norgwyn.com/mmpplus.html>
128. <http://www.vcclab.org>
129. I.V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V.A. Palyulin, E.V. Radchenko, N.S. Zefirov, A.S. Makarenko, V.Yu. Tanchuk, V.V. Prokopenko, Virtual computational chemistry laboratory – design and description, *Journal of Computer-Aided Molecular Design*, 19 (2005) 453–463.
130. S. de Jong, SIMPLS: An alternative approach to partial least square regression, *Chemometrics and Intelligent Laboratory System*, 18 (1993) 251-265.
131. B. Natalini, R. Sardella, E. Camaioni, S. Natalini, R. Pellicciari, Dominant factors affecting the chromatographic behaviour of bile acids, *Chromatographia*, 64 (2006) 343–349.
132. G.L. Biagi, A.M. Barbaro, A. Sapone, M. Recanatini, Determination of lipophilicity by means of reversed-phase thin-layer chromatography. II. Influence of the organic modifier on the slope of the thin-layer chromatographic equation, *Journal of Chromatography A*, 669 (1994) 246–253.
133. G.L. Biagi, A.M. Barbaro, A. Sapone, Determination of lipophilicity by means of reversed-phase thin-layer chromatography. I. Basic aspects and relationship between slope and intercept of TLC equations, *Journal of Chromatography A*, 662 (1994) 341–361.

134. A. Niewiadomy, A. Zabinska, J. Matysiak, J.K. Rozyło, Influence of modifier and molecular structure of some dihydroxythiobenzanilides on retention in reversed-phase high performance thin-layer chromatography, *Journal of Chromatography A*, 791 (1997) 237-243.
135. E.H.M.A. Rabtti, M.M. Natić, D.M. Milojković-Opsenica, J.Đ. Trifković, T. Tosti, I.M. Vučković, V. Vajs, Ž.LJ. Tešić, Quantitative structure-toxicity relationship study of some natural and synthetic coumarins using retention parameters, *Journal of Serbian Chemical Society*, 77 (2012) 1443-1456.
136. Z. Garkani-Nejad, Quantitative structure-retention relationship study of some phenol derivatives in gas chromatography, *Journal of Chromatographic Science*, 48 (2010) 317–323.
137. Z. Garkani-Nejad, M. Ahmadvand, Comparative QSRR modeling of nitrobenzene derivatives based on original molecular descriptors and multivariate image analysis descriptors, *Chromatographia*, 73 (2011)733–742.
138. S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemometrics and Intelligent Laboratory Systems*, 58 (2001) 109–130.
139. I-G. Chong, C-H. Jun, Performance of some variable selection methods when multicollinearity is present, *Chemometrics and Intelligent Laboratory Systems*, 78 (2005) 103 – 112.
140. D.T. Stanton, P.C. Jurs, Computer-assisted study of the relationship between molecular structure and surface tension of organic compounds, *Journal of Chemical Information and Modeling*, 32 (1992) 109-115.
141. R.C. Pasquali, M.P. Taurozzi, C. Bregni, Some considerations about the hydrophilic–lipophilic balance system, *International Journal of Pharmaceutics*, 356 (2008) 44–51.
142. R.M. Baošić, A.D. Radojević, Ž.LJ. Tešić, Prediction of the retention of  $\beta$ -diketonato complexes in TLC systems on silica gel by quantitative structure–retention relationships, *Journal of Serbian Chemical Society*, 75 (2010) 513–521.
143. J.V. Turner, B.D. Glass, S. Agatonović-Kustrin, Prediction of drug bioavailability based on molecular structure, *Analytica Chimica Acta*, 485 (2003) 89–102.

144. J. Kelder, P.D.J. Grootenhuis, D.M. Bayada, L.P.C. Delbressine, J-P. Ploemen, Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs, *Pharmaceutical research*, 16 (1999) 1514-1519.
145. D.E. Clark, Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration, *Journal of Pharmaceutical Science*, 88 (1999) 815-821.
146. Q. Cao, V. Garib, Q. Yu, D.W. Connell, M. Campitelli, Quantitative structure–property relationships (QSPR) for steroidal compounds of environmental importance, *Chemosphere*, 76 (2009) 453–459.
147. G. Camenischa, J. Alsenz, H. van de Waterbeemd, G. Folkers, Estimation of permeability by passive diffusion through Caco-2 cell monolayers using the drugs' lipophilicity and molecular weight, *European Journal of Pharmaceutical Sciences*, 6 (1998) 313–319.
148. O.A. Raevsky, K-J. Schaper, Quantitative estimation of hydrogen bond contribution to permeability and absorption processes of some chemicals and drugs, *European Journal of Medicinal Chemistry*, 33 (1998) 799-807.
149. P. Zakeri-Milani, H. Tajerzadeh, Z. Islambolchilar, S. Barzegar, H. Valizadeh, The relation between molecular properties of drugs and their transport across the intestinal membrane, *DARU*, 14 (2006) 164-171.
150. T. Körtvélyesi, M. Görgényi, K. Héberger, Correlation between retention indices and quantum-chemical descriptors of ketones and aldehydes on stationary phases of different polarity, *Analytica Chimica Acta*, 428 (2001) 73–82.
151. T. Baczek, R. Kaliszan, Predictive approaches to gradient retention based on analyte structural descriptors from calculation chemistry, *Journal of Chromatography A*, 987 (2003) 29–37.
152. R. Ghavami, S. Faham, QSRR models for Kovats' retention indices of a variety of volatile organic compounds on polar and apolar GC stationary phases using molecular connectivity indexes, *Chromatographia*, 72 (2010) 893–903.
153. A. Pyka, Topological indexes and QSRR methodology in pharmaceutical and chemical analysis, *Journal of AOAC International*, 95 (2012) 673-690.



**7**

**Prilog**

## Prilog

### Određivanje kvantitativnog odnosa strukture i retencije arilpiperazina primenom tečne hromatografije i multivarijantnih hemometrijskih metoda

Doktorska disertacija

Sadržaj

**Tabela P1.** Retencioni podaci za hromatografski sistem RP-18 / metanol : voda.

**Tabela P2.** Retencioni podaci za hromatografski sistem RP-18 / dioksan : voda.

**Tabela P3.** Retencioni podaci za hromatografski sistem RP-18 / DMSO : voda.

**Tabela P4.** Hromatografski deskriptori ( $R_M^0$ ,  $b$ ,  $C_0$ , and  $PC1$ ) izvedeni iz retencionih podataka dobijenih za sistem RP18 / Diox:voda.

**Tabela P5.** Hromatografski deskriptori ( $R_M^0$ ,  $b$ ,  $C_0$ , and  $PC1$ ) izvedeni iz retencionih podataka dobijenih za sistem RP18 / MeOH:voda.

**Tabela P6.** Hromatografski deskriptori ( $R_M^0$ ,  $b$ ,  $C_0$ , and  $PC1$ ) izvedeni iz retencionih podataka dobijenih za sistem RP18 / DMSO:voda.

**Tabela P7.** Izračunate log  $P$  vrednosti.

**Tabela P8.** Molekulski deskriptori korišćeni za multivarijantnu analizu.

**Tabela P9.** Korelaciona matrica vrednosti molekulskih deskriptora.

**Tabela P10.** Korelaciona matrica vrednosti molekulskih deskriptora uključenih u MLR model.

**Slika P1.** Grafik zavisnosti broja glavnih komponentata od svojstvenih vrednosti za PCA analizu na retencionim podacima.

**Slika P2.** Grafik zavisnosti broja glavnih komponentata od svojstvenih vrednosti za PCA analizu na vrednostima molekulskih deskriptora.

**Slika P3.** Grafik zavisnosti broja glavnih komponentata od odnosa RMSEC/RMSECV u PCR modelu.

**Slika P4.** Grafik zavisnosti reziduala od  $Y_{\text{modela}}$  za PCR model.

**Slika P5.** Grafik zavisnosti broja glavnih komponentata od odnosa RMSEC/RMSECV u PLS modelu.

**Slika P6.** Grafik zavisnosti reziduala od  $Y_{\text{modela}}$  za PLS model.

**Tabela P1.** Retencioni podaci za hromatografski sistem RP-18 / metanol : voda

Jedinjenje	80% <sup>a</sup>		85%		90%		95%		100%	
	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$
1	0,25	0,477	0,37	0,231	0,44	0,105	0,52	-0,035	0,61	-0,194
2	0,28	0,410	0,40	0,176	0,46	0,070	0,54	-0,070	0,64	-0,250
3	0,15	0,753	0,26	0,454	0,38	0,213	0,46	0,070	0,53	-0,052
4	0,21	0,575	0,29	0,389	0,39	0,194	0,48	0,035	0,56	-0,105
5	0,18	0,659	0,29	0,389	0,40	0,176	0,51	-0,017	0,63	-0,231
6	0,20	0,602	0,29	0,389	0,39	0,194	0,50	0,000	0,60	-0,176
7	0,26	0,454	0,38	0,213	0,44	0,105	0,52	-0,035	0,61	-0,194
8	0,13	0,826	0,24	0,501	0,37	0,231	0,44	0,105	0,52	-0,035
9	0,11	0,908	0,21	0,575	0,36	0,250	0,41	0,158	0,50	0,000
10	0,20	0,602	0,28	0,410	0,37	0,231	0,44	0,105	0,53	-0,052
11	0,20	0,602	0,31	0,347	0,41	0,158	0,52	-0,035	0,65	-0,269
12	0,27	0,432	0,35	0,269	0,45	0,087	0,53	-0,052	0,67	-0,308
13	0,36	0,213	0,46	0,070	0,53	-0,087	0,60	-0,231	0,70	-0,454
14	0,38	0,250	0,46	0,070	0,55	-0,052	0,63	-0,176	0,74	-0,368
15	0,24	0,501	0,37	0,231	0,46	0,070	0,58	-0,140	0,66	-0,288
16	0,20	0,602	0,32	0,327	0,40	0,176	0,51	-0,017	0,63	-0,231
17	0,38	0,213	0,45	0,087	0,54	-0,070	0,62	-0,213	0,70	-0,368
18	0,32	0,327	0,41	0,158	0,50	0,000	0,63	-0,231	0,75	-0,477
19	0,34	0,231	0,43	0,087	0,52	-0,035	0,60	-0,176	0,69	-0,432
20	0,37	0,288	0,45	0,122	0,52	-0,035	0,60	-0,176	0,73	-0,347
21	0,17	0,689	0,35	0,269	0,43	0,122	0,55	-0,087	0,64	-0,250
22	0,15	0,753	0,32	0,327	0,40	0,176	0,52	-0,035	0,61	-0,194
23	0,36	0,250	0,44	0,105	0,49	0,017	0,61	-0,194	0,69	-0,347
24	0,34	0,288	0,42	0,140	0,52	-0,035	0,59	-0,158	0,72	-0,410
25	0,42	0,140	0,53	-0,052	0,64	-0,250	0,70	-0,368	0,78	-0,550
26	0,21	0,575	0,30	0,368	0,45	0,087	0,52	-0,035	0,61	-0,194
27	0,25	0,477	0,34	0,288	0,47	0,052	0,55	-0,087	0,63	-0,231
28	0,41	0,158	0,51	-0,017	0,61	-0,194	0,67	-0,308	0,76	-0,501
29	0,43	0,122	0,54	-0,070	0,64	-0,250	0,69	-0,347	0,76	-0,501
30	0,34	0,288	0,46	0,070	0,56	-0,105	0,63	-0,231	0,74	-0,454
31	0,44	0,105	0,55	-0,087	0,65	-0,269	0,70	-0,368	0,79	-0,575
32	0,41	0,158	0,49	0,017	0,56	-0,105	0,68	-0,327	0,78	-0,550
33	0,43	0,122	0,53	-0,052	0,64	-0,250	0,69	-0,347	0,76	-0,501

<sup>a</sup>Udeo organske komponente u mobilnoj fazi

**Tabela P2.** Retencioni podaci za hromatografski sistem RP-18 / dioksan : voda

Jedinjenje	60% <sup>a</sup>		65%		70%		75%		80%	
	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$
1	0,34	0,288	0,46	0,070	0,54	-0,070	0,61	-0,194	0,70	-0,368
2	0,39	0,194	0,48	0,035	0,55	-0,087	0,65	-0,269	0,76	-0,501
3	0,24	0,501	0,38	0,213	0,49	0,017	0,59	-0,158	0,67	-0,308
4	0,46	0,070	0,55	-0,087	0,61	-0,194	0,70	-0,368	0,85	-0,753
5	0,37	0,231	0,47	0,052	0,57	-0,122	0,66	-0,288	0,77	-0,525
6	0,33	0,308	0,44	0,105	0,54	-0,070	0,60	-0,176	0,69	-0,347
7	0,37	0,231	0,46	0,070	0,55	-0,087	0,63	-0,231	0,75	-0,477
8	0,25	0,477	0,37	0,231	0,48	0,035	0,58	-0,140	0,67	-0,308
9	0,22	0,550	0,34	0,288	0,46	0,070	0,55	-0,087	0,65	-0,269
10	0,45	0,087	0,54	-0,070	0,60	-0,176	0,70	-0,368	0,85	-0,753
11	0,37	0,231	0,48	0,035	0,58	-0,140	0,68	-0,327	0,80	-0,602
12	0,47	0,052	0,56	-0,105	0,62	-0,213	0,71	-0,389	0,84	-0,720
13	0,34	0,288	0,43	0,122	0,58	-0,140	0,64	-0,250	0,75	-0,477
14	0,40	0,176	0,49	0,017	0,55	-0,087	0,69	-0,347	0,82	-0,659
15	0,24	0,501	0,39	0,194	0,49	0,017	0,64	-0,250	0,75	-0,477
16	0,21	0,575	0,35	0,269	0,45	0,087	0,63	-0,231	0,75	-0,477
17	0,47	0,052	0,56	-0,105	0,63	-0,231	0,73	-0,432	0,88	-0,865
18	0,34	0,288	0,44	0,105	0,57	-0,122	0,68	-0,327	0,81	-0,630
19	0,31	0,347	0,42	0,140	0,57	-0,122	0,64	-0,250	0,75	-0,477
20	0,40	0,176	0,48	0,035	0,54	-0,070	0,67	-0,308	0,78	-0,550
21	0,24	0,501	0,37	0,231	0,49	0,017	0,62	-0,213	0,74	-0,454
22	0,20	0,602	0,33	0,308	0,43	0,122	0,62	-0,213	0,71	-0,389
23	0,47	0,052	0,55	-0,087	0,62	-0,213	0,72	-0,410	0,87	-0,826
24	0,30	0,368	0,44	0,105	0,57	-0,122	0,65	-0,269	0,76	-0,501
25	0,44	0,105	0,52	-0,035	0,62	-0,213	0,71	-0,389	0,88	-0,865
26	0,23	0,525	0,35	0,269	0,51	-0,017	0,63	-0,231	0,72	-0,410
27	0,27	0,432	0,40	0,176	0,56	-0,105	0,66	-0,288	0,77	-0,525
28	0,43	0,122	0,51	-0,017	0,60	-0,176	0,70	-0,368	0,86	-0,788
29	0,47	0,052	0,56	-0,105	0,63	-0,231	0,74	-0,454	0,92	-1,061
30	0,42	0,140	0,50	0,000	0,57	-0,122	0,73	-0,432	0,87	-0,826
31	0,48	0,035	0,55	-0,087	0,63	-0,231	0,75	-0,477	0,95	-1,279
32	0,45	0,087	0,55	-0,087	0,64	-0,250	0,73	-0,432	0,91	-1,005
33	0,49	0,017	0,57	-0,122	0,66	-0,288	0,77	-0,525	0,96	-1,380

<sup>a</sup>Udeo organske komponente u mobilnoj fazi

**Tabela P3.** Retencioni podaci za hromatografski sistem RP-18 / DMSO : voda

Jedinjenje	80% <sup>a</sup>		85%		90%		95%		100%	
	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$	$R_F$	$R_M$
1	0,24	0,501	0,56	-0,105	0,68	-0,327	0,77	-0,525	0,88	-0,865
2	0,11	0,908	0,31	0,347	0,42	0,140	0,53	-0,052	0,71	-0,389
3	0,02	1,690	0,04	1,380	0,08	1,061	0,13	0,826	0,21	0,575
4	0,18	0,659	0,32	0,327	0,46	0,070	0,60	-0,176	0,72	-0,410
5	0,37	0,052	0,65	-0,269	0,77	-0,525	0,81	-0,630	0,92	-1,061
6	0,18	0,659	0,49	0,017	0,62	-0,213	0,73	-0,432	0,85	-0,753
7	0,05	1,279	0,15	0,753	0,22	0,550	0,30	0,368	0,48	0,035
8	0,01	1,996	0,04	1,380	0,07	1,123	0,11	0,908	0,19	0,630
9	0,01	1,996	0,03	1,510	0,07	1,123	0,16	0,720	0,25	0,477
10	0,16	0,720	0,30	0,368	0,45	0,087	0,59	-0,158	0,71	-0,389
11	0,21	0,575	0,36	0,250	0,51	-0,017	0,56	-0,105	0,79	-0,575
12	0,29	0,389	0,40	0,176	0,53	-0,052	0,60	-0,176	0,80	-0,602
13	0,17	0,689	0,30	0,368	0,46	0,070	0,55	-0,087	0,78	-0,550
14	0,12	0,865	0,31	0,347	0,41	0,158	0,56	-0,105	0,72	-0,410
15	0,27	0,432	0,59	-0,158	0,77	-0,525	0,89	-0,908	0,92	-1,061
16	0,04	1,380	0,12	0,865	0,25	0,477	0,45	0,087	0,60	-0,176
17	0,20	0,602	0,34	0,288	0,53	-0,052	0,67	-0,308	0,76	-0,501
18	0,23	0,525	0,38	0,213	0,50	0,000	0,60	-0,176	0,80	-0,602
19	0,16	0,720	0,30	0,368	0,45	0,087	0,55	-0,087	0,77	-0,525
20	0,11	0,908	0,30	0,368	0,40	0,176	0,55	-0,087	0,71	-0,389
21	0,05	1,279	0,16	0,720	0,31	0,347	0,52	-0,035	0,59	-0,158
22	0,03	1,510	0,11	0,908	0,24	0,501	0,43	0,122	0,57	-0,122
23	0,19	0,630	0,33	0,308	0,50	0,000	0,66	-0,288	0,75	-0,477
24	0,23	0,525	0,38	0,213	0,52	-0,035	0,60	-0,176	0,80	-0,602
25	0,30	0,368	0,40	0,176	0,55	-0,087	0,65	-0,269	0,82	-0,659
26	0,11	0,908	0,20	0,602	0,29	0,389	0,42	0,140	0,60	-0,176
27	0,01	1,996	0,01	1,996	0,02	1,690	0,03	1,510	0,06	1,195
28	0,32	0,327	0,43	0,122	0,57	-0,122	0,68	-0,327	0,85	-0,753
29	0,37	0,231	0,46	0,070	0,59	-0,158	0,71	-0,389	0,88	-0,865
30	0,14	0,788	0,23	0,525	0,33	0,308	0,45	0,087	0,60	-0,176
31	0,27	0,432	0,40	0,176	0,54	-0,070	0,68	-0,327	0,80	-0,602
32	0,51	-0,630	0,87	-0,826	0,92	-1,061	0,96	-1,380	0,98	-1,690
33	0,36	0,250	0,50	0,000	0,65	-0,269	0,77	-0,525	0,86	-0,788

<sup>a</sup>Udeo organske komponente u mobilnoj fazi

**Tabela P4.** Hromatografski deskriptori ( $R_M^0$ ,  $b$ ,  $C_0$ , and  $PC1$ ) izvedeni iz retencionih podataka dobijenih za sistem **RP18 / Diox:voda**,  $R$  – Pearson-ov korelacioni koeficijent,  $S$  – standardna greška modela,  $P$  – nivo značajnosti,  $n$  - broj merenja.

Jedinjenje	Hromatografski deskriptori i odgovarajući statistički parametri							
	$R_M^0$	$b$	$r$	$s$	$P$	$n$	$C_0$	$PC1$
1	2,15 ± 0,13	0,032 ± 0,002	-0,9952	0,028	0,0001	5	68,3 ± 5,6	-1,24
2	2,25 ± 0,16	-0,033 ± 0,002	-0,9933	0,036	0,0007	5	66,3 ± 6,5	-0,16
3	2,84 ± 0,21	-0,040 ± 0,003	-0,9915	0,048	0,0009	5	71,3 ± 7,6	-2,91
4	2,43 ± 0,42	-0,039 ± 0,006	-0,9662	0,094	0,0074	5	63,1 ± 14,6	1,95
5	2,46 ± 0,10	-0,037 ± 0,001	-0,9979	0,022	0,0001	5	66,5 ± 3,7	-0,03
6	2,19 ± 0,13	-0,032 ± 0,002	-0,9951	0,029	0,0004	5	68,9 ± 5,7	-1,55
7	2,31 ± 0,14	-0,034 ± 0,002	-0,9948	0,032	0,0005	5	67,1 ± 5,7	-0,59
8	2,78 ± 0,13	-0,039 ± 0,002	-0,9968	0,029	0,0002	5	71,5 ± 4,7	-3,08
9	2,93 ± 0,16	-0,040 ± 0,002	-0,9950	0,037	0,0004	5	72,7 ± 5,9	-3,91
10	2,52 ± 0,42	-0,040 ± 0,006	-0,9673	0,095	0,0071	5	63,5 ± 14,3	1,76
11	2,68 ± 0,15	-0,041 ± 0,002	-0,9961	0,033	0,0003	5	66,0 ± 4,9	0,44
12	2,29 ± 0,33	-0,037 ± 0,005	-0,9761	0,074	0,0044	5	62,5 ± 12,1	2,16
13	2,57 ± 0,16	-0,038 ± 0,002	-0,9946	0,036	0,0005	5	67,6 ± 5,9	-0,61
14	2,67 ± 0,34	-0,041 ± 0,005	-0,9793	0,077	0,0036	5	65,6 ± 11,5	0,61
15	3,36 ± 0,14	-0,048 ± 0,002	-0,9975	0,031	0,0002	5	69,9 ± 4,1	-2,00
16	3,69 ± 0,15	-0,052 ± 0,002	-0,9975	0,034	0,0002	5	70,9 ± 4,1	-2,77
17	2,71 ± 0,48	-0,043 ± 0,007	-0,9645	0,108	0,0080	5	62,7 ± 14,9	2,61
18	3,04 ± 0,16	-0,045 ± 0,002	-0,9960	0,037	0,0003	5	67,0 ± 5,0	0,03
19	2,78 ± 0,15	-0,041 ± 0,002	-0,9959	0,034	0,0003	5	68,2 ± 5,1	-0,89
20	2,37 ± 0,25	-0,036 ± 0,004	-0,9856	0,056	0,0021	5	66,0 ± 9,6	0,09
21	3,31 ± 0,07	-0,047 ± 0,001	-0,9994	0,015	<0,0001	5	70,3 ± 2,0	-2,34
22	3,59 ± 0,19	-0,050 ± 0,003	-0,9957	0,042	0,0003	5	71,7 ± 5,4	-3,38
23	2,61 ± 0,47	-0,042 ± 0,007	-0,9636	0,105	0,0083	5	62,9 ± 15,1	2,34
24	2,87 ± 0,15	-0,042 ± 0,002	-0,9960	0,035	0,0003	5	68,0 ± 5,1	-0,68
25	2,93 ± 0,52	-0,046 ± 0,008	-0,9634	0,117	0,0084	5	63,9 ± 15,3	1,99
26	3,34 ± 0,17	-0,047 ± 0,002	-0,9961	0,038	0,0003	5	70,6 ± 5,1	-2,39
27	3,27 ± 0,13	-0,048 ± 0,002	-0,9978	0,029	0,0001	5	68,7 ± 3,7	-1,00
28	2,80 ± 0,45	-0,043 ± 0,006	-0,9696	0,100	0,0064	5	64,4 ± 13,9	1,54
29	3,25 ± 0,76	-0,052 ± 0,011	-0,9394	0,172	0,0177	5	63,0 ± 20,0	2,91
30	3,06 ± 0,49	-0,047 ± 0,007	-0,9686	0,111	0,0067	5	64,8 ± 14,1	1,51
31	3,82 ± 1,10	-0,060 ± 0,016	-0,9125	0,247	0,0307	5	63,2 ± 24,5	3,14
32	3,20 ± 0,66	-0,051 ± 0,009	-0,9524	0,148	0,0124	5	63,3 ± 17,5	2,69
33	4,02 ± 1,15	-0,064 ± 0,016	-0,9146	0,258	0,0296	5	62,8 ± 24,1	3,75

**Tabela P5.** Hromatografski deskriptori ( $R_M^0$ ,  $b$ ,  $C_0$ , and  $PC1$ ) izvedeni iz retencionih podataka dobijenih za sistem **RP18 / MeOH:voda**,  $R$  – Pearson-ov korelacioni koeficijent,  $S$  – standardna greška modela,  $P$  – nivo značajnosti,  $n$  - broj merenja.

Jedinjenje	Hromatografski deskriptori i odgovarajući statistički parametri							
	$R_M^0$	$b$	$r$	$s$	$P$	$n$	$C_0$	$PC1$
1	3,01 ± 0,21	-0,032 ± 0,002	-0,9924	0,037	0,0008	5	93,6 ± 9,3	-1,02
2	2,89 ± 0,20	-0,031 ± 0,002	-0,9928	0,035	0,0007	5	92,1 ± 9,0	-0,36
3	3,88 ± 0,40	-0,040 ± 0,004	-0,9817	0,071	0,0030	5	97,2 ± 14,9	-3,05
4	3,30 ± 0,12	-0,034 ± 0,001	-0,9978	0,021	0,0001	5	96,3 ± 5,1	-2,32
5	4,13 ± 0,14	-0,044 ± 0,002	-0,9981	0,025	<0,0001	5	94,5 ± 4,7	-1,86
6	3,70 ± 0,07	-0,0389 ± 0,0007	-0,9995	0,012	<0,0001	5	95,2 ± 2,5	-2,04
7	2,89 ± 0,21	-0,031 ± 0,002	-0,9916	0,037	0,0009	5	93,5 ± 9,8	-0,92
8	4,14 ± 0,47	-0,042 ± 0,005	-0,9782	0,082	0,0039	5	97,7 ± 16,3	-3,45
9	4,40 ± 0,56	-0,045 ± 0,006	-0,9724	0,098	0,0055	5	98,5 ± 18,5	-4,01
10	3,16 ± 0,13	-0,032 ± 0,001	-0,9972	0,022	0,0002	5	98,0 ± 5,8	-2,90
11	3,98 ± 0,12	-0,042 ± 0,001	-0,9986	0,020	<0,0001	5	93,8 ± 4,0	-1,46
12	3,33 ± 0,18	-0,036 ± 0,002	-0,9953	0,032	0,0004	5	92,4 ± 7,3	-0,59
13	2,84 ± 0,15	-0,033 ± 0,002	-0,9959	0,027	0,0003	5	87,0 ± 6,5	1,29
14	2,61 ± 0,13	-0,030 ± 0,001	-0,9963	0,023	0,0003	5	88,1 ± 6,3	1,86
15	3,58 ± 0,21	-0,039 ± 0,002	-0,9949	0,036	0,0004	5	91,9 ± 7,6	-0,36
16	3,79 ± 0,18	-0,040 ± 0,002	-0,9963	0,032	0,0003	5	94,3 ± 6,5	-1,63
17	2,56 ± 0,05	-0,0292 ± 0,0006	-0,9994	0,009	<0,0001	5	87,6 ± 2,5	1,49
18	3,55 ± 0,21	-0,040 ± 0,002	-0,9950	0,037	0,0004	5	88,9 ± 7,3	1,16
19	2,80 ± 0,26	-0,032 ± 0,003	-0,9881	0,045	0,0016	5	88,0 ± 11,3	0,95
20	2,80 ± 0,05	-0,0314 ± 0,0005	-0,9996	0,008	<0,0001	5	89,1 ± 2,0	1,40
21	4,17 ± 0,49	-0,045 ± 0,005	-0,9782	0,087	0,0039	5	93,3 ± 15,9	-1,19
22	4,27 ± 0,51	-0,045 ± 0,005	-0,9774	0,089	0,0041	5	94,5 ± 16,3	-1,90
23	2,65 ± 0,19	-0,030 ± 0,002	-0,9927	0,033	0,0008	5	88,9 ± 8,9	0,99
24	3,02 ± 0,20	-0,34 ± 0,002	-0,9935	0,035	0,0006	5	89,0 ± 8,4	1,01
25	2,84 ± 0,14	-0,034 ± 0,002	-0,9969	0,025	0,0002	5	83,6 ± 5,6	3,43
26	3,66 ± 0,28	-0,39 ± 0,003	-0,9903	0,050	0,0011	5	94,1 ± 10,5	-1,48
27	3,32 ± 0,20	-0,036 ± 0,002	-0,9946	0,034	0,0005	5	92,8 ± 7,8	-0,73
28	2,72 ± 0,12	-0,32 ± 0,001	-0,9975	0,021	0,0001	5	84,6 ± 5,0	2,87
29	2,53 ± 0,18	-0,030 ± 0,002	-0,9938	0,031	0,0006	5	83,1 ± 7,9	3,36
30	3,13 ± 0,16	-0,036 ± 0,002	-0,9963	0,028	0,0003	5	87,6 ± 6,3	1,73
31	2,72 ± 0,17	-0,033 ± 0,002	-0,9952	0,029	0,0004	5	82,7 ± 6,9	3,72
32	3,01 ± 2,45	-0,035 ± 0,003	-0,9909	0,004	0,0010	5	85,4 ± 9,7	2,68
33	2,57 ± 0,17	-0,031 ± 0,002	-0,9944	0,030	0,0005	5	83,3 ± 7,5	3,32



**Tabela P6.** Hromatografski deskriptori ( $R_M^0$ ,  $b$ ,  $C_0$ , and  $PC1$ ) izvedeni iz retencionih podataka dobijenih za sistem **RP18 / DMSO:voda**,  $R$  – Pearson-ov korelacioni koeficijent,  $S$  – standardna greška modela,  $P$  – nivo značajnosti,  $n$  - broj merenja.

Jedinjenje	Hromatografski deskriptori i odgovarajući statistički parametri							
	$R_M^0$	$b$	$r$	$s$	$P$	$n$	$C_0$	$PC1$
1	5,41 ± 0,74	-0,063 ± 0,008	-0,9757	0,129	0,0045	5	7,3 ± 1,4	0,43
2	5,58 ± 0,66	-0,060 ± 0,007	-0,9781	0,116	0,0039	5	8,4 ± 1,4	-0,36
3	6,12 ± 0,20	-0,056 ± 0,002	-0,9977	0,034	0,0001	5	31,3 ± 1,6	2,65
4	4,85 ± 0,19	-0,053 ± 0,002	-0,9975	0,034	0,0001	5	29,9 ± 1,4	1,36
5	4,17 ± 0,47	-0,052 ± 0,005	-0,9853	0,081	0,0021	5	8,9 ± 1,3	-1,03
6	5,75 ± 0,79	-0,065 ± 0,009	-0,9741	0,139	0,0050	5	7,3 ± 1,4	0,27
7	5,77 ± 0,61	-0,057 ± 0,007	-0,9798	0,107	0,0034	5	9,4 ± 1,5	-0,19
8	6,97 ± 0,77	-0,064 ± 0,009	-0,9742	0,135	0,0049	5	9,0 ± 1,6	1,95
9	8,05 ± 0,43	-0,077 ± 0,005	-0,9941	0,076	0,0005	5	18,6 ± 1,5	3,20
10	5,06 ± 0,25	-0,055 ± 0,003	-0,9961	0,044	0,0003	5	20,1 ± 1,4	0,14
11	4,81 ± 0,54	-0,053 ± 0,006	-0,9813	0,095	0,0030	5	8,9 ± 1,4	-1,10
12	4,15 ± 0,46	-0,047 ± 0,005	-0,9825	0,081	0,0028	5	9,0 ± 1,4	-1,98
13	5,38 ± 0,43	-0,059 ± 0,005	-0,9904	0,075	0,0011	5	12,6 ± 1,4	0,22
14	5,58 ± 0,52	-0,060 ± 0,006	-0,9863	0,092	0,0019	5	10,6 ± 1,4	-0,70
15	6,28 ± 0,78	-0,075 ± 0,009	-0,9805	0,137	0,0032	5	8,0 ± 1,4	3,41
16	7,53 ± 0,45	-0,078 ± 0,005	-0,9938	0,079	0,0006	5	16,6 ± 1,5	4,43
17	5,05 ± 0,31	-0,056 ± 0,003	-0,9945	0,054	0,0004	5	16,4 ± 1,4	1,27
18	4,75 ± 0,41	-0,053 ± 0,005	-0,9891	0,072	0,0014	5	11,6 ± 1,4	-0,57
19	5,41 ± 0,39	-0,059 ± 0,004	-0,9921	0,068	0,0008	5	13,9 ± 1,4	0,48
20	5,68 ± 0,57	-0,061 ± 0,006	-0,9845	0,099	0,0023	5	10,0 ± 1,4	-1,32
21	6,96 ± 0,78	-0,073 ± 0,009	-0,9792	0,137	0,0036	5	8,9 ± 1,5	2,22
22	7,87 ± 0,67	-0,081 ± 0,007	-0,9878	0,117	0,0016	5	11,8 ± 1,5	4,01
23	5,09 ± 0,27	-0,056 ± 0,003	-0,9958	0,047	0,0003	5	19,0 ± 1,4	0,95
24	4,74 ± 0,42	-0,053 ± 0,005	-0,9889	0,073	0,0014	5	11,4 ± 1,4	-1,44
25	4,40 ± 0,36	-0,050 ± 0,004	-0,9908	0,062	0,0011	5	12,3 ± 1,4	-2,55
26	5,11 ± 0,18	-0,053 ± 0,002	-0,9979	0,031	0,0001	5	28,6 ± 1,5	2,19
27	5,43 ± 0,55	-0,042 ± 0,006	-0,9699	0,096	0,0064	5	9,9 ± 1,8	-0,57
28	4,55 ± 0,42	-0,052 ± 0,005	-0,9881	0,074	0,0015	5	10,7 ± 1,4	-3,05
29	4,55 ± 0,60	-0,053 ± 0,007	-0,9772	0,105	0,0041	5	7,6 ± 1,4	-3,91
30	4,57 ± 0,09	-0,047 ± 0,001	-0,9993	0,016	<0,0001	5	49,0 ± 1,5	-1,87
31	4,55 ± 0,05	-0,0514 ± 0,0006	-0,9981	0,009	<0,0001	5	86,6 ± 1,4	-2,79
32	3,70 ± 0,29	-0,054 ± 0,003	-0,9948	0,050	0,0004	5	13,0 ± 1,3	-2,58
33	4,42 ± 0,03	-0,0520 ± 0,0003	-0,9999	0,005	<0,0001	5	163,4 ± 1,4	-3,17

Tabela P7. Izračunate log  $P$  vrednosti

Jedinjenje	Alog $P_s$	AClog $P$	Alog $P$	Mlog $P$	log $P_{KOWWIN}$	Xlog $P_2$	Xlog $P_3$
1	3,97	3,45	4,03	3,19	3,62	3,66	4,01
2	3,94	3,35	4,02	2,90	3,70	3,57	3,98
3	4,50	4,06	4,70	3,68	4,26	4,28	4,63
4	2,45	2,26	2,78	2,39	2,42	2,09	2,62
5	3,81	3,46	3,93	3,24	3,44	3,55	3,83
6	4,21	3,54	4,12	3,30	3,86	3,84	4,07
7	3,98	3,43	4,10	3,01	3,94	3,75	4,04
8	4,66	4,15	4,78	3,79	4,51	4,46	4,70
9	4,45	4,30	5,06	4,12	4,83	4,76	4,95
10	2,11	2,34	2,86	2,51	2,67	2,27	2,68
11	3,84	3,54	4,01	3,35	3,68	3,73	3,90
12	3,03	2,84	3,26	2,74	2,79	2,93	3,21
13	3,99	3,94	4,28	3,05	3,83	3,05	3,68
14	3,99	3,84	4,26	2,76	3,91	2,97	3,65
15	4,55	4,55	4,95	3,55	4,48	3,67	4,31
16	4,45	4,70	5,22	3,87	4,80	3,98	4,57
17	2,34	2,75	3,02	2,25	2,64	1,49	2,30
18	3,46	3,95	4,18	3,10	3,65	2,94	3,51
19	4,11	4,03	4,36	3,17	4,08	3,23	3,74
20	4,05	3,92	4,35	2,87	4,16	3,14	3,72
21	4,58	4,64	5,03	3,65	4,72	3,85	4,37
22	4,63	4,79	5,31	3,98	5,04	4,15	4,63
23	2,39	2,83	3,11	2,37	2,88	1,66	2,36
24	3,72	4,03	4,26	3,21	3,90	3,12	3,57
25	3,32	3,33	3,62	2,56	3,19	2,43	3,05
26	3,67	3,96	4,59	3,01	4,28	3,09	3,78
27	4,35	4,51	4,53	3,29	4,36	3,70	4,30
28	3,36	3,22	3,60	2,27	3,27	2,35	3,02
29	3,42	3,22	3,60	2,27	3,27	2,35	3,02
30	3,95	3,94	4,28	3,05	3,83	3,05	3,68
31	3,49	3,39	3,82	2,94	3,39	2,59	3,15
32	2,80	3,33	3,51	2,61	3,01	2,32	2,88
33	2,94	3,33	3,51	2,61	3,01	2,32	2,88

**Tabela P8.** Molekulski deskriptori korišćeni za multivarijantnu analizu,

Jed.	TD	LUMO Energy	HOMO Energy	SA	Volume	Refractivity	Polarizability	Mass	MW	Log <i>P</i>	HLB	SP	DM
1	3,382	-0,315	-8,850	476,605	987,766	100,688	38,306	340,855	340,855	4,235	8,761	25,801	4,170
2	5,481	-0,307	-8,800	479,257	1027,703	107,151	40,778	370,882	370,882	4,048	8,240	25,623	6,310
3	4,784	-0,081	-8,719	491,920	1022,872	105,493	40,234	375,301	375,300	4,948	7,838	26,503	5,800
4	4,443	-0,234	-8,820	480,721	982,950	98,813	37,597	341,843	341,843	3,287	11,392	26,641	4,510
5	6,442	-1,358	-8,198	465,571	1015,397	108,013	40,147	385,853	401,896	3,838	11,574	24,963	3,290
6	5,048	-0,301	-8,427	481,507	998,839	103,506	39,004	385,306	385,307	4,385	8,582	25,734	5,070
7	3,405	-0,169	-8,775	525,674	1085,575	109,969	41,476	415,333	415,333	4,198	8,086	25,566	4,090
8	3,139	-0,289	-8,816	523,087	1045,888	108,311	40,932	419,751	419,751	5,098	7,679	26,417	4,780
9	6,986	-0,490	-8,563	512,295	1060,356	109,480	40,566	453,305	453,305	5,268	7,529	24,880	6,020
10	3,451	-0,279	-8,733	441,484	966,931	101,631	38,295	386,294	386,294	3,437	11,150	26,546	4,530
11	6,250	-0,557	-8,712	501,145	1056,485	109,826	40,845	430,304	430,304	4,217	10,506	26,144	7,040
12	3,898	-0,460	-8,723	448,995	989,604	102,204	38,219	351,408	351,408	3,354	11,321	25,553	4,630
13	4,784	-0,818	-8,350	468,842	1012,219	109,666	41,969	372,915	372,921	5,715	9,851	25,241	2,240
14	5,769	-0,589	-8,430	521,440	1104,447	116,129	44,441	402,942	402,947	5,527	9,277	25,124	4,100
15	5,429	-0,706	-8,509	500,488	1063,218	114,471	43,897	407,361	407,366	6,428	8,930	25,909	3,420
16	5,764	-0,390	-8,677	499,197	1035,527	106,662	39,868	408,854	408,854	5,118	7,874	24,921	5,740
17	3,213	-0,736	-8,530	447,282	1002,549	107,791	41,260	373,903	373,909	4,766	12,339	25,992	1,870
18	8,212	-0,840	-8,499	521,215	1091,589	115,986	43,810	417,913	417,919	5,546	11,611	25,657	5,460
19	4,913	-0,708	-8,473	495,350	1045,594	112,484	42,667	417,366	417,373	5,865	9,628	25,196	3,680
20	2,693	-0,053	-8,163	513,631	1109,977	114,202	43,536	429,359	429,360	4,843	8,151	25,423	1,850
21	5,069	-0,690	-8,460	514,350	1080,431	117,289	44,595	451,811	453,305	5,268	7,534	24,880	6,730
22	6,576	-0,732	-8,499	579,652	1147,482	118,458	44,229	485,365	485,371	6,748	8,536	24,451	4,400
23	6,201	-0,768	-8,523	463,772	1027,227	110,609	41,958	418,354	418,360	4,916	12,064	25,925	3,530
24	8,630	-1,483	-7,931	487,432	1082,357	118,804	44,508	462,364	462,370	5,696	11,364	25,599	6,740
25	4,934	-0,415	-8,329	476,051	1005,198	104,861	40,041	338,470	338,477	5,002	10,922	24,543	2,060
26	4,556	-0,526	-8,412	500,921	1073,749	114,944	43,711	366,524	366,530	6,300	9,133	23,876	2,080
27	6,459	-0,350	-8,276	490,807	1120,750	121,311	47,311	388,530	388,536	5,976	9,656	25,264	3,170
28	4,217	-0,430	-8,331	503,284	1069,746	111,324	42,513	368,497	368,503	4,814	10,219	24,478	2,730
29	6,145	-0,366	-8,289	526,326	1086,591	111,324	42,513	368,497	368,503	4,921	9,816	24,478	2,980
30	7,181	-1,065	-7,957	472,119	1026,380	109,666	41,969	372,915	372,921	5,715	10,310	25,241	3,450
31	5,584	-0,565	-8,425	443,974	977,394	105,078	39,950	356,461	356,467	5,145	10,665	24,913	2,960
32	7,266	-0,652	-8,453	480,316	1044,414	111,182	41,882	383,468	383,474	4,833	12,663	25,016	4,720
33	4,743	-1,007	-8,541	484,969	1028,742	111,182	41,882	383,468	383,474	4,745	12,256	25,016	3,580

Tabela P8. Nastavak

Jed.	ST	HSA	PSA	CI0	CI1	CI2	CI3	VI0	VI1	VI2	VI3	CI4	CI4
1	61,544	9,548	35,160	37,355	11,793	10,426	8,981	35,335	8,810	6,697	5,035	7,593	3,653
2	64,201	9,957	44,390	40,933	12,742	11,144	9,762	38,666	9,339	7,033	5,339	8,308	3,888
3	62,903	9,147	35,160	37,225	12,204	10,953	9,448	35,390	9,293	7,219	5,425	7,990	3,960
4	61,991	11,846	48,050	36,355	11,793	10,426	8,981	34,205	8,670	6,549	4,887	7,593	3,526
5	77,639	15,840	96,670	44,096	13,490	11,937	9,853	40,290	9,440	7,041	5,207	7,909	3,627
6	63,930	9,557	35,160	37,355	11,793	10,426	8,981	36,164	9,225	7,143	5,351	7,593	3,972
7	66,506	9,965	44,390	40,933	12,742	11,144	9,762	39,495	9,754	7,480	5,655	8,308	4,207
8	65,272	9,146	35,160	37,225	12,204	10,953	9,448	36,219	9,708	7,666	5,741	7,990	4,279
9	54,562	9,536	35,160	39,725	13,398	13,017	10,271	36,721	9,952	7,856	5,768	8,430	4,212
10	64,752	11,828	48,050	36,355	11,793	10,426	8,981	35,034	9,085	6,996	5,203	7,593	3,845
11	107,181	12,427	87,300	38,803	13,114	11,884	10,001	36,281	9,639	7,477	5,589	8,546	4,154
12	104,161	12,426	87,300	38,933	12,704	11,344	9,610	35,398	8,742	6,501	4,933	8,040	3,477
13	72,922	11,230	38,060	38,225	12,187	11,071	9,248	36,532	9,401	7,307	5,348	7,899	3,955
14	74,924	11,657	47,290	41,803	13,135	11,790	10,030	39,863	9,930	7,643	5,652	8,614	4,190
15	73,648	10,838	38,060	38,096	12,597	11,599	9,715	36,587	9,883	7,829	5,739	8,295	4,262
16	48,708	9,946	35,160	39,725	13,398	13,017	10,271	35,892	9,538	7,409	5,452	8,430	3,893
17	73,865	13,521	50,950	37,225	12,187	11,071	9,248	35,402	9,260	7,159	5,200	7,899	3,828
18	119,530	14,132	90,200	39,673	13,508	12,530	10,268	36,650	9,815	7,640	5,586	8,851	4,138
19	75,396	11,205	38,060	38,225	12,187	11,071	9,248	37,361	9,815	7,753	5,675	7,899	4,272
20	80,116	10,538	47,290	43,803	13,135	11,790	10,030	42,178	10,087	7,832	5,851	8,614	4,374
21	54,444	9,548	35,160	39,725	13,398	13,017	10,271	36,721	9,952	7,856	5,768	8,430	4,212
22	63,592	11,203	38,060	40,596	13,792	13,663	10,538	37,918	10,543	8,466	6,092	8,736	4,513
23	76,382	13,492	50,950	37,225	12,187	11,071	9,248	36,231	9,675	7,606	5,527	7,899	4,145
24	123,207	14,078	90,200	39,673	13,508	12,530	10,268	37,478	10,229	8,086	5,913	8,851	4,455
25	71,592	11,695	38,060	38,355	11,776	10,532	8,850	36,478	8,918	6,780	4,994	7,440	3,610
26	66,792	11,270	38,060	44,096	12,597	11,565	9,909	42,323	9,751	7,664	5,789	8,080	4,032
27	63,225	11,693	38,060	42,924	13,759	12,437	10,769	40,632	10,329	7,943	5,980	9,365	4,438
28	73,878	12,110	47,290	41,933	12,725	11,251	9,631	39,808	9,447	7,116	5,299	8,155	3,845
29	73,878	11,692	47,290	41,933	12,708	11,323	9,669	39,808	9,441	7,142	5,310	7,903	3,764
30	72,746	11,694	38,060	38,225	12,170	11,154	9,260	36,532	9,395	7,356	5,316	7,592	3,749
31	71,380	11,662	38,060	38,225	12,170	11,154	9,260	35,778	9,018	6,921	5,065	7,592	3,623
32	120,976	14,568	90,200	39,803	13,097	11,990	9,870	36,595	9,332	7,113	5,232	8,393	3,793
33	121,055	14,146	90,200	39,803	13,081	12,065	9,898	36,595	9,326	7,142	5,224	8,135	3,740

Tabela P9. Korelaciona matrica vrednosti molekulskih deskriptora

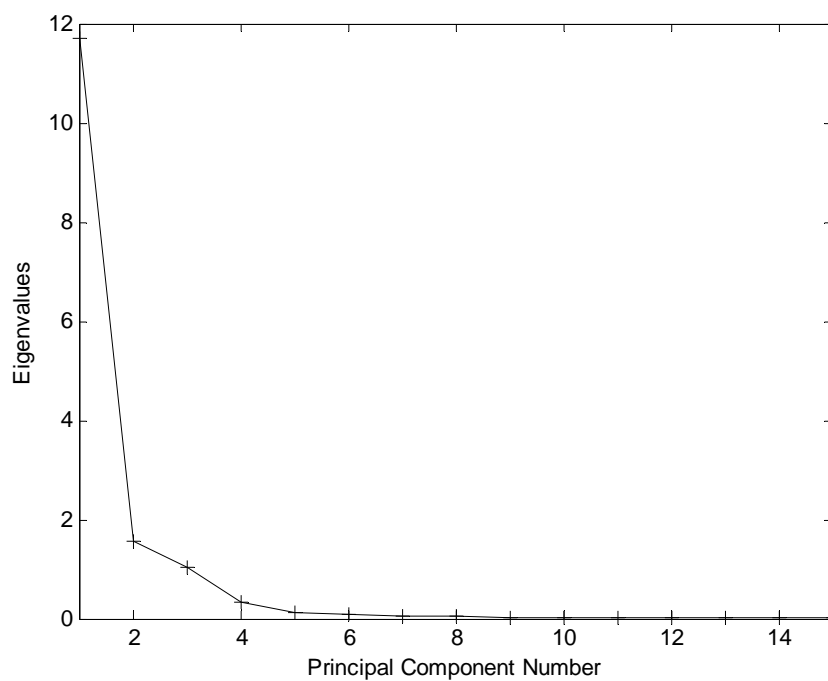
	TD	LUMO Energy	SA	Refractivity	Polarizability	Mass	MW	Log P	HLB	HBA	HBD	DM	ST
<b>TD</b>	1,000												
<b>LUMO Energy</b>	-0,794	1,000											
<b>SA</b>	0,128	0,095	1,000										
<b>Refractivity</b>	0,406	-0,450	0,662	1,000									
<b>Polarizability</b>	0,339	-0,370	0,645	0,980	1,000								
<b>Mass</b>	0,356	-0,336	0,600	0,712	0,601	1,000							
<b>MW</b>	0,375	-0,383	0,579	0,707	0,589	0,996	1,000						
<b>Log P</b>	0,405	-0,340	0,586	0,801	0,833	0,478	0,457	1,000					
<b>HLB</b>	0,231	-0,496	-0,548	-0,255	-0,293	-0,284	-0,253	-0,368	1,000				
<b>HBA</b>	0,381	-0,681	-0,059	0,264	0,147	0,250	0,323	-0,112	0,539	1,000			
<b>HBD</b>	0,336	-0,640	0,154	0,546	0,554	0,138	0,193	0,446	0,293	0,651	1,000		
<b>DM</b>	0,406	-0,097	-0,113	-0,156	-0,259	0,314	0,301	-0,238	-0,048	-0,060	-0,603	1,000	
<b>ST</b>	0,357	-0,501	-0,172	0,301	0,227	0,280	0,283	-0,051	0,532	0,500	0,167	0,273	1,000
<b>HSA</b>	0,427	-0,754	-0,281	0,129	0,042	0,043	0,105	-0,155	0,834	0,895	0,616	-0,075	0,585
<b>PSA</b>	0,319	-0,537	-0,315	0,024	-0,104	0,146	0,194	-0,395	0,643	0,805	0,199	0,321	0,800
<b>CI0</b>	-0,018	-0,157	0,363	0,531	0,507	0,151	0,190	0,185	-0,169	0,516	0,555	-0,437	0,058
<b>CI1</b>	0,412	-0,426	0,506	0,640	0,515	0,681	0,711	0,283	-0,150	0,638	0,324	0,197	0,322
<b>CI2</b>	0,487	-0,422	0,589	0,659	0,540	0,742	0,755	0,507	-0,224	0,451	0,270	0,202	0,173
<b>CI3</b>	0,356	-0,298	0,595	0,703	0,597	0,705	0,714	0,402	-0,318	0,432	0,192	0,223	0,254
<b>VI0</b>	-0,122	-0,066	0,401	0,597	0,614	0,159	0,184	0,278	-0,252	0,326	0,554	-0,544	0,011
<b>VI1</b>	0,336	-0,349	0,722	0,931	0,878	0,889	0,885	0,727	-0,410	0,216	0,387	0,001	0,173
<b>VI2</b>	0,352	-0,335	0,693	0,895	0,851	0,884	0,871	0,799	-0,421	0,088	0,320	0,031	0,131
<b>VI3</b>	0,238	-0,220	0,692	0,883	0,840	0,858	0,842	0,731	-0,505	0,041	0,229	0,054	0,140
<b>CI4</b>	0,249	-0,184	0,569	0,703	0,613	0,787	0,777	0,355	-0,349	0,282	0,032	0,342	0,383
<b>VI4</b>	0,170	-0,141	0,640	0,794	0,753	0,891	0,867	0,627	-0,477	-0,040	0,111	0,150	0,188

Tabela P9. Nastavak

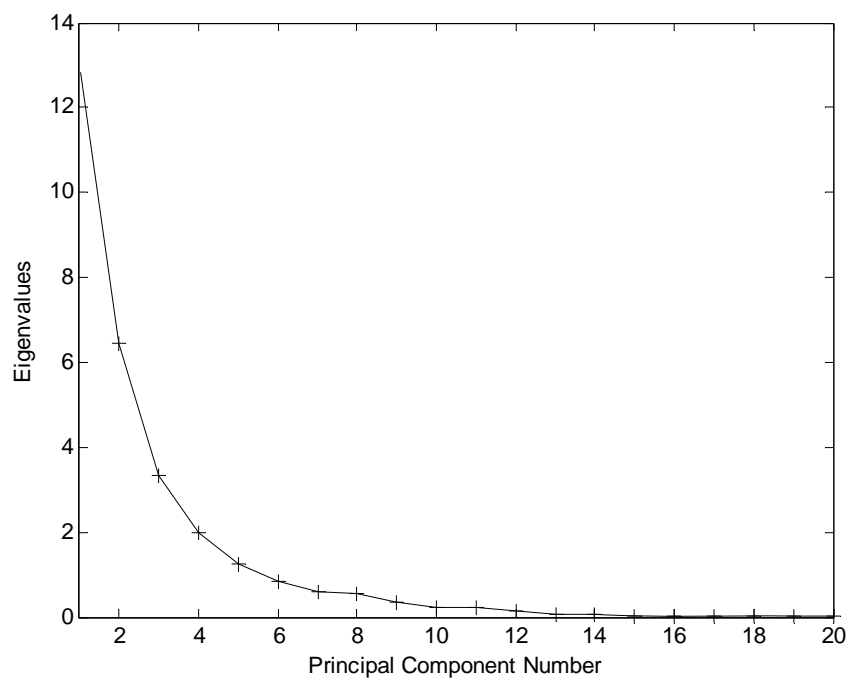
	HSA	PSA	CI0	CI1	CI2	CI3	VI0	VI1	VI2	VI3	CI4	VI4
<b>TD</b>												
<b>LUMO Energy</b>												
<b>SA</b>												
<b>Refractivity</b>												
<b>Polarizability</b>												
<b>Mass</b>												
<b>MW</b>												
<b>Log P</b>												
<b>HLB</b>												
<b>HBA</b>												
<b>HBD</b>												
<b>DM</b>												
<b>ST</b>												
<b>HSA</b>	1,000											
<b>PSA</b>	0,804	1,000										
<b>CI0</b>	0,268	0,202	1,000									
<b>CI1</b>	0,336	0,453	0,623	1,000								
<b>CI2</b>	0,191	0,221	0,422	0,918	1,000							
<b>CI3</b>	0,136	0,292	0,593	0,959	0,924	1,000						
<b>VI0</b>	0,127	0,021	0,941	0,436	0,246	0,445	1,000					
<b>VI1</b>	0,000	-0,028	0,437	0,688	0,733	0,752	0,497	1,000				
<b>VI2</b>	-0,081	-0,128	0,276	0,573	0,692	0,667	0,360	0,976	1,000			
<b>VI3</b>	-0,156	-0,115	0,350	0,575	0,646	0,696	0,447	0,968	0,976	1,000		
<b>CI4</b>	0,019	0,278	0,419	0,854	0,801	0,919	0,342	0,777	0,706	0,757	1,000	
<b>VI4</b>	-0,208	-0,117	0,167	0,465	0,529	0,573	0,299	0,918	0,936	0,958	0,732	1,000

**Tabela P10.** Korelaciona matrica vrednosti molekulskih deskriptora uključenih u MLR model

	<b>TD</b>	<b>LUMO Energy</b>	<b>SA</b>	<b>Polarizability</b>	<b>Mass</b>	<b>HLB</b>	<b>HBA</b>	<b>HBD</b>	<b>DM</b>	<b>Log P</b>	<b>ST</b>	<b>CI0</b>	<b>CI2</b>	<b>VI3</b>	<b>CI4</b>
<b>TD</b>	1,000														
<b>LUMO Energy</b>	-0,607	1,000													
<b>SA</b>	0,174	0,130	1,000												
<b>Polarizability</b>	0,386	-0,305	0,516	1,000											
<b>Mass</b>	0,362	-0,268	0,622	0,510	1,000										
<b>HLB</b>	0,240	-0,505	-0,546	-0,117	-0,320	1,000									
<b>HBA</b>	0,445	-0,688	-0,055	0,186	0,189	0,631	1,000								
<b>HBD</b>	0,334	-0,616	-0,053	0,475	-0,026	0,502	0,645	1,000							
<b>DM</b>	0,386	-0,027	0,200	-0,118	0,454	-0,254	-0,041	-0,621	1,000						
<b>Log P</b>	0,352	-0,308	0,508	0,592	0,718	-0,427	-0,018	0,192	0,119	1,000					
<b>ST</b>	0,386	-0,503	-0,106	0,161	0,085	0,665	0,650	0,325	0,126	-0,244	1,000				
<b>CI0</b>	0,161	-0,093	0,395	0,546	0,161	-0,153	0,429	0,449	-0,232	0,163	0,062	1,000			
<b>CI2</b>	0,556	-0,366	0,555	0,577	0,724	-0,175	0,417	0,150	0,353	0,640	0,161	0,467	1,000		
<b>VI3</b>	0,274	-0,099	0,693	0,778	0,828	-0,488	-0,030	0,081	0,196	0,806	-0,067	0,400	0,630	1,000	
<b>CI4</b>	0,428	-0,124	0,546	0,739	0,635	-0,178	0,283	0,094	0,324	0,463	0,273	0,533	0,780	0,742	1,000
<b>R<sub>M</sub><sup>0</sup></b>	-0,063	0,287	0,576	0,202	0,595	-0,803	-0,447	-0,421	0,324	0,691	-0,623	-0,007	0,473	0,614	0,349

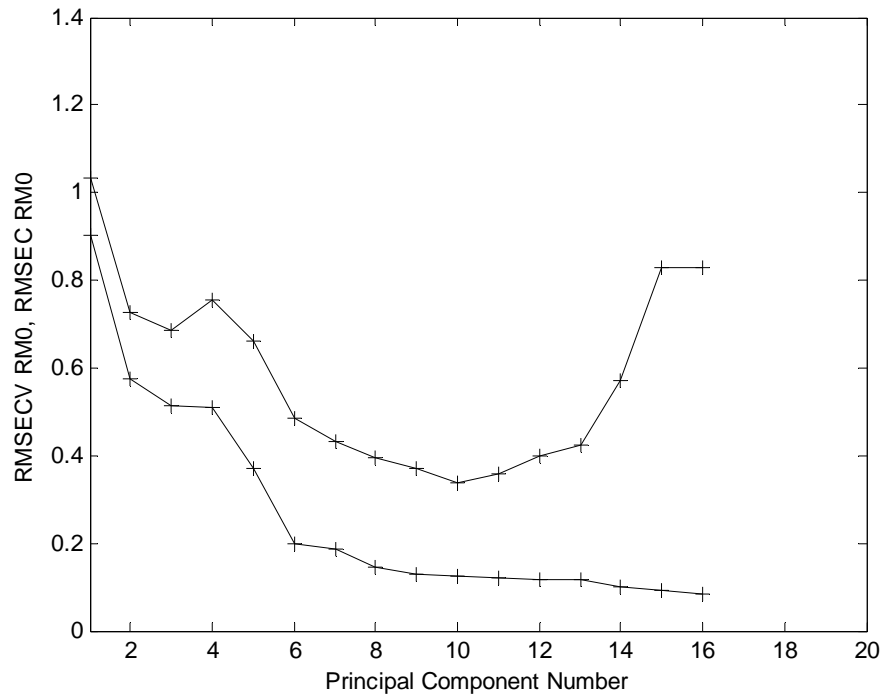


**Slika P1.** Grafik zavisnosti broja glavnih komponenata od svojstvenih vrednosti za PCA primenjenu na retencione podatke

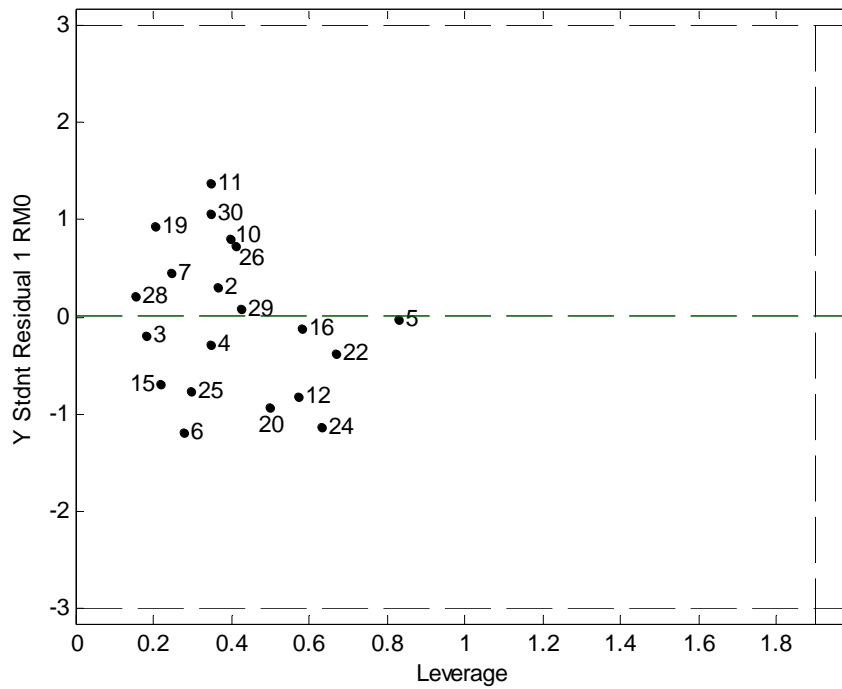


**Slika P2.** Grafik zavisnosti broja glavnih komponenata od svojstvenih vrednosti za PCA primenjenu na molekulske deskriptore

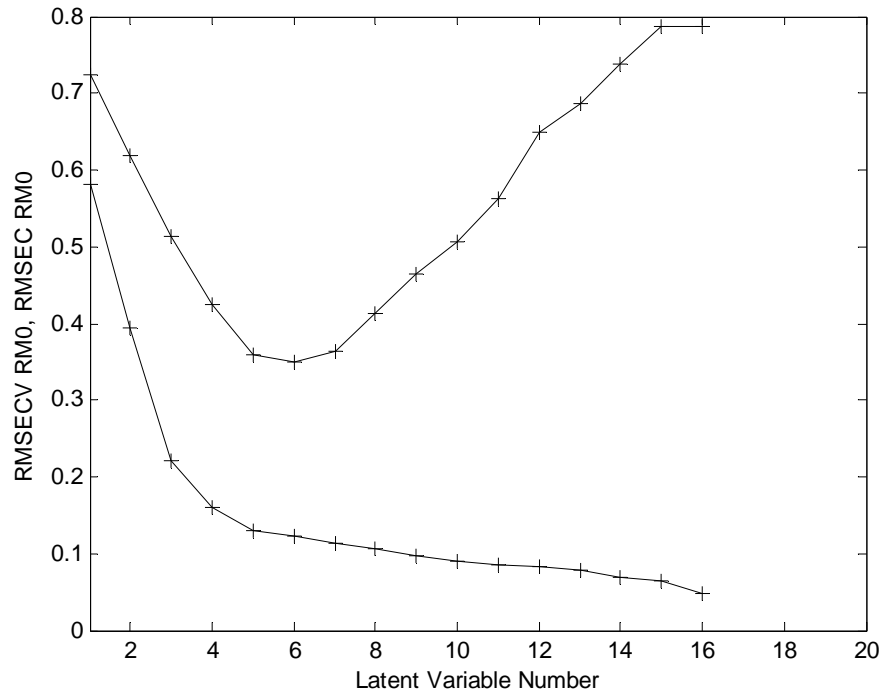




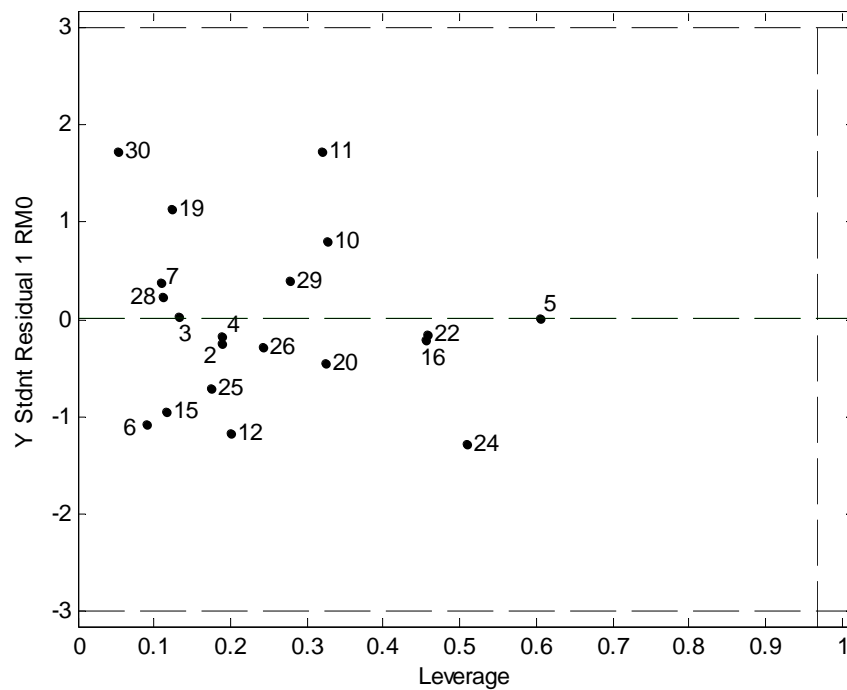
**Slika P3.** Grafik zavisnosti broja glavnih komponenta od odnosa RMSEC/RMSECV u PCR modelu



**Slika P4.** Grafik zavisnosti reziduala od  $Y_{\text{modela}}$  za PCR model



**Slika P5.** Grafik zavisnosti broja glavnih komponentata od odnosa RMSEC/RMSECV u PLS modelu



**Slika P6.** Grafik zavisnosti reziduala od  $Y_{\text{modela}}$  za PLS model

*Jelena Đ. Trifković*

## ***Biografija***

Jelena Đ. Trifković (rođena Veličković) rođena je 03. decembra 1976. godine u Beogradu, Republika Srbija, gde je završila osnovnu i srednju školu. Hemijski fakultet Univerziteta u Beogradu upisala je školske 1995/96. godine, a diplomirala 2001. godine sa prosečnom ocenom 8,27. Poslediplomske studije pri Katedri za analitičku hemiju Hemijskog fakulteta Univerziteta u Beogradu upisala je 2003. godine. Magistarsku tezu pod naslovom „Uticaj supstituenata na aromatičnim jezgrima 1-aralkil-4-aril piperazina na njihovu retenciju u uslovima planarne hromatografije" odbranila je 18. decembra 2006. godine.

Od avgusta 2003. godine bila je zaposlena kao asistent-pripravnik na Katedri za analitičku hemiju Hemijskog fakulteta u Beogradu. U zvanje asistenta za užu naučnu oblast Analitička hemija na Katedri za analitičku hemiju izabrana je 11.05.2007. godine. U tom zvanju se i sada nalazi.

U toku svog nastavnog rada bila je angažovana u izvođenju nastave iz nekoliko kurseva iz oblasti analitičke hemije i hemometrije na sva tri nivoa studija. Koautor je praktikuma koji se koristi u nastavi, „Praktikum iz kvalitativne hemijske analize neorganskih jedinjenja”, Dušanka Milojković-Opsenica, **Jelena Trifković**, Hemijski fakultet, Beograd, 2010, ISBN 978-86-7220-036-2.

Stipendista je Balkanske Ekološke Asocijacije (B.EN.A) i kompanije Carlsberg (2006/2007. godine). Pohađala je obuku "X-ray, IR and Raman Spectroscopy Application" - ISS on Instrumental Analysis 2006 (Sofija, Bugarska, septembar 2006. godine).

### ***Naučni radovi***

1. S. Šegan, **J. Trifković**, T. Verbić, D. Opsenica, M. Zlatović, J. Burnett, B. Šolaja, D. Milojković-Opsenica, Correlation between structure, retention, property, and activity of biologically relevant 1,7-bis(aminoalkyl)diazachrysene derivatives, *Journal of Pharmaceutical and Biomedical Analysis*, 72 (2013) 231-239.

2. Kristina B. Lazarević, Filip Andrić, **Jelena Trifković**, Živoslav Tešić, Dušanka Milojković-Opsenica, Characterisation of Serbian unifloral honeys according to their physicochemical parameters, *Food Chemistry*, 132 (2012) 2060-2064.
3. E.H.M.A. Rabtti, M.M. Natić, D.M. Milojković-Opsenica, **J.Đ. Trifković**, I.M. Vučković, V.E. Vajs, Ž.Lj. Tešić, RP TLC-Based Lipophilicity Assessment of Some Natural and Synthetic Coumarins, *Journal of Brazilian Chemical Society*, 23 (2012) 522-530.
4. E.H.M.A. Rabtti, M.M. Natić, D.M. Milojković-Opsenica, **J.Đ. Trifković**, T. Tosti, I.M. Vučković, V. Vajs, Ž.Lj. Tešić, Quantitative Structure-Toxicity Relationship Study of Some Natural and Synthetic Coumarins Using Retention Parameters, *Journal of Serbian Chemical Society*, 77 (2012) 1443-1456.
5. **J. Trifković**, F. Andrić, P. Ristivojević, D. Andrić, Ž.Lj. Tešić, D.M. Milojković-Opsenica, Structure-retention relationship study of arylpiperazines by linear multivariate modeling, *Journal of Separation Science*, 33 (2010) 2619-2628.
6. F.Lj. Andrić, **J.Đ. Trifković**, A.D. Radoičić, S.B. Šegan, Ž.Lj. Tešić, D.M. Milojković-Opsenica, Determination of the soil-water partition coefficients (log KOC) of some mono- and poly-substituted phenols by reversed-phase thin-layer chromatography, *Chemosphere*, 81 (2010) 299-305.
7. F.Lj. Andrić, **J. Trifković**, Ž.Lj. Tešić, D.M. Milojković-Opsenica, An Approximate Linear Solvation Energy relationships Model Based on Snyder's Selectivity Parameters. Chromatographic Behavior of Some 1-Aralkyl-4-Arylpiperazines, *Chromatographia*, 68 (2008) 453-458.
8. **J. Veličković**, Ž. Tešić, D. Milojković-Opsenica, Evaluation of the Lipophilicity of Some 1-Arylpiperazines by Planar Chromatography, *Journal of Planar Chromatography – Modern TLC*, 18 (2005) 358-363.
9. **J. Veličković**, D. Andrić, G. Roglić, Ž. Tešić, D. Milojković-Opsenica, Planar Chromatography of Some 1-Arylpiperazines Behaving as Dopaminergic Ligands, *Journal of Planar Chromatography – Modern TLC*, 17 (2004) 255-260.

## **Saopštenja**

1. F. Lj. Andrić, **J.Đ. Trifković**, I. Radović, D.M. Milojković-Opsenica, Ž.Lj. Tešić, Characterization and classification of Serbian monofloral honeys according to their mineral profiles by the means of multivariate data analysis. XIII Chemometrics in Analytical Chemistry, CAC-2012, June 2012, Budapest, Hungary, Book of Abstracts P192. **ISBN 978-963-9970-24-3**
2. M.S. Jovetić, **J.Đ. Trifković**, F.Lj. Andrić, S. Škrivanj, Ž.Lj. Tešić, D.M. Milojković-Opsenica, Mineral content of honey as a tool in tracing regional differences of Serbian polyfloral honey using pattern recognition methods, XIII. Chemometrics in Analytical Chemistry Conference, CAC-2012, June 2012, Budapest, Hungary, Book of Abstracts P217. **ISBN 978-963-9970-24-3**
3. Mirjana Stanković, **Jelena Trifković**, Filip Andrić, Nebojša Nedić, Živoslav Tešić, Dušanka Milojković-Opsenica, Determination of polyphenolic compounds in bee pollen from Serbia by UPLC-LTQ-orbitrap MS-MS, 50th Meeting of the Serbian Chemical Society, Belgrade, Serbia, June 2012, Book of Abstracts P16. ISBN 978-86-7132-048-1
4. Filip Andrić, **Jelena Trifković**, Aleksandra Radoičić, Jelena Kečkeš, Živoslav Tešić, Dušanka Milojković-Opsenica, Characterisation of Serbian monofloral honey according to their amino-acid composition. 5th International Symposium on Recent Advances in Food Analysis, Prague, Czech Republic, November 2011, Book of Abstracts P157. ISBN 978-80-7080-795-8.
5. Dušanka Milojković-Opsenica, Sandra Šegan, Jelena Trifković, Mario Zlatović, Dejan Opsenica, Igor Opsenica, Bogdan Šolaja, Quantitative structure-retention relationship study of some 1,7-bis(aminoalkyl) diazachrysenes derivatives, 17th European Symposium on Organic Chemistry, ESOC 2011”, Hersonissos, Crete, Greece, 10-15 July 2011, Book of abstracts (lectronic version, USB), P2.105, p108.
6. K. B. Lazarević, **J. Đ. Trifković**, F. Lj. Andrić, Ž. Lj. Tešić, D. M. Milojković-Opsenica, Characterization of Serbian monofloral honey according to their physico-chemical parameters. Conferentia Chemometrica, September 2011, Sümeg, Hungary, Book of Abstracts P25. ISBN 978-963-9970-15-1

7. K. B. Lazarević, **J. Đ. Trifković**, F. Lj. Andrić, Ž. Lj. Tešić, D. M. Milojković-Opsenica, Physico-chemical parameters in tracing regional differences of Serbian polyfloral honey, Conferentia Chemometrica, September 2011, Sümeg, Hungary, Book of Abstracts P40. ISBN 978-963-9970-15-1
8. P. Ristivojević, **J. Trifković**, N. Terzić, B. Šolaja, D. Milojković-Opsenica, A Quantitative Structure-Retention Relationship Study of Adamantane Derivatives of 4-amino-7-chloroquinoline, EUROANALYSIS 2011-European Conference on Analytical Chemistry, "Challenges in Modern Analytical Chemistry", September 2011, Belgrade, Serbia, CH30.
9. F. Andrić, P. Ristivojević, **J. Trifković**, Ž. Tešić, D. M. Milojković-Opsenica, Determination of the soil-water, octanol-water, and air-water partition coefficients for the twelve benzodiazepines by the means of the reversed-phase thin-layer chromatography, The XXXIV Symposium, Chromatographic Methods of Investigating The Organic Compounds, Katowice - Szczyrk, June 8th-10th, 2011.
10. Marija Nikolić, **Jelena Trifković**, Filip Lj. Andrić, Aleksandra Radoičić, Dragan Vučović, Živoslav Lj. Tešić, Dušanka M. Milojković-Opsenica, Planarna hromatografija hidrofилnih interakcija nekih anestetika, 48. Savetovanje Srpskog hemijskog društva, Novi Sad, 17-18. april 2010., AH-01, p.13. ISBN 978-86-7132-042-9
11. Filip Andrić, Jelena Vlajković, **Jelena Trifković**, Živoslav Tešić, Dušanka M. Milojković-Opsenica, Determination of nine water-soluble food dyes by the means of reversed-phase planar chromatography, 2nd Workshop Specific methods for food safety and quality - Pre event to physical chemistry 2010, Belgrade, 21 September 2010, Book of Abstracts, p.
12. **J. Đ. Trifković**, P. M. Ristivojević, F. Lj. Andrić, D. Andrić, Ž. Lj. Tešić, D. M. Milojković-Opsenica, Relationship Between Structure and Thin-layer Chromatographic Lipophilicity Parameter of Some Arylpiperazines, EUROANALYSIS 2009-European Conference on Analytical Chemistry, "The Impact of Analytical Chemistry on Quality of Life", September 2009, Innsbruck, Austria, Book of Abstracts P121-B1.

13. Aleksandra D. Radoičić, **Jelena Đ. Trifković**, Filip Lj. Andrić, Živoslav Lj. Tešić, Dušanka M. Milojković-Opsenica, Planarna hromatografija hidrofilnih interakcija nekih amino-kiselina, 47. Savetovanje Srpskog hemijskog društva, Beograd, 21. mart 2009., AH-11, p.21. ISBN 978-86-7132-031-2
14. Filip Lj. Andrić, Luka Novković, **Jelena Trifković**, Dušanka M. Milojković-Opsenica, LSER model hromatografskog ponašanja nekih 1-aralkil-4-arilpiperazina, 46. Savetovanje Srpskog hemijskog društva, Beograd, 21. februar 2008., AH 07, p.19. ISBN 978-86-7132-035-1
15. **J. Trifković**, F. Andrić, M. Natić, Ž. Tešić, D. Milojković-Opsenica, Determination of hydrophobic parameters for some arylpiperazines by RP-TLC and PCA, The XXXIst Symposium, Chromatographic Methods of Investigating The Organic Compounds, Katowice - Szczyrk, June 4th-6th, 2007, Book of Abstracts, p.40. ISBN 978-83-925714-0-7
16. Filip Lj. Andrić, **Jelena Đ. Trifković**, Dušanka M. Milojković-Opsenica, Živoslav Lj. Tešić, Optimizacija razdvajanja derivata nekih 1-arilpiperazina na tankom sloju silika-gela, 44. Savetovanje Srpskog hemijskog društva, Beograd, 06-07. februar 2006., AH-P04, p.21. ISBN 86-7132-027-8
17. **Jelena Đ. Veličković**, Deana Andrić, Goran Roglić, Živoslav Lj. Tešić, and Dušanka M. Milojković-Opsenica, Planar Chromatography of Some 1-Arylpiperazines Behaving as Dopaminergic Ligands, Planar Chromatography 2004, Visegrad, Hungary, May 2004., Proceedings of the International Symposium on Planar Separations p. 637-644.
18. **Jelena Đ. Veličković**, Živoslav Lj. Tešić, and Dušanka M. Milojković-Opsenica, Evaluation of the Lipophilicity of Some 1-Arylpiperazines by Planar Chromatography, Planar Chromatography 2005, Siófok, Hungary, May 2005., Proceedings of the International Symposium on Planar Separations p. 649-655.

Прилог 1.

## Изјава о ауторству

Потписани-а Јелена Ђ. Трифковић

број уписа 198

### Изјављујем

да је докторска дисертација под насловом

Одређивање квантитативног односа структуре и ретенције арилпиперазина применом

течне хроматографије и мултиваријантних хеометријских метода

- резултат сопственог истраживачког рада,
- да предложена дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, 04.02.2013.

Јелена Трифковић



Прилог 2.

## Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора \_\_\_\_\_ Јелена Ђ. Трифковић \_\_\_\_\_

Број уписа \_\_\_\_\_ 198 \_\_\_\_\_

Студијски програм \_\_\_\_\_ Доктор хемијских наука \_\_\_\_\_

Наслов рада \_\_\_\_\_ Одређивање квантитативног односа структуре и ретенције арилпиперазина применом течне хроматографије и мултиваријантних хеометријских метода \_\_\_\_\_

Ментор \_\_\_\_\_ др Душанка Милојковић-Опсеница \_\_\_\_\_

Потписани \_\_\_\_\_ Јелена Трифковић \_\_\_\_\_

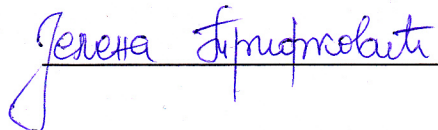
изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, \_\_\_\_\_ 04.02.2013. \_\_\_\_\_



Прилог 3.

## Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Одређивање квантитативног односа структуре и ретенције арилпиперазина применом  
течне хроматографије и мултиваријантних хеометријских метода

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство
2. Ауторство - некомерцијално
3. Ауторство – некомерцијално – без прераде
4. Ауторство – некомерцијално – делити под истим условима
5. Ауторство – без прераде
6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанда

У Београду, 04.02.2013.

