

# Lipofilnost lijekova. IV. Dio: Primjena SMILES označavanja u LogKow programu pri izračunavanju lipofilnosti

---

Medić-Šarić, Marica; Mornar, Ana

Source / Izvornik: **Farmaceutski glasnik, 2003, 59, 491 - 505**

Journal article, Published version

Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:163:684146>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-17**



Repository / Repozitorij:

[Repository of Faculty of Pharmacy and Biochemistry University of Zagreb](#)



## Lipofilnost lijekova

### IV. DIO: Primjena SMILES označavanja u LogKow programu pri izračunavanju lipofilnosti

MARICA MEDICŠARIĆ i ANA MORNAR

Zavod za farmaceutsku kemiju  
Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu

#### 11. RAČUNALNI PROGRAMI U IZRAČUNAVANJU LIPOFILNOSTI

Primjena računala od početka 60-ih godina značajno je utjecala na razvoj metoda izračunavanja lipofilnosti koje zamjenjuju dugotrajne eksperimentalne postupke. Pri izračunavanju log P vrijednosti koriste se brojni programski paketi od kojih su najpoznatiji **Pro\_log P 4.1**, **ClogP 4.0** i **LogKow**. Pokazalo se da su log P vrijednosti za izrazito lipofilne ili izrazito hidrofилne spojeve, izračunate računalnim programima, znatno pouzdanije od eksperimentalno dobivenih. Za nesintetizirane spojeve računalne metode predstavljaju jedinu moguću metodu dobivanja log P vrijednosti. Ipak, rutinska primjena računalnih programa zahtjeva njihovu stalnu provjeru i usporedbu s eksperimentalnim vrijednostima (1–3).

##### 11.1. LogKow PROGRAM

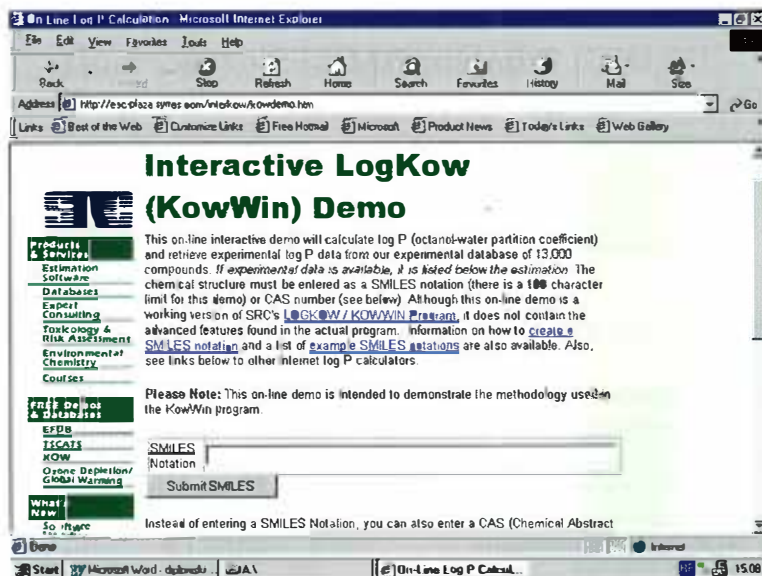
LogKow program se nalazi on-line na Internetu na stranici SRC (Syracuse Research Corporation), a njegova adresa je <http://esc-plaza.syrres.com/interkow/kowdemo.htm> (Slika 1.). Program izračunava log P vrijednost za željene organske strukture koristeći se atom/fragmentarnom metodom (4,5). Vrlo je točan i jednostavan za korištenje. Verzija prilagođena Windows programskom paketu poznata je kao KowWin.

Ispitivana supstancija se unosi u program ili kao SMILES oznaka ili se unese njezin CAS Registry Number.\* Treći način unošenja spoja u LogKow program uključuje komercijalne kemijske programe za crtanje struktura:

1. Alchemy III,
2. ChemDraw,
3. HyperChem,
4. Chembase,

\* CAS Registry Number – brojevana identifikacija svakog pojedinog kemijskog spoja uvedena u Chemical Abstracts CAS – Chemical Abstracts Service

5. ISISDraw,
6. Molecular Presentation Graphics,
7. PC Model.



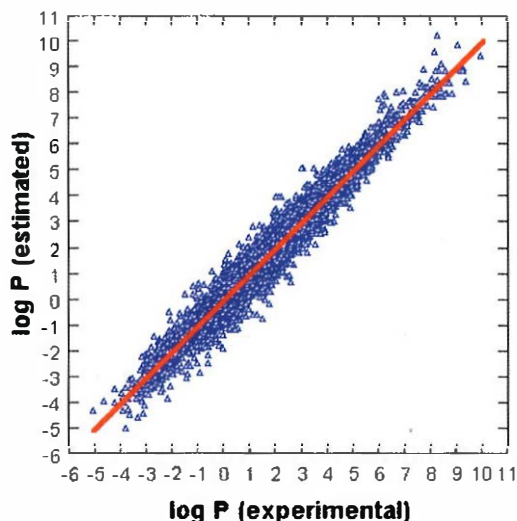
Slika 1. LogKow program on-line na Internetu

Pri izradi ovog rada korišteni su programi ISISDraw i ChemDraw. Bez obzira kojim načinom se unosi kemijska supstancija u LogKow program izračunata vrijednost je jednaka. Korisnik može izabrati onu metodu koja mu je jednostavnija, odnosno dostupna.

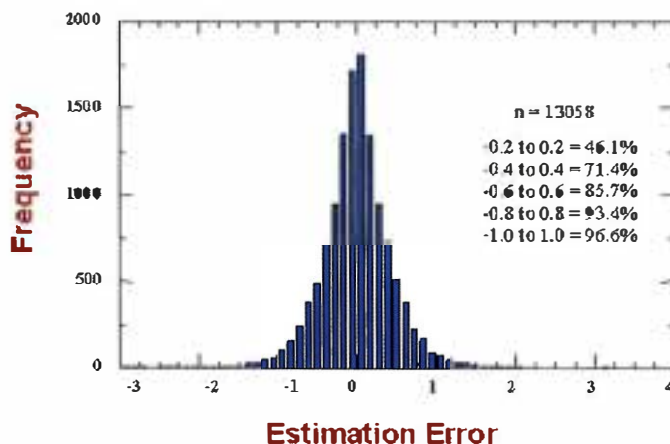
Na stranici Interneta o LogKow programu se mogu pronaći podaci o odnosu izračunatih i eksperimentalnih log P vrijednosti za 13 058 spojeva.

Slika 2. prikazuje odnos izračunatih i eksperimentalno određenih log P vrijednosti te njihov koeficijent korelacije,  $r = 0.95$ , i standardnu devijaciju,  $S_d = 0.436$ . Na slici 3. prikazan je histogram standardne pogreške. Dobiveni statistički parametri ukazuju da je metoda vrlo točna i pouzdana u izračunavanju log P.

Program LogKow se prvenstveno koristi u QSAR (Quantitative Structure-Activity Relationship) istraživanjima. Uporabom računalnog programa korisnik štedi vrijeme i energiju, jer program zamijenjuje dugotrajne, složene i često nepredvidive eksperimentalne postupke. Na ovaj način se već nakon nekoliko sekundi dobivaju izvrsni rezultati. Prednost pred eksperimentalnim postupcima predstavlja i mogućnost izračunavanja točnih log P vrijednosti za vrlo hidrofilne i vrlo lipofilne spojeve, što najčešće nije moguće eksperimentalnim postupcima. Možda je, ipak, njegova najveća vrijednost u mogućnosti izračunavanja log P vrijednosti za još nesintetizirane spojeve. Iz-



Slika 2. Odnos eksperimentalnih i izračunatih  $\log^2 P$  vrijednosti prema LogKow programu



Slika 3. Histogram standardne pogreške u LogKow programu

mjerene  $\log P$  vrijednosti mogu se upotrijebiti osim u QSAR studiju i u drugim istraživanjima u kojima lipofilnost predstavlja značajan parametar.

### 11.2. PRIMJENA SMILES OZNAČAVANJA U LogKow PROGRAMU

Programski paket LogKow koristi SMILES označavanje s linearnim simbolima za predstavljanje strukture molekule. Osnovna pravila označavanja utemeljio je David Weininger. SMILES je skraćenica za **S**implified **M**olecular

**Input Line Entry System.** Prilagođen je farmaceutima i kemičarima za uporabu putem računala. To je način predstavljanja molekule linearnim simbolima. Većina kemičara i farmaceuta može vrlo lako naučiti pisati SMILES oznake, no označavanje složenih struktura može biti dugotrajno.

Udruga SRC kao dodatak na osnovne SRC-programe uvodi i SMILECAS bazu podataka, koja se koristi u LogKow programu. U SMILECAS bazi podataka se nalaze podaci za 103 000 kemijskih spojeva, a podaci uključuju:

1. ime kemijskog spoja,
2. SMILES oznaku kemijskog spoja,
3. molekularnu formulu i molekularnu masu kemijskog spoja,
4. eksperimentalnu vrijednost log P odgovarajućeg kemijskog spoja.

Prema tome, za dobivanje log P vrijednosti neke supstancije, koja se nalazi u SMILECAS bazi podataka, dovoljno je poznavati samo njen CAS Registry Number.

### 11.2.1. Pravila označavanja

SMILES oznake opisuju molekularnu strukturu kao dvodimenzionalnu sliku nacrtanu na listu papira. Označavanje je linearno i postoje određena pravila označavanja. Bitno je istaknuti da se označavanje pojedine kemijske strukture može provesti na mnogo načina. Dakle, pojedina struktura može biti opisana točno na mnogo različitih načina. Štoviše, složene strukture mogu imati neodređeni broj SMILES oznaka koje točno opisuju strukturu.

Bez obzira kojim se načinom provodi označavanje neke molekule, svaka ispravno napisana SMILES oznaka prihvatljiva je za obradu na računalu.

### 11.2.2. Označavanje atoma

Atomi se označavaju njihovim atomskim simbolima:

C – ugljik    N – dušik  
S – sumpor    O – kisik

Prilikom označavanja važno je uočiti razliku između atoma označenih velikim štampanim slovima ili malim pisanim slovima. Svi alifatski atomi se označavaju velikim štampanim slovima dok se aromatski atomi označavaju malim pisanim slovima. SRC program uvažava samo atome ugljika, kisika, sumpora i dušika kao atome u aromatima; trenutno nisu dostupni ostali atomi.

Potrebno je osvrnuti se na označavanje atoma koji u atomskom simbolu sadržavaju dva slova. Prilikom označavanja klora (Cl) i broma (Br) prvo slovo se označava velikim štampanim slovom. Dok je u označavanju broma dopušteno da se slovo »r« označava malim pisanim slovom, u označavanju klora preporuča se slovo »l« označiti velikim štampanim slovom »L«, kako bi se izbjegla moguća zamjena s brojem 1.

Osim u iznimnim slučajevima vodikov atom nije uključen u SMILES oznake. Uobičajno je raditi s H-suprimiranim strukturama, kao pri izračunavanju topologijskih indeksa. U slučaju da korisnik programa označi vodikov atom računalo ga samostalno uklanja (Tablica 1.).

**Tablica 1.**  
SMILES označavanje atoma u H-suprimiranim strukturama

<i>Kemijska supstancija</i>	<i>Molekularna formula</i>	<i>SMILES oznaka</i>
Metan	CH <sub>4</sub>	C
Etan	CH <sub>3</sub> -CH <sub>3</sub>	CC
Propan	CH <sub>3</sub> -CH <sub>2</sub> -CH <sub>3</sub>	CCC
Brometan	CH <sub>3</sub> -CH <sub>2</sub> -Br	CCBr
Etanol	CH <sub>3</sub> -CH <sub>2</sub> -OH	CCO
Etantiol	CH <sub>3</sub> -CH <sub>2</sub> -SH	CCS
n-Propilamin	CH <sub>3</sub> -CH <sub>2</sub> -CH <sub>2</sub> -NH <sub>2</sub>	CCCN

### 11.2.3. Označavanje veza

SMILES označavanje uključuje četiri osnovne veze: jednostruku, dvostruku, trostruku i aromatsku vezu. Jednostruke veze su uglavnom izostavljene. Mogu se označiti simbolom crtice »-« no to samo otežava označavanje. Dvostruka se veza označava simbolom dviju crtica »=«, a trostruka brojevnim simbolom skale »#« (Tablica 2.). Aromatska veza nema oznaku, međutim atomi u aromatu pišu se malim pisanim slovima atomskih simbola: c, n, o, s.

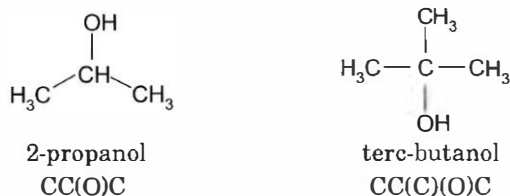
**Tablica 2.**  
SMILES označavanje veza u H-suprimiranim strukturama

<i>Kemijska supstancija</i>	<i>Molekularna formula</i>	<i>SMILES oznaka</i>
Eten	CH <sub>2</sub> =CH <sub>2</sub>	C=C
Propen	CH <sub>2</sub> =CH-CH <sub>3</sub>	C=CC
Etin	CH≡CH	C#C
Propin	CH≡C-CH <sub>3</sub>	C#CC
Butin	CH≡C-CH <sub>2</sub> -CH <sub>3</sub>	C#CCC
Acetonitril	CH <sub>3</sub> -C≡N	CC#N

SRC program prevodi određenu alifatsku strukturu u aromatsku ako uoči aromatičnost. Ako se benzen označi kao C1=CC=CC=C1 program ga prevodi u c1ccccc1. Starija verzija SRC programa ne može prihvatiti neke aromatske strukture kao što je azulen koji se sastoji od sedmeročlanog i peteročlanog prstena. Stoga se azulen označava velikim štampanim slovima pri uporabi starije verzije SRC programa.

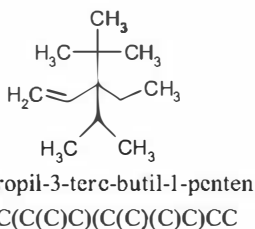
### 11.2.4. Označavanje grananja molekule

Ogranci u molekularnoj strukturi se označavaju simbolom zagrada (Slika 4.). Ogranak ne može započeti SMILES oznaku jer ogranak mora slijediti atom s kojim je vezan. Ako atom ima više od jednog ogranka oni su označeni kao uzastopni parovi zagrada. Također ogranak ne može odmah uslijediti nakon dvostruke niti trostruke veze.



Slika 4. Strukture i SMILES oznake nekih razgranatih molekula

Tzv. »nested branches« ili »ogranci unutar ogranaka« su dopušteni i često potrebni u SMILES označavanju. U SRC programu nije dopuštena uporaba dviju ili više uzastopnih zagrada »(«». Razlog tome je što dvije zagrade s lijeve strane nisu potrebne za predstavljanje niti jedne strukture (Slika 5.).



Slika 5. Struktura i SMILES oznaka strukture s ograncima unutar ogranaka

Nebrojeno različitih, važećih SMILES oznaka postoji za ovu strukturu. Označavanje može započeti na bilo kojem ugljikovom atomu u strukturi. Ako bi se označavanje započelo na centralnom ugljikovom atomu onda bi SMILES oznaka glasila: C(C=C)(CC)(C(C)C)(C(C)C)C. Međutim oznake je najlakše shvatiti u strukturi koja sadrži što je moguće manji broj ogranaka. Nepotrebni ogranci samo otežavaju označavanje. Na primjer, iako je važeća oznaka za butan C(C(C(C))) mnogo jednostavnija je oznaka CCCC.

#### 11.2.5. Označavanje prstenastih struktura

Najsloženiji oblik SMILES oznaka prihvaćen je za višeprstenaste strukture. Sljedeća pravila vrijede za sve prstenaste strukture:

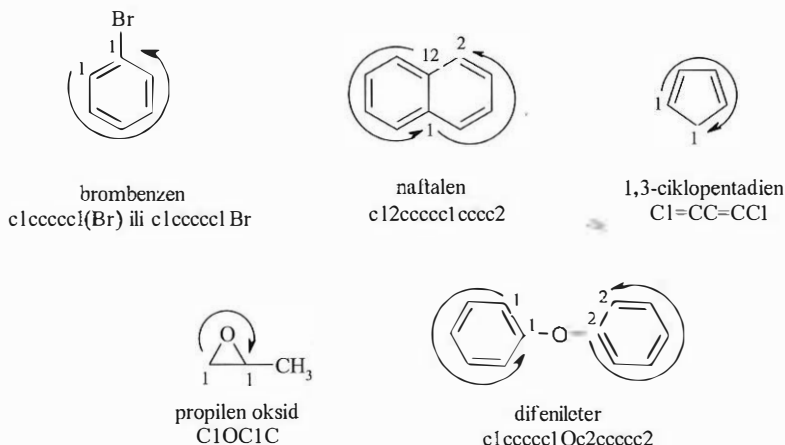
1. istim brojem se označava početni i posljednji atom prstena,
2. početni i posljednji atom moraju biti međusobno povezani,
3. svaki broj koji se koristi mora se pojaviti dva puta u cijeloj oznaci,
4. broj slijedi odmah iza onog atoma koji je označen kao početni ili posljednji.

Ukoliko ima više prstenova u strukturi princip označavanja prstenastih struktura je sljedeći:

1. odabere se jedan prsten iz cijele strukture i označi mu se početni i posljednji atom brojem,

- počevši od početnog atoma povuče se linija kroz prstenastu strukturu tako da linija prođe jednom pored svakog člana prstena i završi na posljednjem atomu,
- popišu se SMILES oznake počevši od početnog atoma i dalje slijedeći liniju
- u SRC programu dopušteno je početni i posljednji atom označiti s dva uzastopna broja, ali tri uzastopna broja nisu dopuštena (Slika 6.).

Za složene višeprstenaste strukture to može biti gotovo zagonetka s mnogo različitih mogućnosti.



Slika 6. Strukture i SMILES oznake nekih prstenastih struktura

### 11.2.6. Označavanje određenih kemijskih skupina

Mnogi novi korisnici SMILES označavanja imaju problema s određenim kemijskim skupinama. Stoga je vrlo korisna lista skupina koje se pri izračunavanju češće javljaju (Tablica 3.).

Tablica 3.  
SMILES označavanje nekih kemijskih skupina

Kemijska skupina	SMILES oznaka
Nitro	<chem>N(=O)(=O)</chem>
Nitratna	<chem>ON(=O)(=O)</chem>
Nitritna	<chem>ON(=O)</chem>
Cijanidna	<chem>C#N</chem>
Azidna	<chem>N=N#N</chem>

### 11.2.7. Označavanje metala

Metali se označavaju njihovim atomskim simbolima u četvrtastim zagradama. Natrij, kalij i litij iznimno mogu biti označeni i bez četvrtastih zagrada.



Nova verzija SRC programa koristi sljedeće metale:

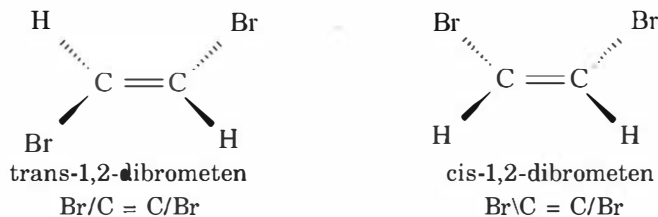
[ Al ] – aluminij	[ Zn ] – cink
[ Bi ] – bizmut	[ Au ] – zlato
[ Hg ] – živa	[ Ca ] – kalcij
[ Na ] – natrij	[ Li ] – litij
[ Sn ] – kositar	[ Pt ] – platina
[ As ] – arsen	[ Be ] – berilij
[ Cd ] – kadmij	[ Fe ] – željezo
[ Ni ] – nikal	[ Sb ] – antimon

### 11.2.8. Označavanje ioniziranih struktura

SRC program ne prihvaća ionizirane strukture. One se označavaju tako da se jednostavno ukloni naboj. MS-Windows verzija SRC programa će odmah ukloniti naboj svim metalima. Na primjer, ako se označi natrij-acetat oznakom [Na+][O-]C(=O)C SRC program ju odmah pretvara u [Na]OC(=O)C.

### 11.2.9. Označavanje izomernih struktura

Za označavanje izomernih struktura koriste se simboli »\« i »/«. Ovi simboli upućuju na relativno usmjerenje međusobno povezanih atoma (Slika 7.).



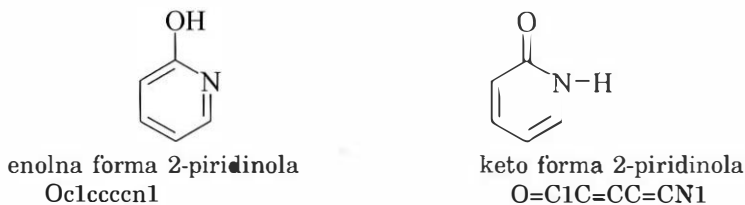
Slika 7. SMILES označavanje cis- i trans- izomera

Kiralni spojevi se označavaju simbolom »@« ali nova verzija programa uklanja simbol za kiralnost iz SMILES oznaka.

### 11.2.10. Označavanje tautomernih struktura

Pri označavanju tautomernih struktura korisnik mora sam unijeti onu formulu koja mu je potrebna.

Dobar primjer je keto i enolna forma 2-piridinola (Slika 8.).



Slika 8. SMILES označavanje tautomernih struktura

2-Piridinol je primjer tautomerne strukture koja sadrži samo jedan prsten, no ista pravila vrijede i za višeprstenaste strukture. U SRC programu keto forma se mora označiti velikim štampanim slovima, a ako korisnik želi da se struktura promatra kao aromatska tada mora označiti enolnu formu.

#### 11.2.11. Označavanje vodika

Kao što je već rečeno vodikov atom je isključen iz SMILES označavanja kako bi struktura bila što jednostavnije zapisana. Ipak, u određenim slučajevima se atom vodika mora koristiti. Primjerice, SMILES oznaka za etilamonijbromid ( $\text{CH}_3\text{CH}_2\text{NH}_3^+\text{Br}^-$ ) glasi CCN(H)(H)(H)Br. Ako bi se izostavili vodikovi atomi tada bi dušik bio shvaćen kao trovalentan, a ne kao peterovalentan.

Dakle, dva su slučaja kada je uvođenje vodika nužno:

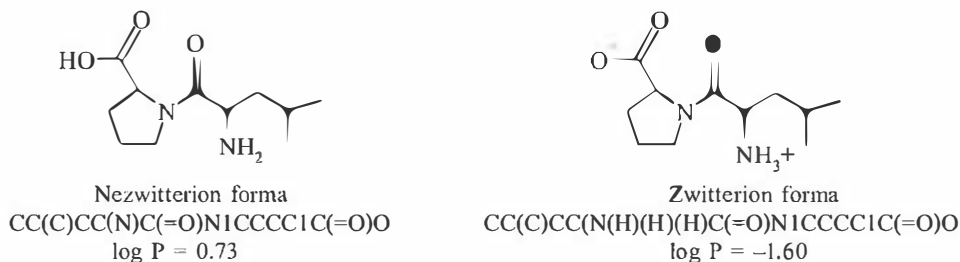
1. za različite organske hidrokloride,
2. za različite zwitterionske strukture.

#### 11.2.12. Označavanje hidroklorida

KowWin (LogKow) program ne prihvaća hidrokloride kao ionizirane strukture. Pri označavanju se ukloni hidroklorid iz strukture ili se označe atomi vodika. Na primjer, benzenpentanamin hidroklorid može se označiti na dva načina kao c1ccccc1CCCCCN (uklonjen je hidroklorid prilikom označavanja) ili kao c1ccccc1CCCCCN(H)(H)(H)Cl (označeni su vodikovi atomi).

#### 11.2.13. Označavanje struktura koje sadrže i kationska i anionska središta (zwitterioni)

KowWin (LogKow) program razlikuje se od CLOGP programa u označavanju zwitteriona. CLOGP program uvijek razmatra aminokiseline i neke lijekove kao što je amoksisilin kao zwitterione, dok korisnik KowWin (LogKow) programa unosi onu formu koju želi. Zwitterion forma se razlikuje od nezwitterion forme po tome što prva uključuje atome vodika u SMILES oznaku (6,7). Kao što se vidi na primjeru L-leucin-L-prolina, razlikuju se i po log P vrijednostima (Slika 9.).



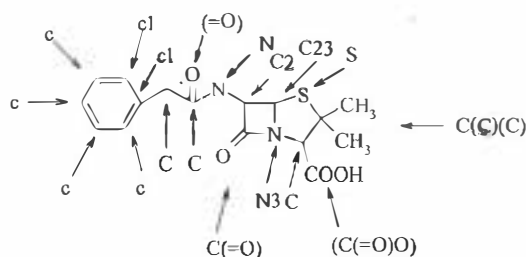
Slika 9. SMILES označavanje strukture u nezwitterion i zwitterion formi

### 11.3. TESTIRANJE LogKow PROGRAMA

#### a) Penicilini

Kako bi se procijenila valjanost predložene metode programom, LogKow izračunate su log P vrijednosti za skupinu penicilina i dobivene vrijednosti su uspoređene s eksperimentalno dobivenim podacima. Penicilini su antibiotici  $\beta$ -laktamske strukture sa snažnim antimikrobnim djelovanjem i vrlo malom toksičnošću. U svojoj strukturi sadrže jedan  $\beta$ -laktamski i jedan tiazolidinski prsten (8).

Na primjeru benzilpenicilina prikazano je SMILES označavanje (Slika 10.) i izračunavanje log P vrijednosti, LogKow programom (Slika 11.).



BENZILPENICILIN

c1ccccc1CC(=O)NC2C(=O)N3C23SC(C)(C)C(C(=O)O)

Slika 10. Struktura i SMILES oznaka benzilpenicilina

SHILES : O=C(NC(C(=O)N(C(=O)O)C(S2)(C(C)C12)C(=O)C1)C(=O)O)  
 CHEN : Benzilpenicilin  
 MOL FOR: C16 H18 N2 O4 S1  
 MOL WT : 334.39

TYPE	NUM	LOGKOW	FRAGMENT DESCRIPTION	COEFF	VALOR
Frag	1		-CH2 (aliphatic carbon)	1	0.5473
Frag	1		-CH2- (aliphatic carbon)	1	0.4911
Frag	1		-CH (aliphatic carbon)	1	0.3614
Frag	1		-NH- (aliphatic attach)	1	-1.4962
Frag	1		-H- (aliphatic attach)	1	-1.8323
Frag	6		Aromatic Carbon	1	0.2940
Frag	1		-COOH (acid, aliphatic attach)	1	-0.6895
Frag	1		-C(=O)N (aliphatic attach)	1	-0.5236
Frag	1		-S- (aliphatic attach)	1	-0.4045
Frag	1		-tert. Carbon (3 or more carbon attach)	1	0.2676
Factor1	1		>N-C(=O)-C(=O)- correction	1	0.8370
Factor1	1		C-C(=O)-N-C(=O)-C-S- correction	1	1.5505
Const	1		Equation Constant		0.2290

LogKow = 1.8483  
 LogKow Estimated Log P: 1.85

Experimental Database Structure Match:  
 Name: Benzilpenicilin  
 CAS Registry Number: 000061-33-6  
 Experimental LogKow: 1.83  
 Experim. Reference: Hansch, C. et al. (1995)

Slika 11. SMILES označavanje i izračunavanje log P vrijednosti LogKow programom za benzilpenicilin

Označavanje benzilpenicilina započinje od benzenskog prstena, koji je označen brojem 1. Aromatičnost benzenskog prstena uočava se označavanjem malim pisanim slovima. Laktamski prsten označen je brojem 2, a tiazolidinski prsten označen je brojem 3. Metilne skupine kao i karboksilna skupina označene su u zagradama jer se smatraju ograncima tiazolskog prstena. Isto vrijedi i za kisikov atom u keto skupinama benzilpenicilina. Po istom principu SMILES oznake su napravljene i za druge peniciline.

Tijekom izračunavanja log P vrijednosti za peniciline korištene su sve tri metode unosa podataka i uočena je jednaka valjanost metoda, te korisnik LogKow programa može izabrati metodu, koja mu je prihvatljivija.

Podaci, dobiveni izračunavanjem, uspoređeni su s eksperimentalno dobivenim vrijednostima koje se nalaze u SMILECAS bazi podataka (Tablica 4.).

**Tablica 4.**  
Odnos eksperimentalno određenih i izračunatih  
log P vrijednosti LogKow programom

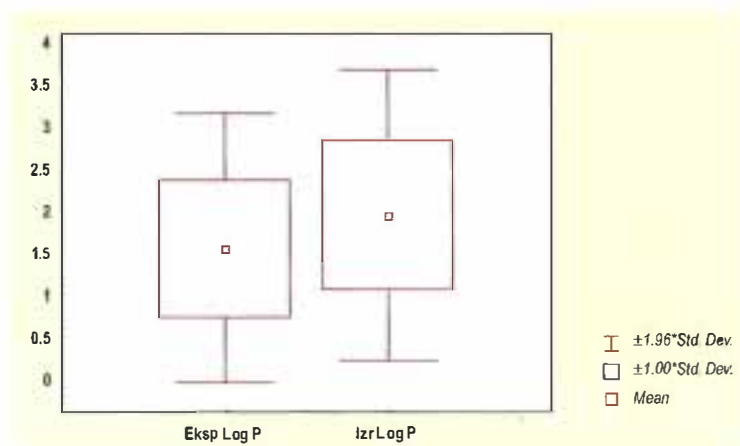
Penicilin	Eksperimentalno određen log P	Izračunati log P
Benzilpenicilin	1.83	1.85
Fenoksimetilpenicilin	2.09	1.87
Feneticilin	2.20	2.29
Propicilin	2.65	2.78
Meticilin	1.22	2.20
Oksacilin	2.38	2.57
Kloksacilin	2.48	3.22
Fluoksacilin	2.58	3.42
Dikloksacilin	2.91	3.86
Ampicilin	1.35	1.45
Amoksicilin	0.87	0.97
Karbenicilin	1.13	1.19
Piperacilin	0.50	1.83
Karfecilin	3.14	3.28

Srednje vrijednosti za obje metode izražene su kao aritmetička sredina  $\pm$  standardna devijacija i prikazane su na slici 12. Srednja vrijednost eksperimentalno određenih log P vrijednosti iznosila je  $1.95 \pm 0.81$ , a LogKow programom izračunatih log P vrijednosti bila je  $2.34 \pm 0.88$ . Izračunate log P vrijednosti bile su nešto veće od eksperimentalno dobivenih.

Za analizu parova vrijednosti eksperimentalnih log P i izračunatih log P korišten je statistički program StatSoft Inc. 1992. Microsoft Corp.

Korištenjem  $\chi^2$ -testa uspoređene su računski dobivene vrijednosti s eksperimentalno dobivenim (Tablica 5.).

$$\chi^2 = \sum \frac{(\log P_{eksp} - \log P_{izr})^2}{\log P_{izr}}$$



Slika 12. Odnos izračunatih i eksperimentalno izmjerenih log P vrijednosti

Tablica 5.

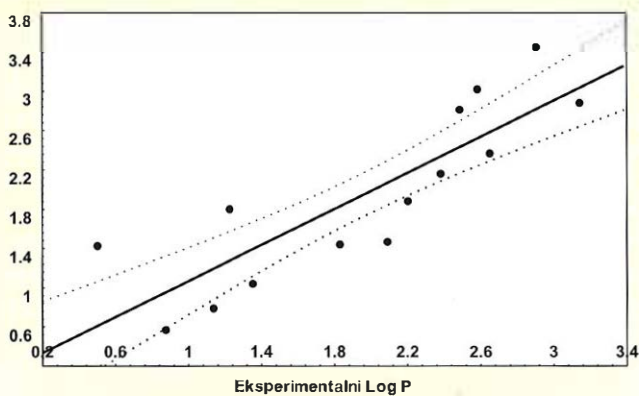
Analiza parova eksperimentalnih i izračunatih log P vrijednosti  $\chi^2$ -testom

	log P <sub>eks</sub>	log P <sub>izr</sub>	$\Delta \log P_{\text{eks}} - \log P_{\text{izr}}$	$\chi^2$ -test
1	1.83	1.85	-0.02	0.00216
2	2.09	1.87	0.22	0.025882
3	2.20	2.29	-0.09	0.003537
4	2.65	2.78	-0.13	0.006079
5	1.22	2.20	-0.98	0.436545
6	2.38	2.57	-0.19	0.014047
7	2.48	3.22	-0.74	0.170062
8	2.58	3.42	-0.84	0.206316
9	2.91	3.38	-0.95	0.233808
10	1.35	1.45	-0.10	0.006897
11	0.87	0.97	-0.10	0.010309
12	1.13	1.19	-0.06	0.003025
13	0.50	1.83	-1.33	0.966612
14	3.14	3.28	-0.14	0.005976
$\Sigma$	27.33	32.78	-5.45	2.089312

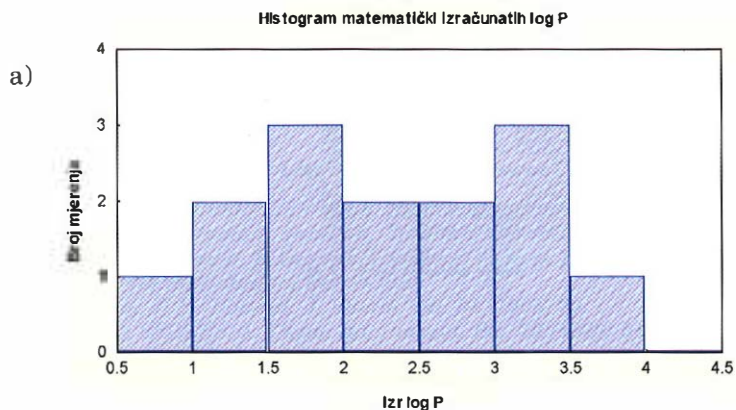
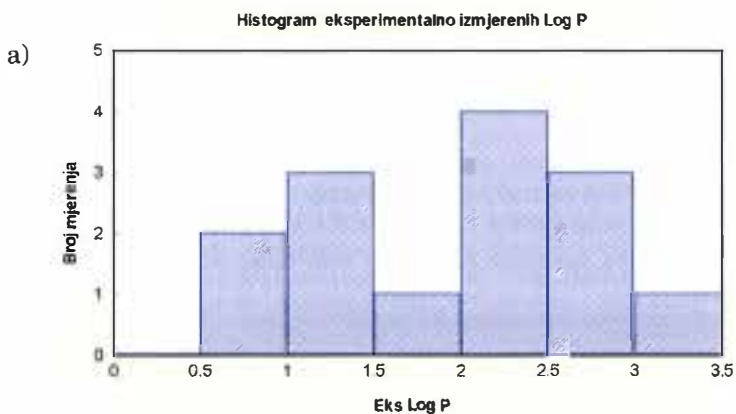
Rezultati nisu pokazali statistički značajne razlike u analizi parova ( $p < 0.999$ ).

Međutim, korelacija (Slika 13.) između izračunatih i eksperimentalno određenih log P vrijednosti bila je značajna ( $r = 0.85$ ,  $p < 0.001$ ).

Slika 14. prikazuje histograme (a) eksperimentalno određenih log P vrijednosti i (b) LogKow programom, izračunatih log P vrijednosti. Analizom raspodjele vrijednosti vidi se odstupanje jedino u log P vrijednostima između 1.5 i 2.0, dok su ostale vrijednosti sukladne.



Slika 13. Korelacija između izračunatih i eksperimentalnih log P vrijednosti

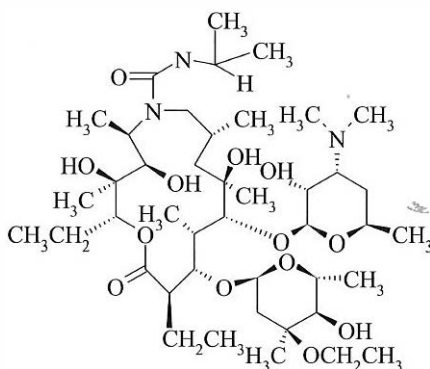


Slika 14. Histogram log P vrijednosti

Statistička obrada podataka istraživane metode pokazala je dobre rezultate, te se izračunate log P vrijednosti mogu koristiti u QSAR istraživanjima.

### b) Nesintetizirani spojevi

Glavna prednost primjenjene metode je mogućnost izračunavanja log P vrijednosti za još nesintetizirane spojeve. Nasumice je nacrtan jedan makrolid nazvan "Makrolid 1" te je napisana njegova SMILES oznaka (Slika 15). Prethodno opisanim SMILES označavanjem uneseni su podaci u LogKow program. Dobiveni rezultat je prikazan na slici 16.



MAKROLID 1

O=C(N(C(C)C(O)C(O)C(C)C(CC)O)C)CC(C)CC(O)C(C)C(OC3OC(C)CC(N(C)C)C3O)C(C)C(OC2CC(C)(OCC)C(O)C(C)O2)C(CC)C1(=O)NC(C)C

Slika 15. Struktura nasumice nacrtanog makrolida sa SMILES oznakom.

**KowWin (LogKow) Log P Calculation:**

SMILES : O=C(N(C(C)C(O)C(O)C(C)C(CC)O)C)CC(C)CC(O)C(C)C(OC3OC(C)CC(N(C)C)C3O)C(C)C(OC2CC(C)(OCC)C(O)C(C)O2)C(CC)C1(=O)NC(C)C

MOLE WT : 640.14

TYPE	SMRT	LOGKOW V1.04 FRAGMENT DESCRIPTION	COEFF	VALUE
Frags	15	-C#N (aliphatic carbon)	0.5473	0.2095
Frags	7	-CH2- (aliphatic carbon)	0.4911	3.4377
Frags	16	(aliphatic carbon)	0.7619	5.7823
Frags	5	-OH (hydroxy, aliphatic attach)	-1.1086	-7.0338
Frags	5	O- (oxygen, aliphatic attach)	-1.2546	-8.2893
Frags	1	(aliphatic attach)	-1.4962	-1.4562
Frags	1	(aliphatic attach)	-1.8223	-3.6646
Frags	1	-C(=O)O (ester, aliphatic attach)	0.5056	-0.5056
Frags	1	NC(=O)- (amide)	1.0455	1.0455
Frags	1	-NH- Catzhen (3 or more carbon attach)	0.7676	0.2829
Factor1		C=O-C=O correction	0.4034	1.6052
Factor2		Multi-olefin correction	0.1064	0.4244
Factor3		C-O-C(=O) D-C correction	0.9500	0.9500
Factor4		H-C-C(=O)-C(=O)-C correction	0.9500	0.9500
Factor5		O=C-C(=O)-C-O-C(=O)-C hydroxy ester part	1.5000	1.5000
Factor6		Equation Constant	1	0.2250

log Pow = 4.6830  
LogKow Estimated Log P: 4.68

Slika 16. SMILES označavanje i izračunavanje log P vrijednosti LogKow programom za nasumice nacrtani makrolid



## 11.4. ZAKLJUČAK

Opisana je metoda izračunavanja log P vrijednosti računalnim programom LogKow koji koristi atom/fragmentarnu metodu. Na skupini penicilinskih preparata ispitana je valjanost metode. Podaci za log P vrijednosti, dobiveni izračunavanjem, uspoređeni su s eksperimentalno dobivenim vrijednostima koje se nalaze u bazi podataka. Statističkom obradom podataka dokazana je točnost metode te sukladnost rezultata s eksperimentalno određenim log P vrijednostima.

Praktična vrijednost programa je predviđanje lipofilnosti za još nesintetizirane spojeve, što je prikazano na izračunavanju log P vrijednosti za jedan nasumice odabran makrolid.

### Lipophilicity of drugs (*Part IV*)

by M. Medić-Šarić and A. Mornar

#### *S u m m a r y*

In this work, we described the LogKow program that estimates the log octanol/water partition coefficient (log P) of organic chemicals using a atom/fragmental method developed at SRC. Chemical structures can be brought into LogKow estimation program using SMILES notation, CAS RN or commercial drawing programs that allow a user to draw a chemical structure and then directly write the SMILES notation. SMILES is an acronym for Simplified Molecular Input Line Entry System. It is a chemical notation system used to represent a molecular structure by a linear string of symbols. A two-dimensional drawing of a single chemical structure is possible in many different forms. That is, a single structure can be depicted correctly by many different SMILES notations. SMILES notations are comprised of atoms (designated by atomic symbols), bonds, parentheses (used to show branching), and numbers (used to designate ring opening and closing positions). The SMILECAS Database is very helpful and time-efficient in obtaining SMILES notations and log P values; all that you need is CAS Registry Number. This database contains the SMILES notations for 103 000 compounds.

In order to evaluate this method estimated log P values were compared with these experimental obtained from the SMILECAS Database. Statistical analysis has been performed using the StatSoft Inc.1992. Microsoft Corp. program. Results show that there was no statistical significant difference,  $p < 0.999$ . Using LogKow, the user can save time and energy because the program can be used as substitute for time-consuming experimental methods. LogKow is even suitable for calculating lipophilicity of very hydrophilic and lipophilic compounds which can't be measured experimentally.

#### Literatura - References

1. M. Medić-Šarić, D. Franić, Ž. Debeljak, Lipofilnost lijekova, I dio, Farm. Glas. 54, (1998) 37.
2. M. Medić-Šarić, D. Franić, Ž. Debeljak, Lipofilnost lijekova, II dio, Farm. Glas. 54 (1998) 74.
3. M. Medić-Šarić, D. Franić, Ž. Debeljak, Lipofilnost lijekova, III dio, Farm. Glas. 54 (1998) 113.
4. <http://esc.syrres.com/interkow/logkow.htm>
5. W. M. Meylan, P. H. Howard, J. Pharm. Sci. 84 (1995) 83.
6. D. Weininger, SMILES, J. Chem. Inf. Comput. Sci. 28 (1988) 31.
7. <http://esc.syrres.com/interkow/docsmile.htm>
8. Goodman and Gilman's THE PHARMACOLOGICAL BASIS OF THERAPEUTICS, 10<sup>th</sup> Edition, The McGraw-Hill Companies, New York 2001, 1073.

Primljeno 9. V. 2003.