

Podstawy obliczeń metodami SAR (ang.: *Structure activity relationship*), we wczesnej fazie badań nad nowymi lekami.

Graficzna i matematyczna implementacja podstawowych zasad kwalifikowania do grona leków - *drugability; druglikeness*.

Tomasz Grabowski

Laboratorium Doświadczalne, Wrocławskie Zakłady Zielarskie Herbapol SA ul. Ks. Witolda 56, 50-203 Wrocław.



Wprowadzenie.

DRUGABILITY, DRUGLIKENESS:

Podobieństwo do leku to jedna z podstawowych cech podlegających opisowi matematycznemu w przedklinicznej fazie badań nad nowymi substancjami leczniczymi NCE (ang.: New chemical entities).

Metody obliczeniowe typu SAR (ang.: *Structure activity relationship*) są standardowymi metodami wczesnej fazy analiz właściwości kandydatów na nowe leki. W chwili obecnej duże znaczenie we wczesnej fazie badań nad nowymi substancjami leczniczymi mają metody obliczeniowe pozwalające symulować parametry farmakokinetyczne przyszłego leku. Metody te opracowywane są na podstawie baz danych właściwości fizykochemicznych substancji już znanych. Pozwalają one na tworzenie teoretycznych (wirtualnych) modeli cząsteczek i symulacji ich charakterystyki farmakokinetycznej (ang.: *Virtual screening*), (Opera T.I. 2002). Obecnie obliczenia i kwalifikowanie do badań SAR wykonywane są za pomocą odpowiednich programów komputerowych i baz danych. Ogół zależności, jakie muszą spełniać związki kandydujące do badań przedklinicznych, pozwala dzielić nowe substancje chemiczne NCE (ang.: *New chemical entities*) na „podobne do leku” posiadające cechy leku lub nie posiadające takich cech. Wstępne zakwalifikowanie substancji do grupy „podobnych do leków” (ang.: *Drug-like; Drugability; Druglikeness*) daje szansę powodzenia w przyszłych badaniach nad wybranymi kandydatami na nowe leki. Większość sformułowanych założeń dotyczących fizjochemicznych parametrów oczekiwanych od NCE w wysokim stopniu koreluje z wartościami takich parametrów farmakokinetycznych jak: dostępność biologiczna, objętość dystrybucji, klirens całkowity $Cl_{(b)}$, biologiczny okres półtrwania $t_{1/2}$ czy średni czas przebywania leku w organizmie MRT. (Lipinski C.A. i wsp. 1997; Proost J.H. i wsp. 1997; Lombardo F. i wsp. 2000; Lombardo F. i wsp. 2002) Podstawą dla tego rodzaju rozważań są wybrane fizjochemiczne i fizykochemiczne cechy wspólne dla większości znanych obecnie leków.

Podstawą dla wytyczenia szeregu zależności wspólnych dla substancji zaliczanych do leków było zdefiniowanie przez Lipińskiego reguły pięciu (ang.: *Rule-of-5*). W regule tej wszystkie wartości numeryczne opisanych cech posiadają wspólny mianownik równy pięć (Lipinski C.A. i wsp. 1997; Lipinski C.A. 2000). Według tej reguły podstawowe cechy fizjochemiczne charakteryzujące każdy lek to: poniżej 5 HBD (ang.: *Hydrogen bond donors*) wyrażonych zwykle jako suma grup OH i NH. Poniżej 10 HBA (ang.: *Hydrogen bond acceptors*) wyrażonych zwykle jako suma grup O i N. Współczynnik rozdziału oktanol / woda LogP o wartości < 5 oraz o masie cząsteczkowej związku mniejszej od 500 Da (Testa B. i wsp. 2001; Waterbeemd H. i wsp. 2001).

Cel pracy:

Wskazanie metody opisu graficznego i matematycznego umożliwiającego szybkie eliminowanie lub kwalifikowanie substancji chemicznej do grona kandydatów na leki.

Modelowanie ADME in silico:

Elementy analizy SAR.

Techniki CADD (ang.: Computer aided drug design), czyli wspomaganie projektowania leków symulacjami komputerowymi, stworzyło podstawy dla nowego środowiska badawczego analizy farmakokinetycznej opartego na technologiach krzemowych - in silico.

W literaturze szeroko prezentowane są prawidłowości dotyczące właściwości fizjochemicznych jakie spełnione są dla około 95 % znanych obecnie leków. (Clark D.E. 2001; Lipinski C.A. 2000; Lipinski C.A. 2001; Egan W.J., Lauri G. 2002; Opera T.I. 2002; Veber D.F. i wsp. 2002; Viswanadhan V.N. i wsp. 2002). Odstępstwa od wyznaczonych wartości zwykle wpływają w sposób niekorzystny na dyspozycję leku w organizmie. Poniżej przedstawiono przykłady zasad wykorzystanych w analizie wybranych substancji leczniczych w obecnej pracy.

MW:	> 100 (160) lecz nie więcej niż 1000 (480)Da
Liczba atomów węgla:	> 3
Liczba wszystkich atomów:	> 20 lecz nie więcej niż 70
Niewielka liczba atomów:	azotu, tlenu, siarki, brak atomów metali
Suma grup OH i NH (HBD):	< 5
Suma atomów O i N(HBA):	< 10
LogP:	mniejszy równy - 0.4 (0) lecz nie większy, równy + 5.0
LogD:	dla leków podawanych <i>per os</i> 0 - 4 dla pozostałych -0.4 - 5.6 (Dressman J.B., Reppas C. 2000)
LogBB:	pomiędzy - 1.5 a + 1.2 (Jain R.K. 1999)
Perm(Caco-2):	< 20 - 34 nm/s mała, > 80 nm/s duża przepuszczalność
MiLogP (Moriguchi LogP):	> -2 (0) < + 4.6 (+ 6)
LogSw:	w zakresie - 6 do 0 (Avdeef A. 2001; Balaz S., Lukacova V. 2002)
PSA:	< (120) 140 Å ²
SASA(total):	pomiędzy 300 a 750 Å ²
MV:	pomiędzy 450 a 1400 Å ³
D:	pomiędzy 0.6 a 10.1

MW - masa cząsteczkowa (ang.: *Molecular weight*)

LogD - współczynnik rozdziału oktanol / woda dla elektrolitów (ang.: *Effective partition coefficient*)

LogP(o/w) - współczynnik rozdziału oktanol / woda (ang.: *Partition coefficient octanol / water*)

LogBB - współczynnik rozdziału krew / mózg (ang.: *Partition coefficient blood / brain*)

Perm(Caco-2) - Przepuszczalność błon komórkowych Caco-2 (ang.: *Caco-2 permeability*)

Sw - rozpuszczalność w wodzie (ang.: *Water solubility*)

PSA - polarna powierzchnia cząsteczki (ang.: *Polar surface area*)

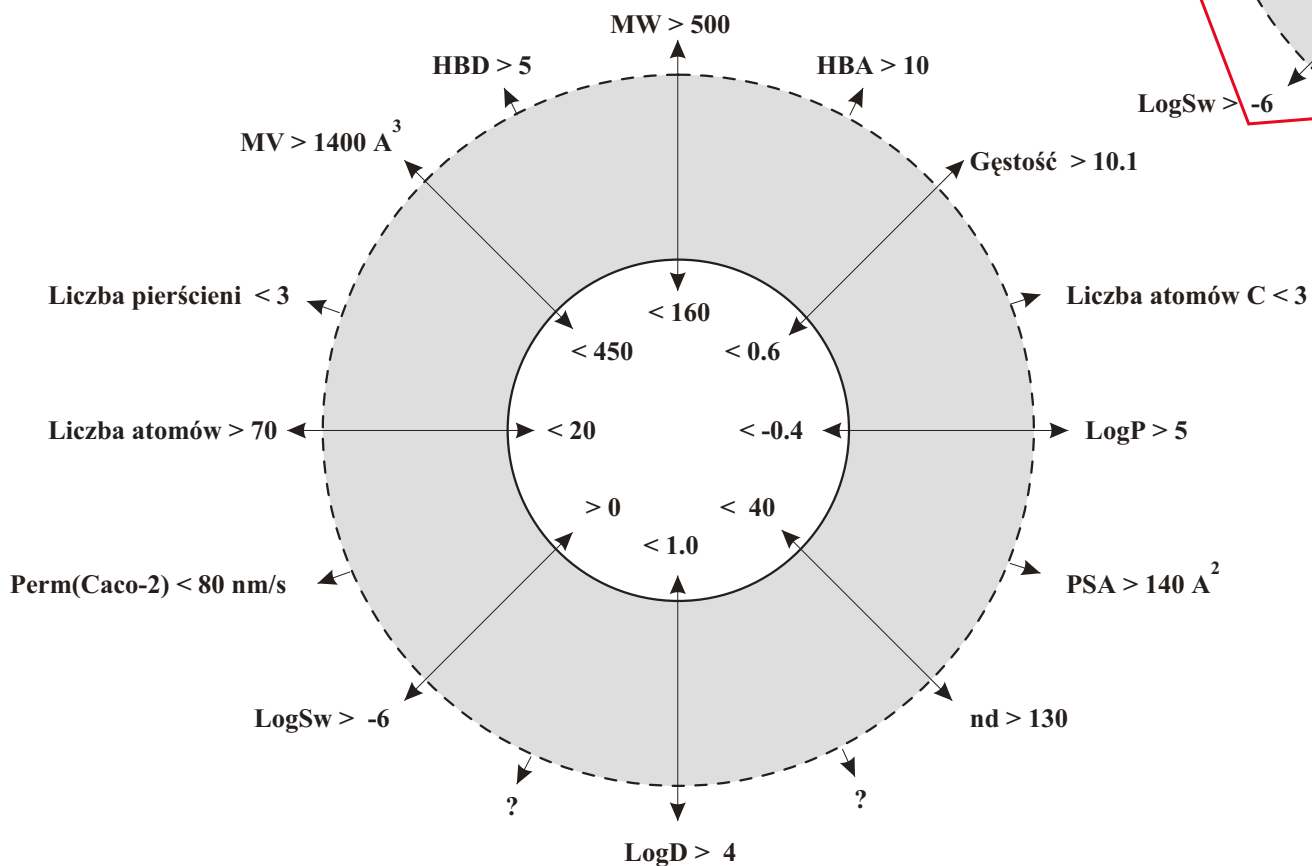
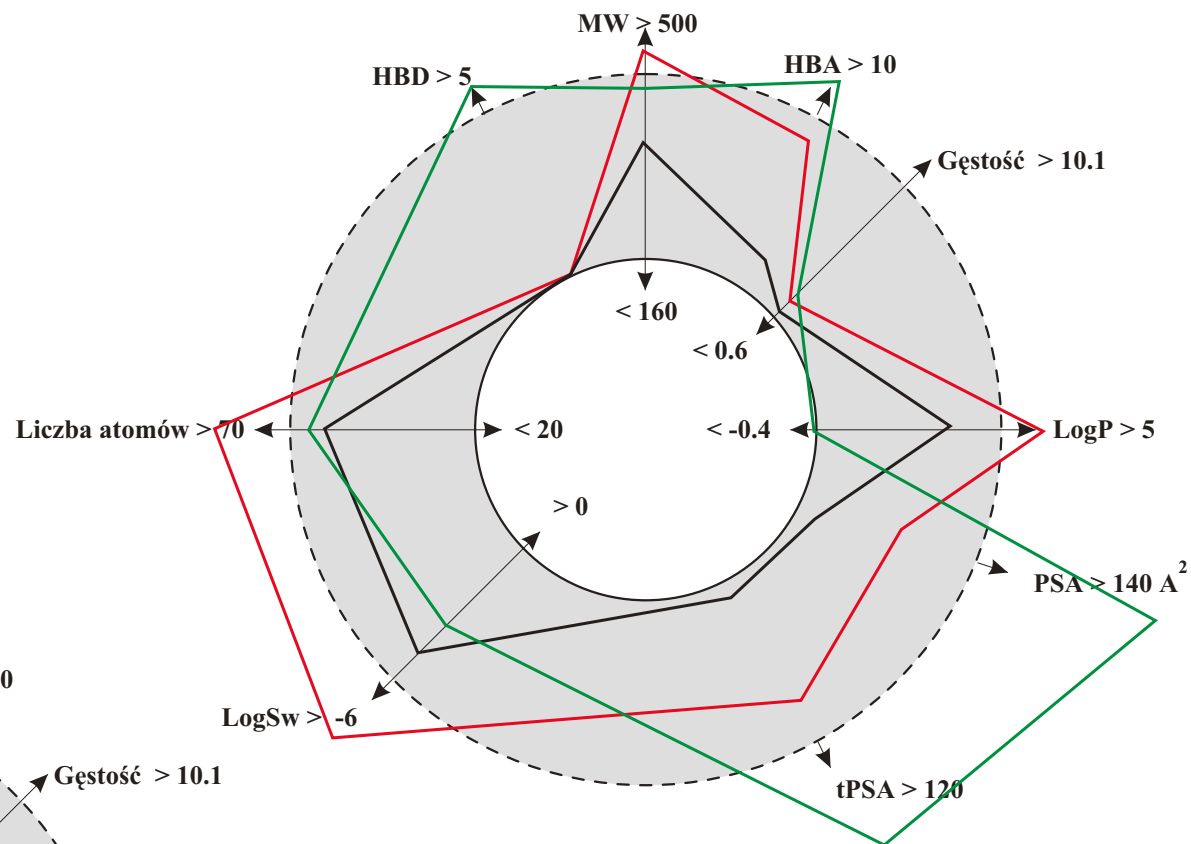
SASA - całkowita powierzchnia cząsteczki dostępna dla procesów rozpuszczania (ang.: *Solvent accesible surface area*)

MV - objętość cząsteczki (ang.: *Molar volume*)

D - moment dipolowy (ang.: *Dipole moment*)

Graficzna analiza SAR.

- tetrandryna
- myrtekaina
- minocyklina



Tarcza trafności wyboru.

Przedstawiony wykres skumulowany umożliwiłby graficzny opis podobieństwa badanej substancji do leku. 95 % znanych obecnie leków lokuje wartości wskazanych na wykresie parametrów w strefie wyróżnionej szarym kolorem.

Metody i wyniki.

SMILES

Kod SMILES (ang.: Simplified molecular input entry system) pozwala przetransponować wzór chemiczny związku do zapisu liniowego. Stanowi parametr możliwy do odczytania przez większość specjalistycznych aplikacji funkcjonujących w ramach technik CADD.

Analizie SAR poddano substancje lecznicze stosowane w zarejestrowanych obecnie lekach. Wybrano substancje z różnych grup terapeutycznych o zróżnicowanej budowie chemicznej. Analizy przeprowadzono wg. schematu:

- opis struktury chemicznej
- transformacja wzoru chemicznego do kodu SMILES

Substancje

poddane analizie: Kod SMILES związku:

1.	geraniol	<chem>C(=C(/CC\C=C(\C)C)C)/CO</chem>
2.	metronidazol	<chem>c1c(n(c(n1)C)CCO)[N+](=[O-])=O</chem>
3.	ibuprofen	<chem>c1(ccc(cc1)CC(C)C)C(C)C(O)=O</chem>
4.	propranolol	<chem>c12ccccc2ccc1OCC(CNC(C)C)O</chem>
5.	myrtekaina	<chem>C=1(C2C(C(CC=1)C2)(C)C)CCOCCN(CC)CC</chem>
6.	amoksycylina	<chem>C1(C(N2C1SC(C2C(=O)O)(C)C)=O)NC(C(c3ccc(cc3)O)N)=O</chem>
7.	minocyklina	<chem>c21c(ccc(c1C(C=3C(C2)CC4C(\C(=C(/C(C4(C=3O)O)=O)C(=O)N)O)N(C)C)=O)O)N(C)C</chem>
8.	tetrandryna	<chem>c21CCN(C6c1cc(c(c2)OC)Oc3c4C(N(Cc4cc(c3OC)OC)C)Cc5cc(c(cc5)OC)Oc7ccc(C6)cc7)C</chem>
9.	taxoter	<chem>C\1C3(CC(/C(=C/1CC(C(C2C(CC5C(C2C3OC(=O)c4cccc4)(OC(=O)C)CO5)O)=O)O)C)OC(C(O)C(NC(=O)OC(C)(C)C)c6cccc6)=O)O</chem>

- obliczenia podstawowych parametrów fizjochemicznych
- analiza graficzna wybranych związków
- sformułowanie i obliczenie wskaźnika podobieństwa do leku - DA

Do analizy wybranych parametrów zastosowano kilka programów komputerowych:

ACD Chem/Sketch - <http://www.acdlabs.com>
Molinspiration - <http://www.molinspiration.com>
Vega OpenGL- <http://www.ddl.unimit.it>
EPI Suite- <http://www.epa.gov>

Duża ilość informacji płynących z analizy SAR badanej cząsteczki wymaga stosowania programów decyzyjnych, statystycznej analizy danych oraz specjalistycznego oprogramowania. Jednak w trakcie selekcji NCE można także kierować się wynikami prostych równań i algorytmów. Jednym z prostszych może być przedstawiony niżej algorytm, opisujący podobieństwo do leku DA - (ang.: *Drugability*). Wynik większy od zera charakteryzuje leki o optymalnych parametrach e-ADME. Mniejszy od zera wskazuje na duże prawdopodobieństwo trudności w optymalizacji farmakokinetycznej leku *in vivo*.

$$DA = \{[(\text{LogP}(o/w) \times 4) - (\text{LogP}(o/w) / 4)] + [(12 - (\text{HBA} + \text{HBD})) / 5] + [(80 / \text{PSA}) \times 10]\} / \sqrt{\text{MW}}$$

Ryc. Analiza danych związków w oparciu o hipotetyczne parametry fizjochemiczne -
DA - indeks podobieństwa do leku dla związków o: znacznej masie cząsteczkowej, dużej liczbie donorów i akceptorów wodoru, niewielkiej wartości współczynnika rozdziału oktanol woda i dużej wartości pola polarnego cząsteczki, oscyluje wokół lub jest mniejszy od zera.

DA	MW	LogP	tPSA	HBA	HBD
1,00	100	3,0	60	1	1
0,59	150	2,5	70	2	2
0,38	200	2,0	80	3	3
0,26	250	1,5	90	4	4
0,17	300	1,0	100	5	5
0,11	350	0,5	110	6	6
0,07	400	0,0	120	7	7
0,03	450	-0,5	130	8	8
0,01	500	-1,0	140	9	9
-0,02	550	-1,5	150	10	10
-0,03	600	-2,0	160	11	11
-0,09	700	-2,5	200	12	12

		geraniol	metronidazol	ibuprofen	propranolol	myrtekaina	amoksycylina	minocyklina	tetrandryna	taxoter
Masa cząsteczkowa	Da	154,2	171,2	206,3	259,3	265,4	365,4	457,5	622,8	779,8
Gęstość	g/cm	0,866	1,450	1,029	1,093	0,930	1,540	1,550	1,172	1,400
Polaryzacja		19,7	16,2	24,1	31,3	32,4	36,3	46,0	70,6	77,7
MiLogP		3,3	-0,3	3,3	3,3	3,6	-0,6	-2,3	5,0	3,7
LogP(o/w)		3,6	0,0	4,0	3,5	4,6	1,0	-0,1	7,6	2,0
SR		4,4	2,1	4,6	4,3	5,1	2,7	2,0	7,0	3,4
LogSw		-2,8	-0,8	-3,7	-3,1	-4,1	-2,0	-3,8	-10,4	-6,6
tPSA	A ²	20,2	83,9	37,3	41,5	12,5	133,0	164,6	61,9	224,5
PSA	A ²	57,3	164,9	86,4	35,8	10,1	252,1	214,4	64,7	205,4
SASA	A ²	360,9	323,0	397,3	453,0	495,6	443,4	575,8	882,3	934,0
fPSA	%	15,9	51,1	21,7	7,9	2,0	56,9	37,2	7,3	22,0
HBA		1,0	6,0	2,0	3,0	2,0	8,0	10,0	8,0	15,0
HBD		1,0	1,0	1,0	2,0	0,0	5,0	6,0	0,0	5,0
DA (drugability)		0,91	0,41	0,59	0,45	0,56	0,04	-0,05	0,15	-0,11

Ryc. Analiza danych dla gupy różnorodnych pod względem budowy chemicznej oraz zaszeregowania terapeutycznego substancji leczniczych.

Wnioski.

Mapowanie leków (ang.: Drug clustering)

Analiza podstawowych parametrów fizjochemicznych leku - SAR jest punktem wyjścia do mapowania leków (mapowania ich struktury). Obecnie analiza SAR dotyczy także badań porównawczych grup farmakoforowych i toksykoforowych leków z uwzględnieniem podobieństw i różnic ich ułożenia w przestrzeni.

I.

Przedstawione zasady ilustrujące zależność między budową chemiczną a parametrami farmakokinetycznymi stanowią obecnie bardzo szeroką grupę. Stosowanie ich w praktyce optymalizacji farmakokinetycznej leku w fazie przedklinicznej wymaga tworzenia odpowiednich algorytmów i równań matematycznych. Składowymi dla proponowanych równań mogą być zarówno dane ujęte w postaci zasad (ilość atomów mniejsza niż...) jak i danych liczbowych opisujących konkretną cechę ($\text{LogP}_{(o/w)}$). Przedstawione wyżej równanie odnosi się zarówno do pierwszej jak i drugiej grupy cech.

II.

Opisana wyżej analiza dotyczy jedynie wąskiej grupy związków chemicznych. Jednak o zdecydowanie różnej budowie chemicznej, parametrach fizykochemicznych oraz zaszeregowaniu terapeutycznym. Zastosowany algorytm zaszeregował większość substancji poddanych analizie do grupy "podobnych" do leków, wskazując jednocześnie na znaczne różnice e-ADME.

III.

Precyzyjna odpowiedź na pytanie o stopniu podobieństwa badanego związku chemicznego do typowego leku możliwa jest dzięki zastosowaniu kryteriów statystycznych. W ramach tych kryteriów matematyczny opis substancji chemicznej stanowi podstawę do opisu leku wyrażonego w postaci klastrów czyli ciągów złożonych z pojedynczych bitów (ang.: *Drug clustering*). Wymaga to jednak zastosowania złożonego oprogramowania. Analiza podstawowych parametrów fizjochemicznych badanej molekuly za pomocą testowanego algorytmu ułatwia wstępne selekcjonowanie związków - kandydatów w fazie badań przedklinicznych.

IV.

Proponowane równanie służące określaniu stopnia podobieństwa do leku może być stosowane w dostępnych arkuszach kalkulacyjnych. Modyfikowane odpowiednio do potrzeb stanowi łatwe w obsłudze narzędzie wstępnej analizy typu SAR.

1. Dressman J.B., Reppas C.: In vitro - in vivo correlations for lipophilic poorly water -soluble drugs. *Eur J Pharm Sci* 2000, 11:2, 73-80
2. Avdeef A.: Physiochemical profiling (solubility, permeability, and charge state). *Curr Top in Med Chem.* 2001, 1:4, 1-76
3. Balaz S., Lukacova V.: Subcellular pharmacokinetics and its potential for library focusing. *J of Mol Graph and Mod* 2002, 20, 479-490
4. Clark D.E.: Prediction of intestinal absorption and blood brain barrier penetration by computational methods. *Comb Chem & High Throug Scree* 2001, 4, 477-496
5. Egan W.J., Lauri G.: Prediction of intestinal permeability. *Adv Drug Del Rev* 2002, 54 s. 273-289
6. Grabowski T.: Farmakokinyka i biofarmacja. www.biokinetica.pl 2002-2003
7. Jain R.K.: Understanding barriers to drug delivery: High Resolution in vivo imaging is key. *Clin Canc Res* 1999, 5, 1605-1606
8. Jorgensen W.L., Duffy E.M.: Prediction of drug solubility from Monte Carlo Simulations. *J Pharm Sci* 1994, 83, 948-951
9. Kramer S.D., Wunderli-Allespach H.: Physiochemical properties in pharmacokinetics lead optimization. *IL Farmaco* 2001, 56, 145-148
10. Kratochwil N.A., Huber W., Müller F., Kansy M., Gerber P.R.: Predicting plasma protein binding of drugs: a new approach. *Biochem Pharm* 2002, 64, 1355-1374
11. Lipinski C.A., Lombardo F., Dominy B.W., Feeney P.J.: Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Del Rev* 1997, 23, 3-25
12. Lipinski C.A.: Drug-like properties and the causes of poor solubility and poor permeability. *J of Pharm and Tox Met* 2000, 44, 235-249
13. Lipinski C.A.: The design of drug-like properties. *Lipinski ACD European Workshop* 2001, 1-26
14. Lombardo F., Obach R.S., Shalaeva M.Y., Gao F.: Prediction of volume distribution values in humans for neutral and basic drugs using physiochemical measurements and plasma protein binding. *J Med Chem* 2002, 45:13, 2867-2876
15. Lombardo F., Shalaeva M.Y., Tupper K.A., Gao F., Abraham M.H.: ELogPoc: A tool for lipophilicity determination in drug discovery. *J Med Chem* 2000, 43:15, 2922-2928
16. Opera T.I.: Virtual screening in lead discovery: A viewpoint. *Molec* 2002, 7, 51-62
17. Proost J.H., Roggeveld J., Wierda J.M.K.H., Meijer D.K.F.: Relationship between chemical structure and physiochemical properties of series of bulky organic cations and their hepatic uptake and biliary excretion rates. *J of Pharm and Exp Ther* 1997, 282:2, 715-726
18. Testa B., Waterbeemd H., Folkers G., Guy R.H.: Pharmacokinetic optimization in drug research. *WILEY-VCH, Verlag Helvetica Chimica Acta Zurich*, 2001, 257- 591
19. Veber D.F., Johnson S.R., Cheng H.Y., Smith B.R., Ward K.W. Kopple K.D.: Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 2002, 6:45, 2615-2623
20. Viswanadhan V.N., Balan C., Hulme C., Cheetham J.C., Sun Y. Knowledge based approaches in the design and selection of compound libraries for drug discovery.: *Curr Opp in Drug Disc and Dev* 2002, 5:3, 400-406
21. Waterbeemd H., Smith D.A., Jones B.C.: Lipophilicity in PK design: methyl, ethyl, futile. *J Comput Aided Mol Des* 2001,15:3, 273-286