IVth International Symposium Computational Methods in Toxicology and Pharmacology Integrating Internet Resources (CMTPI-2007)

H-Bonding Parametrization in Quantitative Structure–Activity Relationships and Drug Design (Plenar lecture)

Oleg Raevsky

**Department of Computer-Aided Molecular Design,
Institute of Physiologically Active Compounds,
Russian Academy of Sciences, 142432, Chernogolovka,
Moscow region, Russia**

H-bond Functional Groups in Biological Molecules

Donors	Acceptors
O—H N—H N^+—H S—H C—H P—H	O^-—P O^-—C O^-—S O=C O=P O=S S=C O N N

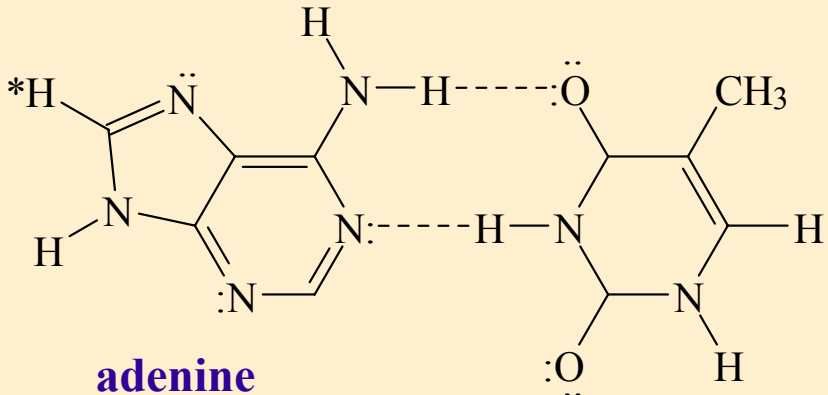
Most common H-bonds in biological molecules

N—H····O—C	Amino acids, zwitterions, proteins
O—H····O—C	Carboxyl acids, acid nitrates
H····O—C—O	Peptides, proteins, nucleosides
N—H····O=C	Nucleotides, nucleic acids
O—H····O—H	Hydrates
O—H····O(H)—C O—H····O(C)—H	Carbohydrates
O—H····N(C)=C N—H····N(C)=C O—H····O^-—P N—H····O^-—P	Proteins, nucleosides, nucleotides, nucleic acids

G.A.Jeffard, W.Saenger. "H-bonding in Biological Structures", 1991, Berlin: Springer

DNA and proteins are held together in their defined three-dimensional structures primarily by hydrogen bonds. The double helix of DNA and RNA structures, the peptide and protein secondary structures like α-helices, β-sheets, β- and γ-loops, and the tertiary structures of proteins are formed by H-bonds (enthalpy contributions) and by hydrophobic contacts (primarily entropy cotribution).

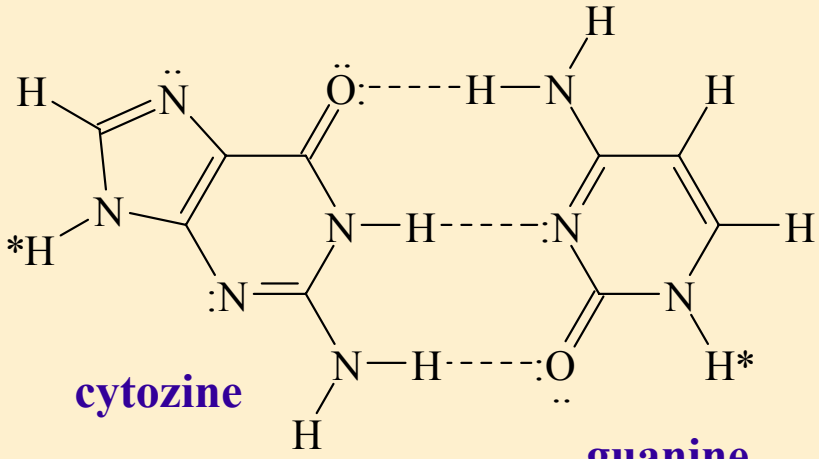
H-bonding in DNA Replication Process



adenine

40 kJ/mol

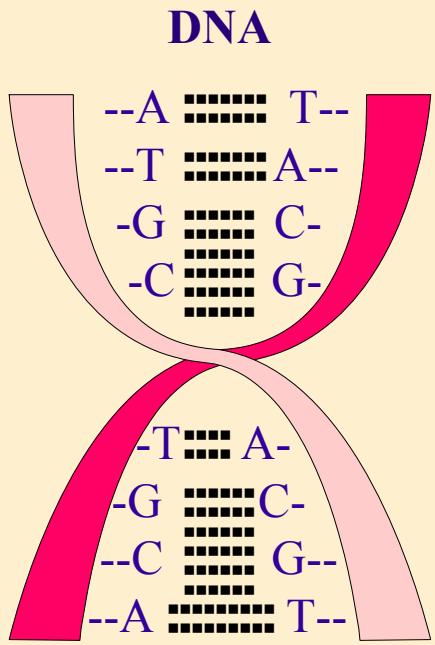
thymine



cytozine

guanine

60 kJ/mol



Right Recognition Probability = 10^7

L.Pauling, P.Pauling. Chemistry, 1975/ Freeman and C., San Francisco

"Despite of all attempts to arrive at a better understanding of the role of water and of hydrogen bonds in biological systems and of all the individual enthalpy and entropy terms that are involved in disolvation, hydrogen-bond formation, and hydrophobic interactions we are far from a satisfactory situation".

Kubinyi, Hydrogen Bonding, the Last Mystery in Drug Design. in:*Pharmacokinetic Optimization in Drug Research. Biological, Physicochemical, and Computational Strategies,* B. Testa, H. van de Waterbeemd, G. Folkers and R. Guy, Eds., Helvetica Chimica Acta and Wiley-VCH, Zürich, **2001**, 513-524.

2D Indirect and Indicator H-bonding Descriptors

Name or meaning	Type	Symbol	Author(s)
The difference between octanol-water and cyclohexane-water logP values	Indirect	IH	P.Seiler (1974)
The difference logP octanol-water for polar and nonpolar chemicals with the same molecular weight	indirect	EW	I. Moriguchi **(1982)**
The atomic charge on the hydrogen atom	indirect	QH	L.YWilson G.R .Famini (1991) J.C.Dearden (1997) E.Gancia (2001)
The energy of the lowest unoccupied molecular orbital	indirect	ELUMO	
The electron donor superdelocalizability	indirect	DE	
The self-atom polarizability	Indirect	PE	
The charge on the most negatively charge atom	indirect	QMH	
The energy of the highest occupied molecular orbital	indirect	EHOMO	
Surface electrostatic potential maxima	indirect	Vs,max	J.S.Murray (1992)
Electrostatic potential minima	indirect	Vmin	
Numbers of O and N Types and numbers of H-bonds The sum of OHs and NHs (donors), the sum of Ns and Os (acceptors)	Indicator Indicator indicator	Na, Nd	Fujita (1977) Charton (1982) C.Lipinski (1997)

O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

Thermodynamics of H-bond Complexation

$$\Delta G = -RT\ln K = \Delta H - T\Delta S$$

Where ΔH is a change in enthalpy;

ΔG is a change in free energy;

ΔS is a change in entropy;

K is a binding constant ;

T is an absolute temperature (in Kelvin).

R is the universal gas constant.

Binding constants of H-bond complexes, ΔH or (and) ΔG are normally used to create thermodynamic hydrogen bond scales.

2D Direct Thermodynamics H-bonding Descriptors

Name or meaning	Type	Symbol	Author(s)
Infrared or NMR Spectral Shifts	thermodynamic	$\Delta\nu$OH, δ H	M.K.Kroger R.S.Drago A.Joganssen (1970 ths)
Parameter of acidity	thermodynamic	**A**	R.W.Taft, M.J.Kamlet M.Abraham (1980 ths)
Parameter of basicity	thermodynamic	**B**	
H-bond acceptor enthalpy and free energy factors	thermodynamic	Ea , Ca	O.A.Raevsky (1980 ths)

O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

Enthalpy and Free Energy H-bond Factors

$\Delta H = 4.96$ (kJ/mole) E_a E_d

$\Delta G = 2.43$ (kJ/mole) C_a C_d + 5.70 (kJ/mole)

E_a is H-bond acceptor enthalpy factor

E_d is H-bond donor enthalpy factor

C_a is H-bond acceptor free energy factor

C_d is H-bond donor free energy factor

$\Delta H_{calc} = -0.27(\pm 0.45) + 1.00(\pm 0.02)\Delta H_{exp}$

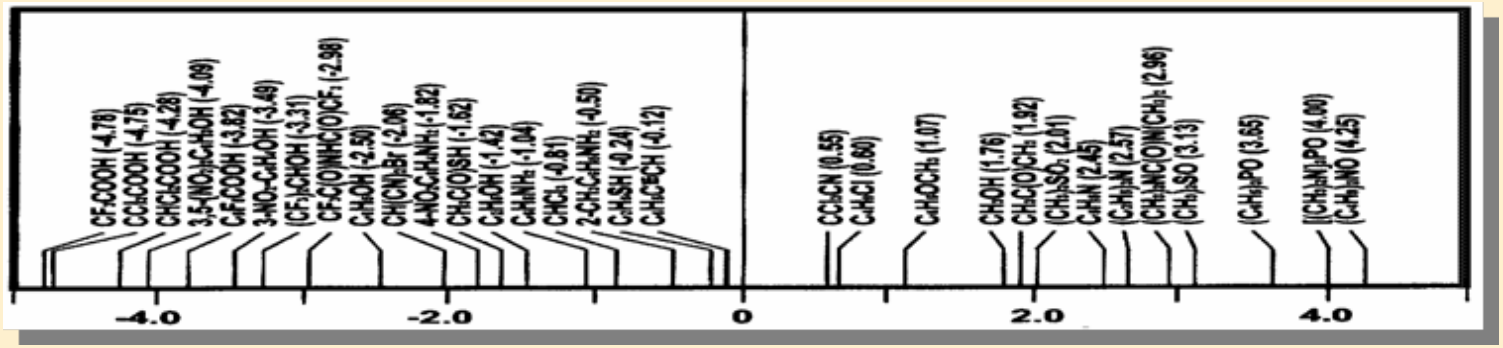
n = 936, r2 = 0.91, s = 2.70, F = 9553

$\Delta G_{calc} = -0.07(\pm 0.12) + 1.00(\pm 0.01)\Delta G_{exp}$

n=936, r2=0.97, s=1.11, F=28556

Raevsky, O.A., Grigor'ev, V.Ju., Kireev, D., Zefirov, N.S. Complete Thermodynamic Description of H-Bonding in the Framework of Multiplicative Approach. *Quant. Struct.-Act. Relat.,* **1992***, .11* 49-64.

FREE ENERGY H-BOND FACTOR SCALE

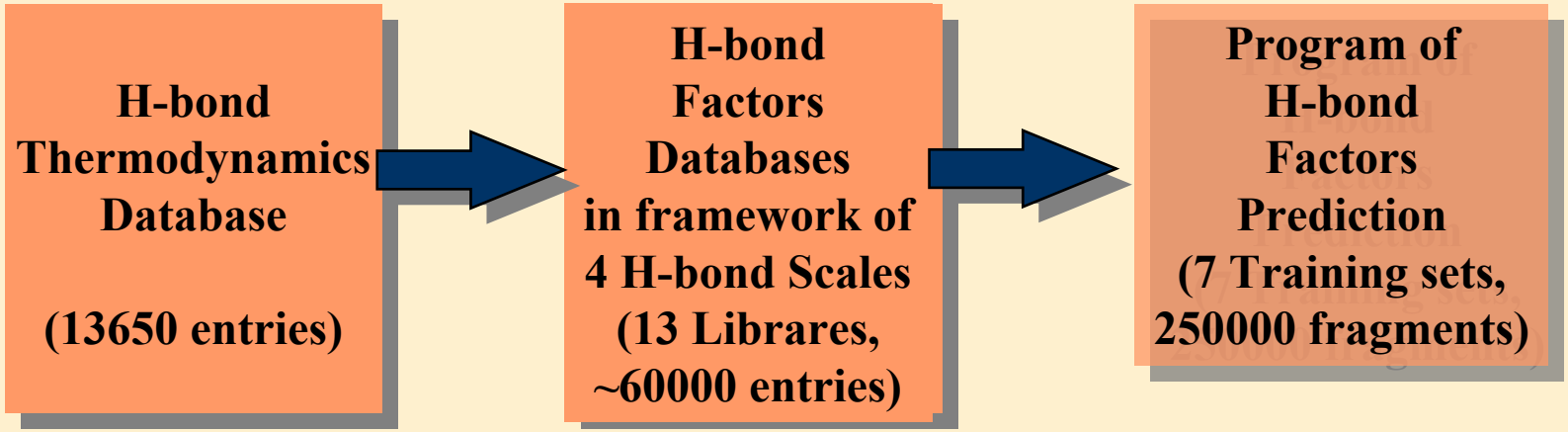


H-bond donors

H-bond acceptors

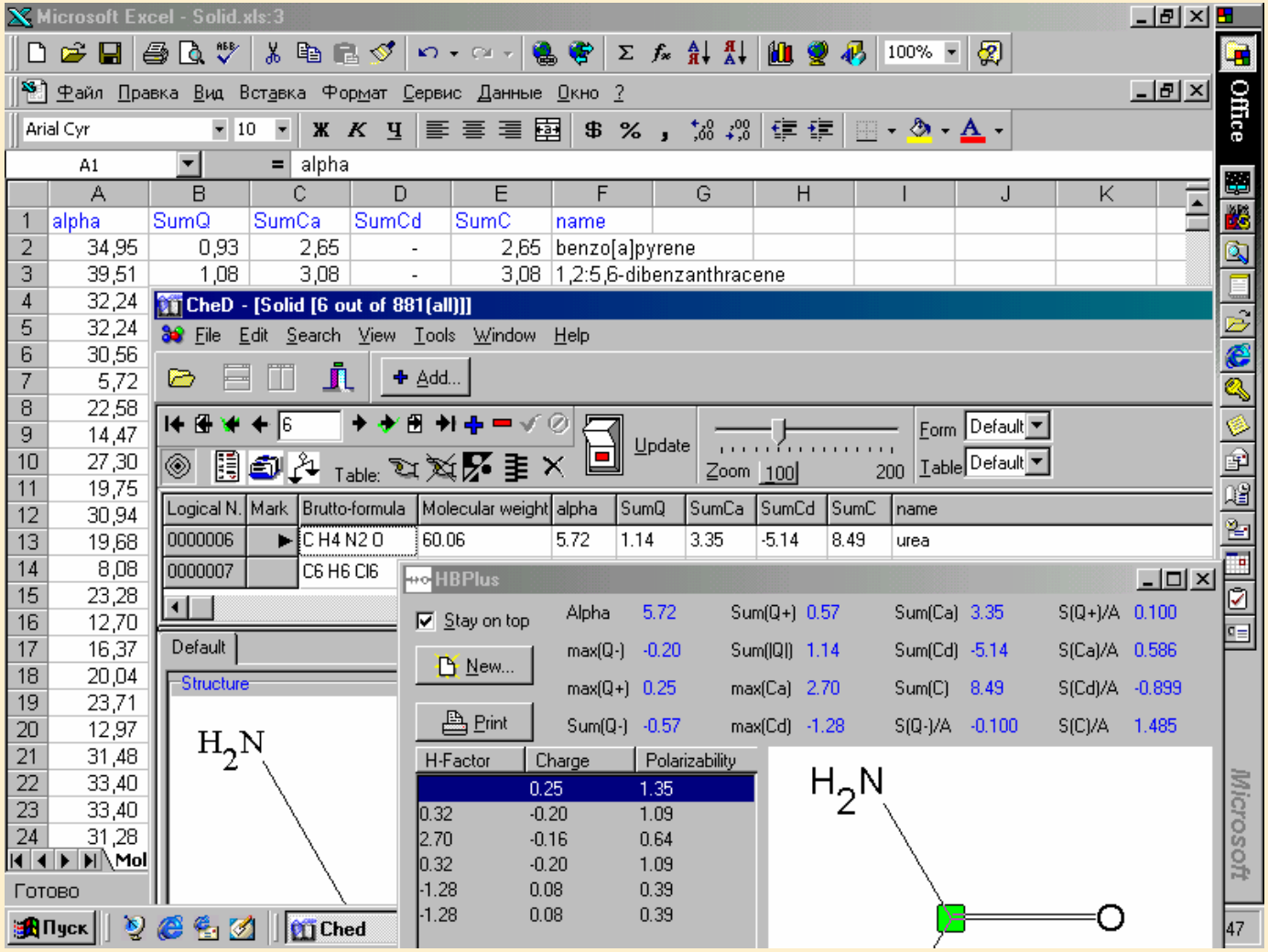
Raevsky, O.A., Grigor'ev, V.Ju., Solov'ev, V.P., Kireev, D.B., Sapegin, A.M., Zefirov, N.S. Drug Design H-Bonding Scale. In *QSAR: Rational Approaches in the Design of Bioactive Compounds*. Silipo, C., Vittoria, A. (eds.), Elsevier, Amsterdam, **1991**.

The Common Scheme of HYBOT Program Package

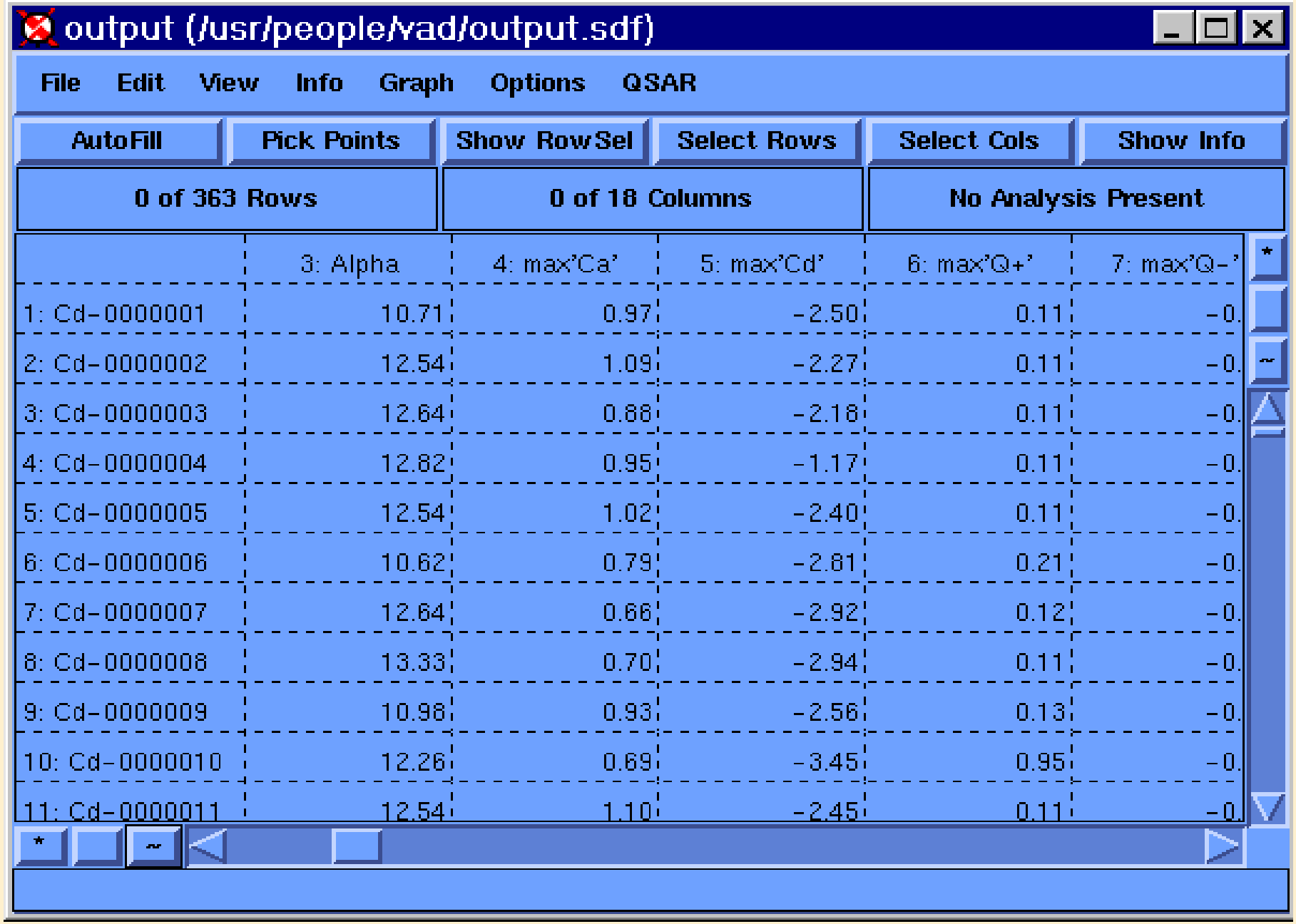


Raevsky, O.A. Hydrogen Bond Strength Estimation by means of HYBOT. In *Computer-Assisted Lead Finding and Optimization*. Van de Waterbeemd, H., Testa, B., Folkers, G. (eds.), Basel: Verlag, Basel, **1997**.

Predicted QSAR Descriptors for Set of Structures in HYBOT-PLUS-2006/CHED

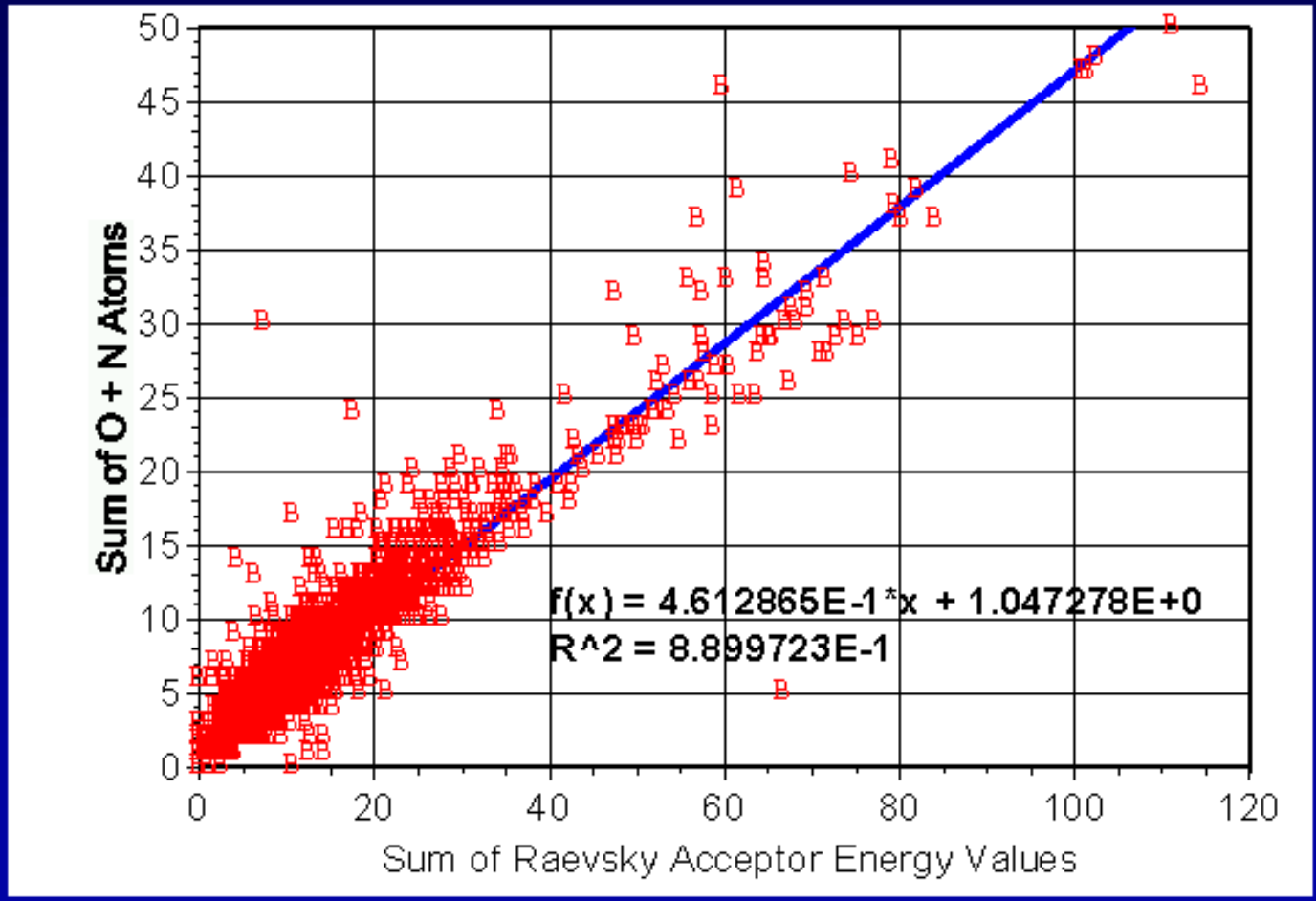


Raevsky, O.A., Grigor'ev, V.Ju., Trepalin, S.V. HYBOT (Hydrogen Bond hermodynamics) program package. *Registration by Russian State Patent Agency* N 990090 *of 26.02.99.*



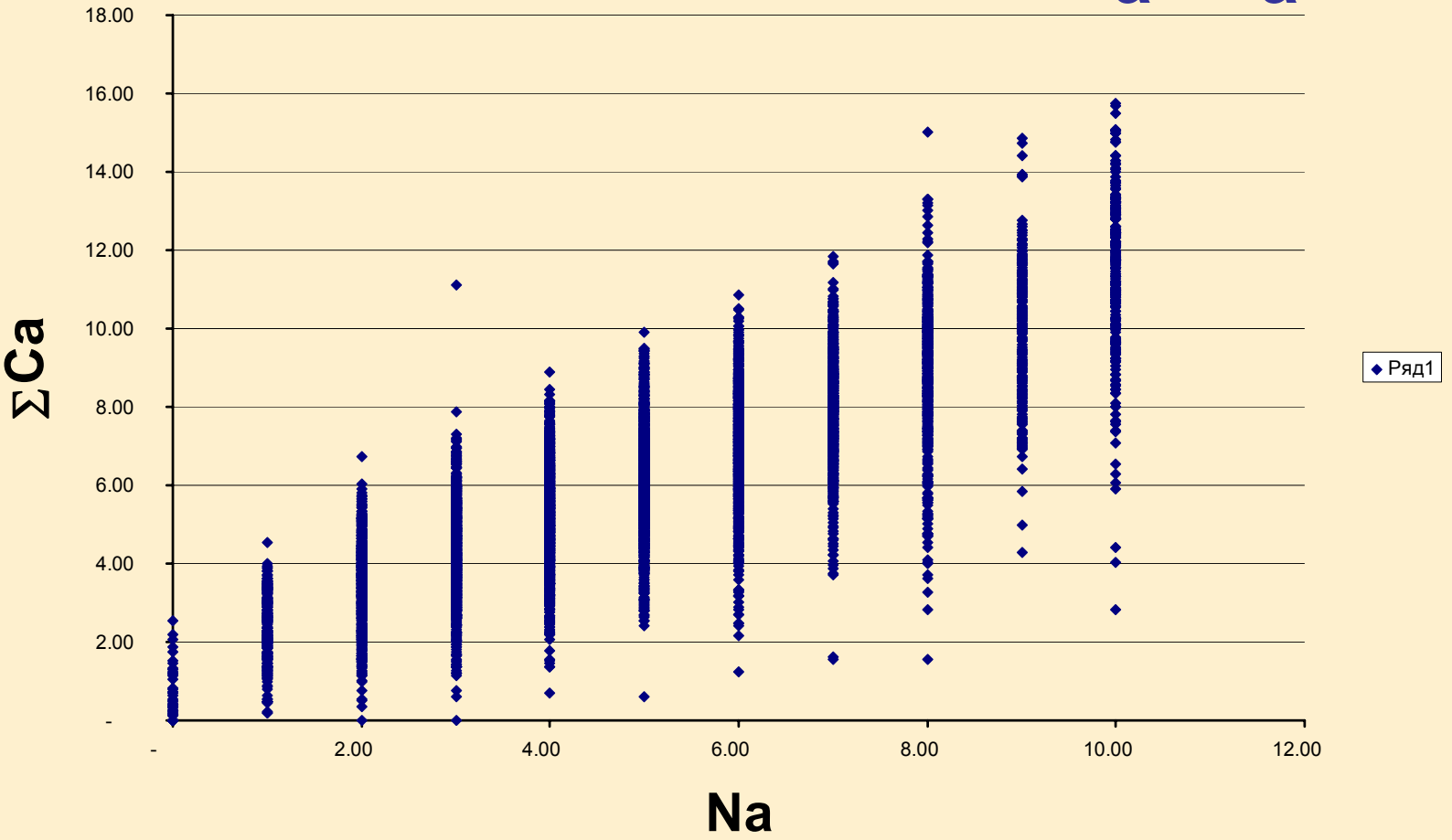
Raevsky, O.A., Skvortsov, V.S., Grigor'ev, V.Ju., Trepalin, S.V. HYBOT in UNIX/LINUX.*Registration by Russian State Patent Agency* N 2002610496 *of 05.02.02*.

Simple H-bond calcs - not as good for acceptors



Lipinski, C.A. Drugs Structure and Properties, Past and Present. Can We Design Drugs with Beautiful Properties? http://www.iainm.demon.co.uk/spring99/lipins_n.pdf

Correlation $\Sigma C_a/N_a$



O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

3D H-bonding Descriptors

Name or meaning	Symbol & Formule	Authors
Polar surface area	PSA, PSAd, TPSA	H.Van de Waterbeemd (1996) K.Palm (1997)
Optimum H-bond enthalpy for three types of complexes	Em(OHO),Em(OHN) Em(NHN)	P.Goodford (1989)
3D H-bond distance descriptors	HB+ +, HB- -, HB+ -	O.Raevsky (1987)
Similarity Indexes of Spectra of H-bond interactions	SIS+ +, SIS - -, SIS + -	O.Raevsky (1999)
Hydrogen bonding potential	MHBP	G.Caron (2001)

Molecular Drug Properties. Measurement and Prediction. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

3D HYBOT Potentials Calculations

$$\Delta H = k_1(\text{kcal/mol})E_a E_d = E_m \qquad (1)$$

$$E_r = C/r^8 - D/r^6 , \qquad (2)$$

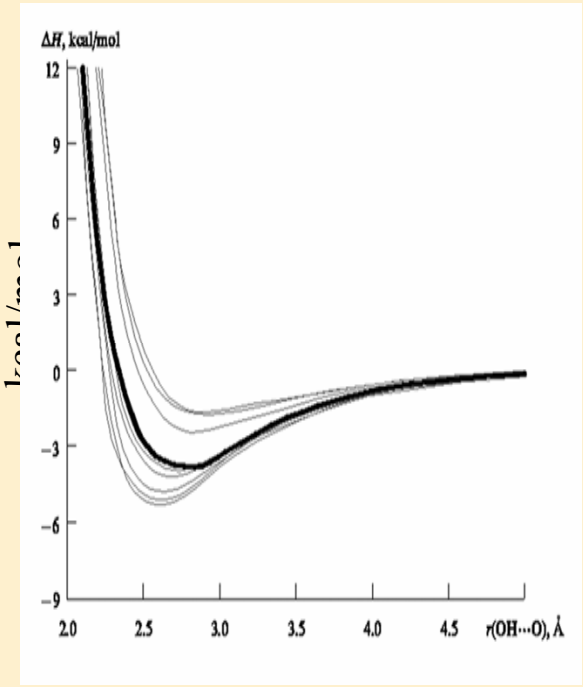
where E_m is optimum H-bond enthalpy in kcal/mole; $C = -3E_m r^8 E^8/\text{mol}$; $D = -4E_m r_m{}^6 E^6/\text{mol}$; r is the distance between the acceptor and the donor heavy atoms.

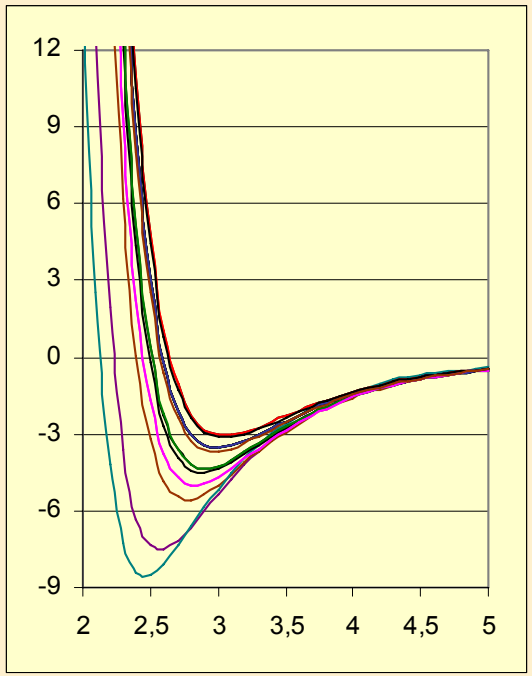
$$r_m = k'\log[(k''-E_m)/(E_m)] + k_0 \qquad (3)$$

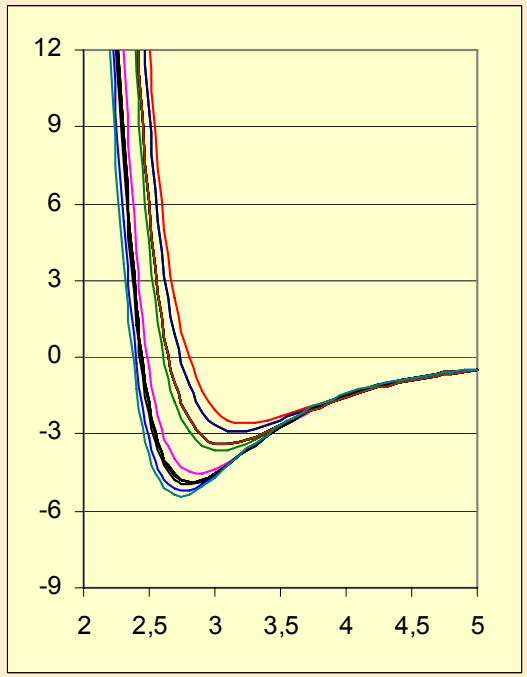
Thus, for each specific pair of atoms participating in an H-bond, its H-bonding potential can be calculated on the basis of eqs. (1), (2) and (3).

O.A.Raevsky, V.S.Skvortsov, 3-D hydrogen bond thermodynamics (HYBOT) potentials in molecular modelling, *Journal Computer-Aided Molecular Design*, **16, N1, pp. 1-10 (2002).**

3D HYBOT (HYdrogen BOnd Thermodynamics) Potentials





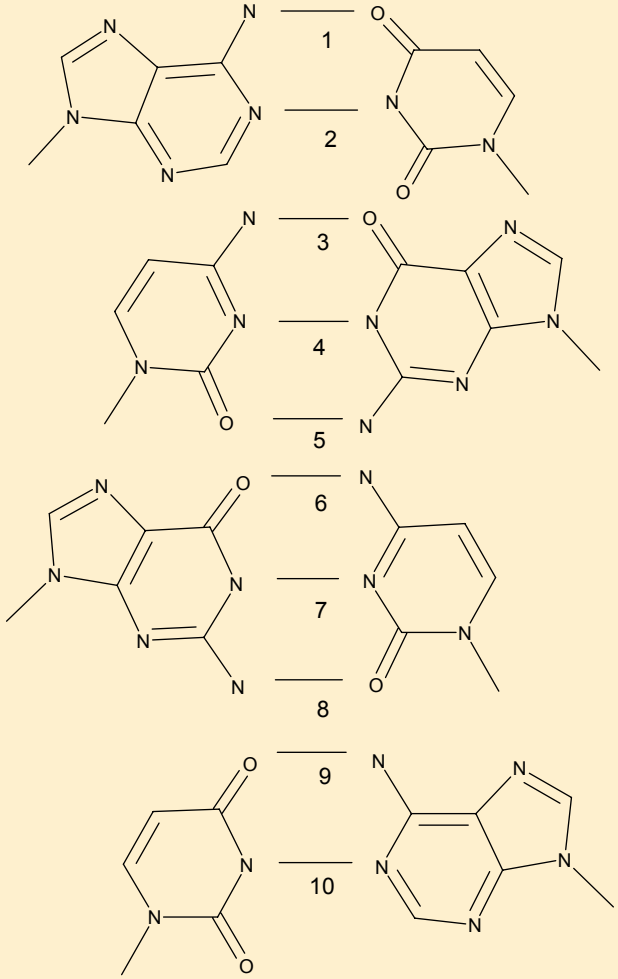


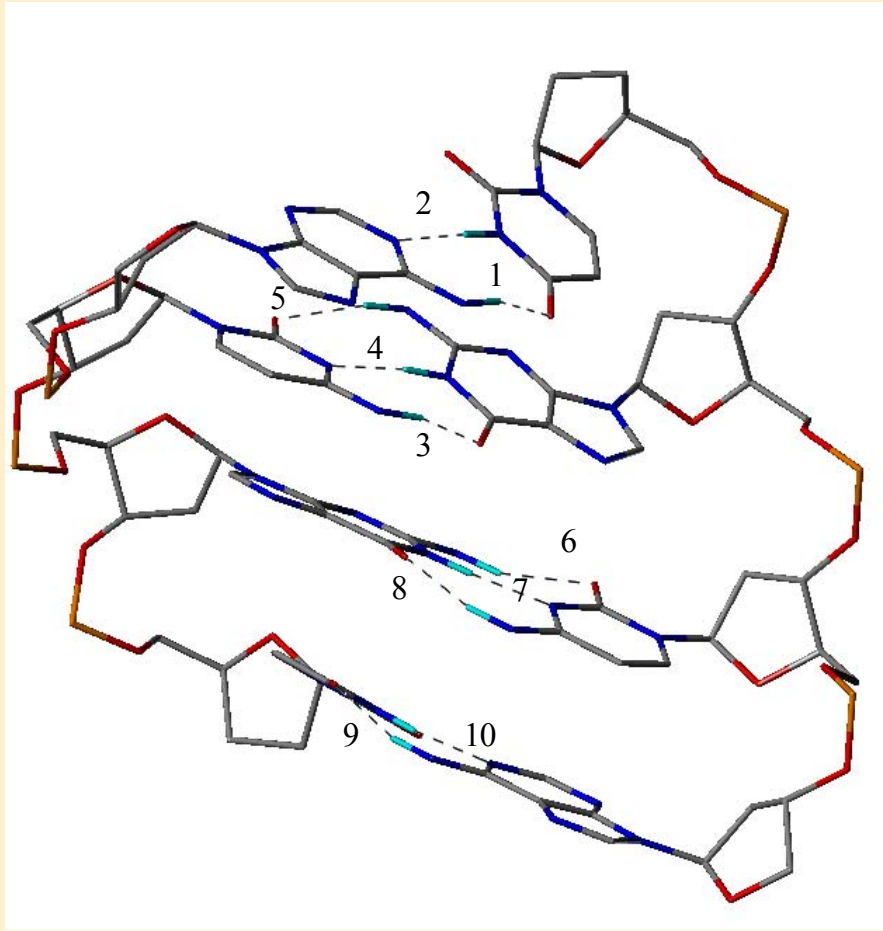
r (Å)

OH…O N(H)…O NH…N

O.A.Raevsky, V.S.Skvortsov, 3-D hydrogen bond thermodynamics (HYBOT) potentials in molecular modelling, *Journal Computer-Aided Molecular Design*, **16, N1, pp. 1-10 (2002).**

Hydrogen Bonds System in Fragment of Double Helix of RNA in A Form





O.A.Raevsky, V.S.Skvortsov, 3-D hydrogen bond thermodynamics (HYBOT) potentials in molecular modelling, *Journal Computer-Aided Molecular Design*, **16, N1, pp. 1-10 (2002).**

The Calculated Energy Values for Bases Interactions in Double Helix of RNA in A Model

Donor	Cd	Accep -tor	Ca	Em (this work) (kcal/ mol)	Em [21] (kcal/ mol)	Rm (this work)(A)	Rm [21] (A)	R (A)	Er (this work) (kcal/m ol)	Er (as in [21] (kcal/m ol)	Ehb(this work) kcal/mol)	Ehb(as [21]) (kcal/mol)
N(39)-H(56)	-1.83	N(11)	1.99	-4.84	-2.00	2.840	3.20	2.953	-4.69	-1.53	-4.64	-1.53
N(29)-H(51)	-1.83	N(20)	1.19	-2.90	-2.00	3.186	3.20	2.909	-2.01	-1.31	-1.96	-1.27
N(30)-H(52)	-1.29	O(21)	2.06	-3.53	-2.80	2.995	3.00	3.078	-3.48	-2.76	-2.81	-2.23
N(18)-H(48)	-1.29	O(27)	1.78	-3.06	-2.80	3.057	3.00	2.724	-1.35	-1.80	-1.20	-1.61
N(9)-H(43)	-1.29	O(37)	1.78	-3.06	-2.80	3.057	3.00	2.833	-2.45	-2.51	-2.18	-2.23
N(39)-H(56)	-1.83	N(11)	1.99	-4.84	-2.00	2.840	3.20	2.954	-4.69	-1.53	-4.63	-1.53
N(29)-H(51)	-1.83	N(20)	1.19	-2.90	-2.00	3.186	3.20	2.909	-2.01	-1.31	-1.96	-1.27
N(30)-H(52)	-1.29	O(21)	2.06	-3.53	-2.80	2.995	3.00	3.078	-3.48	-2.76	-2.81	-2.23
N(18)-H(48)	-1.29	O(27)	1.78	-3.06	-2.80	3.057	3.00	2.724	-1.35	-1.80	-1.20	-1.61
N(9)-H(43)	-1.29	O(37)	1.78	-3.06	-2.80	3.057	3.00	2.833	-2.45	-2.51	-2.18	-2.23
									ΣEr = -27.96	ΣEr = -19.82	ΣEhb = -25.55	ΣEhb= -17.73

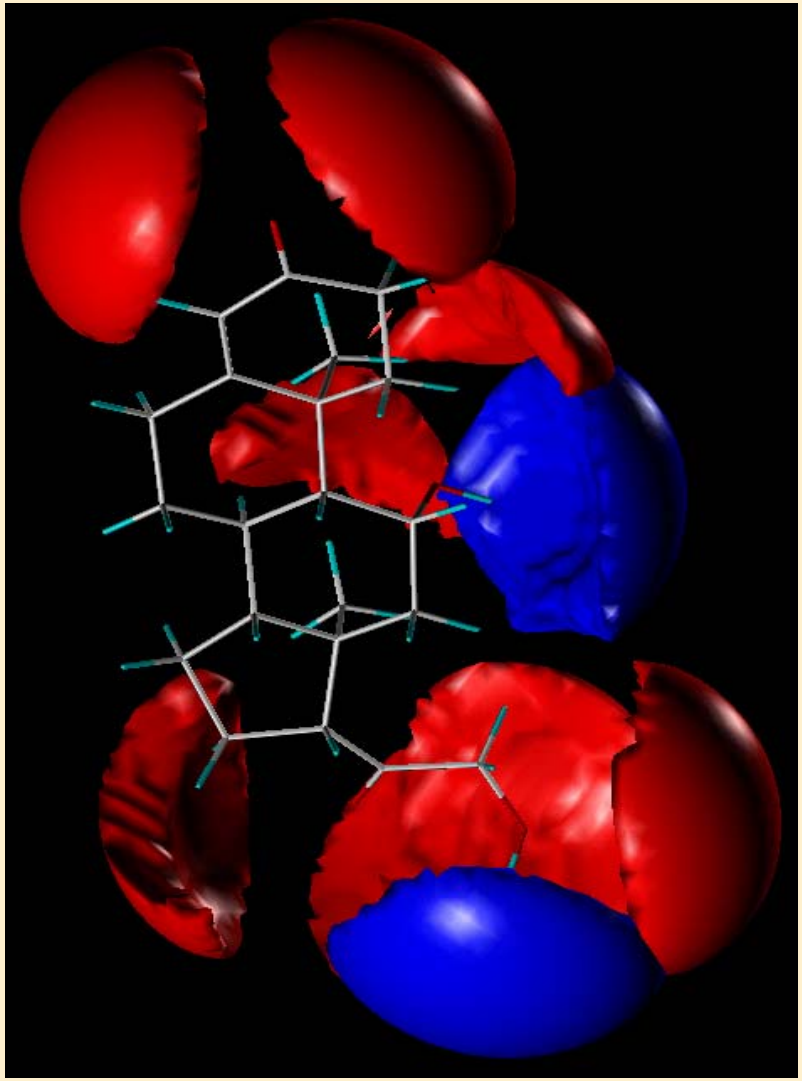
O.A.Raevsky, V.S.Skvortsov, 3-D hydrogen bond thermodynamics (HYBOT) potentials in molecular modelling, *Journal Computer-Aided Molecular Design*, **16, N1, pp. 1-10 (2002).**

3D H-bonding Descriptors

Name or meaning	Symbol & Formule
Van der Waal's acceptor surface area which is proportional to Ea	WEASA
Van der Waal's acceptor surface area which is proportional to Ca	WFEASA
Van der Waal's donor surface area which is proportional to Ed	WEDSA
Van der Waal's donor surface area which is proportional to Cd	WFEDSA
Surface area around a molecule where optimum enthalpy of interactions of acceptor atoms with H-bond donor probe is realized	OEASAprobe
Surface area around a molecule where optimum free energy of interactions of acceptor atoms with H-bond donor probe is realized	OFEASAprobe
Surface area around a molecule where optimum enthalpy of interactions of donor atoms with H-bond acceptor probe is realized	OEDSAprobe
Surface area around a molecule where optimum free energy of interactions of donor atoms with H-bond acceptor probe is realized	OFEDSAprobe
Sum of enthalpy values(kcal/m Å2) of interactions between the acceptor atoms in a molecule and donor probe on the surface OEASA	SIEAprobe
Sum of enthalpy values(kcal/m Å2) of interactions between the donor atoms in a molecule and an acceptor probe on the surface OEDSA	SIEDprobe

Raevsky, O.A., Skvortsov, V.S. Quantifying hydrogen bonding in QSAR and molecular modeling.*SAR & QSAR in Environmental Research,* **2005**, *12*, 1-14 .

OFEASA (red) and OFEDSA (blue)



Raevsky, O.A., Skvortsov, V.S. Quantifying hydrogen bonding in QSAR and molecular modeling.*SAR & QSAR in Environmental Research,* **2005**, *12*, 1-14 .

HYBOT Program Package

"Calculation of many different descriptors is possible using a range of commercially available software packages, such as Sybyl, Cerius2, Tsar, Malconn-Z, HYBOT etc."

H.Waterbeemd, D.S.Smith, K.Beaumont, D.K.Walker. *J. Med. Chem.,* **44**, 1313-1333 (2001)

"A further refinement in PSA approach is expected to come from taking into account the strength of the hydrogen bonds, which in principle already is the basis of the HYBOT approach"

H.Waterbeemd. *Drug Bioaviability.,* Wiley-VCH, Weinheim, 3-20 (2003)

The Components of the «Critical Quartet» log K_{ow} Values Assessed by Four Commercial Software Packages

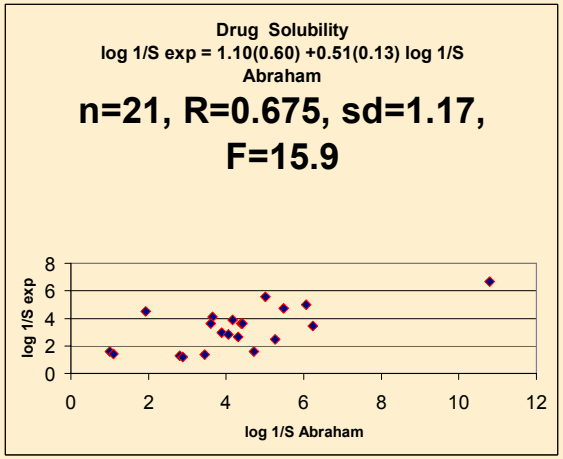
Partition coefficient (log Kow) values of 103 compounds in four different solvent-water systems (alkane-water, octanol-water, chloroform-water and propylene glycol dipelargonate (PGDP)-water --- the «critical quartet») were published by Leahy et al. (1992). J. Chem. Soc., Perkin Trans., 2, pp. 723-731.

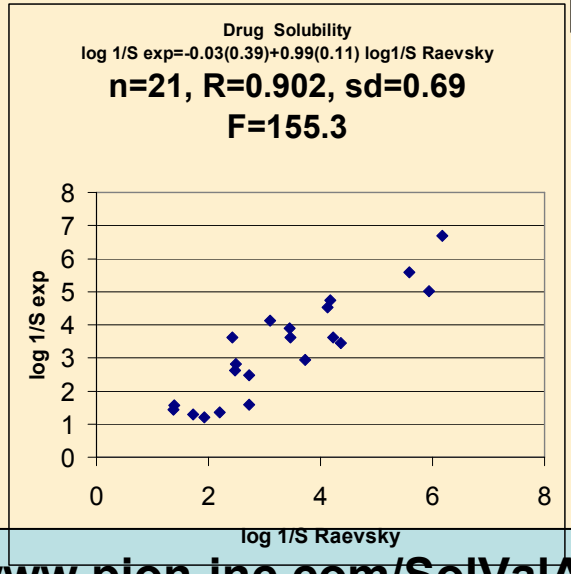
We have used ABSOLV and three other commercial software packages (HYBOT-PLUS, Molecular modeling Pro and QsarIS) that calculate descriptors of similar classes, to model partitioning in the four solvent-water systems.

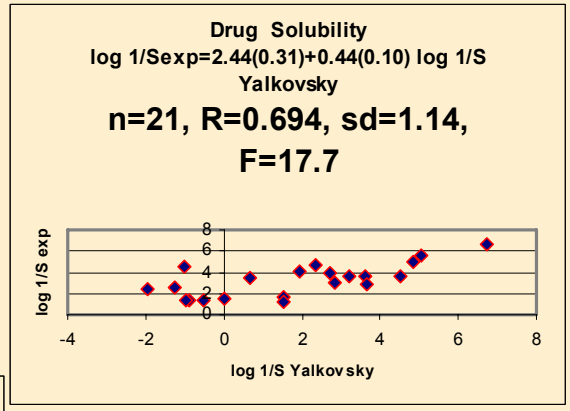
Overall, HYBOT-PLUS gave the best results.

J.C.Dearden, D.Bentley.
SAR QSAR Environ Res., **13** (1), pp. 185–197 (2002)

Experemental solubility values measured by pSOL and the calculated values using the Abraham, Yalkovsky-Valvani and Raevsky descriptors

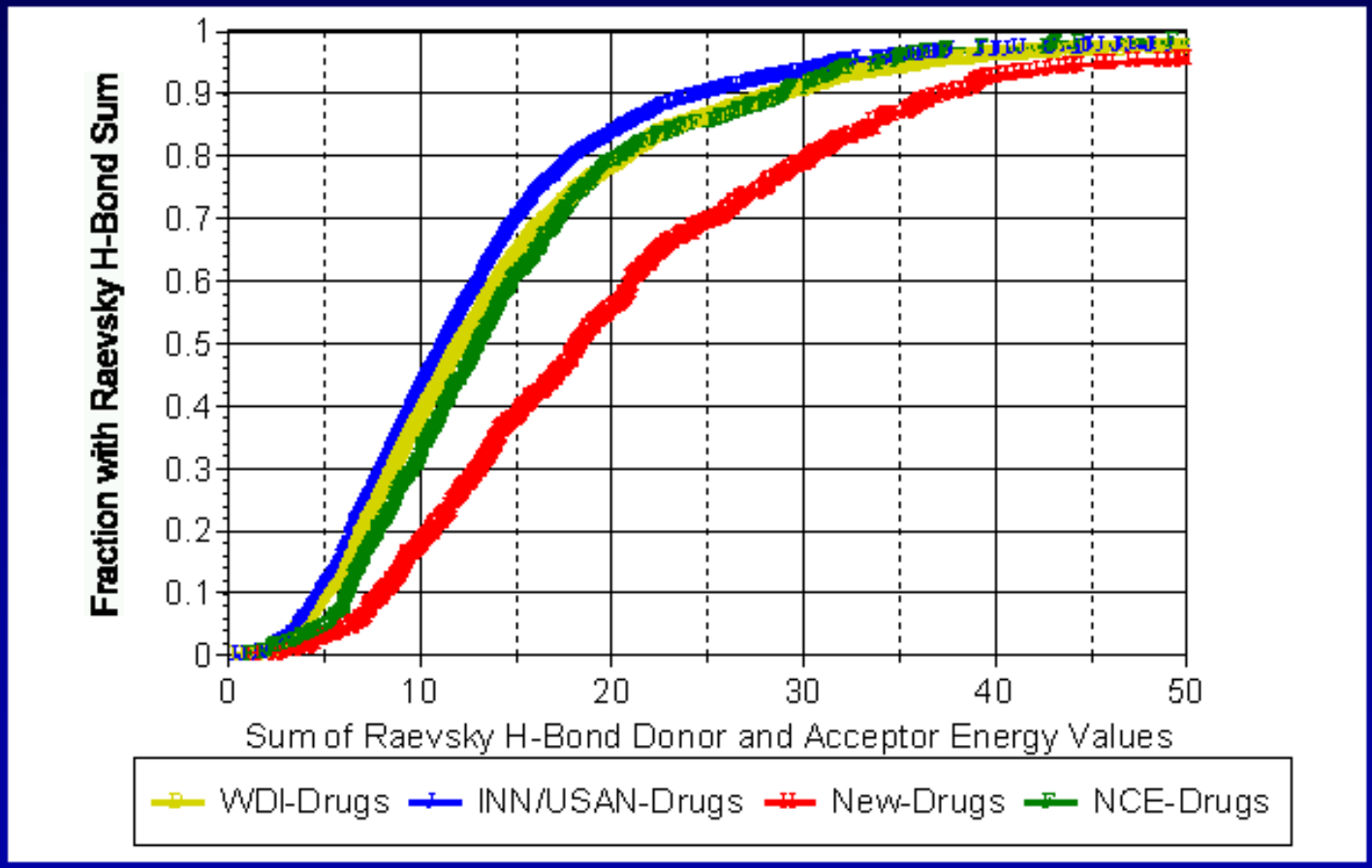






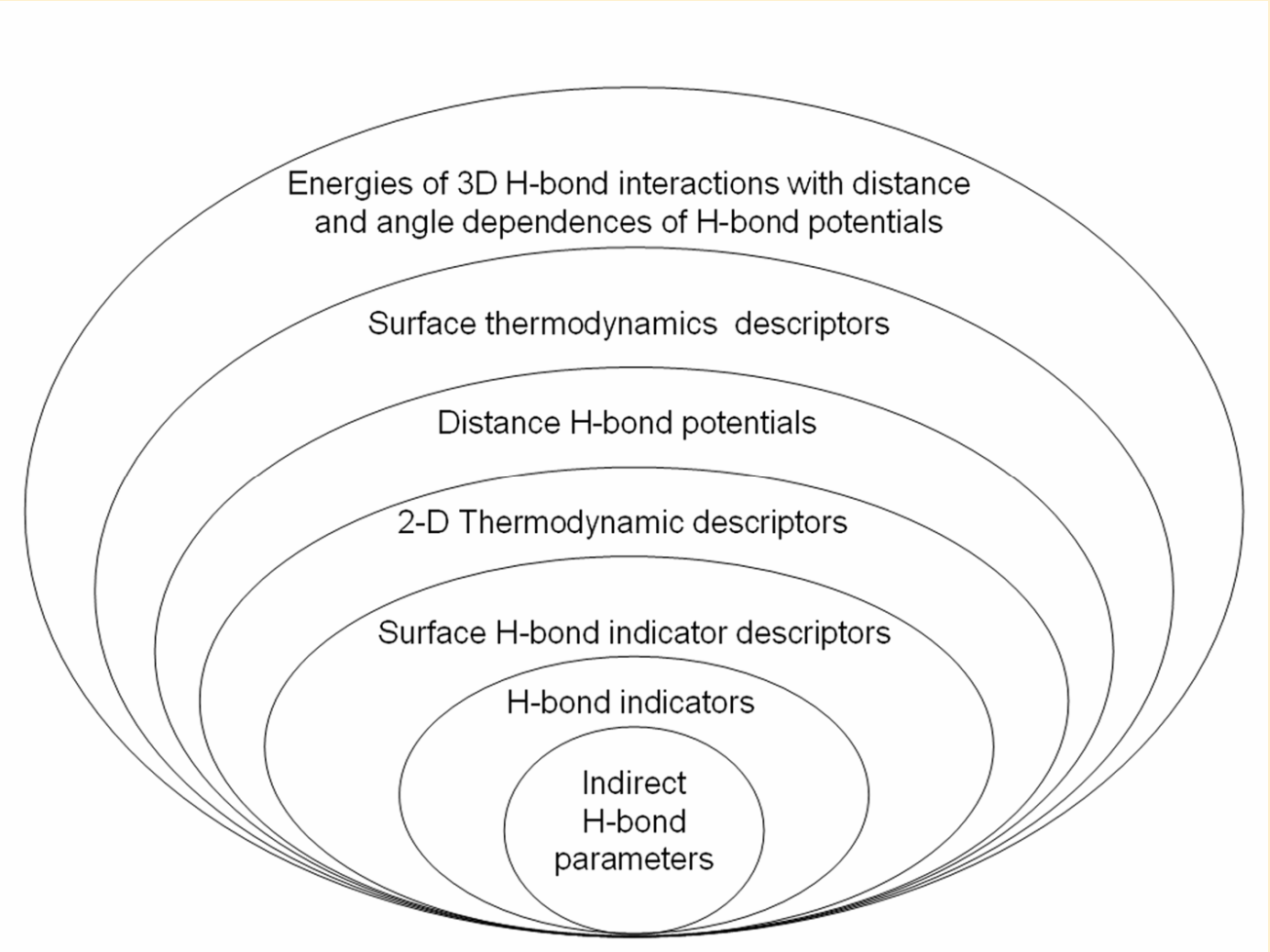
(http://www.pion-inc.com/SolValAAPS01.pdf

Newer drugs are less permeable



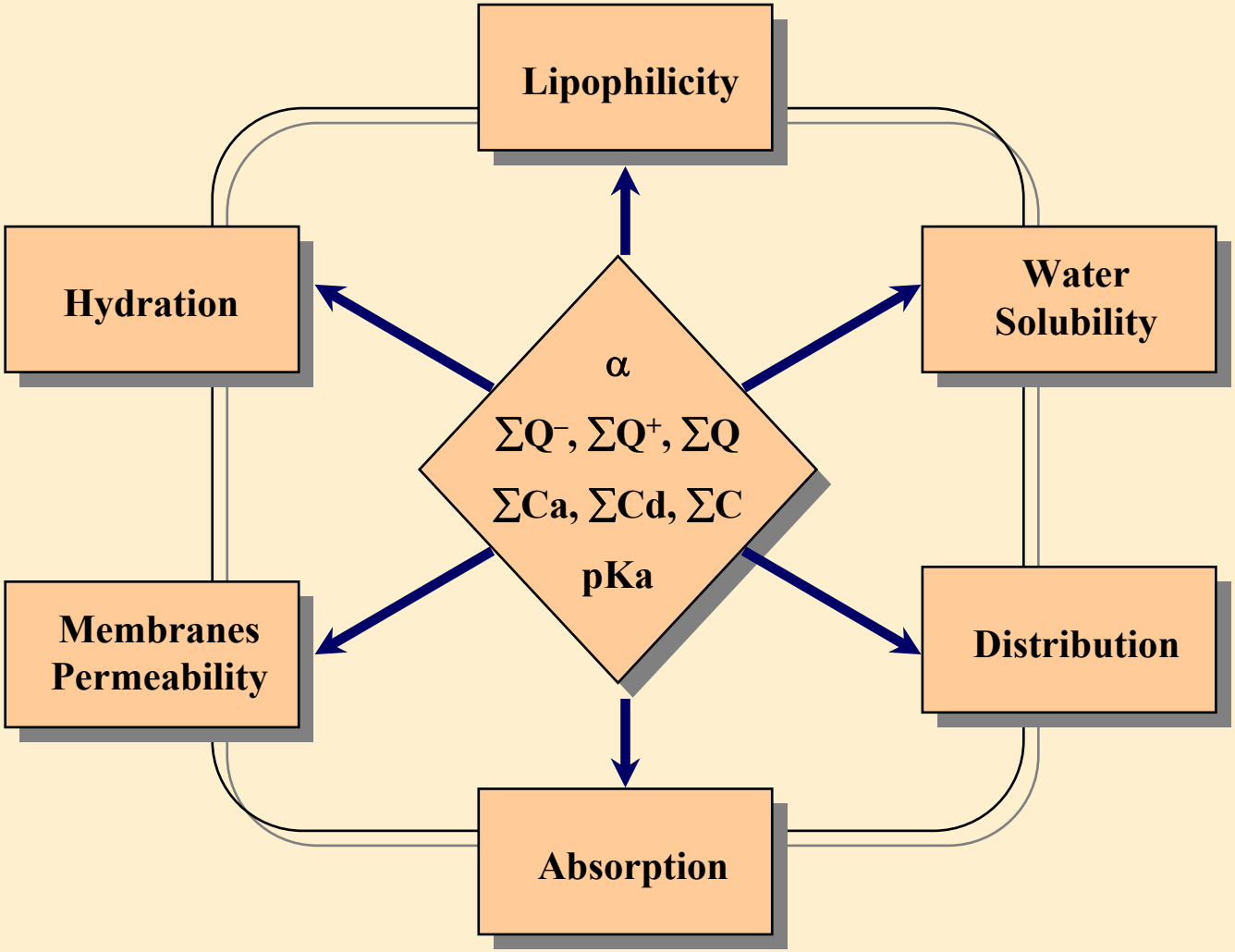
Lipinski, C.A. Drugs Structure and Properties, Past and Present. Can We Design Drugs with Beautiful Properties? http://www.iainm.demon.co.uk/spring99/lipins_n.pdf

Information Content and Relationships of H-bond Descriptors

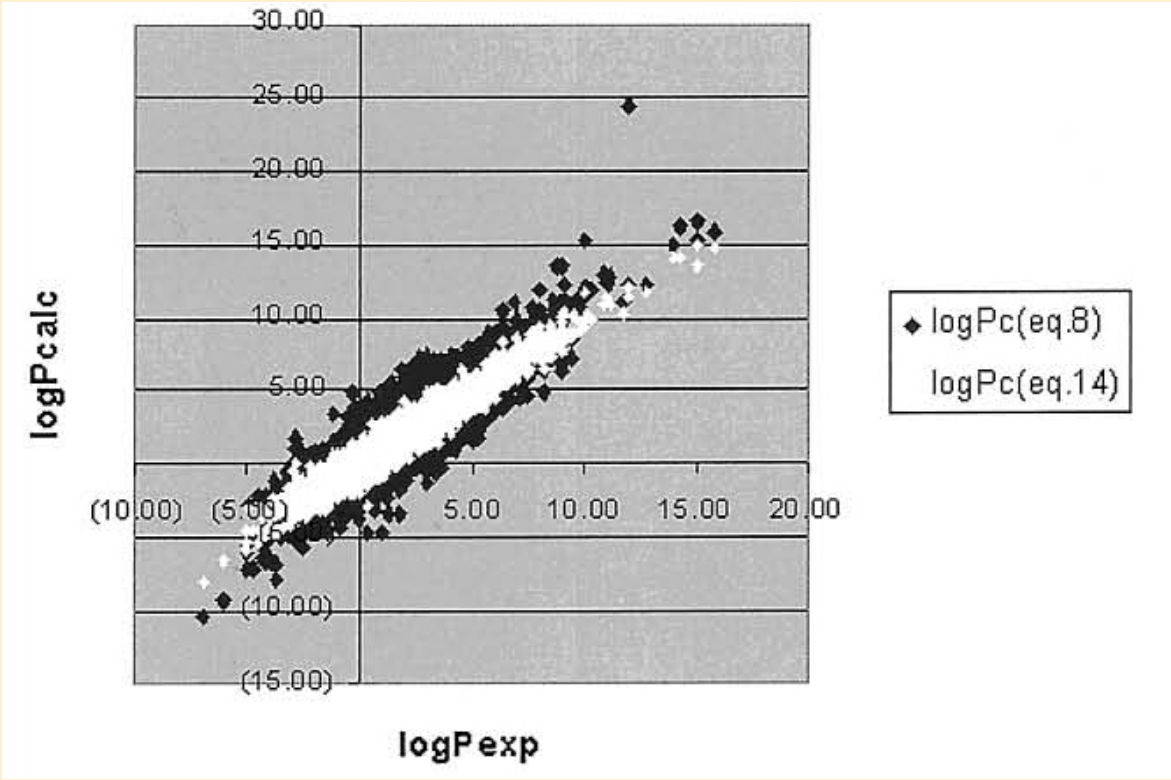


O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

H-bond Descriptors in QSAR



log P Calculation



$$\log P = 0.267\alpha - 1.00\ \Sigma C_a$$

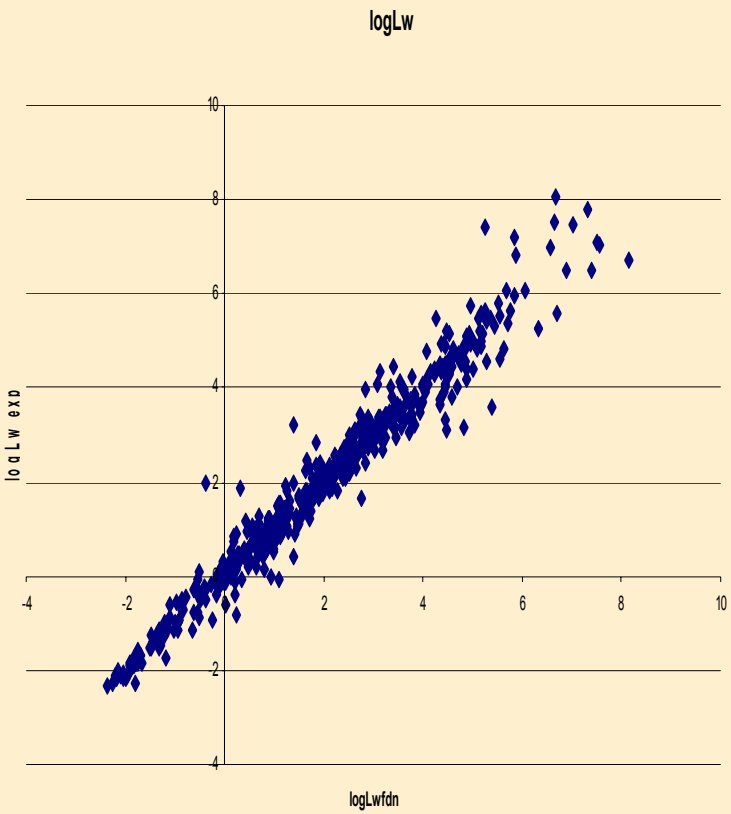
(N = 2850, R = 0.970, S = 0.23)

$$\log P_{i(calc)} = \log P_{nn} + 0.267(\alpha_i - \alpha_{nn}) - 1.00(\Sigma C_{ai} - \Sigma C_{nn})$$

(N = 10937, R = 0.971, S = 0.26)

O.A.Raevsky, S.V.Trepalin, E.P.Trepalina, V.A.Gerasimenko. Chem. Inf. And comput. Sci., 2002, v. 42 , pp. 540-549.

Chemicals Distribution in Water-Air System



$$\log L^w = 1.74\Sigma Ca$$
(n = 426, r^2 = 0.725, s = 1.21)

$$\log L^w = 1.38\Sigma Ca - 0.83\Sigma Cd$$
(n = 426, r^2 = 0.854, s = 0.93)

$$\log L^w = -0.052\alpha + 1.73\Sigma Ca - 0.82\Sigma Cd$$
(n = 426, r^2 = 0.851, s = 0.84)

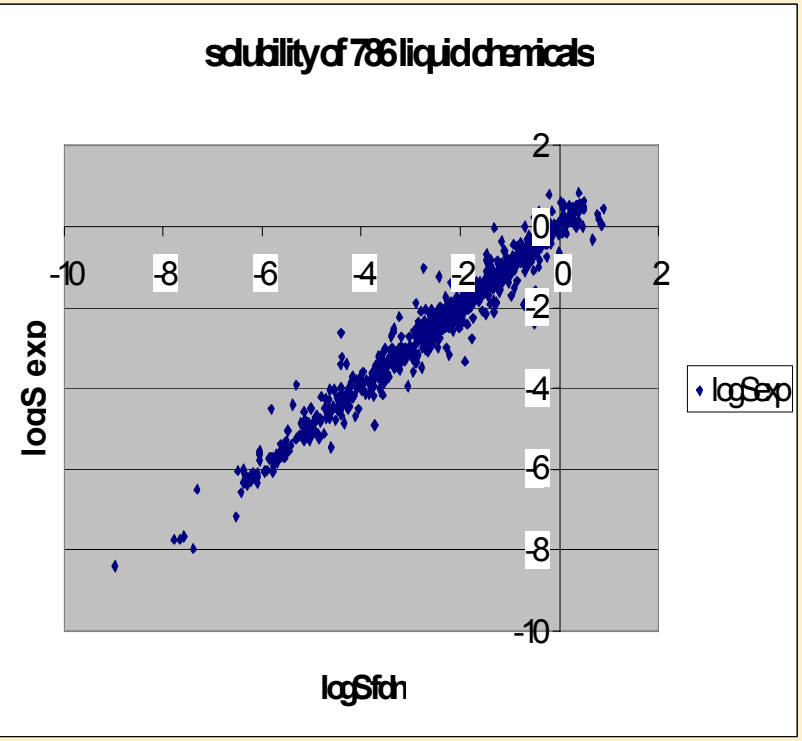
$$\log L^w = -0.052\alpha + 1.73\Sigma Ca - 0.82\Sigma Cd + 7\ I$$
(n = 426, r^2 = 0.929, s = 0.58)

$$\log L^w = \text{Log } L^w_{nn} - 0.032(\alpha_i - \alpha_{nn}) + 1.63(\Sigma Ca_i - \Sigma Ca_{nn}) - 1.04(\Sigma Cd_i - \Sigma Cd_{nn})$$

$$\text{Log } L^w_{exp} = 1.00(\pm 0.01) * \log L^w_{fdn}$$ (n = 558, r^2 = 0.962, s = 0.40).

O.A.Raevsky, K.J.Schaper. *QSAR & Comb. Sci.*, **22**, p. 926-942, 2003

Water Solubility of Liquid Pollutants



$$\log S = 0.578 - 0.305\alpha + 1.155\Sigma Ca$$
($n = 630$, $r^2 = 0.876$, $s = 0.586$)

$$\log S = 0.434 - 0.298\alpha + 1.090\Sigma Ca + + 0.304|\Sigma Cd|$$
($n = 630$, $r^2 = 0.897$, $s = 0.536$)

$$\log S = 0.523 - 0.274\alpha + 0.882\Sigma Ca + + 0.318|SCd| + 6\ I$$
($n = 787$, $r^2 = 0.935$, $s = 0.47$)

$$\log S_i = \text{Log } S_{nn} - 0.275(\alpha_i - \alpha_{nn}) + 0.90(\Sigma Ca_i - \Sigma Ca_{nn}) + 0.33(|\Sigma Cd|_i - |\Sigma Cd|_{nn})$$
$$\text{Log } S_{exp} = 0.98(\pm 0.01)^* \log S_{fdn}$$ ($n = 786$, $r^2 = 0.970$, $s = 0.32$).

K.J.Schaper, B.Kunz, O.A.Raevsky. *QSAR & Comb. Sci.*, **22**, p. 943-958, 2003

Chemicals Partitioning and Solubility

$$\log P_{ow} = 0.267\alpha - 1.00\ \Sigma C_a$$
(N = 2850, R = 0.970, S = 0.23)

$$\log K_{air/w} = 0.032\alpha - 1.63\Sigma C_a + 1.04\Sigma C_d$$
(N = 322, R = 0.954, S = 0.65)

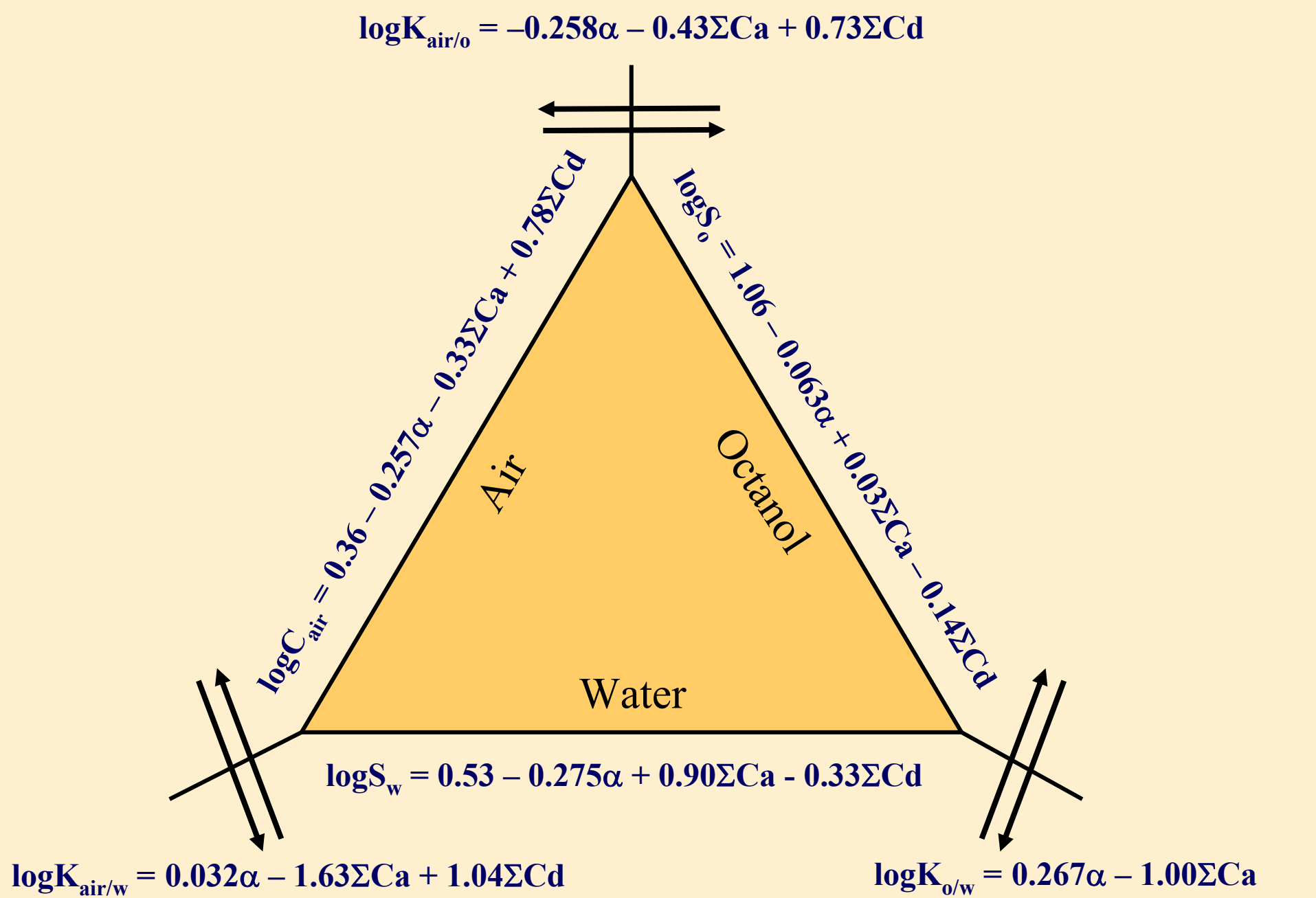
$$\log K_{air/o} = -0.258\alpha - 0.43\Sigma C_a + 0.73\Sigma C_d$$
(N = 98, R = 0.928, S = 0.61)

$$\log S_o = 1.06 - 0.063\alpha + 0.03\Sigma C_a - 0.14\Sigma C_d$$
(n = 23, R = 0.957, S = 0.16)

$$\log S_w = 0.434 - 0.298\alpha + 1.090\Sigma C_a - 0.304\Sigma C_d$$
(n = 630, r2 = 0.897, s = 0.536)

$$\log C_{air} = 0.36 - 0.257\alpha - 0.33\Sigma C_a + 0.78\Sigma C_d$$
(n = 90, R = 0.951, S = 0.40, F = 272.4

O.A.Raevsky, K.-J.Schaper, O.E.Raevskaya. *QSAR & Comb. Sci., SAR & QSAR in Environ. Res*., 2002-2007



O.A.Raevsky, K.-J.Schaper. *QSAR & Comb. Sci.*, 2007 (in press)

Parameters and Statistics of Equation $\log P = a_0 + a_1\alpha + a_2\sum Q + a_3\sum Ca + a_4\sum Cd$

$P = a_0 + a_1\alpha + a_2\sum Q + a_3\sum Ca + a_4\sum Cd$											
System	a_0	a_1	a_2	a_3	a_4	**n**	R^2	R_g^2	**R**	**SD**	**F**
Hexadecane-water	–0.05	0.298	–0.69	–1.56	0.58	372	0.923	0.922	0.961	0.60	1098.9
Octanol-water	0.30	0.244	–0.72	–1.13	0.04	560	0.959	0.959	0.979	0.32	1279.8
Bu_2O-water	0.10	0.310	–0.72	–1.49	0.18)	101	0.932	0.929	0.966	0.43	330.2
PGDP-water	0.63	0.243	–0.36	–1.27	0.14	52	0.898	0.889	0.947	0.32	102.9
air-water	0.35	–0.034	–2.24	–1.91	0.83	432	0.881	0.880	0.938	0.77	784.6
Hexadecane-air	–0.40	0.325	1.43	0.37	–0.20	390	0.935	0.934	0.967	0.38	1373.7
Octanol-air	–0.23	0.298	1.27	0.64	–0.83	407	0.876	0.875	0.936	0.73	711.6
Bu_2O-air	0.37	0.298	0.91	0.20	-0.83	91	0.892	0.887	0.944	0.75	177.1
PGDP-air	–0.16	0.317	0.09	0.17	–0.96	32	0.799	0.769	0.894	0.63	26.8

O.A.Raevsky. O.E.Raevskaya. *SAR & QSAR* in Environ. Res., (submitted)

PERMEABILITY

PARTITION OF PHENOLS IN LECITHIN LIPOSOME/WATER SYSTEM

O.A.Raevsky (unpublished data)

$$\log Per = 0.86(\pm 1.13) + 0.126(\pm 0.071)\alpha$$

N = 26, R = 0.598, SD = 0.52, F = 13.3

$$\log Per = 1.32(\pm 0.51) + 0.158(\pm 0.032)\alpha - 0.68(\pm 0.14)\Sigma C_a$$

N = 26, R = 0.939, SD = 0.23, F = 85.4

$$\log Per = 0.78(\pm 0.79) + 0.171(\pm 0.034)\alpha - 0.69(\pm 0.13)\Sigma C_a - 0.15(\pm 0.18)\ \Sigma C_d$$

N = 26, R = 0.947, SD = 0.22, F = 63.3

$$\log Per = 2.04(\pm 0.59) + 0.160(\pm 0.026)\alpha - 0.74(\pm 0.12)\Sigma C_a - 0.07(\pm 0.04)pK_a$$

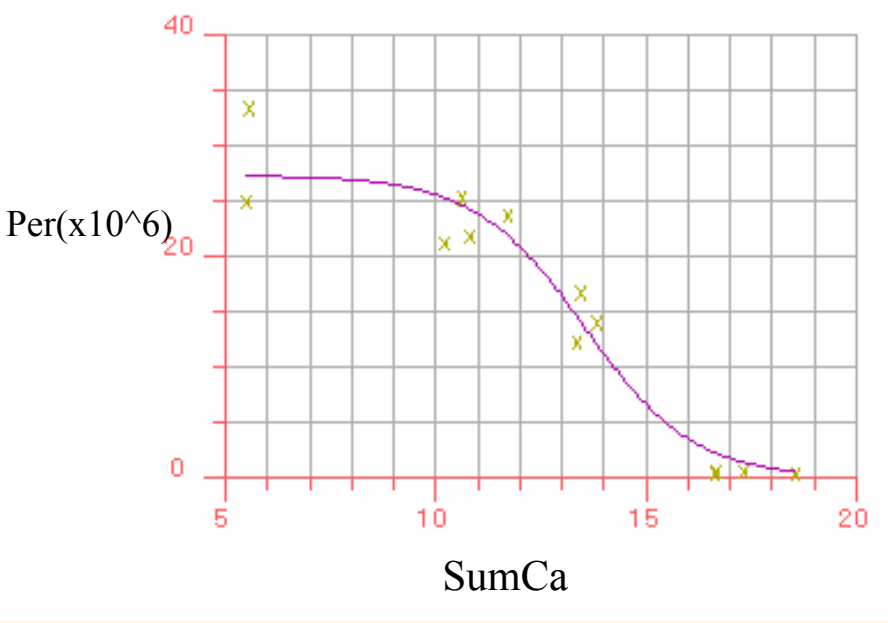
N = 26, R = 0.962, SD = 0.18, F = 90.3

$$\log Per = 1.50(\pm 0.72) + 0.173(\pm 0.026)\alpha - 0.75(\pm 0.11)\Sigma C_a - 0.15(\pm 0.14)\ \Sigma C_d - 0.07(\pm 0.04)pK_a$$

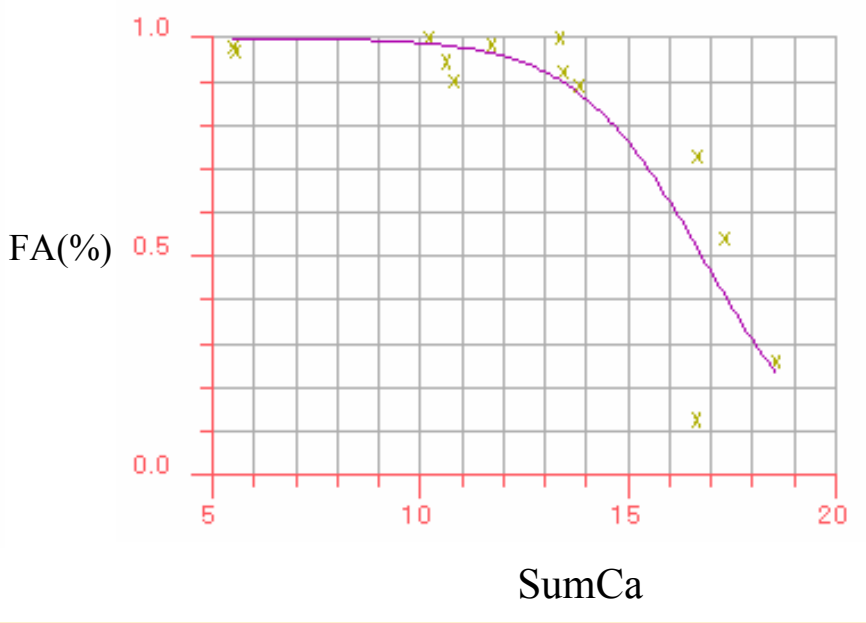
N = 26, R = 0.970, SD = 0.17, F = 82.6

Biological data are taken from H.Miyoshi, T.Nishioka, T.Fujita, *Bull. Chem. Soc. Jpn.*, v.59, 1099–1107 (1986)

Caco-2 Permeability for 13 Compounds



Oral Absorption in Human for 13 Compounds



$Per\ (10^6) = 27.37/(1+10^{-[4.52-0.33\ SumC]})$
N=13, R=0.97, S=3.07, F=77.9

$Per\ (10^6) = 27.34/(1+10^{-[4.47-0.34\ SumCa+0.32\ SumCd]})$
N=13, R=0.97, F=48.0

$FA = 1/(1+10^{-[4.79-0.29\ SumC]})$
N=13, R=0.88, S=0.15, F=37.6

$FA = 1/(1+10^{-[6.25-0.44\ SumCa+0.32\ SumCd]})$
N=13, R=0.98, S=0.06, F=1269

O.A.Raevsky (unpublished data, 1992)

Absorption in Human for 31 Passively Transported Drugs

$$FA = 1/(1 + 10^{-[5{,}02 - 0{,}31 \Sigma Cad]})$$

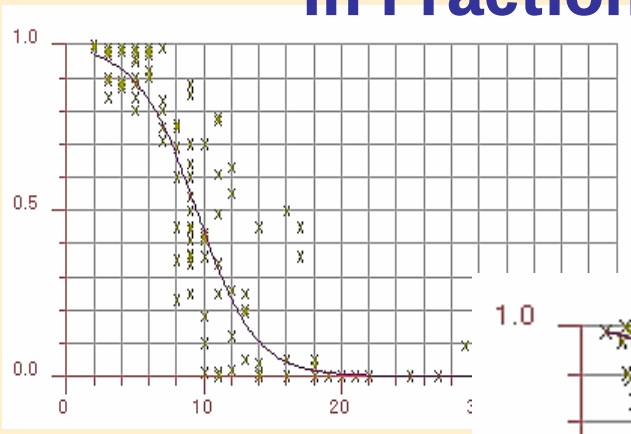
$n = 31$, $r^2 = 0.89$, $s = 0.12$

$$FA = 1/(1 + 10^{-[5{,}05 - 0{,}36 \Sigma Ca + 0{,}26 \Sigma Cd]})$$

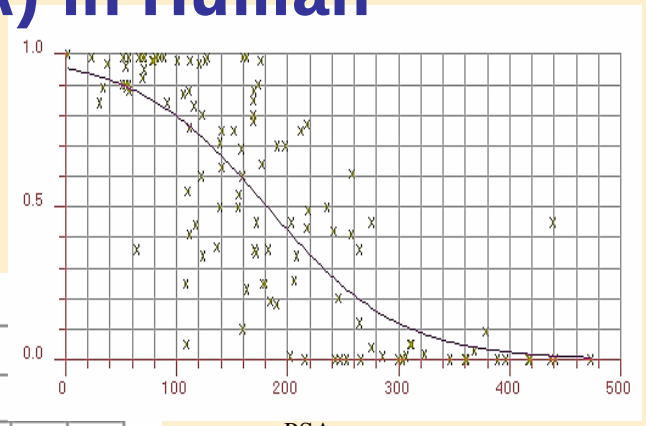
$n = 31$, $r^2 = 0.95$, $s = 0.09$

Raevsky, O.A., Fetisov, V.I., Trepalina, E.P., McFarland, J.W., Schaper, K.-J. Quantitative Estimation of Drug Absorption in Human for Passively transpoetred Compounds on the Basis of Their Physico-chemical Parameters. *Quant. Struct.-Act.Relat.*, **2000**, *19*, 366-374.

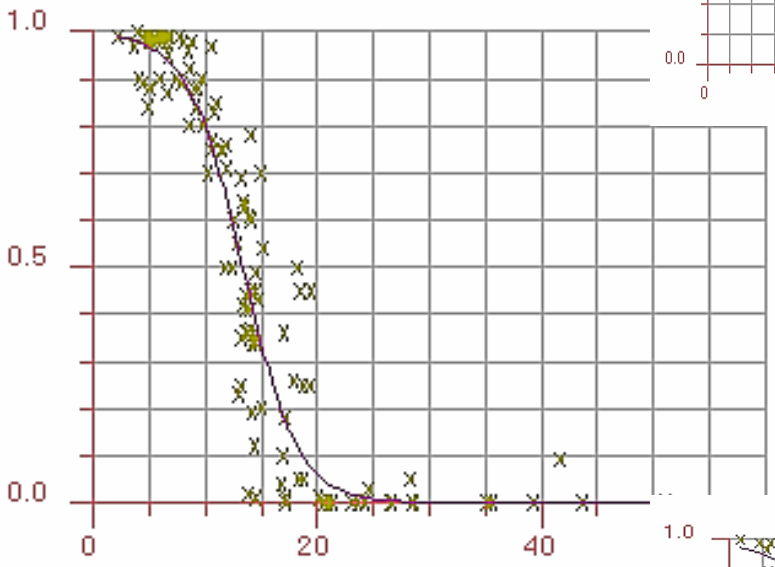
H-Bond Descriptors Contributions in Fraction Absorption (FA) in Human



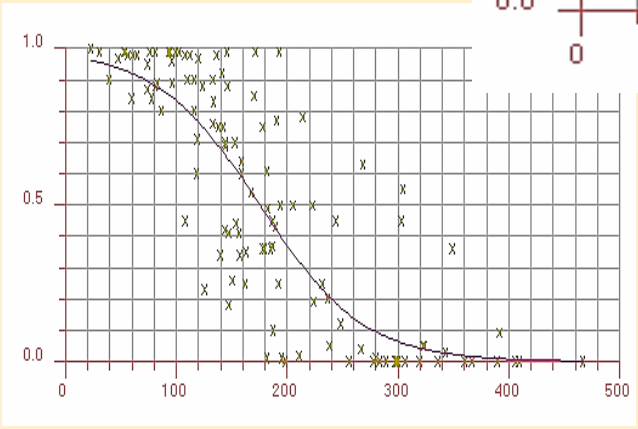
Nad



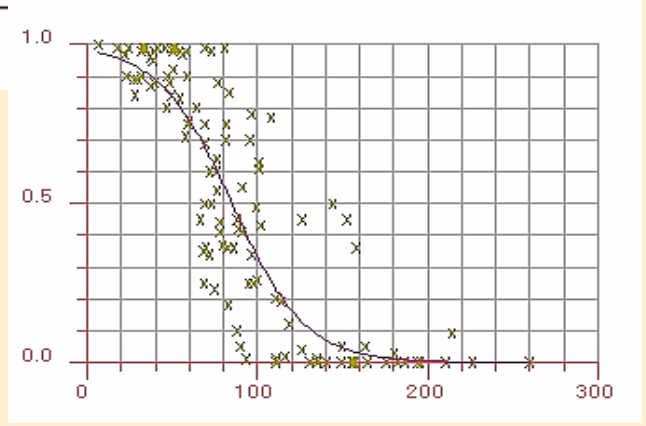
PSA



ΣCad



VNOH



SAREANOH

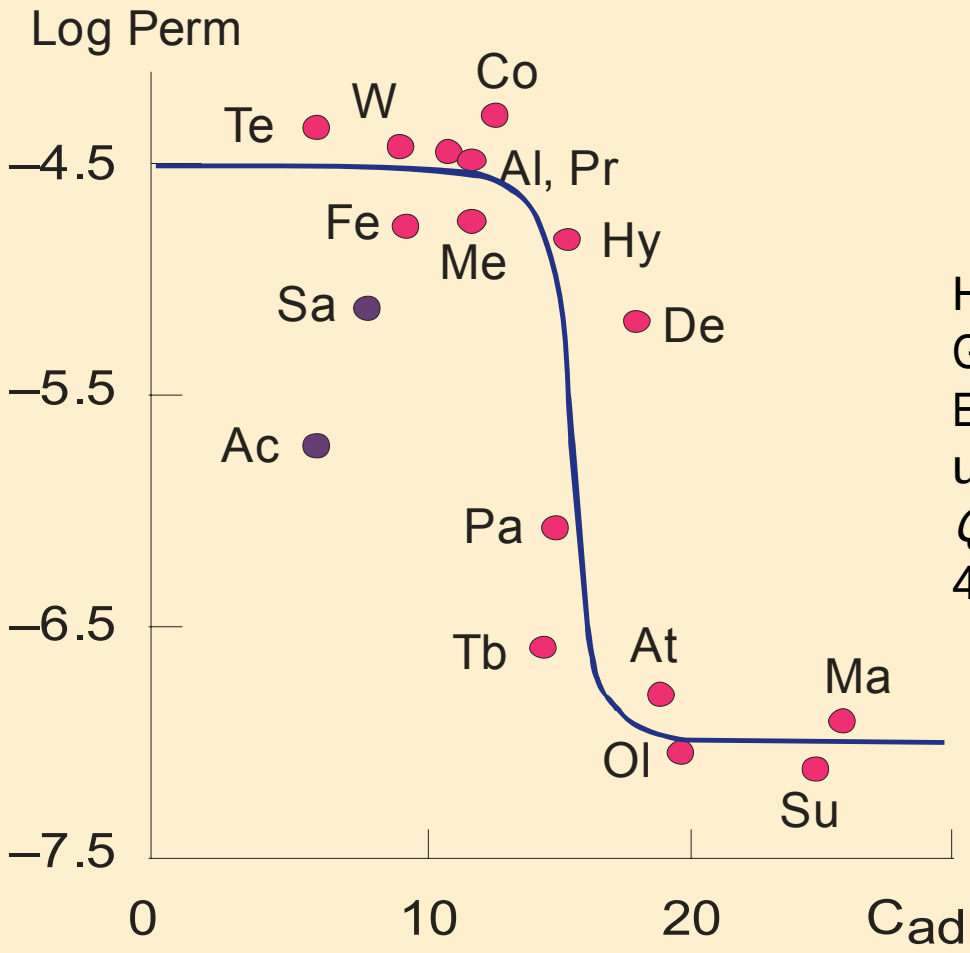
Absorption in Human for 154 Passively Transported Drugs

$$FA = 1/(1 + 10^{-[2,2 - 0,016\ (\mathbf{OFEASA+OFEDSA})]})$$

n = 154, r2 = 0.83, s = 0.15

Raevsky, O.A., Skvortsov, V.S. Quantifying hydrogen bonding in QSAR and molecular modeling.*SAR & QSAR in Environmental Research,* **2005**, *3*, 287-300.

Correlation of Permeability and Total Hydrogen Bonding Capacity (C_{ad})



H. van de Waterbeemd, G. Camenisch, G. Folkers, **O. Raevsky.** Estimation of Caco-2 Cell Permeability using Calculated Molecular Descriptors. *Quant. Struct.-Act. Relat.*, **15**, 480–490 (1996)

BBB+/BBB-Classification

The influence of physicochemical properties, including lipophilicity, H-bonding capacity and molecular size and shape descriptors on brain uptake has been investigated using a selection of 45 known CNS-active and 80 CNS-inactive drugs. A combination of a H-bonding and a molecular size descriptor, i.e., the major components of lipophilicity and permeability, avoiding knowledge of distribution coefficients, is proposed to estimate the blood-brain barrier penetration potential of new drug candidates

Van de Waterbeemd, H., Camenisch, G., Folkers, G., Chretien, J.R., Raevsky, O.A. Estimation of Blood-Brain Barrier Crossing of Drugs Using Molecular Size and Shape, and H-Bonding Descriptors. *J. Drug Targeting*, **1998**, *6*, 151-165.

Chemicals BBB^+/BBB^- Classification on the Basis of LDA

		Descriptor	Classification (right/wrong)		Recognition(%)		
			BBB-	BBB+	BBB-	BBB+	**Total**
		MW	206/104	254/56	66.5	81.9	74.2
		Sumq-	184/126	270/40	59.4	87.1	73.2
		SumCad(o)	223/87	302/8	71.9	97.4	84.7
		PSA	218/92	284/26	70.3	91.6	81.0
		WFEASA	241/69	293/17	77.7	94.5	86.1
		(WFEASA(o)+WFEDSA)	**238/72**	**302/8**	**76.8**	**97.4**	**87.1**
		(NHD+NHA)	224/86	297/13	72.3	95.8	84.0
		$logP_{ow}$	207/103	231/79	66.8	74.5	70.6
		logD	207/103	271/39	66.8	87.4	77.1

O.Raevsky, V.Grigoriev, V.Skvortsov, R.Mannhold. Unpublished recent results

Chemicals BBB^+/BBB^- Structure Classification on the Basis of Binary Approach

	Chemicals	1 nearest neighbour (max cosine)	
	Number	BBB-	BBB+
	310	279	31
		90.0%	10.0%
	310	24	286
		7.7%	92.3%
Total	620	Right recognition is equal to 91.1%	

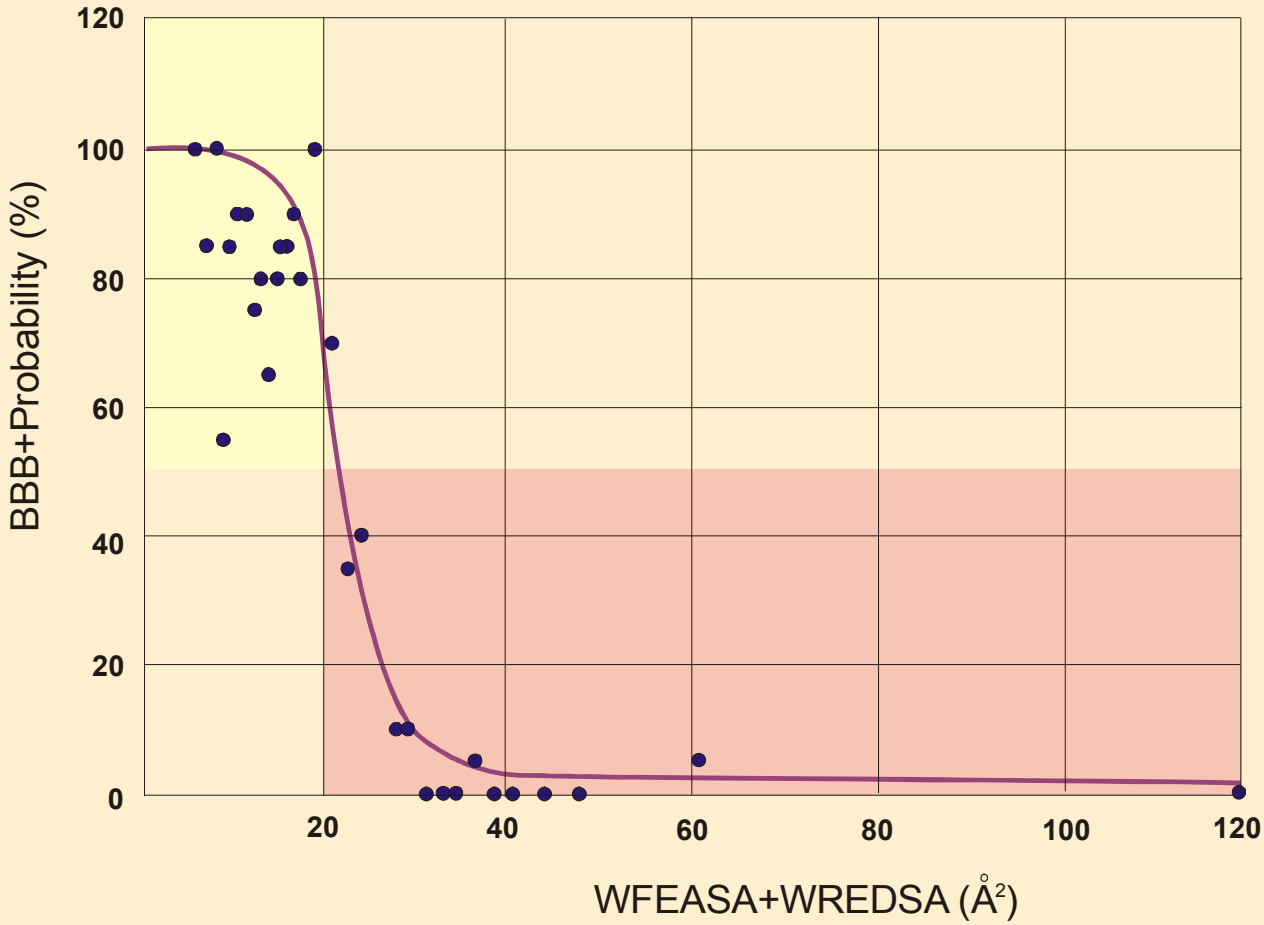
O.Raevsky, V.Grigoriev, V.Skvortsov, R.Mannhold. Unpublished recent results

Chemicals BBB+/BBB- Classification on the Basis of Structure and Physicochemical Similarity

MOA	Chemicals Number	WFEASA+WFEDSA neighbour	
		BBB–	BBB+
BBB$^-$	310	285	25
		91.1%	8.9%
BBB$^+$	310	17	293
		94.5%	93.2%
Total	620	Right recognition 93.2%	

O.Raevsky, V.Grigoriev, V.Skvortsov, R.Mannhold. Unpublished recent results

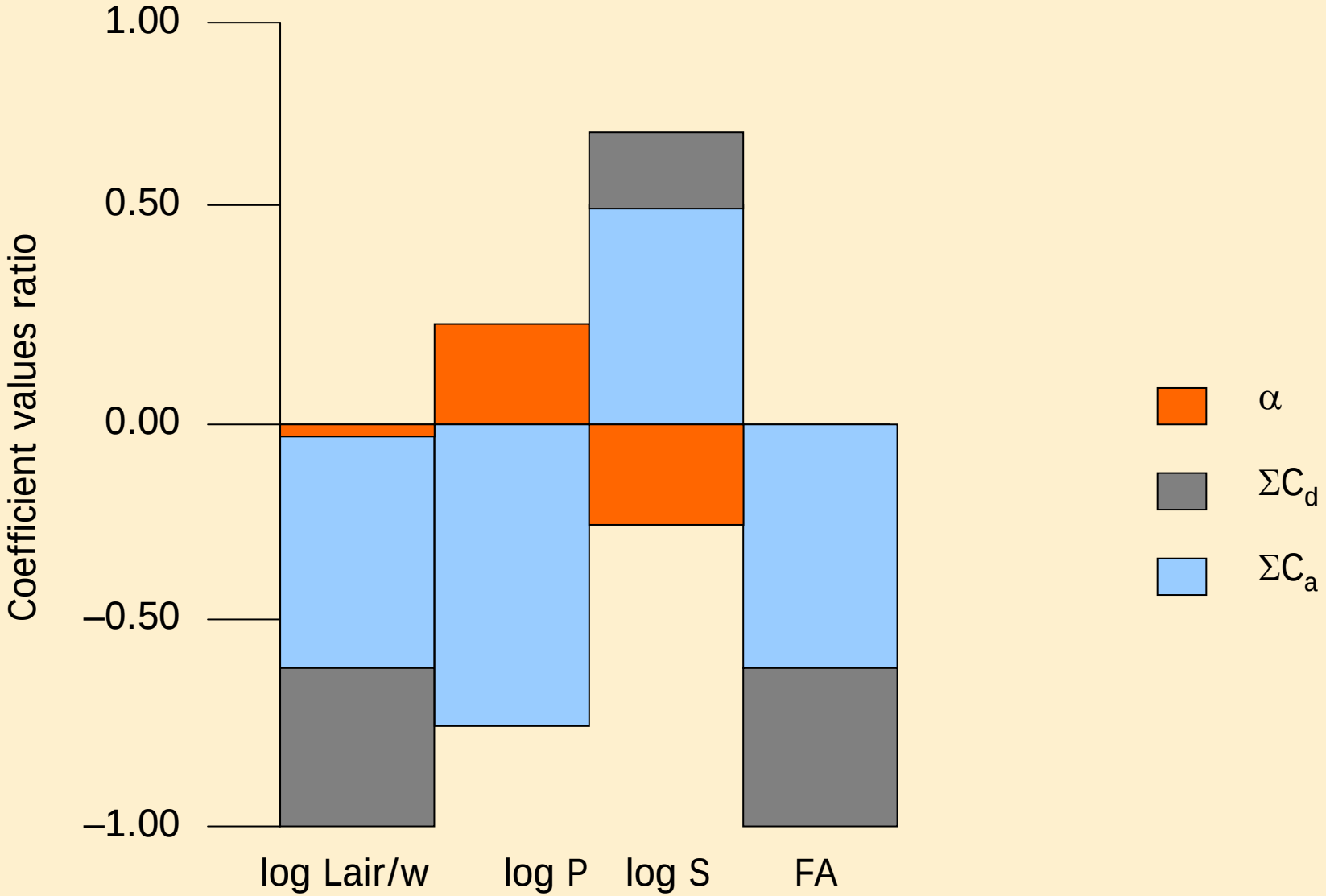
BBB+Probability as a Function of WFEASA+WFEDSA H-bond Descriptor



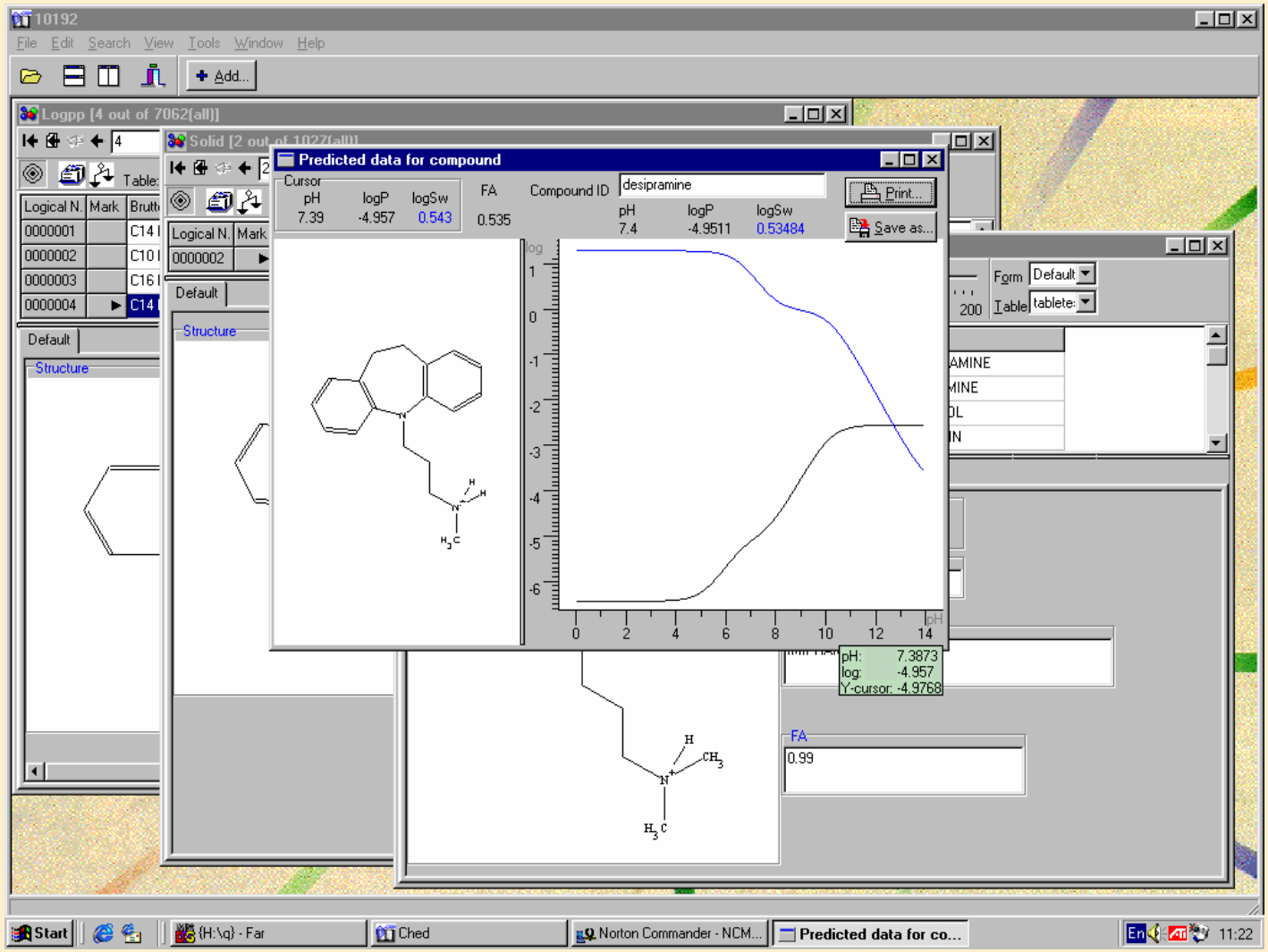
O.Raevsky, V.Grigoriev, V.Skvortsov, R.Mannhold. Unpublished recent results

WFEASA-WFEDSA	Chemical number	BBB^+	BBB^-
0.00–18.10	310	265 (85.5%)	45 (14.5%)
>18.10	310	45 (14.5%)	265 (85.5%)
Total	620	310	310

Ratio of Molecular Size and H-bond Contributions

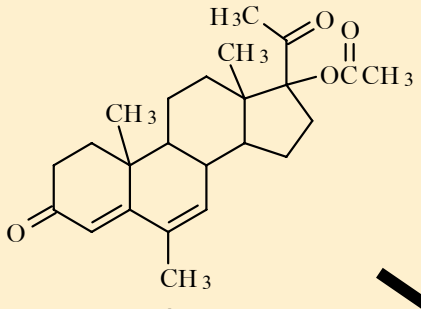


SLIPPER-2000. Results of Properties Prediction (Monitor View)



O.A.Raevsky, S.V.Trepalin, E.P.Trepalina, V.A.Gerasimenko. Chem. Inf. And comput. Sci., 2002, v. 42 , pp. 540-549.

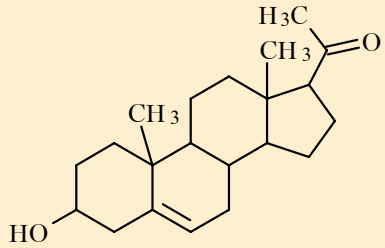
Lead Development Summary



Megestrol acetate

α = 42.07 log P = 3.90
ΣC$_a$ = 5.29 log S = –5.35
ΣC$_d$ = 0.00 S = 1.72 μg/mL

Δ log P = 0.08
Δ log S = 0.58
log P$_c$ = 3.98
log S$_c$ = –4.77
S$_c$ = 5.37 μg/mL



Pregnenolone

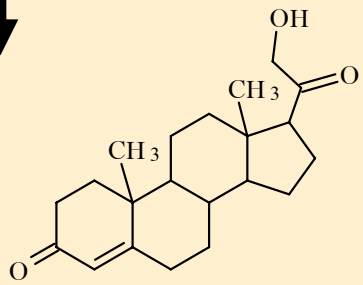
α = 36.81 log P = 4.22
ΣC$_a$ = 3.81 log S = –4.65
ΣC$_d$ = –1.37 S = 7.09 μg/mL

Δ log P = –1.10
Δ log S = 1.79
log P$_c$ = 2.80
log S$_c$ = –3.56
S$_c$ = 91.02 μg/mL

Δ log P = –1.18
Δ log S = 1.22
log P$_c$ = 3.04
log S$_c$ = –3.43
S$_c$ = 122.18 μg/mL

Δ log P = –1.31
Δ log S = 2.03
log P$_c$ = 2.59
log S$_c$ = –3.33
S$_c$ = 128.35 μg/mL

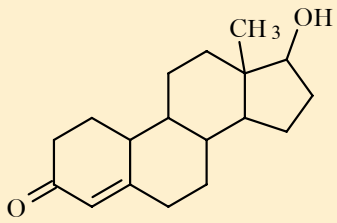
Δ log P = –1.39
Δ log S = 1.45
log P$_c$ = 2.83
log S$_c$ = –3.20
S$_c$ = 173.13 μg/mL



Desoxycartone

α = 36.97 log P = 2.88
ΣC$_a$ = 5.03 log S = –3.45
ΣC$_d$ = –1.92 S = 117.26 μg/mL

Δ log P = –0.21
Δ log S = 0.24
log P$_c$ = 2.67
log S$_c$ = –3.21
S$_c$ = 166.73 μg/mL



Nandrolone

α = 31.31 log P = 2.62
ΣC$_a$ = 3.73 log S = –3.02
ΣC$_d$ = –1.43 S = 262.05 μg/mL

О.А.Раевский. *Росс. Хим. Журн.*, **50** (2). С. 97-107, 2006

Biopharmaceutics Classification (BPC)

BPC	Solubility	FA	$\alpha/\Sigma C_a$	ΣC_{ad}
class 1	≥50 μg/mL	≥0.50	≤ 2.0	<15.0
class 2	<50 μg/mL	≥0.50	>2.0	<15.0
class 3	≥50 μg/mL	<0.50	≤ 2.0	≥15.0
class 4	<50 μg/mL	<0.50	>2.0	≥15.0
«alert»	<5 μg/mL	<0.05	>5.0	>25.0

Strong H-bond complexes in crystal lattice can move drug from class 1 to class 2 or from class 3 to class 4.
Active transport can move drug from class 4 to class 2 or from class 3 to class 1.

O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

Biopharmaceutics Classification for 254 Drugs

class		1	2	3	4
number		156	74	17	7
%		61.4	29.1	6.7	2.8
right	S(N/%)	115 (73.7)	61 (83.4)	16 (94.1)	5(74.1)
recog-nition	FA(N/%)	135 (86.5)	69(93.2)	13 (76.5)	3 (42.9)

Total right recognition ~75%

O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) Copyright © 2007 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 978-3-527-31755-4, pp. 127-154

Aquatic Toxicity (Subject: *Poecilia reticulata* (guppy))

Inert chemicals; non polar narcosis

Classical QSAR	*Similarity + QSAR*
$\log LC_{50} = 5.14 - 0.258\ \alpha + 0.77\ \Sigma Ca - 0.02\ \Sigma Cd$	$\log LC_{50}(exp) = 1.02 \log LC_{50}(3nn)$
$n = 90 \quad R = 0.976 \quad s = 0.32$	$n = 90 \quad R = 0.987 \quad s = 0.23$

Less inert chemicals; polar narcosis

$\log LC_{50} = 4.14 - 0.200\ \alpha + 0.69\ \Sigma Ca + 0.11\ \Sigma Cd$	$\log LC_{50}(exp) = 1.01 \log LC_{50}(3nn)$
$n = 121 \quad R = 0.931 \quad s = 0.37$	$n = 121 \quad R = 0.974 \quad s = 0.23$

Reactive chemicals

$\log LC_{50} = 2.44 - 0.120\ \alpha + 0.29\ \Sigma Ca - 0.11\ \Sigma Cd$	$\log LC_{50}(exp) = 0.94 \log LC_{50}(3nn)$
$n = 51 \quad R = 0.663 \quad s = 0.70$	$n = 51 \quad R = 0.934 \quad s = 0.32$

Specifically acting compounds

$\log LC_{50} = 2.57 - 0.120\ \alpha + 0.11\ \Sigma Ca - 0.05\ \Sigma Cd$	$\log LC_{50}(exp) = 0.84 \log LC_{50}(3nn)$
$n = 31 \quad R = 0.772 \quad s = 0.90$	$n = 31 \quad R = 0.861 \quad s = 0.68$

GENERAL MODEL

$$log\ LC_{50} = log\ LC_{50}nn - 0.(207\ \alpha_I - \alpha_{nn}) + 0.44\ (\Sigma Ca_i - \Sigma Ca_{nn}) - 0.06\ (\Sigma Cd_i - \Sigma C)d_{nn}$$

$\log LC_{50} = 4.25 - 0.207\ \alpha + 0.44\ \Sigma Ca - 0.06\ \Sigma Cd$	$LC_{50}(exp) = 1.00 \log LC_{50}(3nn)$
$n = 293 \quad R = 0.800 \quad s = 0.86$	$n = 293 \quad R = 0.973 \quad s = 0.33$

O.A.Raevsky and J.C.Dearden, Creation of Predictive Models of Aquatic Toxicity of Environmental Pollutants with Different Mechanisms of Action on the Basis of Molecular Similarity and HYBOT Descriptors, *SAR & QSAR in Environmental Research,* 2004, **15**, 433-448.

Relationships of Nonpolar Narcosis with Molecular Polarizability and H-Bond Factors

Guppy

$\log(1/LC50) = -2.09\ (\pm 0.11) + 0.26\ (\pm 0.01)\ \alpha - 0.81\ (\pm 0.03)\ \Sigma C_a$

N = 90, R = 0.978, S = 0.31, F = 938

Fathead minnow

$\log(1/LC50) = -1.92\ (\pm 0.13) + 0.26\ (\pm 0.01)\ \alpha - 0.93(\pm 0.05)\ \Sigma C_a$

N = 66, R = 0.965, S = 0.32, F = 425

Rainbow trout

$\log(1/LC50) = -1.75(\pm 0.27) + 0.26(\pm 0.02)\ \alpha - 0.93(\pm 0.07)\ \Sigma C_a$

N = 28; R = 0.971; S = 0.39; F = 203

Total (Guppy + *Fathead minnow* + *Rainbow trout*)

$\log(1/LC_{50}) = -1.94(0.09) + 0.26(0.01)\ \alpha - 0.87(0.02)\Sigma C_a$

N = 184; R = 0.970; S = 0.34; F = 1422

O.A.Raevsky, V.D.Tonkopii, J.C.Dearden, E.Weber. Aquatic Toxicity QSPR models. Abstracts of the 12th International Workshop on Quantitative Structure-Activity Relationship in Environmental Toxicology (Lyon, 8-12 May, 2006). P. 114

Recent progress in H-Bonding Parametrization in Quantitative Structure–Activity Relationships and Drug Design

- **Structure of liquid water,**
- **New X-ray data for specific H-bonding complexes,**
- **Quantitative estimation of contribution of H-bond acceptor and donor factors and volume-related terms in chemicals solvation processes, partitioning in water/solvent/air systems,**
- **A refinement in the PSA approach,**
- **Improvement of Grid potentials,**
- **Calculation schemes of optimum H-bonding potential values for any concrete H-bonding atoms in any complexes which consider the nature of interacting atoms and the influence of substituents.**

O.A.Raevsky. In ***Molecular Drug Properties. Measurement and Prediction***. R.Mannhold (Ed.) ISBN: 978-3-527-31755-4, pp. 127-154

Acknowledgements

This investigation was supported by the International Science & Technology Centre (Project # 888) and the Russian Ministry of Education & Sciences (contract N 02.434.11.1014).

Co-authors and co-workers:

V.Yu.Grigoriev, S.V.Trepalin, V.A.Gerasimenko, O.E.Raevskaya, E.V.Bovina, V.S.Skvortsov, E.Andreeva, A.N.Rasdolsky.