

The Block Relevance (BR) analysis to aid medicinal chemists to determine and interpret lipophilicity

Giulia Caron, Maura Vallaro and Giuseppe Ermondi*

CASMedChem research group, BMSS at the Centre for Innovation, Università degli Studi di Torino, via

Quarello 15, 10135 Torino, Italy.

E-mail: giuseppe.ermondi@unito.it

Telephone: +39 011 6708337

Supporting Information

List of Supporting Information Material

Tables

Table S1. Data used in the paper.

Table S2. Blocks definition.

Table S3. The final composition of the six blocks. Some changes were made relatively to the original block composition. In particular we moved W1, WO1 and WN1 (the volume of the molecular envelope which is accessible to and attractively interact respectively with OH2, O and N1 probes computed for molecular field respectively of -0.2, -1.0 and -1.0 kcal/mol) to the block of the corresponding probe. Moreover PSA and PSAR (the polar surface area and the ratio between PSA and the total surface) were moved from the *Others* to the *Water* block, whereas HSA and HSAR (the hydrophobic surface and the ratio between HAS and the total surface) were moved from the *Others* to the DRY block.

Figures

Figure S1. The variation of blocks values vs the number of LVs of the PLS model. The optimal number of LVs is generally associated to the highest value of Q^2 . However it is often verified that more than 3 LVs poorly improved the explained variance of Y and X and thus 3 is generally chosen by many authors as the final number of LVs, at least for modeling physicochemical and ADME properties. In other words, the choice of the optimal number of LVs to include in a PLS model is therefore somewhat subjective. To take this point into account, we define the uncertainty associated to blocks as the difference between block values calculated for 3LVs and the corresponding registered for the largest number of LVs. This strategy is probably rough but consistent with the meaning of VIPs.

Figure S2. Examples of HPLC results: Linear regressions are shown for 3-Bromoquinoline and lorazepam.

Figure S3. Correlation matrix of lipophilicity indexes.

Figure S4. BR plots without sign calculated for a series of log k' values determined using the same column with different mobile phase composition (see text for details).

Annexes

Annex S1. A deeper insight in blocks significance.

Table S1.

Compound	Elog k'w	ElogPoct	logPoct	log kw	log k'80	log k'70	log k'60	log k'50	log k'40	log k'30	log k'55	log k'45	log k'35
3,5-dichlorophenol	3.4	3.88	3.68	3.50	0.24	0.60	1.03	1.46					
3-bromoquinoline	2.54	2.93	3.03	2.70	-0.22	0.07	0.41	0.78	1.21	1.63			
3-chlorophenol	2.49	2.88	2.5	2.49	-0.24	0.05	0.40	0.74	1.13	1.46			
acetaminophen	-0.02	0.11	0.51	0.69				-1.02	-0.73		-0.13	-0.81	-0.60
acetophenone	1.31	1.58	1.58	1.56	-0.67	-0.48	-0.19	0.10	0.43	0.73			
allopurinol	-0.89	-0.85	-0.55	0.02					-0.94	-0.76	-1.37	-1.13	
antipyrine	0.09	0.23	0.38	1.05			-0.86	-0.55	-0.22	0.12	-0.69	-0.40	-0.09
bifonazole	4.37	4.95	4.77	4.80	-0.04	0.50	1.11	1.79					
bromazepam	1.13	1.38	1.65	2.18	-0.83	-0.47	-0.15	0.20	0.65	1.10			
caffeine	-0.31	-0.21	-0.07	0.73			-0.98	-0.75	-0.39		-0.85	-0.61	-0.27
carbamazepine	1.7	2.01	2.19	2.43	-0.84	-0.47	-0.12	0.29	0.75	1.25			
chloramphenicol	1.18	1.43	1.14	1.90	-0.95	-0.54	-0.20	0.12	0.50	0.84			
clotrimazole	4.13	4.69	5.2	4.39	-0.09	0.38	0.92	1.62					
dexamethasone	2.03	2.37	1.83	3.15	-0.72	-0.28	0.14	0.62	1.16	1.76			
diazepam	2.63	3.03	2.79	3.05	-0.43	-0.06	0.35	0.80	1.35				
diethylstilbestrol	4.16	4.72	5.07	4.51	-0.01	0.50	1.04	1.71					
estradiol	3.34	3.82	4.01	3.86	0.12	0.54	1.01	1.53					
fluconazole	0.37	0.54	0.5	1.50		-1.10	-0.76	-0.41	0.00	0.41	-0.60	-0.23	0.17
griseofulvin	2.21	2.57	2.18	3.17	-0.88	-0.43	0.00	0.49	1.07	1.73			
hydrocortisone	1.43	1.71	1.55	2.63	-0.72	-0.35	-0.01	0.42	0.90	1.44			
hydrocortisone-21-acetate	1.93	2.26	2.19	3.06	-0.58	-0.19	0.24	0.71	1.28				
lorazepam	2.36	2.74	2.51	3.03	-0.60	-0.21	0.20	0.64	1.18	1.71			
lormetazepam	2.4	2.78	2.72	3.09	-0.55		0.22	0.67	1.23	1.78			
methylthioinosine	0.19	0.34	0.09	1.25			-0.81	-0.50	-0.13	0.24	-0.65	-0.36	0.01
metronidazole	-0.32	-0.22	-0.02	0.34	-1.18	-1.01	-0.90	-0.68	-0.43	-0.21			
naphthalene	3.06	3.51	3.37	3.21	0.052	0.401	0.795	1.202	1.647				
nifedipine	2.46	2.85	3.17	3.06	-0.56	-0.19	0.24	0.71	1.30				
nifuroxime	1.11	1.36	1.28	1.48	-0.61	-0.36	-0.08	0.16	0.45	0.69			

nitrofurazone	0.1	0.24	0.23	0.97	-1.13	-0.81	-0.68	-0.39	-0.08	0.22			
pentoxifylline	0.16	0.31	0.29	1.71			-0.82	-0.44	-0.01	0.48	-0.61	-0.24	0.20
prednisolone	1.54	1.83	1.6	2.65	-0.74	-0.40	-0.02	0.39	0.89	1.44			
prednisone	1.21	1.47	1.46	2.47	-0.93	-0.55	-0.21	0.20	0.71	1.26			
quinoline	1.56	1.85	2.03	1.80	-0.68	-0.46	-0.19	0.15	0.53	0.90			
testosterone	2.74	3.15	3.29	3.44	-0.25	0.13	0.57	1.05	1.64				
thiamphenicol	-0.24	-0.13	-0.27	1.07			-1.12	-0.69	-0.35	0.00	-0.86	-0.54	-0.20
tolnaftate	4.55	5.15	5.4	4.79	0.17	0.71	1.33						

Table S2.

Block	Definition	Color code
Size	descriptors that characterize the size and shape of the solute	green
OH2 (Water)	descriptors that express the solute's interaction with water molecules (= with the GRID OH2 probe)	light blue
N1*	descriptors that describe the solute's ability to form hydrogen bond interactions with the GRID N1 probe (that mimics the system)	blue
O*	descriptors expressing the solute's ability to form hydrogen bond interactions with the GRID O probe (that mimics the system)	red
DRY	descriptors describing the solute's propensity of the solute to participate in hydrophobic (= with the GRID probe DRY) interactions	yellow
Others	descriptors mainly describing the imbalance between hydrophilic and hydrophobic regions	grey

* For the sake of clarity, to identify hydrogen bonding (HB) interactions (Hydrogen Bond Acceptor capability (HBA and Hydrogen Bond Donor capability (HBD)) we refer to the probe's properties and not to the solute (see following scheme).

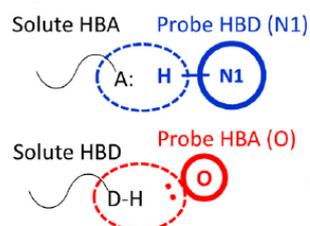


Table S3.

Block	Descriptors	Sum Descr
Size	V, S, R, G	4
OH2 (Water)	W1-8, IW1-4, CW1-8; PSA, PSAR	22
DRY	D1-8, DD1-8, ID1- 4, CD1-8, HSA, PHSAR	30
O	WO1-6	6
N1	WN1-6	6
Others	HL1-2, A, CP, DRDRDR, DRDRAC, DRDRDO, DRACAC, DRACDO, DRDODO, ACACAC, ACACDO, ACDODO, DODODO	14

Figure S1.

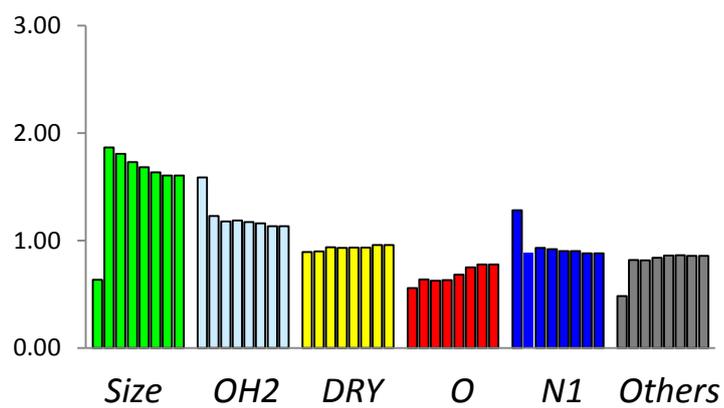


Figure S2.

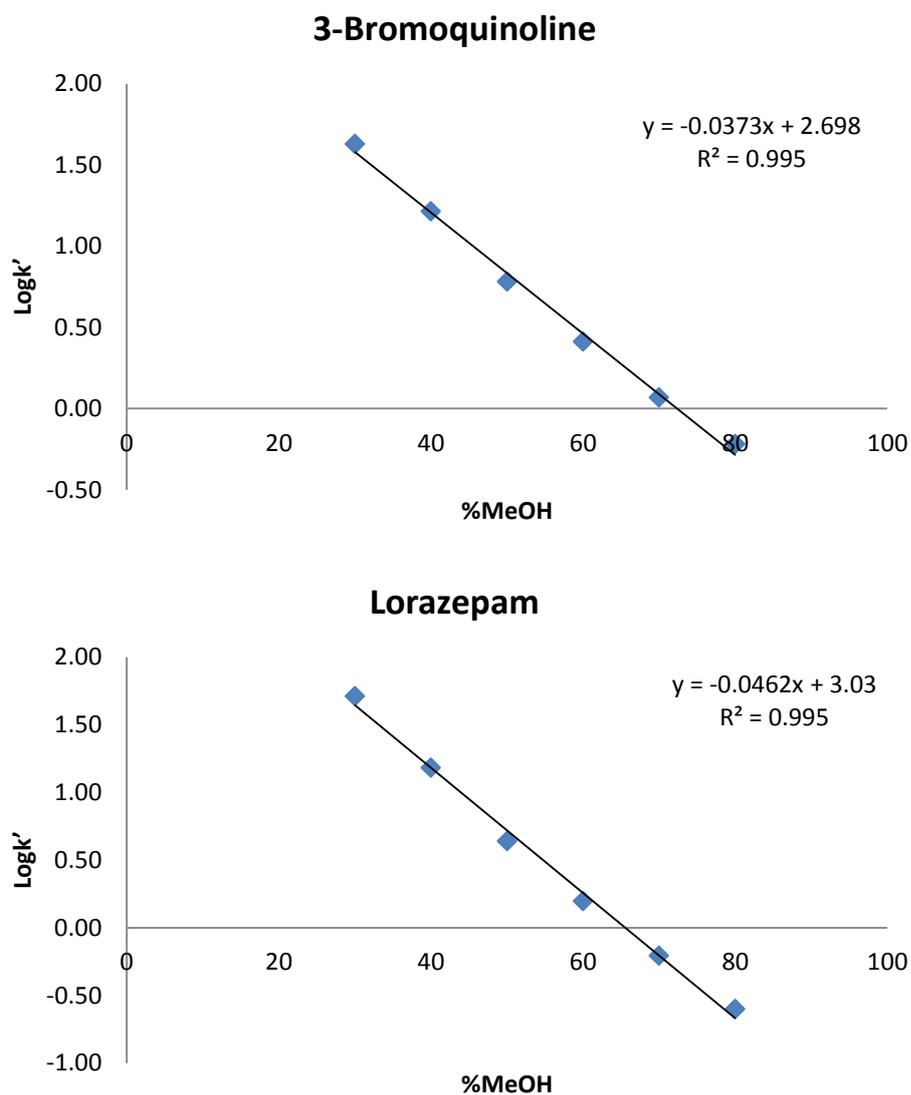
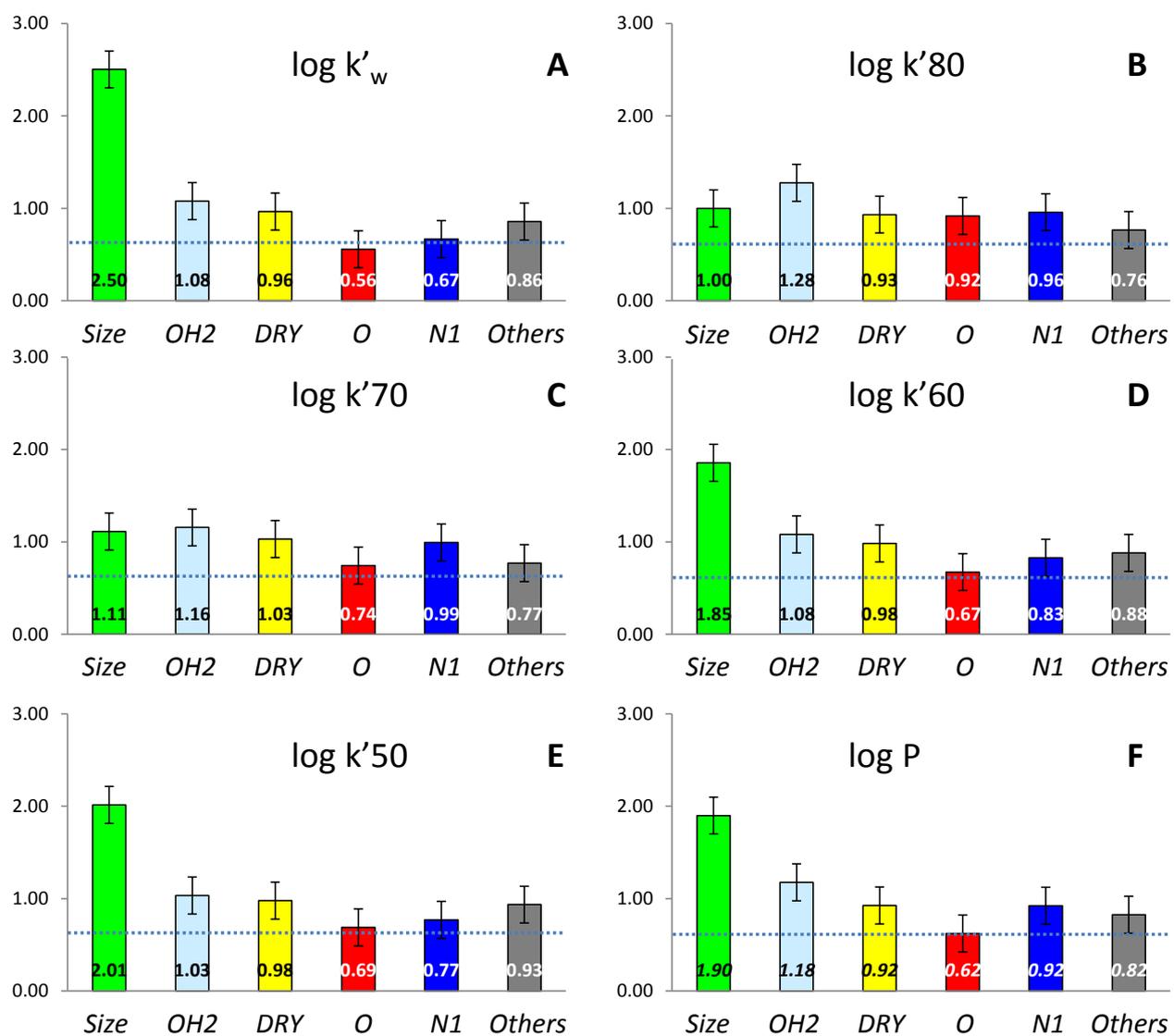


Figure S3. Relationship between lipophilicity descriptors.

	ElogP	logP_{oct}	log k'_w	log k'80	log k'70	log k'60	log k'50	log k'40	log k'30
ElogP	1	0.963	0.872	0.797	0.914	0.957	0.970	0.957	0.913
logP_{oct}	0.963	1	0.795	0.817	0.891	0.916	0.917	0.896	0.848
log k'_w	0.872	0.795	1	0.490	0.718	0.826	0.909	0.961	0.993
log k'80	0.797	0.817	0.490	1	0.947	0.883	0.803	0.709	0.589
log k'70	0.914	0.891	0.718	0.947	1	0.981	0.941	0.881	0.791
log k'60	0.957	0.916	0.826	0.883	0.981	1	0.986	0.949	0.882
log k'50	0.970	0.917	0.909	0.803	0.941	0.986	1	0.988	0.949
log k'40	0.957	0.896	0.961	0.709	0.881	0.949	0.988	1	0.985
log k'30	0.913	0.848	0.993	0.589	0.791	0.882	0.949	0.985	1

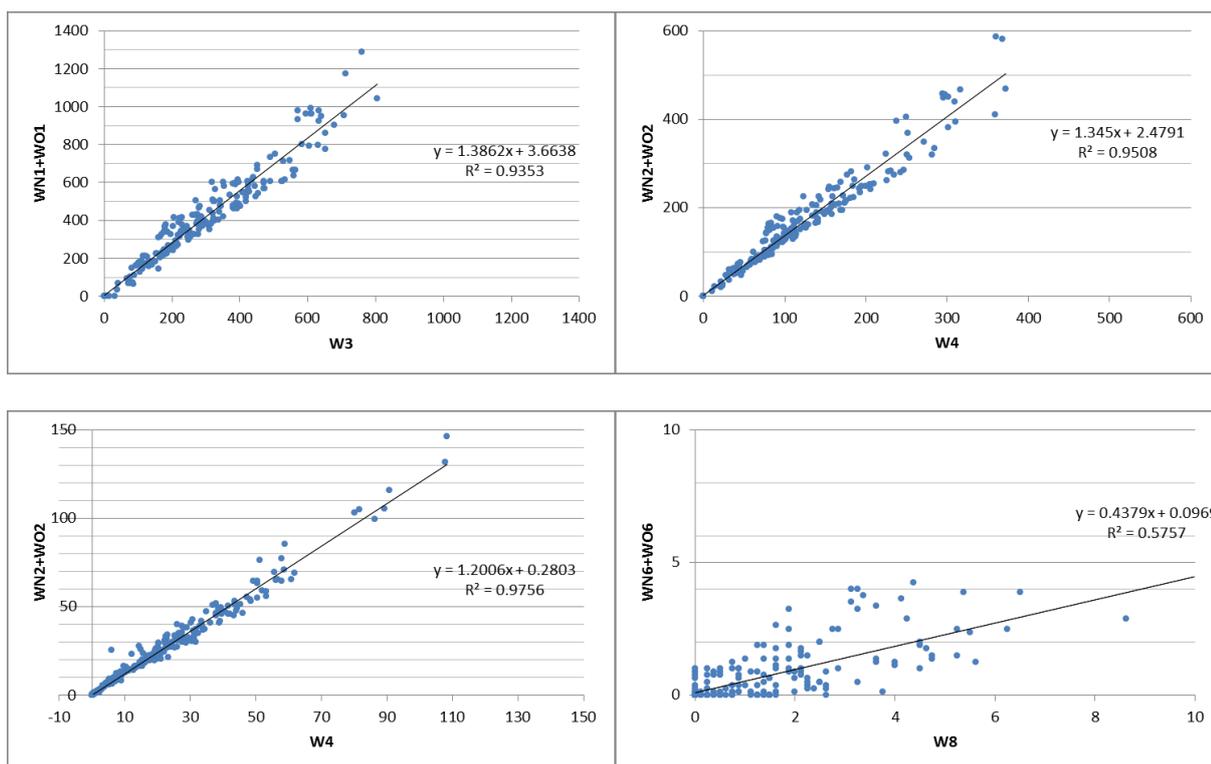
Figure S4.



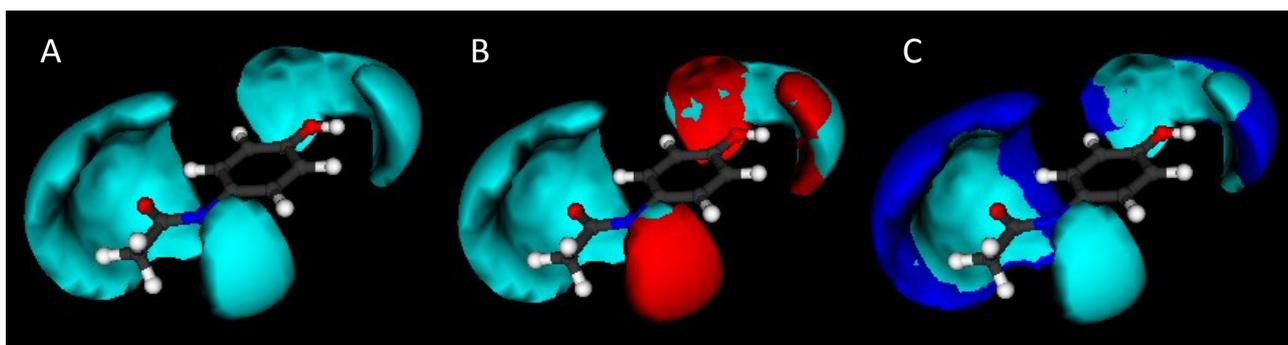
Annex S1

Apolar (*Size* and *DRY*) and polar (*Water*, *N1* and *O*) blocks can be roughly split in two parts: the first refers to the bulk properties of the molecule (*Size* and *Water*) whereas the second is related to the presence of functional moieties localized in some regions of the solute (*DRY*, *N1* and *O*). For example the interaction with the OH2 probe takes into account of a general tendency of the molecule to interact with water beyond the hydrogen bond interaction. In fact it is often verified that the sum of the volumes of the molecular envelope which is accessible to, and interacts attractively with the N1 and O probes does not give the analogous volume due to the interaction with OH2 probe. Numerical results and a graphical example are shown in the figures below

Plots of BR descriptors to explain differences between the Water and N1/O blocks.



Paracetamol is taken as an example to show differences among selected VS+ descriptors due to different polar probes. Isosurfaces at -1.0 kcal/mol for OH2 probe (cyan), N1 probe (blue) and O probe (red). A) isosurface for OH2 probe is shown; B) isosurface for OH2 probe is shown with the correspondent O isosurface; C) isosurface for OH2 probe is shown with the correspondent N1 isosurface. B) and C) show how superposition regions between isosurfaces are not extended to the whole molecule



These considerations reflect the tridimensional nature of the VS+ descriptors and do not allow to obtain a clear superposition between VS+ and Abraham's parameters. Nevertheless some analogies between blocks and solvatochromic parameters (in the format used by Lombardo and coworkers: R_2 in the excess molar refraction, π^H_2 is the dipolarity/polarizability, $\Sigma\alpha^H_2$ and $\Sigma\beta^H_2$ are the hydrogen bond acidity and basicity, respectively, and V_x is the McGowan's volume) can be found and are summarized below

- Size block can be associated with V_x and R_2
- N1 block can be associated $\Sigma\beta^H_2$
- O block can be associated with $\Sigma\alpha^H_2$
- The negative portion of the DRY block can be associated with $\Sigma\pi^H_2$.