Hierarchical quantitative structure-activity relationships (HiQSARs) for the prediction of physicochemical and toxicological properties of chemicals using computed molecular descriptors

Subhash C. Basak

International Society of Mathematical Chemistry and Natural Resources Research Institute & Department of Chemistry and Biochemistry, University of Minnesota Duluth, USA

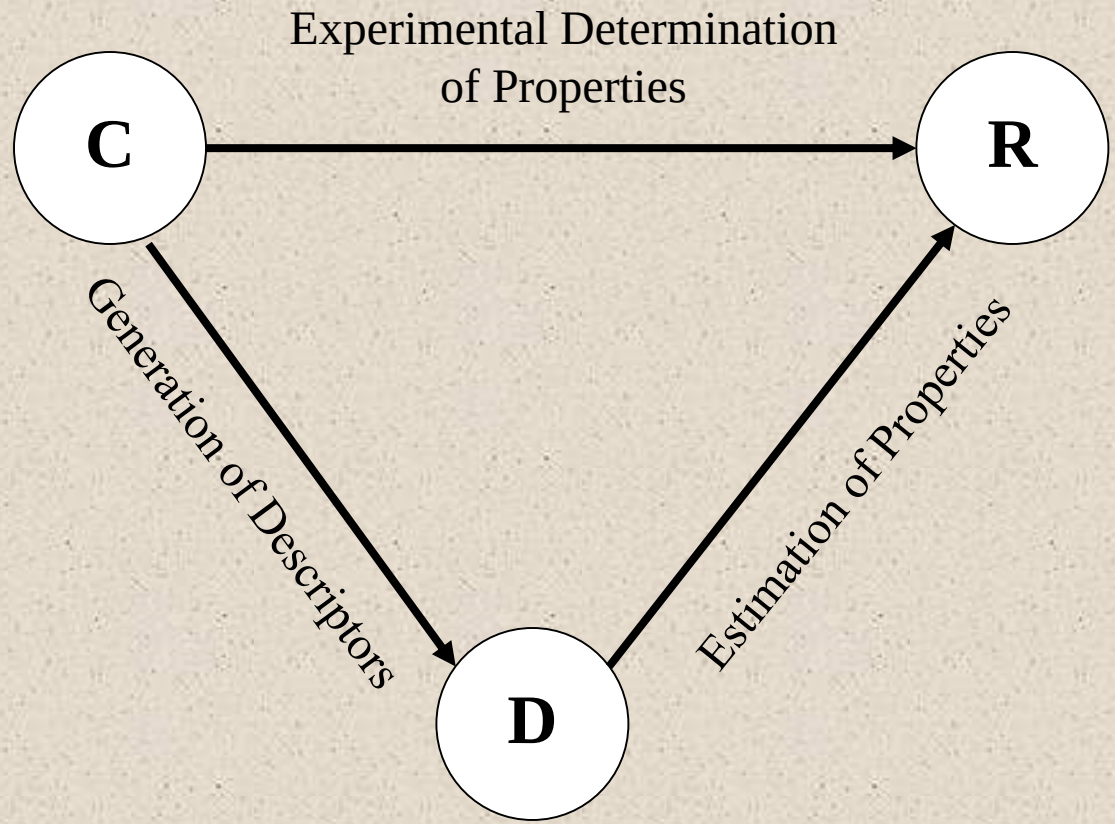
Subhabrata Majumdar

School of Statistics University of Minnesota Twin Cities Minneapolis, MN, USA

Need for chemical evaluation

We need to evaluate chemicals for various purposes, e.g., new drug discovery, risk assessment of environmental pollutants, specialty chemical design, medical diagnostics

Experimental *vs in silico* structural approach



C = a set of chemicals

R = the set of real numbers

D = a set of structural descriptors

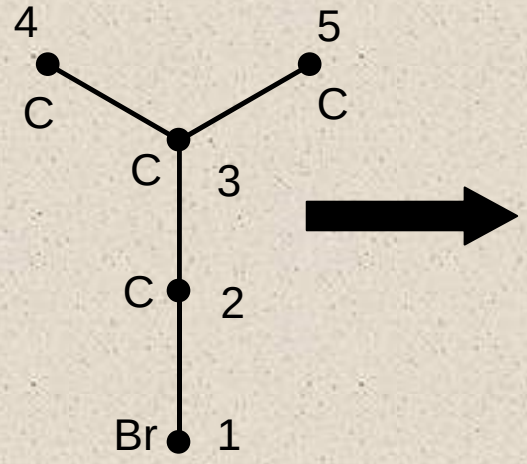
$$P = f(S)$$

Characterization of Molecular Structure

Wiener Index, W

$$W = 1/2 \sum_{ij} d_{ij}$$

where d_{ij} is the distance between vertices v_i and v_j in G_1



	1	2	3	4	5	Row Sum
1	0	1	**2**	3	**3**	9
2	1	0	1	2	2	6
3	2	1	0	1	1	5
4	3	2	1	0	2	8
5	3	2	1	2	0	8
						36

W = 36 / 2
= 18

QSAR development

Topostructural (TS), topochemical, (TC), geometrical (3-D),and quantum chemical (QC) indices have been used for QSAR

Ridge regression has been used for QSAR formulation

Interrelated two way clustering (ITC) was used for variable selection

HiQSARs for the prediction of Vapor Pressure

	Training Set (342)			Test Set (134)	
• Parameter	F	R2	S	R2	S
TS	104	.48	.56	.58	.46
TC	126	.79	.36	.86	.27
3-D	169	.52	0.53	.62	.44
All Indices	117	.80	0.35	.84	.28

Conclusion

HiQSAR studies of vapor pressure, 508 diverse mutagen data, and other QSARs (reference given below) indicate that in many cases a combination of TS + TC descriptors gives reasonably good QSAR. The addition of 3-D or QC descriptors after the use of TS and TC descriptors does not make much improvement in model quality.

REFREENCES

Gute, B. D.; Basak, S. C. Predicting acute toxicity of benzene derivatives using theoretical molecular descriptors: a hierarchical QSAR approach, SAR QSAR Environ. Res., 1997, 7, 117–131.

Gute, G. D.; Grunwald, G. D.; Basak, S. C. Prediction of the dermal penetration of polycyclic aromatic hydrocarbons (PAHs): A hierarchical QSAR approach, B.D. SAR QSAR Environ. Res., 1999, 10, 1–15.

Basak, S. C.; Mills, D. R.; Balaban, A. T.; Gute, B. D. Prediction of mutagenicity of aromatic and heteroaromatic amines from structure: A hierarchical QSAR approach, , J. Chem. Inf. Comput. Sci., 2001, 41, 671–678

REFREENCES

Basak, S. C.; Majumdar, S. Current landscape of hierarchical QSAR modeling and its applications: Some comments on the importance of mathematical descriptors as well as rigorous statistical methods of model building and validation, in Advances in Mathematical Chemistry and Applications, volume 1, pp. 251-281, Basak, S. C., Restrepo, G. and Villaveces, J. L., Editors, Bentham eBooks, Bentham Science Publishers, 2015.

Gute, B. D.; Basak, S. C.; Balasubramanian, K.; Geiss, K.; Hawkins, D. M. Prediction of halocarbon toxicity from structure: A hierarchical QSAR approach, Environ. Toxicol. Pharmacol., 2004, 16, 121–129.

[Basak, S. C.; Natarajan, R.; Mills, D. Structure-activity relationships for mosquito repellent aminoamides using the hierarchical QSAR method based on calculated molecular descriptors, Conference proceedings, WSEAS Transactions on Information Science and Applications, 2005, 7, 958–963.

Basak, S. C. Philosophy of Mathematical Chemistry: A Personal Perspective,

HYLE--International Journal for Philosophy of Chemistry, 2013, 19, 3-17.