

**Read-across-based intelligent learning: Development
of a global q-RASAR model for the efficient
quantitative predictions of skin sensitization potential
of diverse organic chemicals**

Arkaprava Banerjee, Kunal Roy*

**Drug Theoretics and Cheminformatics Laboratory, Department of Pharmaceutical
Technology, Jadavpur University, Kolkata 700032, India**

Email: kunal.roy@jadavpuruniversity.in

Supplementary Material SI-2

The PLS QSAR model

pEC_3

$$= -11.497 + 0.165 \times SCBO - 0.248 \times nCsp2 + 2.165 \times Eta_sh_p - 0.1 \times nArCO + 10.370 \times PDI + 3.110 \times Eta_epsi_2 + 0.152 \times max_conj_path \\ \times -0.146 \times F04[N-O] - 0.316 \times F09[O-O]$$

$$R_{Train}^2 = 0.696$$

$$Q_{LOO}^2 = 0.644$$

$$MAE_{Train} = 0.410$$

$$Q_{F1}^2 = 0.526$$

$$Q_{F2}^2 = 0.524$$

$$MAE_{Test} = 0.562$$

SCBO: sum of conventional bond orders (H-depleted)

nCsp2: number of sp2 hybridized Carbon atoms

Eta_sh_p: eta p shape index

X2sol: solvation connectivity index of order 2

nArCO: number of ketones (aromatic)

PDI: packing density index

Eta_epsi_2: eta electronegativity measure 2

max_conj_path: maximum number of atoms that can be in conjugation with each other

F04[N-O]: frequency of N – O at topological distance 4

F09[O-O]: frequency of O – O at topological distance 9

Coefficient Plot CoeffCS[Comp. 8](YVar pEC3 (LLNA))

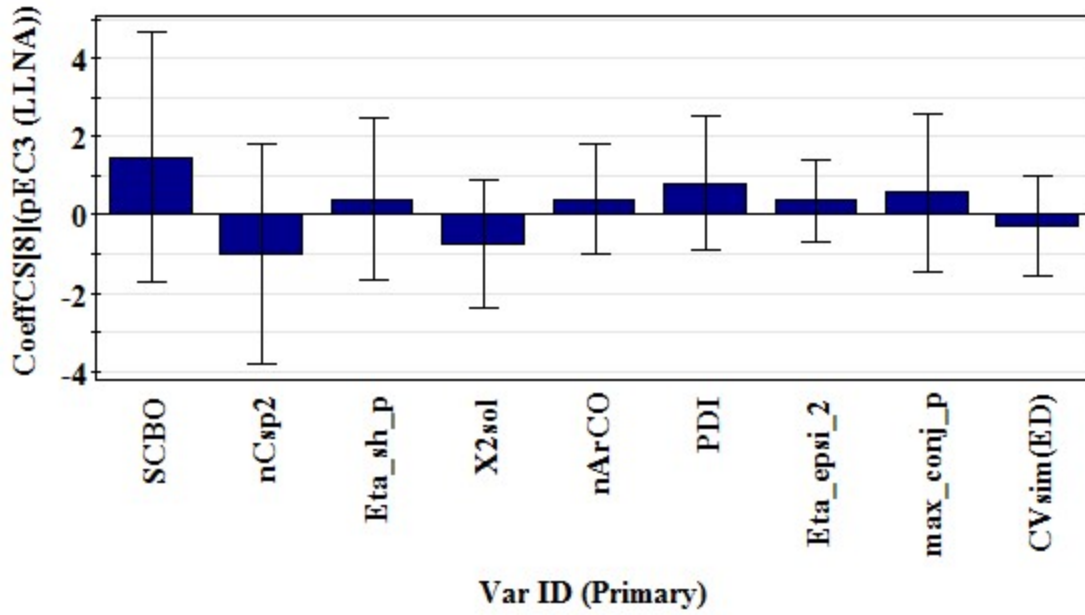
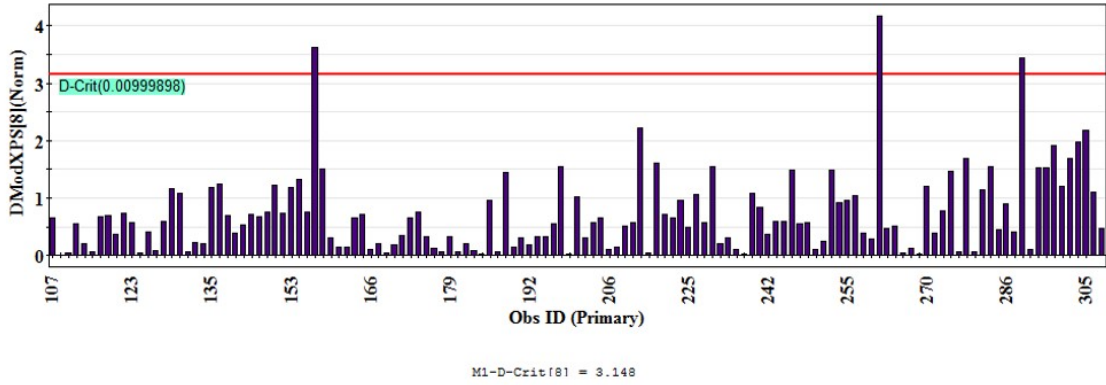


Figure S1. Regression Coefficient Plot

**DModX AD Plot (TRAIN)
DModXPS[Comp. 8]**



**DModX AD Plot (TEST)
DModXPS[Comp. 8]**

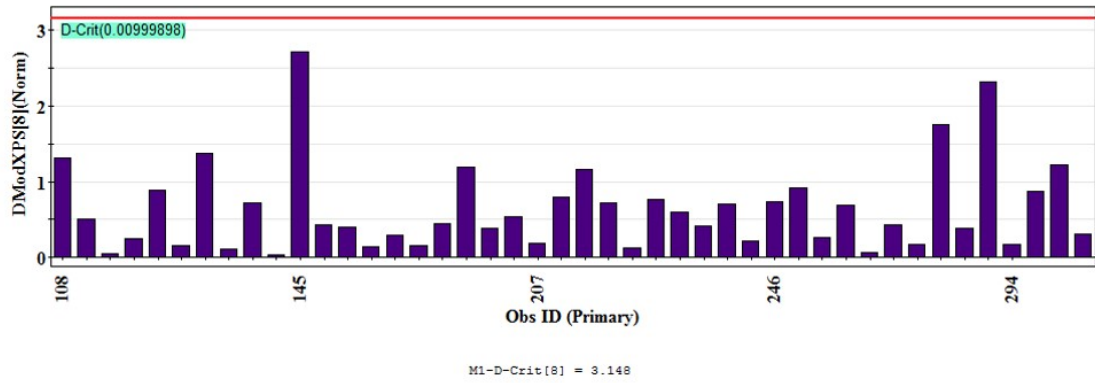


Figure S2. The DModX AD Plots

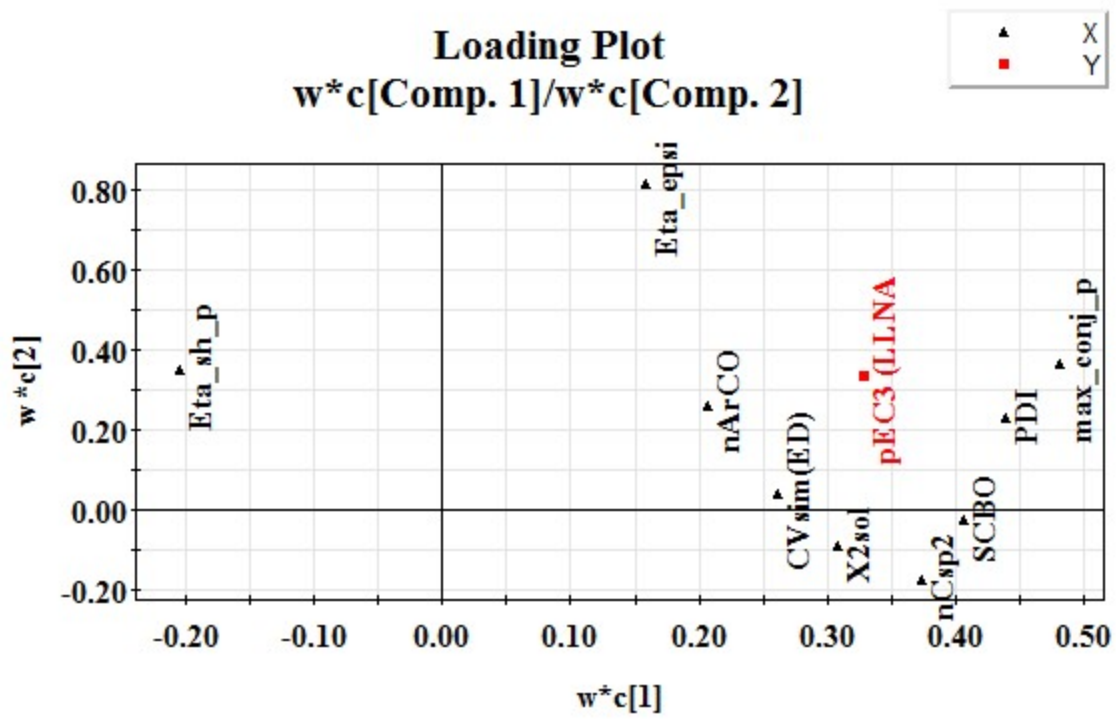


Figure S3. Loading Plot

Score Plot
t[Comp. 1]/t[Comp. 2]

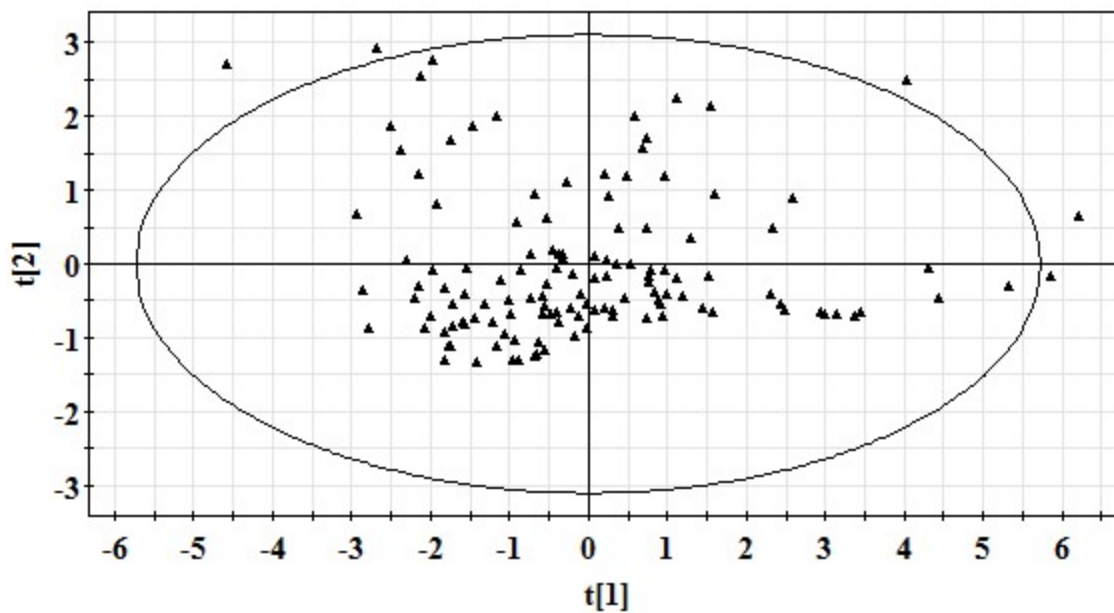


Figure S4. Score Plot

Variable Importance Plot VIP[Comp. 8]

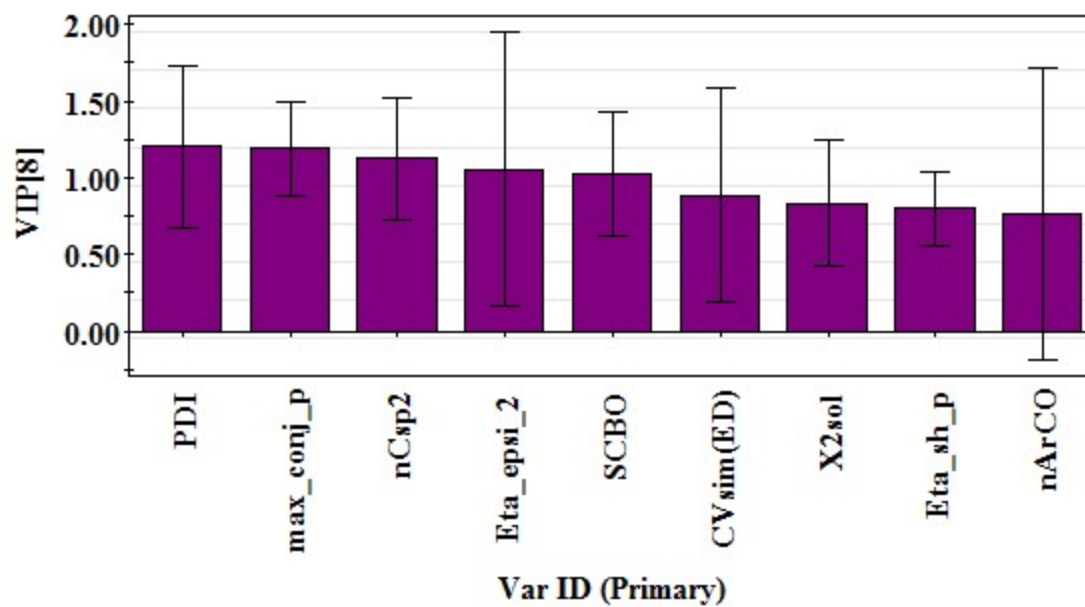


Figure S5. Variable Importance Plot

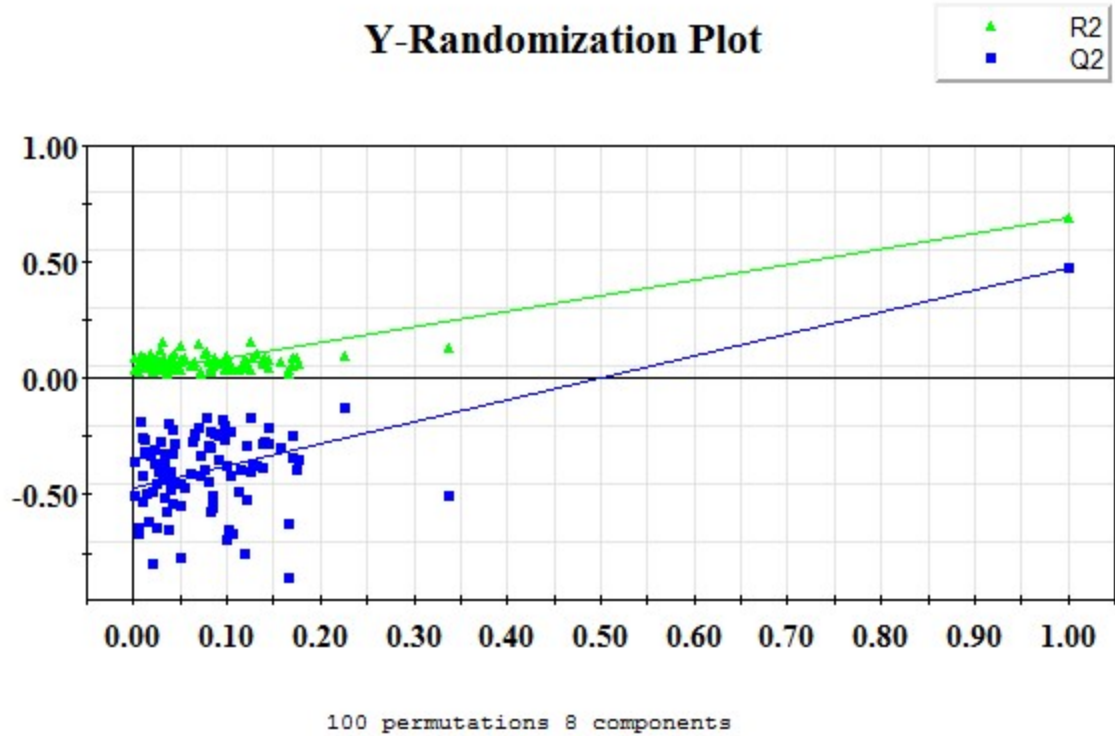


Figure S6. Y-Randomization Plot