Quantitative read-across structure-property relationship (q-RASPR): A novel approach to estimate the bioaccumulative potential for diverse classes of industrial chemicals in aquatic organisms

Prodipta Bhattacharyya, Pabitra Samanta, Ankur Kumar, Shubha Das, Probir Kumar Ojha*

Drug Discovery and Development Laboratory (DDD Lab),
Department of Pharmaceutical Technology, Jadavpur University,
Kolkata 700032, India
*Corresponding author: P K Ojha, Email: probirojha@yahoo.co.in
ORCID id: 0000-0003-4796-3915, Phone: +91 8777677004

Contents:

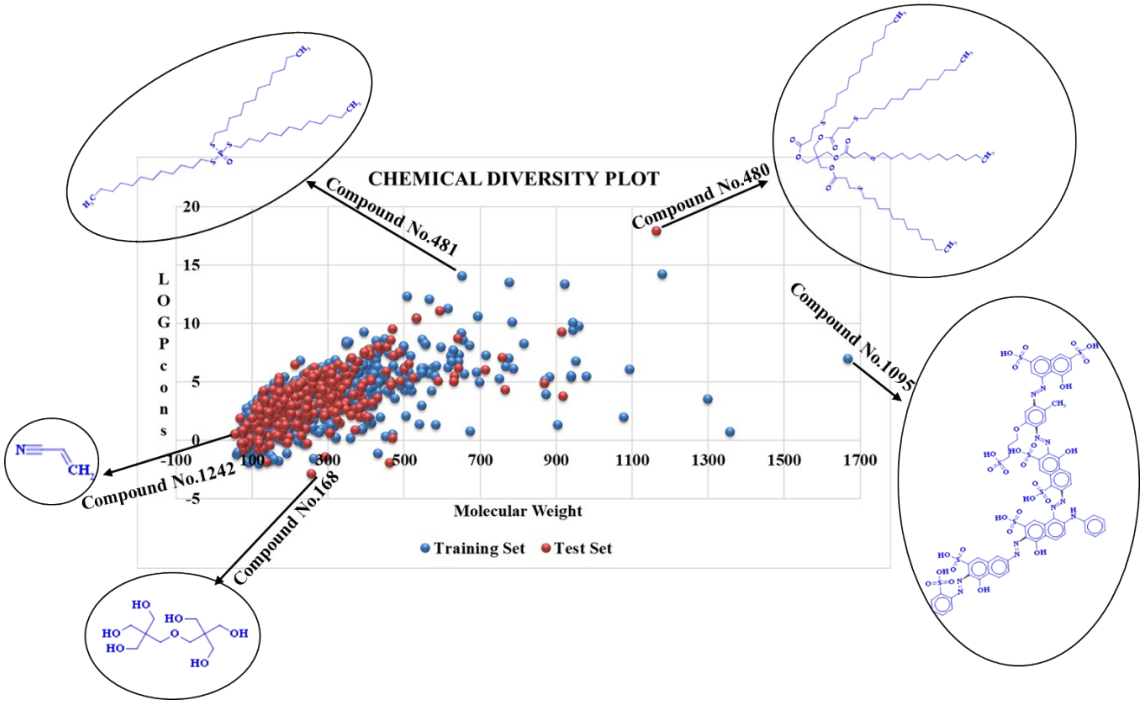

Fig. S1. Chemical diversity plot (molecular weight vs LOGPcons) of the training and test sets to check the heterogeneity of the dataset.

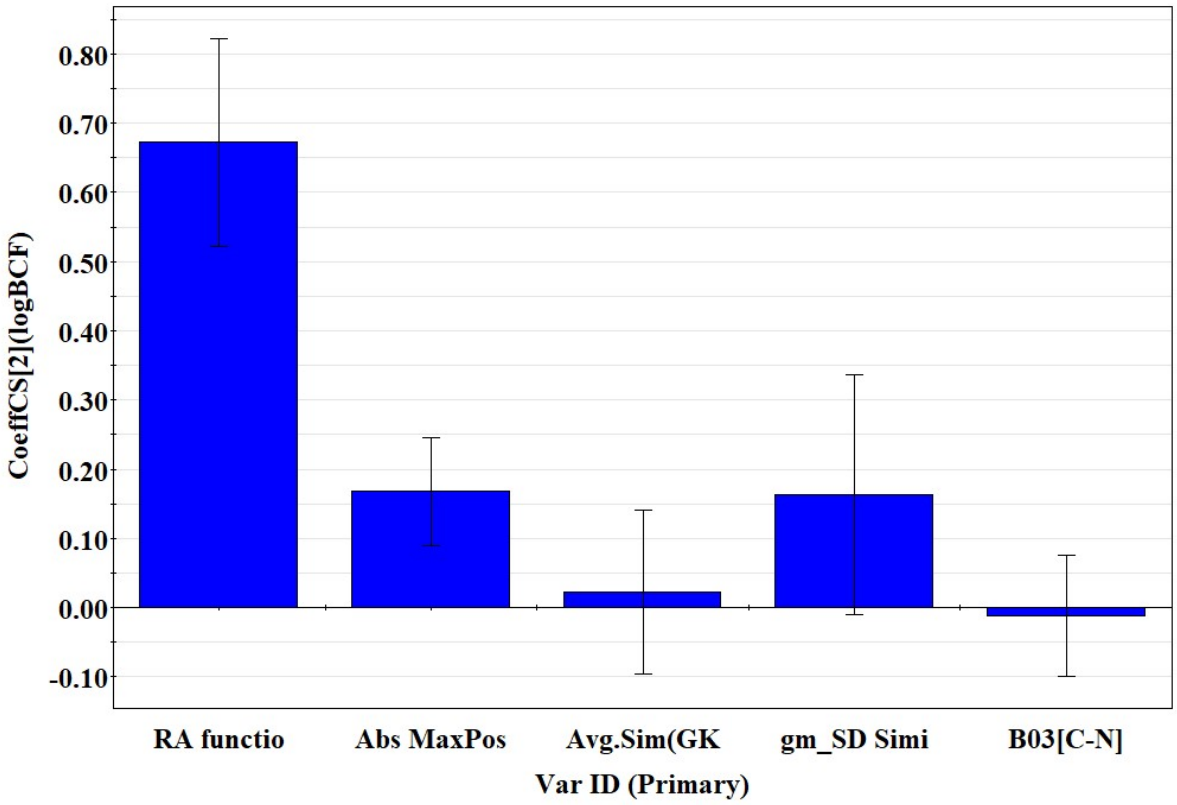

Fig. S2. Regression coefficient plot of PLS-based q-RSAPR model.

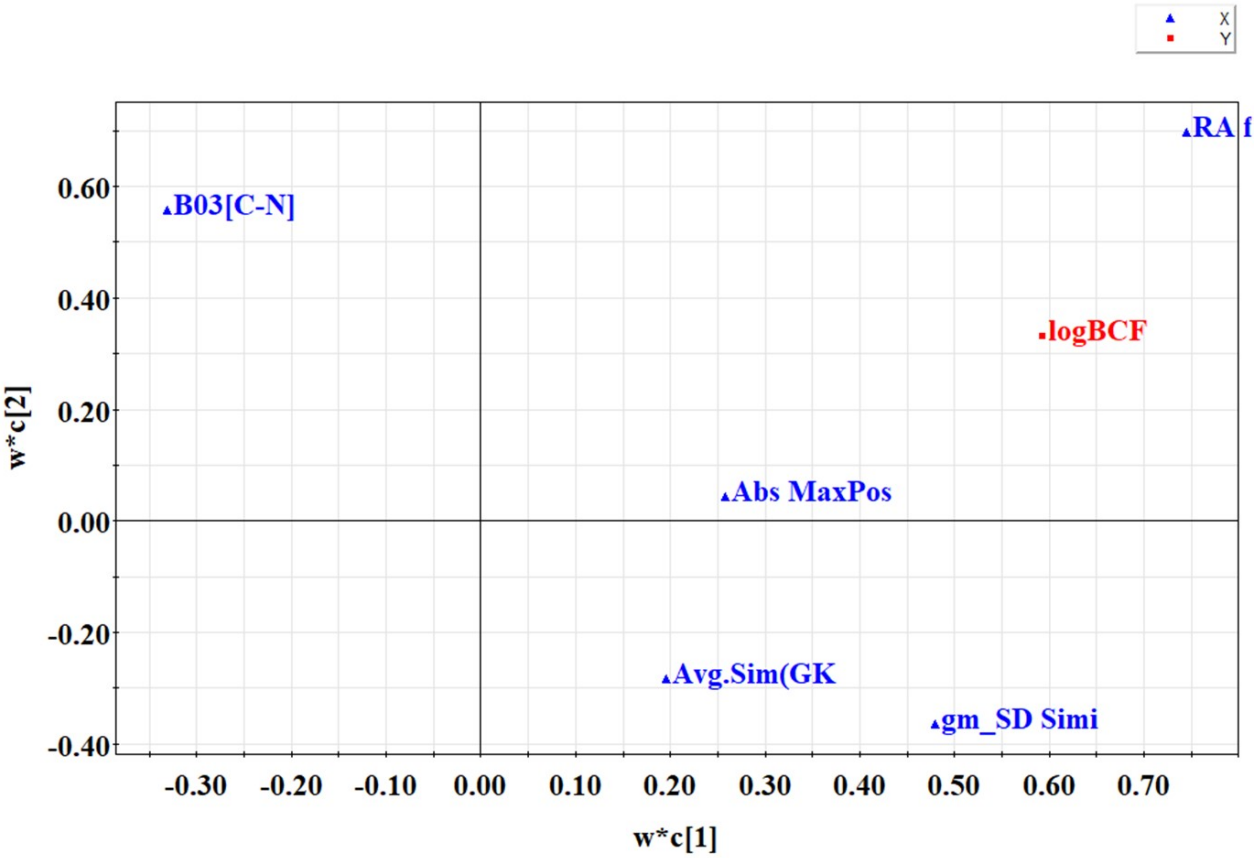

Fig S3. Loading plot of PLS-based q-RSAPR model.

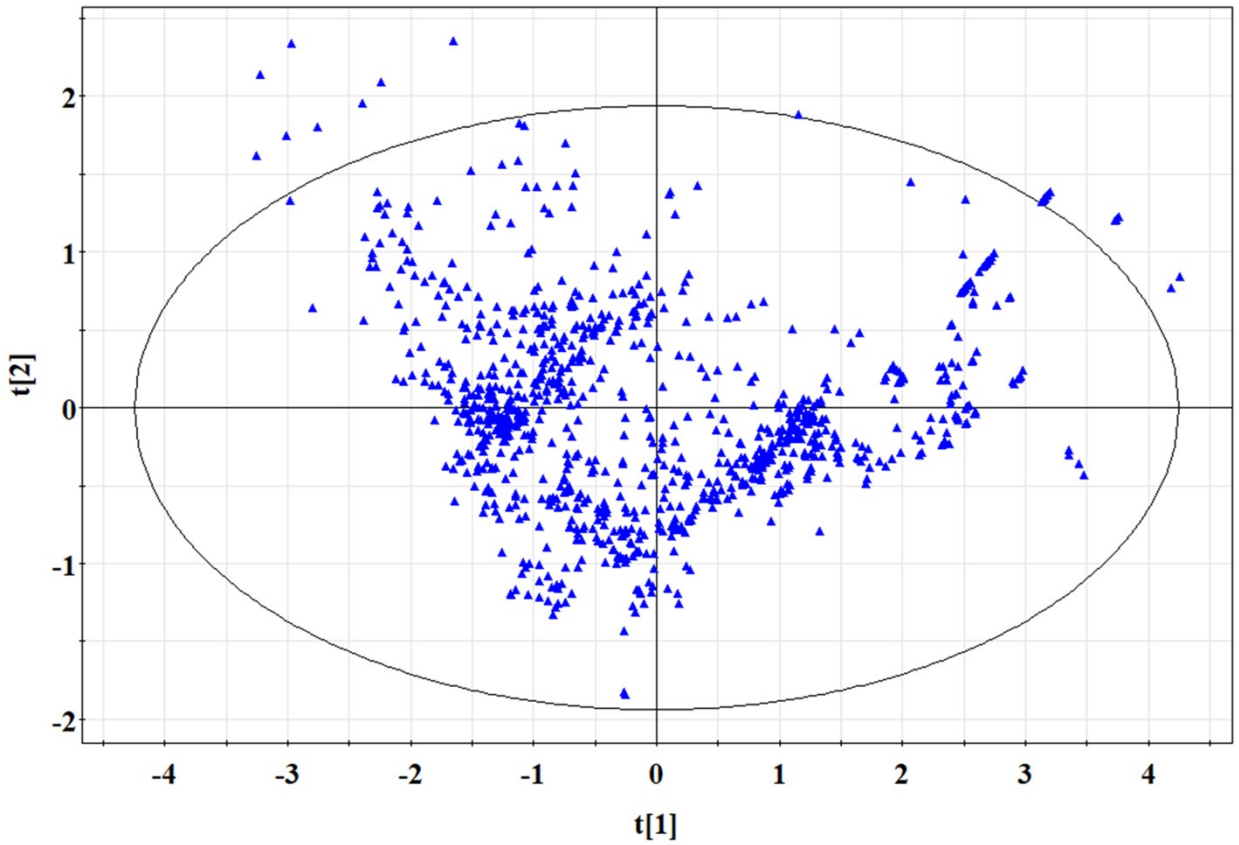

Fig. S4. Score plot of PLS-based q-RSAPR model.

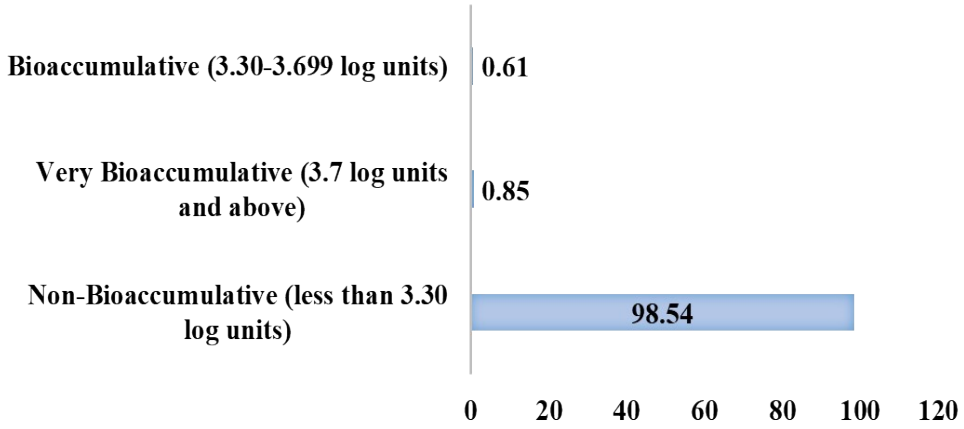

Fig. S5. Ranking of the compounds of PPDB based on their bioaccumulative potential.

Table S1. Detailed list of RASPR descriptors and their definition.

SL.NO.	RASPR DESCRIPTORS	DEFINITIONS
1	*RA function*	A composite function derived from Read-Across.
2	*MaxPos*	Similarity score of the closest positive source compound (with an observed response value greater than the mean activity of the training set).
3	*MaxNeg*	Similarity score of the closest negative source compound (with an observed response value less than the mean activity of the training set).
4	*Abs Maxpos-MaxNeg*	Absolute difference between the *MaxPos* and *Maxneg* levels.
5	*SE*	Weighted standard error of the close source compounds' response values.
6	*CVact*	Coefficient of variation of the close source compounds' observed response values.
7	*SD_Activity*	Weighted standard deviation of the close source compounds' observed response values.
8	*CVsim*	Coefficient and variation of the similarity values of the close source compounds.
9	*SD_similarity*	The standard deviation of the close source compounds' similarity levels.
10	*Pos.Avg.Sim*	The positive close source compounds' average similarity levels.
11	*Neg.Avg.Sim*	The negative close source compounds ' average similarity levels.
12	*Avg.Sim*	Average similarity level of the close source compounds.
13	g_m	A novel concordance measure also known as *Banerjee-Roy Coefficient*
14	*gm*SD_Similarity*	Product of the g_m and *SD similarity* levels
15	*gm*Avg.Sim*	Product of the g_m and *Avg.Sim* levels

Table S2: Top and least 20 compounds based on PLS-based q-RASPR model prediction of PPDB.

SL. NO.	CHEMICAL	PREDICTION	REAL-WORLD DATA	SOURCE
Top 20 compounds based on PLS-based q-RASPR model prediction				
1	Hexachlorobenzene	5.033 (Very Bioaccumulative)	4.544 (Very Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
2	Dienochlor	4.893 (Very Bioaccumulative)	6.4313 (Very Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
3	2,4-DEP	4.573 (Very Bioaccumulative)	Bioaccumulative	https://pubchem.ncbi.nlm.nih.gov/compound/2_4-Dimethylphenol
4	Tetrasul	4.440 (Very Bioaccumulative)	Bioaccumulative	https://www.coastalwiki.org/wiki/Tetrasul
5	Dieldrin	4.269 (Very Bioaccumulative)	4.5440 (Very Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
6	Endrin	4.269 (Very Bioaccumulative)	3.6 (Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
7	DDT	4.248 (Very Bioaccumulative)	3.501 (Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
8	O,P'-DDT	4.245 (Very Bioaccumulative)	Bioaccumulative	https://pubchem.ncbi.nlm.nih.gov/compound/O_P_-Ddt
9	O,O'-DDT	4.241 (Very Bioaccumulative)	No data found	-
10	Isobenzan	4.153 (Very Bioaccumulative)	Bioaccumulative	https://pubchem.ncbi.nlm.nih.gov/compound/9271#section=Environmental-

11	DDD	4.120 (Very Bioaccumulative)	Very Bioaccumulative	Bioconcentration&fullscreen=true https://pubchem.ncbi.nlm.nih.gov/compound/6294#section=Environmental-Bioconcentration&fullscreen=true
12	Bromocyclen	4.057 (Very Bioaccumulative)	No data found	-
13	Chlorbenside	3.919 (Very Bioaccumulative)	No data found	-
14	Tridiphane	3.781 (Very Bioaccumulative)	3.703 (Very Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
15	Chlorbicyclen	3.692 (Bioaccumulative)	No data found	-
16	Leptophos	3.663 (Bioaccumulative)	3.782 (Very Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
17	EPBP	3.660 (Bioaccumulative)	No data found	-
18	Fenchlorphos	3.595 (Bioaccumulative)	Bioaccumulative	https://pubchem.ncbi.nlm.nih.gov/compound/Fenchlorphos
19	Trichloronate	3.580 (Bioaccumulative)	Bioaccumulative	https://pubchem.ncbi.nlm.nih.gov/compound/9477#section=Environmental-Fate-Exposure-Summary&fullscreen=true
20	Chlorphonium	3.493 (Bioaccumulative)	No data found	-

Least 20 compounds based on PLS-based q-RASPR model prediction

SL. NO.	CHEMICAL	PREDICTION	REAL-WORLD DATA	SOURCE
1	2-methoxyethanol	0.091 (Non-Bioaccumulative)	0.477 Non-Bioaccumulative	https://sitem.herts.ac.uk/aeru/ppdb/
2	Izopamfos	0.166 (Non-Bioaccumulative)	No data found	-
3	Butan-1-ol	0.189 (Non-Bioaccumulative)	Non-Bioaccumulative	https://labequipsupply.com/product/butan-1-ol-ar/
4	DAEP	0.268 (Non-Bioaccumulative)	No data found	-
5	2-aminobutane	0.296 (Non-Bioaccumulative)	Non-Bioaccumulative	https://www.fishersci.ch/store/msds?partNumber=11459923&countryCode=CH&language=en
6	Methylene bisthiocyanate	0.304 (Non-Bioaccumulative)	Non-Bioaccumulative	https://chempak.net/msds/TOLCIDE%20MBT.PDF
7	Ethoate-methyl	0.308 (Non-Bioaccumulative)	Non-Bioaccumulative	https://sitem.herts.ac.uk/aeru/ppdb/
8	Mevinphos	0.318 (Non-Bioaccumulative)	0.505 Non-Bioaccumulative	https://www.rivm.nl/bibliotheek/rapporten/601714004.pdf
9	Mepiquat	0.321 (Non-Bioaccumulative)	0.301 Non-Bioaccumulative	https://sitem.herts.ac.uk/aeru/ppdb/
10	Gliftor	0.322 (Non-Bioaccumulative)	No data found	-
11	Omethoate	0.322 (Non-Bioaccumulative)	1.875 (Non-Bioaccumulative)	https://sitem.herts.ac.uk/aeru/ppdb/
12	Fenuron-TCA	0.327 (Non-Bioaccumulative)	No data found	-
13	Amidithion	0.328 (Non-Bioaccumulative)	No data found	-
14	Dimethoate	0.330	0.903	https://sitem.

		(Non-Bioaccumulative)	(Non-Bioaccumulative)	herts.ac.uk/aeru/ppdb/
15	Oxydeprofos	0.332 (Non-Bioaccumulative)	No data found	-
16	Methacrifos	0.338 (Non-Bioaccumulative)	Non-Bioaccumulative	https://sitem.herts.ac.uk/aeru/ppdb/
17	Oxydemeton-methyl	0.343 (Non-Bioaccumulative)	0.491 (Non-Bioaccumulative)	https://pubchem.ncbi.nlm.nih.gov/compound/4618#section=Environmental-Fate-Exposure-Summary
18	Methylacetophos	0.346 (Non-Bioaccumulative)	No data found	-
19	Acetophos	0.361 (Non-Bioaccumulative)	0.301 (Non-Bioaccumulative)	https://pubchem.ncbi.nlm.nih.gov/compound/17041
20	Lime sulphur	0.361 (Non-Bioaccumulative)	Non-Bioaccumulative	https://brandt.co/media/6248/brandt-lime-sulfur-sds.pdf