

Table 2. Data types and ranges of 13 input variables

Name	Description	Data type	Value range
h-CLAT	h-CLAT majority call	Categorical	0/1
DPRA	DPRA majority call	Categorical	0/1
KeratinoSens	KeratinoSens majority call	Categorical	0/1
Toolbox	Read-across prediction using QSAR Toolbox	Categorical	0/1
Avg.Lys.Cys	Average lysine and cysteine depletion measurements from DPRA	Numeric	0–95.0
Lys	Average lysine depletion from DPRA	Numeric	0–91.0
Cys	Average cysteine depletion from DPRA	Numeric	0–100
Log P	Octanol/water partition coefficient	Numeric	–8.28 to 6.46 ^a
Log S	Water solubility (mol l ⁻¹)	Numeric	–6.39 to 1.92 ^a
Log VP	Vapor pressure (mmHg)	Numeric	–28.47 to 5.89 ^a
MP	Melting point (°C)	Numeric	–148.5 to 288.0
BP	Boiling point (°C)	Numeric	–19.1 to 932.2
MW	Molecular weight (g mol ⁻¹)	Numeric	30.03– 581.57

DPRA, direct peptide reactivity assay; h-CLAT, human cell line activation test.
^a a Range for base 10 logarithm of these measurements.