Table 1: Descriptors used for the development of structure based models on natural dataset.

S. No.	Descriptors	Class	Description
1	PubchemFP149	Pubchem Fingerprint	>= 1 unsaturated non-aromatic heteroatom-containing ring size 5
2	PubchemFP36	Pubchem Fingerprint	>= 8 S
3	PubchemFP375	Pubchem Fingerprint	C(~N)(~N)
4	PubchemFP700	Pubchem Fingerprint	O-C-C-C-C-C-O-C
5	EStateFP7	Estate fingerprint	[OD2H0](-*)-*
6	ExtFP228	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
7	ExtFP236	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
8	ExtFP816	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
9	ExtFP875	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
10	ExtFP992	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
11	FP1021	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
12	FP12	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
13	FP276	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
14	FP286	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
15	FP648	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
16	FP702	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
17	FP704	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
18	FP773	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
19	FP825	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
20	FP924	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
21	FPSA-3	Charged Partial Surface Area (CPSA)	PPSA-3 / total molecular surface area (3D)
22	GraphFP167	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
23	GraphFP359	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
24	GraphFP752	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
25	GraphFP888	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
26	GraphFP919	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account

27	KRFP2410	Klekota-Roth fingerprint	[!#1]N1[CH2][CH2][CH2][CH]1C(=O)[OH]
28	KRFP3767	Klekota-Roth fingerprint	CCN(CCO)C=O
29	KRFP480	Klekota-Roth fingerprint	[!#1][CH2][NH]C(=O)[!#1]
30	KRFP669	Klekota-Roth fingerprint	[!#1][CH3]
31	KRFP681	Klekota-Roth fingerprint	[!#1][NH][CH]([!#1])[CH2][CH2][CH3]
32	KRFP691	Klekota-Roth fingerprint	[!#1][NH][CH]([CH2][CH]([CH3])[CH3])C(=O)[OH]
33	KRFP7	Klekota-Roth fingerprint	[!#1][CH]([!#1])[CH2][CH2][CH3]
34	KRFPC3054	Klekota-Roth fingerprint count	Count of chemical substructures
35	MACCSFP104	MACCS fingerprint	('[!#6;!#1;!H0]~*~[CH2]~*',0), # QHACH2A
36	MACCSFP109	MACCS fingerprint	('*~[CH2]~[#8]',0), # ACH2O
37	MACCSFP160	MACCS fingerprint	('[C;H3,H4]',0), #CH3
38	MACCSFP66	MACCS fingerprint	('[#6]~[#6](~[#6])(~[#6])~*',0), # CC(C)(C)A
39	MACCSFP90	MACCS fingerprint	('[$([!#6;!#1;!H0]~*~*~[CH2]~*),$([!#6;!#1;!H0;R]1@[R]@[R]@[CH2;R]1),$([!#6;!#1;!H0]~[R]1@[R]@[CH2;R]1)]',0), # QHAACH2A
40	maxdssC	Maximum atom-type E-State: =C<	Maximum atom-type E-State: =C< (2D)
41	maxssCH2	Maximum atom-type E-State: -CH2-	Maximum atom-type E-State: -CH2- (2D)
42	RNCS	Charged partial surface area	Charged partial surface area (2D)
43	VC-5	Chi cluster	Chi cluster(2D) / Valence cluster, order 5
44	Weta3.eneg	WHIMDescriptor	Directional WHIM descriptor weighted by Mulliken atomic electronegativites
45	Wnu1.eneg	WHIMDescriptor	Directional WHIM descriptor weighted by Mulliken atomic electronegativites

Table 2: Descriptors used for the development of structure based models on modified dataset.

S. No.	Descriptors	Class	Description
1	ExtFP28	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
2	ExtFP816	Extended Fingerprint	Extends the Fingerprinter with additional bits describing ring features
3	FP286	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
4	FP307	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
5	FP841	CDK fingerprint	Fingerprint of length 1024 and search depth of 8
6	FPSA-3	Charged Partial Surface Area (CPSA)	PPSA-3 / total molecular surface area (3D)
7	GraphFP48	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
8	GraphFP546	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
9	GraphFP752	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
10	GraphFP839	CDK graph only fingerprint	Specialized version of the Fingerprinter which does not take bond orders into account
11	KRFP2264	Klekota-Roth fingerprint	[!#1]N([!#1])[CH3]
12	KRFP309	Klekota-Roth fingerprint	[!#1][CH2][CH]([CH3])[NH]C(=O)[!#1]
13	KRFP3375	Klekota-Roth fingerprint	CC(C)C=O
14	KRFP3668	Klekota-Roth fingerprint	CCC=C(C)C
15	KRFP382	Klekota-Roth fingerprint	[!#1][CH2][CH2][CH3]
16	KRFP4664	Klekota-Roth fingerprint	OC(=O)CCc1c[nH]c2ccccc12
17	KRFP608	Klekota-Roth fingerprint	[!#1][CH2]N([!#1])[CH3]
18	KRFP630	Klekota-Roth fingerprint	[!#1][CH2]N([CH3])C(=O)[!#1]
19	KRFP758	Klekota-Roth fingerprint	[!#1][NH][CH2][CH]([!#1])[!#1]
20	KRFPC3054	Klekota-Roth fingerprint count	Count of chemical substructures
21	KRFPC309	Klekota-Roth fingerprint count	Count of chemical substructures
22	KRFPC3375	Klekota-Roth fingerprint count	Count of chemical substructures
23	KRFPC466	Klekota-Roth fingerprint count	Count of chemical substructures
24	KRFPC4664	Klekota-Roth fingerprint count	Count of chemical substructures
25	KRFPC608	Klekota-Roth	Count of chemical substructures

		fingerprint count	
26	KRFPC630	Klekota-Roth fingerprint count	Count of chemical substructures
27	KRFPC822	Klekota-Roth fingerprint count	Count of chemical substructures
28	MACCSFP104	MACCS fingerprint	('[!#6;!#1;!H0]~*~[CH2]~*',0), # QHACH2A
29	MACCSFP109	MACCS fingerprint	('*~[CH2]~[#8]',0), # ACH2O
30	MACCSFP68	MACCS fingerprint	('[!#6;!#1;!H0]~[!#6;!#1;!H0]',0), # QHQH
31	MACCSFP90	MACCS fingerprint	('[$([!#6;!#1;!H0]~*~*~[CH2]~*),$([!#6;!#1;!H0;R]1@[R]@[R]@[CH2;R]1),$([!#6;!#1;!H0]~[R]1@[R]@[CH2;R]1)]',0), # QHAACH2A
32	maxssNH	Atom-type E-state (Electrotopological state)	Maximum atom-type E-State: -NH- (2D)
33	minHdsCH	Atom-type E-state (Electrotopological state)	Minimum atom-type H E-State: =CH- (2D)
34	PubchemFP151	Pubchem Fingerprint	>= 2 saturated or aromatic carbon-only ring size 5
35	PubchemFP172	Pubchem Fingerprint	>= 5 saturated or aromatic carbon-only ring size 5
36	PubchemFP193	Pubchem Fingerprint	>= 3 saturated or aromatic carbon-only ring size 6
37	PubchemFP194	Pubchem Fingerprint	>= 3 saturated or aromatic nitrogen-containing ring size 6
38	PubchemFP36	Pubchem Fingerprint	>= 2 saturated or aromatic nitrogen-containing ring size 8
39	PubchemFP690	Pubchem Fingerprint	O-C-C-C-C-C-O
40	PubchemFP700	Pubchem Fingerprint	O-C-C-C-C-C-O-C
41	Weta3.eneg	WHIMDescriptor	Directional WHIM descriptor weighted by Mulliken atomic electronegativites
42	WK.volume	WHIMDescriptor	Non-directional WHIM weighted by van der Waals volumes
43	Wnu1.eneg	WHIMDescriptor	Directional WHIM descriptor weighted by Mulliken atomic electronegativites