Supporting information for: Metabolite collision cross section prediction without energy-minimized structures, by Soper-Hopper, *et al*.

Table S1: Comparison of previous approaches used for CCS value prediction

Approach Name	Training Set	Validation Approach	Regression Method	Average Prediction Errors	Reference
Ghent University CCS prediction	Molecular descriptors for 56 deprotonated phenolics	Leave-one-out crossvalidation and 30% of the initial dataset used as test set.	Stepwise Multiple Linear Regression (SMLR), Principal Components regression (PCR) and PLS	90% of compounds predicted with <5% error via SMLR. 93% predicted with errors better than 5% by PCR and PLS.	Anal. Chim. Acta, 2016, 924, 68-76
MetCCS	Molecular descriptors for 396 and 400 metabolite standards in positive and negative modes, respectively. Creation of a predicted CCS database of 35203 metabolites in the HMDB.	Ten-fold crossvalidation and external validation with 78 and 79 metabolites in positive and negative modes, respectively.	Support Vector Regression (SVR)	3% median relative error.	Anal. Chem, 2016, 88, 11084-11091. Bioinformatics, 2017, 33, 2235-2237. (Web Server)
LipidCCS	Molecular descriptors for n=329 (positive mode) and n=129 (negative mode) lipid adduct ions.	30% of the initial dataset used as test set.	Support Vector Regression (SVR)	0.4-0.5% median relative errors.	Anal. Chem., 2017, 89, 9559-9566.
Georgia Tech Collision Cross Section Prediction (CCSP)	Molecular descriptors for 146 lipids and 14 depsipeptides.	Ten-split Venetian Blinds crossvalidation and 25% of the initial dataset used as test set.	Partial Least Squares (PLS) linear multivariate regression	0.5-1.0% RMSEP following genetic algorithm variable selection	Chem. Commun., 2017,53, 7624-7627
Jaume I University CCS Prediction	Molecular descriptors for 205 pesticides	131 pesticides	Artificial Neural Networks (ANNs)	6% for 95% of tested analytes, with a median error of 2%	Anal. Chem., 2017, 89, 6583-6589.
University of Copenhagen CCS and retention time (RT) prediction.	Molecular descriptors for 357 pharmaceuticals, drugs of abuse and their metabolites.	External validation (n=36)	ANNs	2 min for RT and 5% for CCS for 91% of compounds	J. Chromatography A, 2018, 1542, 82-88.
DeepCCS	SMILES characters for compounds in the MetCCS, Astarita, Baker, McLean and CBM2018 databases. Total of 2439 molecules	Training set of n=1637, validation set of n=182 and test set of n=620.	Deep Neural Networks (DNNs)	1.49-4.86% median error. Absolute median error of 2.6%.	Anal. Chem., 2019, 91, 5191-5199.

Table S2: Genetic algorithm settings for variable selection:

Parameter	Value
Population Size	256
Window Width	1
% initial terms	10
Target Max	100
Penalty Slope	0.01
Max Generations	200
% at Convergence	50
Mutation Rate	0.005
Crossover	Double
Regression	PLS
# of LVs	5
Cross-Validation	Random
# of Splits	5
# of iterations	5
Replicate Runs	3

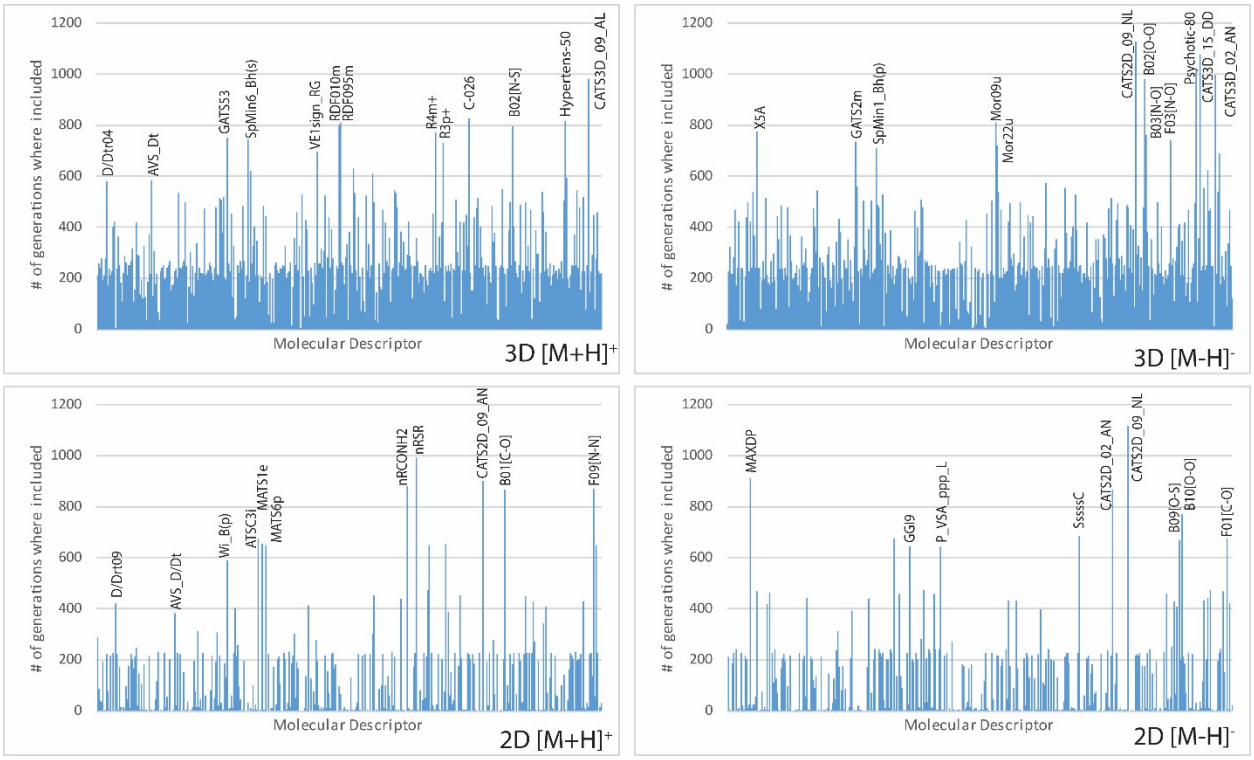

Figure S1. Frequency Plots describing the number of times each descriptor was chosen for use in the GA modeling, over the course of up to 200 generations in each of 5 iterations with 3 replicates. Labeled are the descriptors which were more frequently included. Abbreviation codes/ full descriptor names can be found at the end of supplemental information.

Models prepared using 3D conformer files, without molecular dynamics were also used to compare the 2D files. Models agree with similar errors.

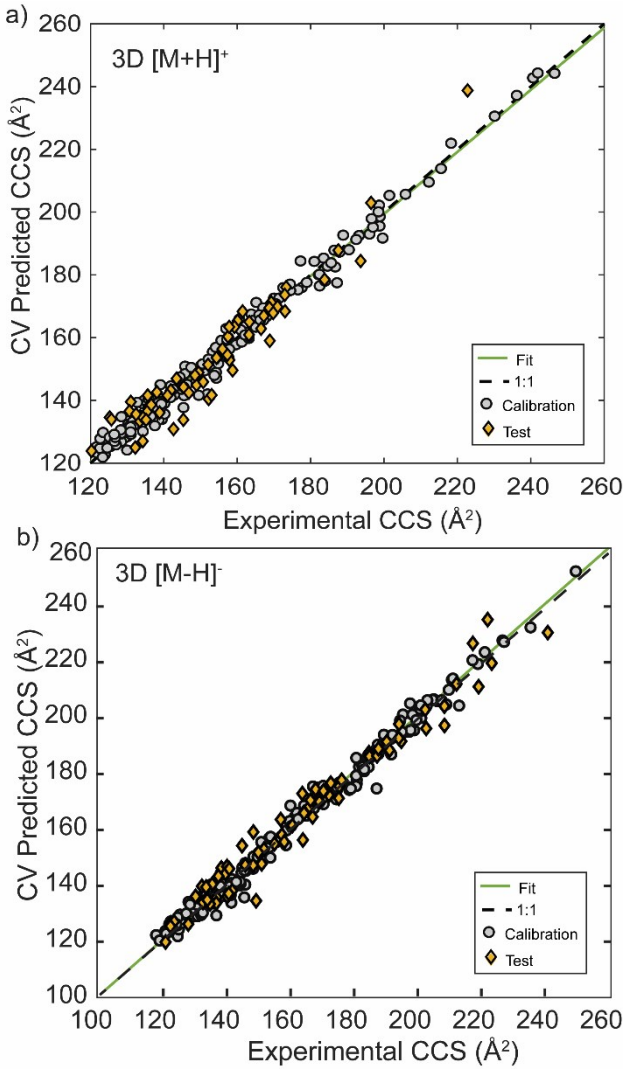

Figure S2. Measured *vs*. cross-validation (CV)-predicted CCS for metabolites developed using a) 3D structural files to calculate molecular descriptors for the metabolites with the $[M+H]^+$ adducts, and b) 3D conformer files to calculate molecular descriptors for the metabolites with the $[M-H]^-$ adducts. Optimal molecular descriptors were selected by a genetic algorithm.

The test set for each model was assessed for percent error on the prediction of the individual metabolites. Below are histograms demonstrating the percentage of each test set within a given percent error range. In yellow is the cumulative % of metabolites in the test set. The majority of metabolites in each model are predicted with less than a 3% error. Errors of less than 8% can be expected on nearly all metabolites predicted.

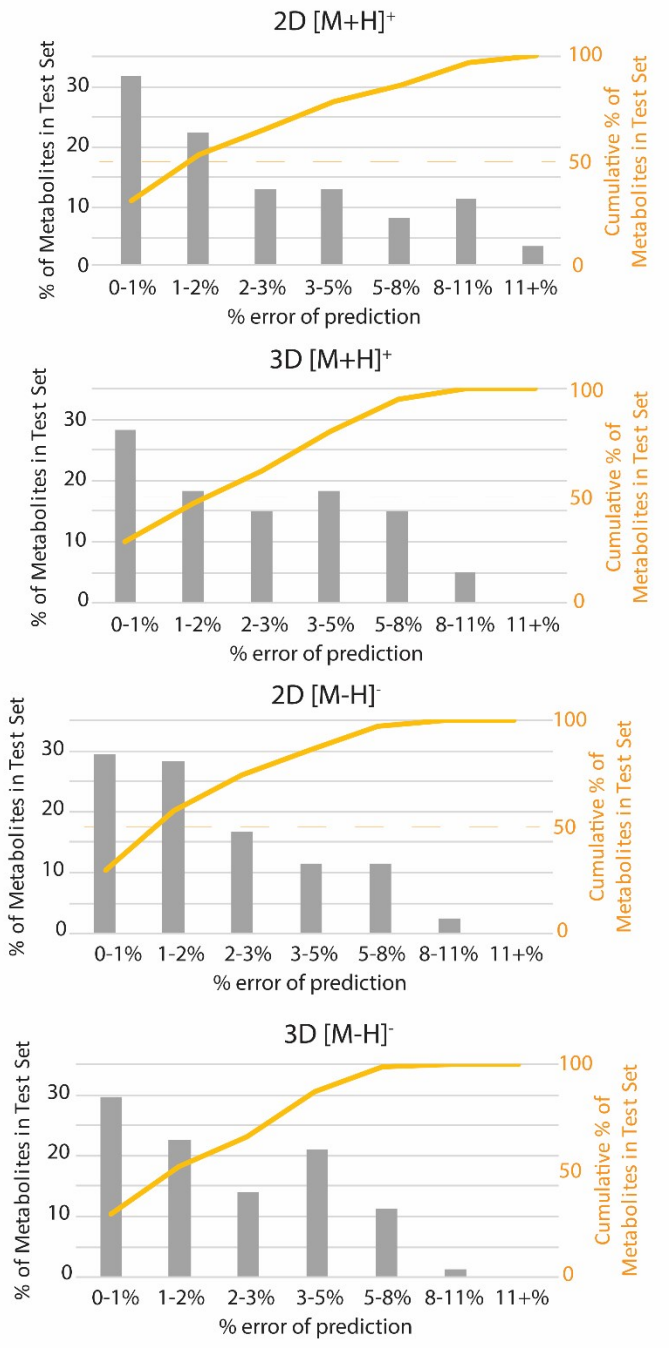

Figure S3: Breakdown of the observed accuracies for all PLS models developed to predict CCS.

Represented below are the models prior to genetic algorithm selection, along with the accompanying results of root mean square error in cross validation (RMSECV) and root mean square error in prediction (RMSECP). Units are in Å^2.

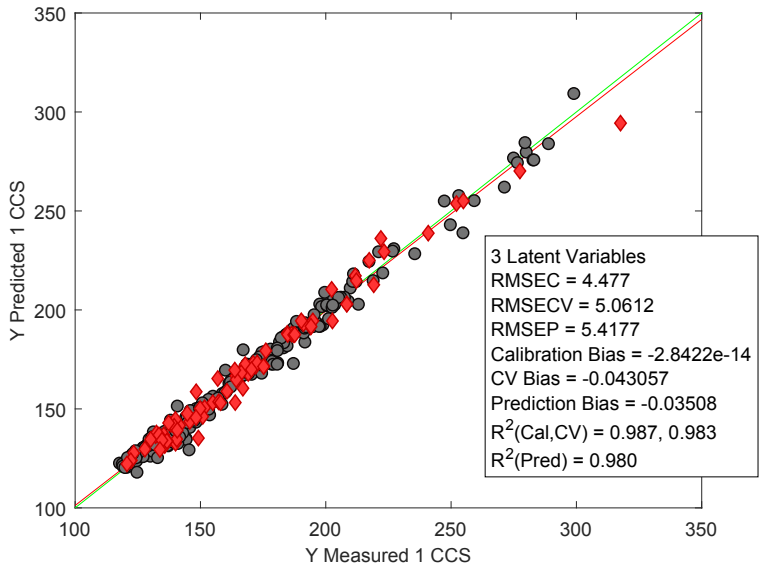

Figure S4: [M-H]$^-$ 2D model before genetic algorithm.

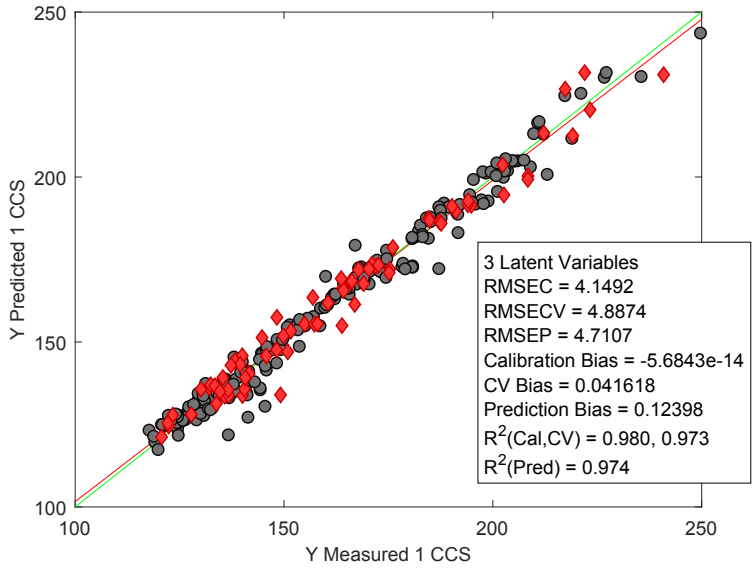

Figure S5: [M-H]$^-$ 3D model before genetic algorithm.

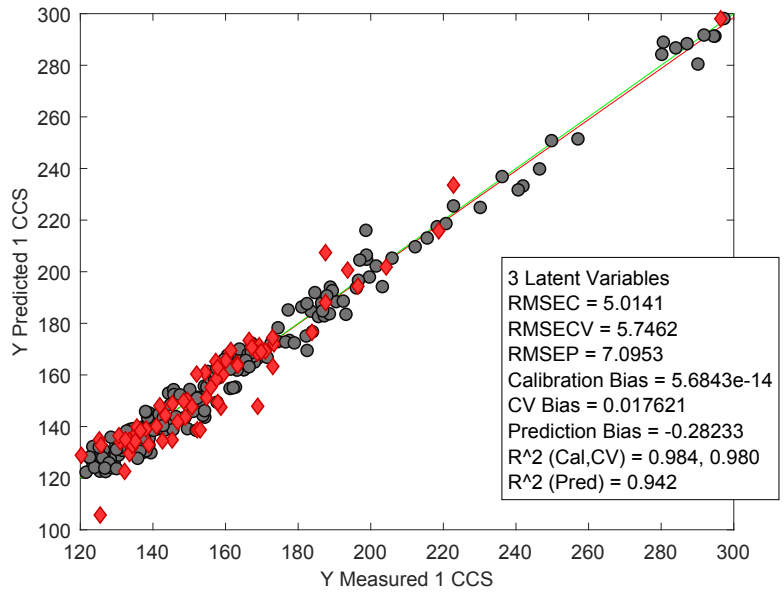

Figure S6: $[M+H]^+$ 2D model before genetic algorithm.

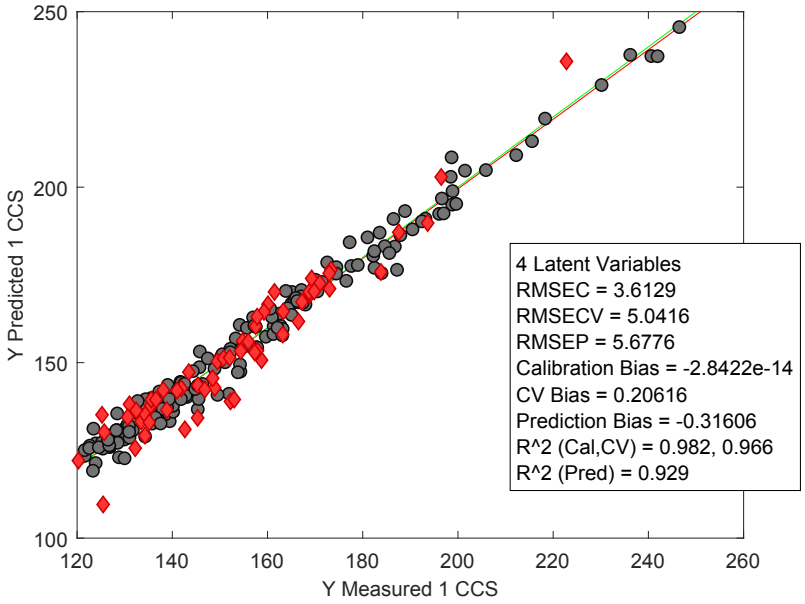

Figure S7: $[M+H]^+$ 3D model before genetic algorithm.

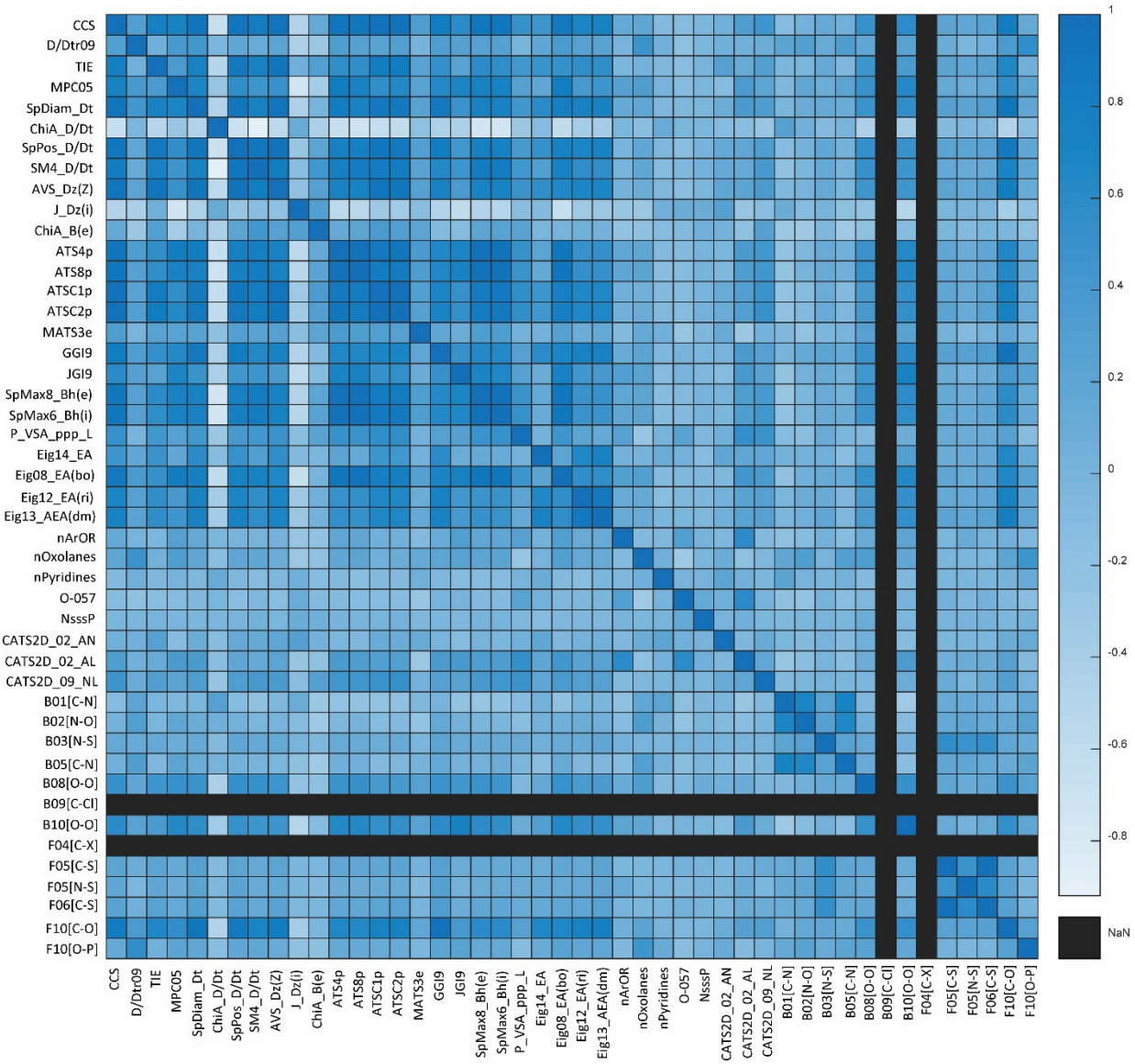

Figure S8: Correlation plot for molecular descriptors selected for use in the 2D $[M-H]^-$ model. Darker blue indicates more positively correlated variables, while more white indicates negatively correlated values.

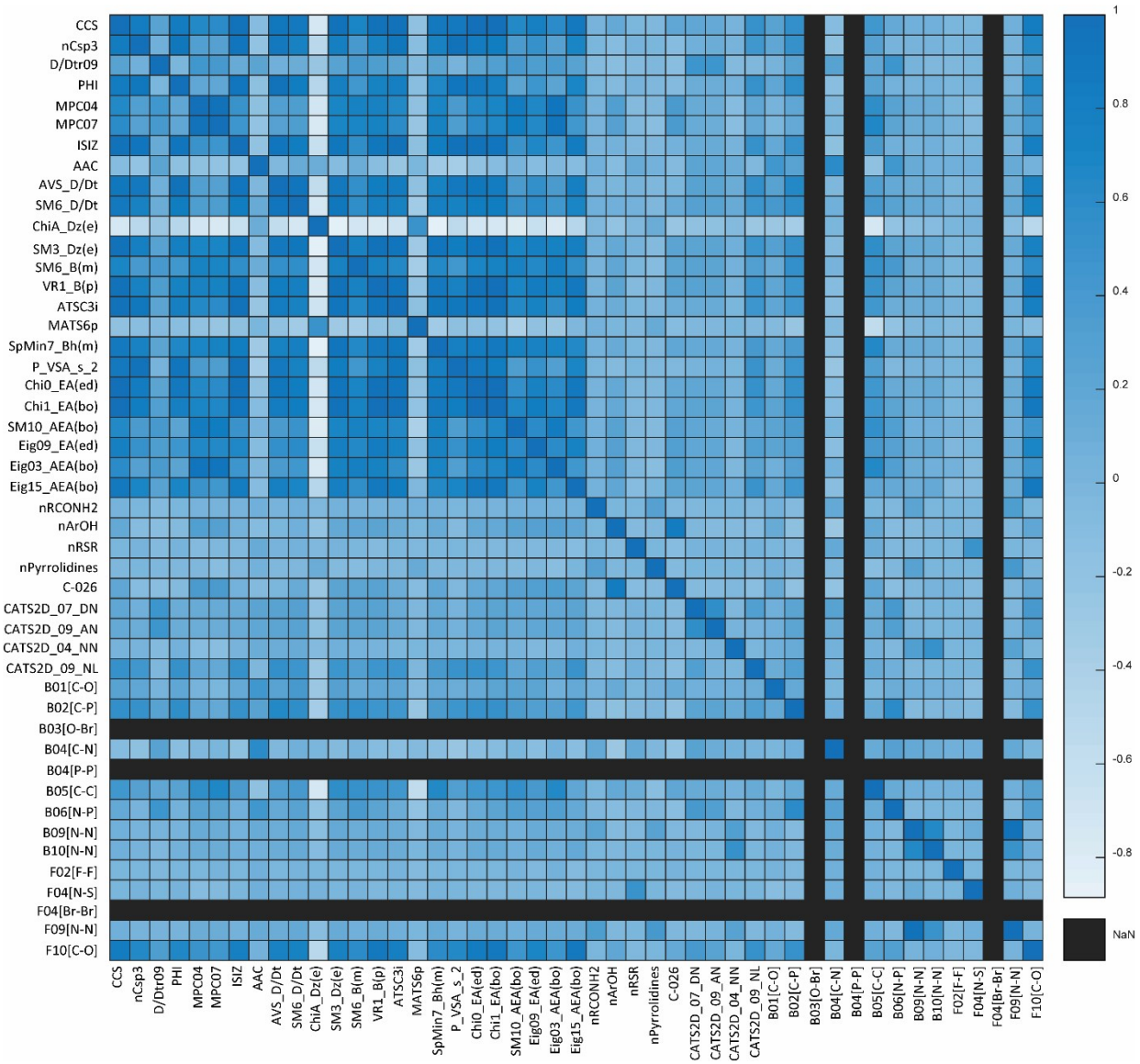

Figure S9: Correlation plot for molecular descriptors selected for use in the 2D $[M+H]^+$ model. Darker blue indicates more positively correlated variables, while more white indicates negatively correlated values.

Several molecules were predicted better by the 3D models than the 2D models. Figures S10 and S11 below show the 2D structures for molecules that were predicted with >0.5% error difference between the models, where 3D predicted more accurately than 2D. Limited features are shared between molecules. A detailed look at the Q residuals and Hotelling T^2 values did not reveal an explaination for why these specific molecules had better predictions from the 3D models.

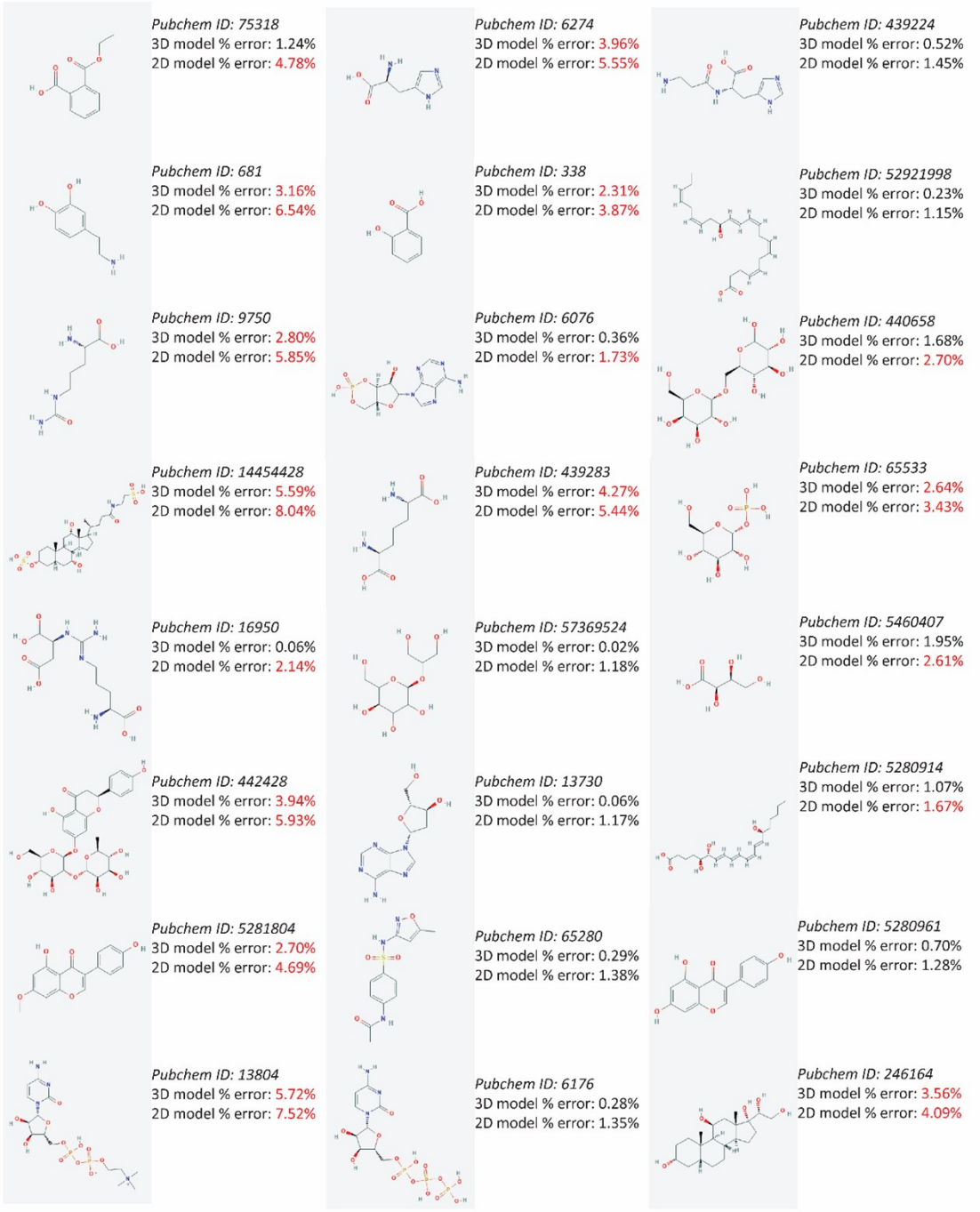

Figure S10: $[M-H]^-$ model predictions for metabolites which 3D descriptors predicted more accurately than 2D descriptors. Absolute % error in prediction from each molecule is reported. Errors in red text are above the median error for that model.

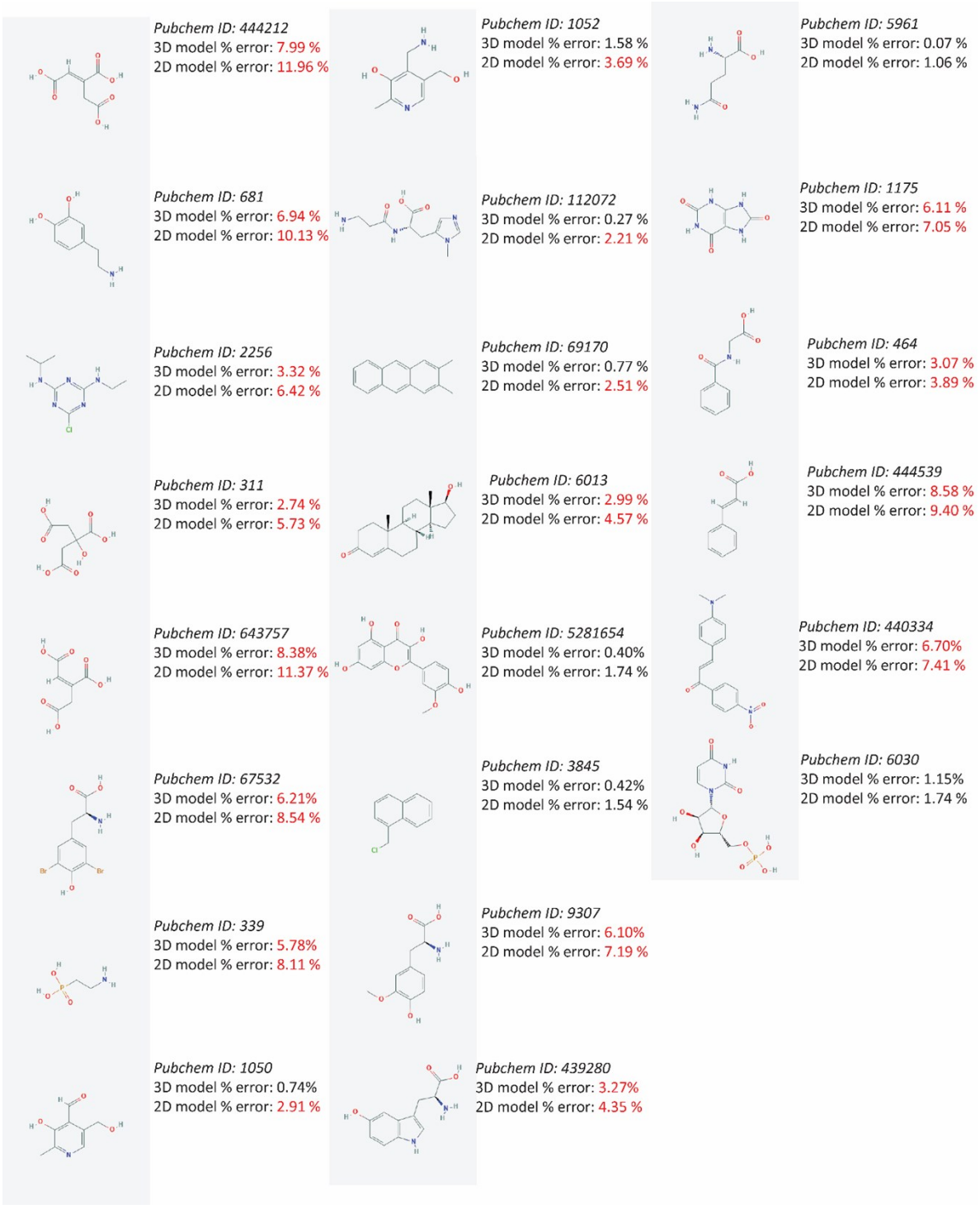

Figure S11: $[M+H]^+$ model predictions for metabolites which 3D descriptors predicted more accurately than 2D descriptors. Absolute % error in prediction from each molecule is reported. Errors in red text are above the median error for that model.

Of the 490 unique molecular descriptors, only 28 appeared in 2 models and 2 appeared in 3 models. nCsp3, the number of sp3 hybridized carbon atoms, appears in the 2D [M+H]+ and 3D [M-H]- models developed in this work and in the lipid work previously published. CATS2D_09_NL (CATS2D Negative-Lipophilic at lag 09) was selected in the 2D [M+H]+, and both 2D and 3D [M-H]- models. This suggests that a few singular molecular descriptors are not, on their own, most correlated with CCS but that it is a combination that is most important to the prediction.

Table S3. Molecular descriptors selected by genetic algorithm and used in each model presented in the current work. Additionally, a comparison to the molecular descriptors found in previous work with lipids only. Descriptors in 2 models are in green text, descriptors in 3 models are shaded yellow.

Molecular Descriptor	$[M+H]^+$ 2D	$[M+H]^+$ 3D	$[M-H]^-$ 2D	$[M-H]^-$ 3D	Lipid, Soper-Hopper *et al*, ChemComm, 2017
AAC	X				
ALOGP					x
ALOGP2					x
ASP		X			
ATS3i		X			
ATS3m		X			x
ATS4e				X	
ATS4p			X		
ATS5e		X			
ATS7v				X	
ATS8p			X		
ATSC1p			X		
ATSC2p			X		
ATSC2v		X			
ATSC3i	X				
ATSC3m		X			
ATSC5p				X	
ATSC5s					
ATSC7p				X	
ATSC7s		X			
ATSC8i				X	
AVS_D/Dt	X				
AVS_Dz(v)		X			
AVS_Dz(Z)			X		
B01[C-C]		X			
B01[C-N]			X		
B01[C-O]	X	X			
B01[C-P]				X	

B01[O-P]		X			
B02[C-P]	X				
B02[N-O]			X		
B02[N-S]		X			
B02[O-F]				X	
B02[O-O]				X	x
B03[C-F]		X			
B03[N-Cl]				X	
B03[N-O]				X	
B03[N-S]			X		
B03[O-Br]	X				
B04[C-Cl]		X		X	
B04[C-F]		X			
B04[C-N]	X				x
B04[P-P]	X				
B05[C-C]	X				
B05[C-N]			X		
B05[C-P]		X			
B05[O-P]				X	
B06[N-Cl]		X			
B06[N-O]				X	
B06[N-P]	X				
B06[O-P]		X		X	
B07[C-F]		X			
B07[O-P]				X	
B08[C-F]		X			
B08[O-O]			X		
B09[C-Cl]			X		
B09[C-P]		X			
B09[N-N]	X				
B09[N-O]		X			
B10[N-N]	X				
B10[O-O]			X		
BAC		X			x
BID				X	
C-002					x
C-003		X			
C-004		X			
C-005		X			
C-007		X			
C-012		X			

C-026	X				
C-027				X	
CATS2D_00_AN					x
CATS2D_01_LL				X	
CATS2D_02_AL			X		
CATS2D_02_AN			X		
CATS2D_03_DA				X	
CATS2D_03_PP				X	
CATS2D_04_NN	X				
CATS2D_05_DL					x
CATS2D_05_NL		X			
CATS2D_06_AP				X	
CATS2D_06_LL				X	
CATS2D_06_PN				X	
CATS2D_07_AA		X			
CATS2D_07_AP		X			
CATS2D_07_DN	X				
CATS2D_08_DN		X			
CATS2D_09_AN	X				
CATS2D_09_NL	X		X	X	
CATS2D_09_PL		X			
CATS2D_09_PP		X			
CATS3D_00_AA		X			
CATS3D_01_NN		X			
CATS3D_02_AN				X	
CATS3D_02_LL		X			
CATS3D_03_LL				X	
CATS3D_04_DN				X	
CATS3D_05_AP		X			
CATS3D_05_DL				X	
CATS3D_06_AL				X	
CATS3D_06_AN				X	
CATS3D_06_DN		X			
CATS3D_07_DA		X			
CATS3D_07_DL				X	
CATS3D_07_DN				X	
CATS3D_07_PP				X	
CATS3D_09_AL		X			
CATS3D_09_PL				X	
CATS3D_09_PN				X	
CATS3D_11_DP				X	

CATS3D_12_DD		X			
CATS3D_12_DN				X	
CATS3D_13_DL				X	
CATS3D_13_DP				X	
CATS3D_14_DA		X			
CATS3D_15_DD				X	
CATS3D_15_PL				X	
CATS3D_15_PN				X	
CATS3D_17_DP		X			
CENT					x
Chi_D/Dt					x
Chi_Dz(i)		X			
Chi_Dz(m)		X			
Chi_H2				X	
Chi_RG				X	
Chi0_EA(ed)	X				
Chi1_AEA(dm)		X			
Chi1_EA				X	
Chi1_EA(bo)	X				
ChiA_B(e)			X		
ChiA_B(i)				X	
ChiA_D/Dt			X		
ChiA_Dz(e)	X				
ChiA_G/D				X	
CSI				X	
D/Dtr07		X			
D/Dtr08		X			
D/Dtr09	X		X		
De		X			
DECC					x
DLS_04		X			
DLS_05				X	
DLS_07		X			
Dp		X			
DP11				X	
DP19				X	
EE_B(m)				X	
EE_D/Dt				X	
EE_Dz(e)		X			
Eig02_EA(ri)				X	
Eig03_AEA(bo)	X	X			

Eig03_EA(dm)					x
Eig03_EA(ri)		X			
Eig04_AEA(bo)					x
Eig04_AEA(ed)				X	
Eig04_EA(bo)					x
Eig04_EA(ed)				X	
Eig06_EA(ri)					x
Eig08_EA(bo)			X	X	
Eig08_EA(dm)					x
Eig08_EA(ed)				X	
Eig09_EA(dm)					x
Eig09_EA(ed)	X	X			
Eig09_EA(ri)				X	
Eig10_EA(ri)		X			
Eig12_AEA(dm)					x
Eig12_EA(bo)				X	
Eig12_EA(ri)			X		
Eig13_AEA(dm)			X		
Eig13_EA(bo)				X	
Eig14_AEA(bo)				X	
Eig14_AEA(ri)				X	
Eig14_EA			X		
Eig15_AEA(bo)	X				
Eta_sh_p				X	
F02[C-Br]		X			
F02[C-S]		X			
F02[F-F]	X				
F03[C-C]				X	
F03[Cl-Cl]		X			
F03[N-O]				X	
F04[Br-Br]	X				
F04[C-X]			X		
F04[N-O]				x	
F04[N-S]	X				
F05[C-N]		X		X	
F05[C-S]			X		
F05[N-S]			X		
F05[O-Cl]				X	
F05[S-P]				X	
F06[C-S]			X	X	
F06[N-Br]		X			

F06[N-O]				X	
F07[C-Cl]		X			
F08[C-F]		X			
F-081				X	
F09[N-N]	X				
F10[C-C]				X	
F10[C-O]	X		X		
F10[N-P]				X	
F10[O-P]			X	X	
G(N..P)				X	
G(S..S)		X			
G3p		X			
G3v				X	
GATS1i		X			
GATS1s				X	
GATS2e					x
GATS3e		X			
GATS5e		X			
GATS5m				X	
GATS6v		X			
GATS7v		X			
GGI10				X	
GGI4				X	
GGI9			X		
GMTI					x
H_D/Dt					x
H-051				X	
H0v				X	
HATS4e		X			
HATS6m				X	
Ho_A				X	
Ho_D					x
Ho_Dz(e)				X	
Ho_Dz(m)		X			
Ho_Dz(Z)					x
HOMT		X			
HTm		X		X	
HTp				X	
HTs		X			
HTv				X	
Hypertens-50		X			

Hypnotic-80				X	
HyWi_B(s)				X	
HyWi_Dz(e)				X	
IAC		X			
IC3					x
IDET					x
IDMT					x
ISIZ	X				x
IVDM				X	
J_D/Dt		X			
J_Dz(i)			X		
J_X		X			
JGI1		X		X	
JGI8					x
JGI9			X	X	
Km				X	
L1p		X			
L2u		X			
LLS_02		X		X	
MATS3e			X		
MATS3m		X			
MATS6i		X			
MATS6p	X				x
MATS6v		X			
MATS7m					x
MATS8p		X			
MAXDN				X	
MAXDP				X	
MCD					x
MLOGP				X	
Mor01p		X			
Mor06s		X			
Mor08i		X			
Mor09u				X	
Mor10v		X			
Mor11m				X	
Mor14s		X			
Mor15e				X	
Mor15i				X	
Mor16e				X	
Mor16i		X			

Mor21i		X			
Mor22m		X		X	
Mor22s				X	
Mor22u				X	
Mor23m		X			
Mor24s		X			
Mor24u				X	
Mor27m				X	
Mor28m				X	
Mor31s		X			
MPC04	X				
MPC05			X		
MPC07	X				
MWC04		X			
N-066					X
N-072		X			
N-074		X		X	
nAB		X			
nArCHO				X	
nArCN				X	
nArOH	X				
nArOR			X		
nAT				X	
nBM		X			
nCsp3	X			X	x
Neoplastic-50		X			
nF				X	
nHBonds				X	
nHet		X			
nN+		X			
nOHp					x
nOxiranes				X	
nOxolanes			X		
nPyrazoles				X	
nPyridines			X	X	
nPyrrolidines	X				
nR=Cp				X	
nRCONH2	X				
nRSR	X				
nS				X	
NssNH					x

NssS				X	
NsssP			X		
NtN				X	
O-057			X	X	
P_VSA_LogP_4				X	x
P_VSA_LogP_6		X			
P_VSA_m_4		X			
P_VSA_MR_4		X		X	
P_VSA_p_4					X
P_VSA_ppp_L			X		
P_VSA_s_1					X
P_VSA_s_2	X				
P_VSA_v_1		X			
PHI	X				
piPC02				X	
piPC04		X			
Psychotic-80				X	
PW2		X			
PW4					X
Qindex		X			
QW_L					X
QZZp				X	
R1i+		X			
R1s+		X			
R2e				X	
R2i+				X	
R2p		X			
R3e+				X	
R3i+				X	
R3s+				X	
R7u+				X	
RDF010e		X			
RDF010m		X			
RDF010s				X	
RDF020i				X	
RDF025s		X		X	
RDF030p		X			
RDF030v		X			
RDF035p				X	
RDF050p		X			
RDF055s				X	

RDF060i		X			
RDF070u		X			
RDF080e		X			
RDF095m		X			
RDF100p		X			
RDF105m		X			
RDF105p		X			
RDF115m		X			
RDF125u		X			
RDF145e		X			
RDF145i		X			
RDF145p		X			
Ro5		X			
RTp		X			
S0K		X			
S2K		X			
SaaCH		X			
SaaO				X	
SdsN				X	
SIC2					X
SM03_AEA(ri)				X	
SM03_EA(ed)					X
SM06_AEA(bo)		X			x
SM07_AEA(bo)					x
SM09_EA(ed)					x
SM1_Dz(p)		X			
SM10_AEA(bo)	X	X			
SM11_AEA(ri)		X			
SM12_EA(ed)					X
SM12_EA(ri)		X			
SM14_AEA(bo)				X	
SM14_EA(dm)				X	
SM15_AEA(dm)		X			
SM2_Dz(Z)		X			
SM2_RG		X			
SM2_X				X	
SM3_Dt				X	
SM3_Dz(e)	X				
SM3_Dz(v)		X			
SM3_G				X	
SM4_D/Dt			X		

SM4_G/D				X	
SM5_Dz(m)				X	
SM6_B(m)	X				
SM6_D/Dt	X				
SM6_Dz(e)				X	
SM6_L		X			
SMTI					X
SNar		X			
SpAbs_Dz(p)		X			
SpAD_EA(bo)		X			
SpDiam_Dt			X		
SpDiam_Dz(e)					X
SpDiam_EA(ri)		X			
SpMAD_AEA(bo)		X			
SpMAD_AEA(ri)				X	
SpMAD_B(i)		X			
SpMAD_B(v)		X			
SpMAD_D/Dt					X
SpMAD_Dt		X			
SpMAD_L				X	
SpMAD_X				X	
SpMax_B(m)					X
SpMax_D/Dt					X
SpMax_L					X
SpMax1_Bh(v)				X	
SpMax2_Bh(s)				X	
SpMax3_Bh(e)				X	
SpMax3_Bh(i)		X			
SpMax3_Bh(v)					X
SpMax5_Bh(s)					X
SpMax6_Bh(i)			X		
SpMax6_Bh(v)				X	
SpMax7_Bh(e)				X	
SpMax8_Bh(e)			X		
SpMax8_Bh(s)		X			
SpMaxA_A				X	
SpMaxA_Dz(v)		X			
SpMin1_Bh(e)		X			
SpMin1_Bh(i)				X	
SpMin1_Bh(m)		X			
SpMin1_Bh(p)				X	

SpMin5_Bh(i)					X
SpMin6_Bh(i)				X	
SpMin6_Bh(s)		X			
SpMin7_Bh(m)	X	X			
SpMin8_Bh(p)		x			
SpPos_D/Dt			X		
SpPos_Dz(v)		X			
SpPosA_Dz(i)				X	
SpPosLog_Dz(v)				X	
SRW02				X	
SRW03		X			x
SRW06				X	
SRW08				X	
SsCH3				X	
SsNH2					x
StN				X	
T(N..Cl)				X	
TDB02m		X			
TDB04p				X	
TDB05p		X			
TDB06e		X			
TDB07m		X			
TDB07r				X	
TDB07u				X	
TDB09e				X	
TIC0				X	
TIE			X		
VE1sign_D					x
VE1sign_D/Dt					x
VE1sign_G				X	
VE1sign_RG		X			
VE2_D					x
VE2_Dz(i)					x
VE2_Dz(v)					x
VE2sign_B(v)				X	
VR1_B(p)	x				
VR3_B(e)		X			
Wi_B(v)				X	
Wi_D/Dt					x
Wi_Dz(i)					X
Wi_Dz(p)					X

WiA_D/Dt					X
X0A				X	
X0sol				X	
X1Per		X			
X5A				X	
ZM1Kup				X	
ZM2MulPer		X			
ZM2Per		X			

Variable Importance in Projection (VIP) scores summarize the contribution a variable makes to the model, and is calculated as a weighted sum of the squared correlations between the components and the predicted variable. Scores greater than 1 can be considered more important in the given model, however these values should not be compared between columns as they are relative only to the other variables in their model.

Table S4. VIP scores for the 25 most important molecular descriptors in each model. The description for each, and the block they are a part of can be found at the end of the supplemental information.

$[M+H]^+$ 2D		$[M+H]^+$ 3D		$[M-H]^-$ 2D		$[M-H]^-$ 3D	
Molecular Descriptor	**VIP score**	**Molecular Descriptor**	**VIP score**	**Molecular Descriptor**	**VIP score**	**Molecular Descriptor**	**VIP score**
Chi1_EA(bo)	1.66	S0K	1.67	ATSC1p	1.60	Chi_RG	1.66
Chi0_EA(ed)	1.65	Ho_Dz(m)	1.66	ATSC2p	1.60	Ho_Dz(e)	1.65
ISIZ	1.63	VR3_B€	1.66	SpDiam_Dt	1.58	HTp	1.65
VR1_B(p)	1.63	Chi1_AEA(dm)	1.64	SpPos_D/Dt	1.57	HTv	1.65
SM3_Dz(e)	1.61	SM3_Dz(v)	1.63	ATS4p	1.56	Chi_H2	1.65
ATSC3i	1.61	X1Per	1.62	AVS_Dz(Z)	1.55	SM2_X	1.64
AVS_D/Dt	1.52	SM2_Dz(Z)	1.61	ATS8p	1.53	BID	1.64
P_VSA_s_2	1.46	EE_Dz€	1.60	SpMax6_Bh(i)	1.52	IVDM	1.63
Eig15_AEA(bo)	1.46	SNar	1.60	SpMax8_Bh(e)	1.51	nAT	1.63
SpMin7_Bh(m)	1.46	IAC	1.60	Eig08_EA(bo)	1.48	Wi_B(v)	1.63
nCsp3	1.45	Mor01p	1.58	SM4_D/Dt	1.43	Ho_A	1.63
SM6_D/Dt	1.44	SpPos_Dz(v)	1.58	F10[C-O]	1.42	SM4_G/D	1.63
F10[C-O]	1.44	AVS_Dz(v)	1.58	GGI9	1.37	SM3_G	1.63
PHI	1.43	RTp	1.58	TIE	1.33	X0sol	1.62
Eig09_EA(ed)	1.30	SpMin8_Bh(p)	1.56	MPC05	1.23	SM5_Dz(m)	1.62
ChiA_Dz(e)	1.27	SpAbs_Dz(p)	1.55	Eig13_AEA(dm)	1.20	HyWi_Dz(e)	1.62
SM6_B(m)	1.19	SpAD_EA(bo)	1.54	Eig12_EA(ri)	1.18	Chi1_EA	1.62
SM10_AEA(bo)	1.08	ATS3m	1.53	ChiA_D/Dt	1.13	SM6_Dz(e)	1.62
MPC04	1.07	SM2_RG	1.52	B10[O-O]	1.05	SpPosLog_Dz(v)	1.62
Eig03_AEA(bo)	1.07	ZM2MulPer	1.52	JGI9	1.03	SpMin6_Bh(i)	1.62
MPC07	1.03	SpMax8_Bh(s)	1.51	P_VSA_ppp_L	0.91	TIC0	1.61
B02[C-P]	0.90	RDF080e	1.51	B08[O-O]	0.88	SRW02	1.61
CATS2D_09_NL	0.87	ZM2Per	1.51	J_Dz(i)	0.81	ATSC8i	1.60
B05[C-C]	0.80	SpMaxA_Dz(v)	1.50	CATS2D_09_NL	0.75	ATS7v	1.59

MATS6p	0.49	SM15_AEA(dm)	1.50	Eig14_EA	0.75	ATS4e	1.58

List of selected PubChem IDs

2D $[M+H]^+$ Calibration PubChem ID				
89	5202	9131	72924	445580
190	5429	9154	73201	493570
199	5570	9700	78821	638278
469	5610	10250	89034	638678
500	5754	10690	91451	643976
588	5810	11050	91552	736715
597	5839	11317	92258	2761491
643	5962	11328	92729	5280353
673	5994	11396	96215	5280442
763	6006	11936	96373	5280805
785	6022	12599	97536	5280863
790	6047	13450	104987	5280896
846	6076	13730	105024	5281600
932	6128	13804	105085	5281647
936	6137	13945	119258	5281672
938	6249	14160	122356	5281708
970	6262	14180	145742	5281804
1014	6267	14985	156391	5282367
1044	6287	15304	160814	5283205
1054	6305	15965	165271	5284421
1066	6306	16500	171548	5364680
1110	6508	16562	193653	5379265
1131	6613	16950	261491	5381226
1135	6697	30479	439182	5497181
1150	6723	31423	439183	9546744
1174	6724	34755	439224	9547100
1188	6758	37766	439227	9547127
2153	6802	38479	439283	9548602
2181	6804	39965	439285	10100010
2519	7002	42030	439377	16219771
2537	7055	60961	439406	24778933
3120	7127	64959	439774	24779450
3220	7243	65059	439917	44251425
3614	8113	65280	440624	46891763
3715	8582	68077	440735	52927248
4276	8742	69520	442428	54675776
4687	8955	70914	444412	131879552
4794	9064	72281	444493	

2D $[M+H]^+$ Test	PubChem ID
311	9307
339	11361
464	21236
681	21387
764	37880
802	54738
931	64969
967	66535
1048	67532
1050	69170
1051	75546
1052	91763
1175	92918
1988	98792
2256	112072
2336	152323
3845	439176
4169	439213
5961	439280
6013	440334
6021	440473
6030	444212
6083	444539
6106	637511
6131	643757
6238	5280961
6274	5281654
6616	7018721
6780	11306073
9108	24778922
9117	44146693
9158	46907933
9161	

3D $[M+H]^+$ Calibration PubChem ID				
89	5570	11396	145742	54675776
134	5610	11936	156391	
190	5754	12599	160814	
199	5810	13450	165271	
247	5839	13730	171548	
469	5942	13804	193653	
500	5962	13945	261491	
588	5994	14160	439182	
597	6006	14180	439183	
643	6047	14985	439224	
673	6076	15304	439227	
763	6128	16500	439283	
764	6137	16562	439285	
785	6249	16950	439377	
790	6262	30479	439406	
846	6267	31423	439774	
932	6287	34755	439917	
936	6305	37766	440624	
938	6306	38479	440735	
970	6508	39965	442428	
1014	6613	42030	444412	
1044	6697	60961	445580	
1054	6723	64959	493570	
1066	6724	65059	638278	
1110	6758	65280	638678	
1131	6802	68077	643976	
1135	6804	69170	736715	
1136	7002	69520	2761491	
1150	7055	70914	5280353	
1174	7127	72281	5280442	
1188	7243	72924	5280805	
2153	8113	73201	5280863	
2181	8582	78821	5280896	
2519	8742	89034	5281600	
2537	8955	91451	5281647	
3120	9064	91552	5281672	
3220	9131	92258	5281708	
3614	9154	92729	5281804	
3715	9700	96215	5282367	
4276	10250	96373	5283205	
4687	10690	104987	5284421	
4794	11050	105085	5379265	
5202	11317	119258	9548602	
5429	11328	122356	16219771	

3D $[M+H]^+$ Test	PubChem ID
311	69170
339	91763
464	92918
681	98792
764	112072
802	152323
931	439176
967	439213
1048	439280
1050	440334
1051	440473
1052	444212
1175	444539
1988	637511
2256	643757
2336	5280961
3845	5281654
4169	7018721
5961	44146693
6013	
6021	
6030	
6083	
6106	
6131	
6238	
6274	
6616	
6780	
9108	
9117	
9158	
9161	
9307	
11361	
21236	
21387	
37880	
54738	
64969	
66535	
67532	

2D [M-H]- Calibration PubChem ID				
51	5958	33032	222656	5280581
89	5960	33037	439162	5280805
127	5961	38479	439182	5280896
190	5988	54738	439184	5280934
204	6036	60961	439193	5281426
311	6047	64959	439216	5281600
464	6106	64969	439242	5281647
476	6128	65246	439280	5281654
516	6137	66535	439302	5281708
588	6249	67532	439531	5282263
785	6262	70914	439559	5282768
790	6267	72281	439606	5282944
892	6287	72924	439668	5282947
932	6305	73201	439917	5283137
938	6306	78168	440015	5283154
967	6508	78821	440080	5283156
979	6749	83887	440624	5283204
985	6758	89034	440995	5283205
1044	6802	90301	441422	5283852
1050	6804	91493	444212	5283853
1052	6912	91637	444412	5284421
1054	6926	91763	444899	5379265
1066	8582	92258	445638	5381226
1081	9064	92729	446495	5459879
1110	9307	92797	447315	5460935
1175	9903	92805	459070	6441567
1188	10250	92817	493570	6971040
1198	10253	92918	638678	9546744
1202	10349	96373	643757	9547100
2153	10393	97536	643975	9547127
3120	10465	99288	643976	10855600
3496	10690	99715	644066	16219771
3806	11197	108965	736715	24778922
3893	11266	112072	2734157	35022907
4794	11333	115245	3080745	42607309
5202	12599	119258	3586141	44146693
5281	13849	119411	5174584	44251425
5312	13945	122336	5280353	52927248
5333	14985	122340	5280360	52940142
5430	15047	124886	5280373	54670067
5789	15965	161276	5280378	54675776
5816	16562	171548	5280442	

2D [M-H]- Test PubChem ID					
5839	18396	192826	5280450	
5957	31401	221493	5280492	

338	439227
469	439283
643	439524
681	439538
1135	439750
3220	439774
3845	440658
4276	440993
6006	442428
6029	445580
6057	445639
6076	451489
6176	638278
6274	3035456
6322	5280733
6613	5280863
7091	5280886
8742	5280914
9700	5280961
9750	5281804
10467	5283155
11361	5283203
13712	5287993
13730	5460407
13804	5497181
16500	7018721
16950	11046191
18189	13489094
21236	14454428
65059	20056755
65280	52921998
65533	52927160
75318	53477669
77982	57369524
86398	71308662
92824	92132303
122357	
151838	
246164	
439176	
439183	
439224	

3D [M-H]- Calibration			PubChem ID	
51	5958	33032	439182	5281708
89	5960	33037	439184	5282263
127	5961	38479	439193	5282944
190	5988	54738	439216	5282947
204	6036	60961	439242	5283137
311	6047	64959	439280	5283154
464	6106	64969	439531	5283156
476	6128	65246	439559	5283204
516	6137	66535	439668	5283205
588	6249	67532	439917	5283852
785	6262	70914	440015	5283853
790	6267	72281	440080	5284421
892	6287	72924	440624	5379265
932	6288	73201	440995	5459879
938	6305	78168	441422	5460935
967	6306	78821	444212	6441567
979	6508	83887	444412	6971040
985	6749	89034	444899	10855600
1044	6758	90301	445638	16219771
1050	6802	91493	446495	35022907
1052	6804	91637	447315	42607309
1054	6912	91763	459070	44146693
1066	6926	92258	493570	54670067
1081	8582	92729	638678	54675776
1110	9064	92797	643757	
1175	9307	92805	643976	
1188	9903	92817	736715	
1198	10250	92918	2734157	
1202	10253	96373	3080745	
2153	10349	99288	5280353	
3120	10393	99715	5280360	
3496	10465	112072	5280373	
3806	10690	115245	5280378	
3893	11266	119258	5280442	
4794	11333	119411	5280450	
5202	12599	122336	5280492	
5312	13849	122340	5280581	
5333	13945	124886	5280805	
5430	14985	161276	5280896	
5789	15047	171548	5280934	
5816	15965	192826	5281426	
5839	16562	221493	5281600	
5951	18396	222656	5281647	
5957	31401	439162	5281654	

3D [M-H]$^{-}$ Test	PubChem ID
338	439524
469	439538
643	439750
681	439774
1135	440658
3220	440993
3845	442428
4276	445580
6006	445639
6029	451489
6057	638278
6076	3035456
6176	5280733
6274	5280863
6322	5280886
6613	5280914
7091	5280961
8742	5281804
9700	5283155
9750	5283203
11361	5287993
13712	5460407
13730	7018721
13804	11046191
16500	14454428
16950	20056755
18189	52921998
21236	53477669
65059	57369524
65280	
65533	
75318	
77982	
86398	
92824	
122357	
246164	
439176	
439183	
439224	
439227	
439283	

Abbreviations for molecular descriptors, their description and block as calculated by Dragon 7.0 from Kode Chemoinformatics.

This is the subset of molecular descriptors chosen from genetic algorithm variable selection. A full list of all molecular descriptors that can be calculated via Dragon 7.0 can be found at https://chm.kode-solutions.net/products_dragon_descriptors.php

Name	Description	Block
nAT	number of atoms	Constitutional indices
nBM	number of multiple bonds	Constitutional indices
nAB	number of aromatic bonds	Constitutional indices
nS	number of Sulfur atoms	Constitutional indices
nF	number of Fluorine atoms	Constitutional indices
nHet	number of heteroatoms	Constitutional indices
nCsp3	number of sp3 hybridized Carbon atoms	Constitutional indices
MCD	molecular cyclized degree	Ring descriptors
D/Dtr07	distance/detour ring index of order 7	Ring descriptors
D/Dtr08	distance/detour ring index of order 8	Ring descriptors
D/Dtr09	distance/detour ring index of order 9	Ring descriptors
ZM1Kup	first Zagreb index by Kupchik vertex degrees	Topological indices
ZM2Per	second Zagreb index by perturbation vertex degrees	Topological indices
ZM2MulPer	second Zagreb index by multiplicative perturbation vertex degrees	Topological indices
Qindex	quadratic index	Topological indices
SNar	Narumi simple topological index (log function)	Topological indices
DECC	eccentric	Topological indices
CENT	centralization	Topological indices
SMTI	Schultz Molecular Topological Index (MTI)	Topological indices

Name	Description	Block
GMTI	Gutman Molecular Topological Index	Topological indices
CSI	eccentric connectivity index	Topological indices
S2K	2-path Kier alpha-modified shape index	Topological indices
PHI	Kier flexibility index	Topological indices
PW2	path/walk 2 - Randic shape index	Topological indices
PW4	path/walk 4 - Randic shape index	Topological indices
MAXDN	maximal electrotopological negative variation	Topological indices
MAXDP	maximal electrotopological positive variation	Topological indices
TIE	E-state topological parameter	Topological indices
BAC	Balaban centric index	Topological indices
MWC04	molecular walk count of order 4	Walk and path counts
SRW02	self-returning walk count of order 2	Walk and path counts
SRW03	self-returning walk count of order 3	Walk and path counts
SRW06	self-returning walk count of order 6	Walk and path counts
SRW08	self-returning walk count of order 8	Walk and path counts
MPC04	molecular path count of order 4	Walk and path counts
MPC05	molecular path count of order 5	Walk and path counts
MPC07	molecular path count of order 7	Walk and path counts
piPC02	molecular multiple path count of order 2	Walk and path counts
piPC04	molecular multiple path count of order 4	Walk and path counts
BID	Balaban ID number	Walk and path counts
X0A	average connectivity index of order 0	Connectivity indices

Name	Description	Block
X5A	average connectivity index of order 5	Connectivity indices
X0sol	solvation connectivity index of order 0	Connectivity indices
X1Per	perturbation connectivity index	Connectivity indices
ISIZ	information index on molecular size	Information indices
IAC	total information index on atomic composition	Information indices
AAC	mean information index on atomic composition	Information indices
IDET	total information content on the distance equality	Information indices
IDMT	total information content on the distance magnitude	Information indices
IVDM	mean information content on the vertex degree magnitude	Information indices
S0K	Kier symmetry index	Information indices
IC3	Information Content index (neighborhood symmetry of 3-order)	Information indices
TIC0	Total Information Content index (neighborhood symmetry of 0-order)	Information indices
SIC2	Structural Information Content index (neighborhood symmetry of 2-order)	Information indices
SpMaxA_A	normalized leading eigenvalue from adjacency matrix	2D matrix-based descriptors
Ho_A	Hosoya-like index (log function) from adjacency matrix	2D matrix-based descriptors
Ho_D	Hosoya-like index (log function) from topological distance matrix	2D matrix-based descriptors
VE2_D	average coefficient of the last eigenvector (absolute values) from topological distance matrix	2D matrix-based descriptors
VE1sign_D	coefficient sum of the last eigenvector from topological distance matrix	2D matrix-based descriptors
QW_L	quasi-Wiener index (Kirchhoff number) from Laplace matrix	2D matrix-based

Name	Description	Block
		descriptors
SpMax_L	leading eigenvalue from Laplace matrix	2D matrix-based descriptors
SpMAD_L	spectral mean absolute deviation from Laplace matrix	2D matrix-based descriptors
SM6_L	spectral moment of order 6 from Laplace matrix	2D matrix-based descriptors
J_X	Balaban-like index from chi matrix	2D matrix-based descriptors
SpMAD_X	spectral mean absolute deviation from chi matrix	2D matrix-based descriptors
SM2_X	spectral moment of order 2 from chi matrix	2D matrix-based descriptors
Chi_H2	Randic-like index from reciprocal squared distance matrix	2D matrix-based descriptors
SpDiam_Dt	spectral diameter from detour matrix	2D matrix-based descriptors
SpMAD_Dt	spectral mean absolute deviation from detour matrix	2D matrix-based descriptors
SM3_Dt	spectral moment of order 3 from detour matrix	2D matrix-based descriptors
Wi_D/Dt	Wiener-like index from distance/detour matrix	2D matrix-based descriptors
WiA_D/Dt	average Wiener-like index from distance/detour matrix	2D matrix-based descriptors
AVS_D/Dt	average vertex sum from distance/detour matrix	2D matrix-based descriptors

Name	Description	Block
H_D/Dt	Harary-like index from distance/detour matrix	2D matrix-based descriptors
Chi_D/Dt	Randic-like index from distance/detour matrix	2D matrix-based descriptors
ChiA_D/Dt	average Randic-like index from distance/detour matrix	2D matrix-based descriptors
J_D/Dt	Balaban-like index from distance/detour matrix	2D matrix-based descriptors
SpPos_D/Dt	spectral positive sum from distance/detour matrix	2D matrix-based descriptors
SpMax_D/Dt	leading eigenvalue from distance/detour matrix	2D matrix-based descriptors
SpMAD_D/Dt	spectral mean absolute deviation from distance/detour matrix	2D matrix-based descriptors
EE_D/Dt	Estrada-like index (log function) from distance/detour matrix	2D matrix-based descriptors
SM4_D/Dt	spectral moment of order 4 from distance/detour matrix	2D matrix-based descriptors
SM6_D/Dt	spectral moment of order 6 from distance/detour matrix	2D matrix-based descriptors
VE1sign_D/Dt	coefficient sum of the last eigenvector from distance/detour matrix	2D matrix-based descriptors
Ho_Dz(Z)	Hosoya-like index (log function) from Barysz matrix weighted by atomic number	2D matrix-based descriptors
SM2_Dz(Z)	spectral moment of order 2 from Barysz matrix weighted by atomic number	2D matrix-based descriptors
Chi_Dz(m)	Randic-like index from Barysz matrix weighted by mass	2D matrix-based

Name	Description	Block
		descriptors
Ho_Dz(m)	Hosoya-like index (log function) from Barysz matrix weighted by mass	2D matrix-based descriptors
SM5_Dz(m)	spectral moment of order 5 from Barysz matrix weighted by mass	2D matrix-based descriptors
AVS_Dz(v)	average vertex sum from Barysz matrix weighted by van der Waals volume	2D matrix-based descriptors
SpPos_Dz(v)	spectral positive sum from Barysz matrix weighted by van der Waals volume	2D matrix-based descriptors
SpPosLog_Dz(v)	logarithmic spectral positive sum from Barysz matrix weighted by van der Waals volume	2D matrix-based descriptors
SpMaxA_Dz(v)	normalized leading eigenvalue from Barysz matrix weighted by van der Waals volume	2D matrix-based descriptors
SM3_Dz(v)	spectral moment of order 3 from Barysz matrix weighted by van der Waals volume	2D matrix-based descriptors
VE2_Dz(v)	average coefficient of the last eigenvector (absolute values) from Barysz matrix weighted by van der Waals volume	2D matrix-based descriptors
ChiA_Dz(e)	average Randic-like index from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
HyWi_Dz(e)	hyper-Wiener-like index (log function) from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
SpDiam_Dz(e)	spectral diameter from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
Ho_Dz(e)	Hosoya-like index (log function) from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
EE_Dz(e)	Estrada-like index (log function) from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors

Name	Description	Block
SM3_Dz(e)	spectral moment of order 3 from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
SM6_Dz(e)	spectral moment of order 6 from Barysz matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
Wi_Dz(p)	Wiener-like index from Barysz matrix weighted by polarizability	2D matrix-based descriptors
SpAbs_Dz(p)	graph energy from Barysz matrix weighted by polarizability	2D matrix-based descriptors
SM1_Dz(p)	spectral moment of order 1 from Barysz matrix weighted by polarizability	2D matrix-based descriptors
Wi_Dz(i)	Wiener-like index from Barysz matrix weighted by ionization potential	2D matrix-based descriptors
Chi_Dz(i)	Randic-like index from Barysz matrix weighted by ionization potential	2D matrix-based descriptors
J_Dz(i)	Balaban-like index from Barysz matrix weighted by ionization potential	2D matrix-based descriptors
SpPosA_Dz(i)	normalized spectral positive sum from Barysz matrix weighted by ionization potential	2D matrix-based descriptors
VE2_Dz(i)	average coefficient of the last eigenvector (absolute values) from Barysz matrix weighted by ionization potential	2D matrix-based descriptors
SpMax_B(m)	leading eigenvalue from Burden matrix weighted by mass	2D matrix-based descriptors
EE_B(m)	Estrada-like index (log function) from Burden matrix weighted by mass	2D matrix-based descriptors
SM6_B(m)	spectral moment of order 6 from Burden matrix weighted by mass	2D matrix-based descriptors
Wi_B(v)	Wiener-like index from Burden matrix weighted by van der Waals volume	2D matrix-based

Name	Description	Block
		descriptors
SpMAD_B(v)	spectral mean absolute deviation from Burden matrix weighted by van der Waals volume	2D matrix-based descriptors
VE2sign_B(v)	average coefficient of the last eigenvector from Burden matrix weighted by van der Waals volume	2D matrix-based descriptors
ChiA_B(e)	average Randic-like index from Burden matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
VR3_B(e)	logarithmic Randic-like eigenvector-based index from Burden matrix weighted by Sanderson electronegativity	2D matrix-based descriptors
VR1_B(p)	Randic-like eigenvector-based index from Burden matrix weighted by polarizability	2D matrix-based descriptors
ChiA_B(i)	average Randic-like index from Burden matrix weighted by ionization potential	2D matrix-based descriptors
SpMAD_B(i)	spectral mean absolute deviation from Burden matrix weighted by ionization potential	2D matrix-based descriptors
HyWi_B(s)	hyper-Wiener-like index (log function) from Burden matrix weighted by I-State	2D matrix-based descriptors
ATS3m	Broto-Moreau autocorrelation of lag 3 (log function) weighted by mass	2D autocorrelations
ATS7v	Broto-Moreau autocorrelation of lag 7 (log function) weighted by van der Waals volume	2D autocorrelations
ATS4e	Broto-Moreau autocorrelation of lag 4 (log function) weighted by Sanderson electronegativity	2D autocorrelations
ATS5e	Broto-Moreau autocorrelation of lag 5 (log function) weighted by Sanderson electronegativity	2D autocorrelations
ATS4p	Broto-Moreau autocorrelation of lag 4 (log function) weighted by polarizability	2D autocorrelations
ATS8p	Broto-Moreau autocorrelation of lag 8 (log function) weighted by polarizability	2D autocorrelations
ATS3i	Broto-Moreau autocorrelation of lag 3 (log function) weighted by ionization potential	2D autocorrelations

Name	Description	Block
ATSC3m	Centred Broto-Moreau autocorrelation of lag 3 weighted by mass	2D autocorrelations
ATSC2v	Centred Broto-Moreau autocorrelation of lag 2 weighted by van der Waals volume	2D autocorrelations
ATSC1p	Centred Broto-Moreau autocorrelation of lag 1 weighted by polarizability	2D autocorrelations
ATSC2p	Centred Broto-Moreau autocorrelation of lag 2 weighted by polarizability	2D autocorrelations
ATSC5p	Centred Broto-Moreau autocorrelation of lag 5 weighted by polarizability	2D autocorrelations
ATSC7p	Centred Broto-Moreau autocorrelation of lag 7 weighted by polarizability	2D autocorrelations
ATSC3i	Centred Broto-Moreau autocorrelation of lag 3 weighted by ionization potential	2D autocorrelations
ATSC8i	Centred Broto-Moreau autocorrelation of lag 8 weighted by ionization potential	2D autocorrelations
ATSC5s	Centred Broto-Moreau autocorrelation of lag 5 weighted by I-state	2D autocorrelations
ATSC7s	Centred Broto-Moreau autocorrelation of lag 7 weighted by I-state	2D autocorrelations
MATS3m	Moran autocorrelation of lag 3 weighted by mass	2D autocorrelations
MATS7m	Moran autocorrelation of lag 7 weighted by mass	2D autocorrelations
MATS6v	Moran autocorrelation of lag 6 weighted by van der Waals volume	2D autocorrelations
MATS3e	Moran autocorrelation of lag 3 weighted by Sanderson electronegativity	2D autocorrelations
MATS6p	Moran autocorrelation of lag 6 weighted by polarizability	2D autocorrelations
MATS8p	Moran autocorrelation of lag 8 weighted by polarizability	2D autocorrelations
MATS6i	Moran autocorrelation of lag 6 weighted by ionization potential	2D autocorrelations
GATS5m	Geary autocorrelation of lag 5 weighted by mass	2D autocorrelations
GATS6v	Geary autocorrelation of lag 6 weighted by van der Waals volume	2D autocorrelations
GATS7v	Geary autocorrelation of lag 7 weighted by van der Waals volume	2D autocorrelations
GATS2e	Geary autocorrelation of lag 2 weighted by Sanderson electronegativity	2D autocorrelations
GATS3e	Geary autocorrelation of lag 3 weighted by Sanderson electronegativity	2D autocorrelations

Name	Description	Block
GATS5e	Geary autocorrelation of lag 5 weighted by Sanderson electronegativity	2D autocorrelations
GATS1i	Geary autocorrelation of lag 1 weighted by ionization potential	2D autocorrelations
GATS1s	Geary autocorrelation of lag 1 weighted by I-state	2D autocorrelations
GGI4	topological charge index of order 4	2D autocorrelations
GGI9	topological charge index of order 9	2D autocorrelations
GGI10	topological charge index of order 10	2D autocorrelations
JGI1	mean topological charge index of order 1	2D autocorrelations
JGI8	mean topological charge index of order 8	2D autocorrelations
JGI9	mean topological charge index of order 9	2D autocorrelations
SpMax1_Bh(v)	largest eigenvalue n. 1 of Burden matrix weighted by van der Waals volume	Burden eigenvalues
SpMax3_Bh(v)	largest eigenvalue n. 3 of Burden matrix weighted by van der Waals volume	Burden eigenvalues
SpMax6_Bh(v)	largest eigenvalue n. 6 of Burden matrix weighted by van der Waals volume	Burden eigenvalues
SpMax3_Bh(e)	largest eigenvalue n. 3 of Burden matrix weighted by Sanderson electronegativity	Burden eigenvalues
SpMax7_Bh(e)	largest eigenvalue n. 7 of Burden matrix weighted by Sanderson electronegativity	Burden eigenvalues
SpMax8_Bh(e)	largest eigenvalue n. 8 of Burden matrix weighted by Sanderson electronegativity	Burden eigenvalues
SpMax2_Bh(s)	largest eigenvalue n. 2 of Burden matrix weighted by I-state	Burden eigenvalues
SpMax5_Bh(s)	largest eigenvalue n. 5 of Burden matrix weighted by I-state	Burden eigenvalues
SpMax8_Bh(s)	largest eigenvalue n. 8 of Burden matrix weighted by I-state	Burden eigenvalues
SpMin1_Bh(m)	smallest eigenvalue n. 1 of Burden matrix weighted by mass	Burden eigenvalues
SpMin7_Bh(m)	smallest eigenvalue n. 7 of Burden matrix weighted by mass	Burden eigenvalues
SpMin1_Bh(e)	smallest eigenvalue n. 1 of Burden matrix weighted by Sanderson electronegativity	Burden eigenvalues
SpMin1_Bh(p)	smallest eigenvalue n. 1 of Burden matrix weighted by polarizability	Burden eigenvalues

Name	Description	Block
SpMin8_Bh(p)	smallest eigenvalue n. 8 of Burden matrix weighted by polarizability	Burden eigenvalues
SpMin1_Bh(i)	smallest eigenvalue n. 1 of Burden matrix weighted by ionization potential	Burden eigenvalues
SpMin5_Bh(i)	smallest eigenvalue n. 5 of Burden matrix weighted by ionization potential	Burden eigenvalues
SpMin6_Bh(i)	smallest eigenvalue n. 6 of Burden matrix weighted by ionization potential	Burden eigenvalues
SpMin6_Bh(s)	smallest eigenvalue n. 6 of Burden matrix weighted by I-state	Burden eigenvalues
P_VSA_LogP_4	P_VSA-like on LogP, bin 4	P_VSA-like descriptor
P_VSA_LogP_6	P_VSA-like on LogP, bin 6	P_VSA-like descriptor
P_VSA_MR_4	P_VSA-like on Molar Refractivity, bin 4	P_VSA-like descriptor
P_VSA_m_4	P_VSA-like on mass, bin 4	P_VSA-like descriptor
P_VSA_v_1	P_VSA-like on van der Waals volume, bin 1	P_VSA-like descriptor
P_VSA_p_4	P_VSA-like on polarizability, bin 4	P_VSA-like descriptor
P_VSA_s_1	P_VSA-like on I-state, bin 1	P_VSA-like descriptor
P_VSA_s_2	P_VSA-like on I-state, bin 2	P_VSA-like descriptor
P_VSA_ppp_L	P_VSA-like on potential pharmacophore points, L - lipophilic	P_VSA-like descriptor
Eta_sh_p	eta p shape index	ETA indices
SpAD_EA(bo)	spectral absolute deviation from edge adjacency mat. weighted by bond order	Edge adjacency indices
SpDiam_EA(ri)	spectral diameter from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
SpMAD_AEA(bo)	spectral mean absolute deviation from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
SpMAD_AEA(ri)	spectral mean absolute deviation from augmented edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Chi1_EA	connectivity-like index of order 1 from edge adjacency mat.	Edge adjacency indices

Name	Description	Block
Chi0_EA(ed)	connectivity-like index of order 0 from edge adjacency mat. weighted by edge degree	Edge adjacency indices
Chi0_EA(bo)	connectivity-like index of order 0 from edge adjacency mat. weighted by bond order	Edge adjacency indices
Chi1_AEA(dm)	connectivity-like index of order 1 from augmented edge adjacency mat. weighted by dipole moment	Edge adjacency indices
SM03_EA(ed)	spectral moment of order 3 from edge adjacency mat. weighted by edge degree	Edge adjacency indices
SM09_EA(ed)	spectral moment of order 9 from edge adjacency mat. weighted by edge degree	Edge adjacency indices
SM12_EA(ed)	spectral moment of order 12 from edge adjacency mat. weighted by edge degree	Edge adjacency indices
SM14_EA(dm)	spectral moment of order 14 from edge adjacency mat. weighted by dipole moment	Edge adjacency indices
SM12_EA(ri)	spectral moment of order 12 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
SM06_AEA(bo)	spectral moment of order 6 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
SM07_AEA(bo)	spectral moment of order 7 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
SM10_AEA(bo)	spectral moment of order 10 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
SM14_AEA(bo)	spectral moment of order 14 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
SM15_AEA(dm)	spectral moment of order 15 from augmented edge adjacency mat. weighted by dipole moment	Edge adjacency indices
SM03_AEA(ri)	spectral moment of order 3 from augmented edge adjacency mat. weighted by resonance integral	Edge adjacency indices
SM11_AEA(ri)	spectral moment of order 11 from augmented edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig14_EA	eigenvalue n. 14 from edge adjacency mat.	Edge adjacency indices
Eig04_EA(ed)	eigenvalue n. 4 from edge adjacency mat. weighted by edge degree	Edge adjacency indices
Eig08_EA(ed)	eigenvalue n. 8 from edge adjacency mat. weighted by edge degree	Edge adjacency indices

Name	Description	Block
Eig09_EA(ed)	eigenvalue n. 9 from edge adjacency mat. weighted by edge degree	Edge adjacency indices
Eig04_EA(bo)	eigenvalue n. 4 from edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig08_EA(bo)	eigenvalue n. 8 from edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig12_EA(bo)	eigenvalue n. 12 from edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig13_EA(bo)	eigenvalue n. 13 from edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig03_EA(dm)	eigenvalue n. 3 from edge adjacency mat. weighted by dipole moment	Edge adjacency indices
Eig08_EA(dm)	eigenvalue n. 8 from edge adjacency mat. weighted by dipole moment	Edge adjacency indices
Eig09_EA(dm)	eigenvalue n. 9 from edge adjacency mat. weighted by dipole moment	Edge adjacency indices
Eig02_EA(ri)	eigenvalue n. 2 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig03_EA(ri)	eigenvalue n. 3 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig06_EA(ri)	eigenvalue n. 6 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig09_EA(ri)	eigenvalue n. 9 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig10_EA(ri)	eigenvalue n. 10 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig12_EA(ri)	eigenvalue n. 12 from edge adjacency mat. weighted by resonance integral	Edge adjacency indices
Eig04_AEA(ed)	eigenvalue n. 4 from augmented edge adjacency mat. weighted by edge degree	Edge adjacency indices
Eig03_AEA(bo)	eigenvalue n. 3 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig04_AEA(bo)	eigenvalue n. 4 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig14_AEA(bo)	eigenvalue n. 14 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig15_AEA(bo)	eigenvalue n. 15 from augmented edge adjacency mat. weighted by bond order	Edge adjacency indices
Eig12_AEA(dm)	eigenvalue n. 12 from augmented edge adjacency mat. weighted by dipole moment	Edge adjacency indices
Eig13_AEA(dm)	eigenvalue n. 13 from augmented edge adjacency mat. weighted by dipole moment	Edge adjacency indices
Eig14_AEA(ri)	eigenvalue n. 14 from augmented edge adjacency mat. weighted by resonance integral	Edge adjacency indices

Name	Description	Block
ASP	asphericity	Geometrical descriptors
HOMT	HOMA total	Geometrical descriptors
QZZp	quadrupole z-component value / weighted by polarizability	Geometrical descriptors
SM3_G	spectral moment of order 3 from geometrical matrix	3D matrix-based descriptors
VE1sign_G	coefficient sum of the last eigenvector from geometrical matrix	3D matrix-based descriptors
Chi_RG	Randic-like index from reciprocal squared geometrical matrix	3D matrix-based descriptors
SM2_RG	spectral moment of order 2 from reciprocal squared geometrical matrix	3D matrix-based descriptors
VE1sign_RG	coefficient sum of the last eigenvector from reciprocal squared geometrical matrix	3D matrix-based descriptors
ChiA_G/D	average Randic-like index from distance/distance matrix	3D matrix-based descriptors
SM4_G/D	spectral moment of order 4 from distance/distance matrix	3D matrix-based descriptors
TDB07u	3D Topological distance based descriptors - lag 7 unweighted	3D autocorrelations
TDB02m	3D Topological distance based descriptors - lag 2 weighted by mass	3D autocorrelations
TDB07m	3D Topological distance based descriptors - lag 7 weighted by mass	3D autocorrelations
TDB06e	3D Topological distance based descriptors - lag 6 weighted by Sanderson electronegativity	3D autocorrelations
TDB09e	3D Topological distance based descriptors - lag 9 weighted by Sanderson electronegativity	3D autocorrelations
TDB04p	3D Topological distance based descriptors - lag 4 weighted by polarizability	3D autocorrelations
TDB05p	3D Topological distance based descriptors - lag 5 weighted by polarizability	3D autocorrelations

Name	Description	Block
TDB07r	3D Topological distance based descriptors - lag 7 weighted by covalent radius	3D autocorrelations
RDF070u	Radial Distribution Function - 070 / unweighted	RDF descriptors
RDF125u	Radial Distribution Function - 125 / unweighted	RDF descriptors
RDF010m	Radial Distribution Function - 010 / weighted by mass	RDF descriptors
RDF095m	Radial Distribution Function - 095 / weighted by mass	RDF descriptors
RDF105m	Radial Distribution Function - 105 / weighted by mass	RDF descriptors
RDF115m	Radial Distribution Function - 115 / weighted by mass	RDF descriptors
RDF030v	Radial Distribution Function - 030 / weighted by van der Waals volume	RDF descriptors
RDF010e	Radial Distribution Function - 010 / weighted by Sanderson electronegativity	RDF descriptors
RDF080e	Radial Distribution Function - 080 / weighted by Sanderson electronegativity	RDF descriptors
RDF145e	Radial Distribution Function - 145 / weighted by Sanderson electronegativity	RDF descriptors
RDF030p	Radial Distribution Function - 030 / weighted by polarizability	RDF descriptors
RDF035p	Radial Distribution Function - 035 / weighted by polarizability	RDF descriptors
RDF050p	Radial Distribution Function - 050 / weighted by polarizability	RDF descriptors
RDF100p	Radial Distribution Function - 100 / weighted by polarizability	RDF descriptors
RDF105p	Radial Distribution Function - 105 / weighted by polarizability	RDF descriptors
RDF145p	Radial Distribution Function - 145 / weighted by polarizability	RDF descriptors
RDF020i	Radial Distribution Function - 020 / weighted by ionization potential	RDF descriptors
RDF060i	Radial Distribution Function - 060 / weighted by ionization potential	RDF descriptors
RDF145i	Radial Distribution Function - 145 / weighted by ionization potential	RDF descriptors
RDF010s	Radial Distribution Function - 010 / weighted by I-state	RDF descriptors
RDF025s	Radial Distribution Function - 025 / weighted by I-state	RDF descriptors

Name	Description	Block
RDF055s	Radial Distribution Function - 055 / weighted by I-state	RDF descriptors
Mor09u	signal 09 / unweighted	3D-MoRSE descriptors
Mor22u	signal 22 / unweighted	3D-MoRSE descriptors
Mor24u	signal 24 / unweighted	3D-MoRSE descriptors
Mor11m	signal 11 / weighted by mass	3D-MoRSE descriptors
Mor22m	signal 22 / weighted by mass	3D-MoRSE descriptors
Mor23m	signal 23 / weighted by mass	3D-MoRSE descriptors
Mor27m	signal 27 / weighted by mass	3D-MoRSE descriptors
Mor28m	signal 28 / weighted by mass	3D-MoRSE descriptors
Mor10v	signal 10 / weighted by van der Waals volume	3D-MoRSE descriptors
Mor15e	signal 15 / weighted by Sanderson electronegativity	3D-MoRSE descriptors
Mor16e	signal 16 / weighted by Sanderson electronegativity	3D-MoRSE descriptors
Mor01p	signal 01 / weighted by polarizability	3D-MoRSE descriptors
Mor08i	signal 08 / weighted by ionization potential	3D-MoRSE descriptors
Mor15i	signal 15 / weighted by ionization potential	3D-MoRSE descriptors
Mor16i	signal 16 / weighted by ionization potential	3D-MoRSE descriptors
Mor21i	signal 21 / weighted by ionization potential	3D-MoRSE descriptors
Mor06s	signal 06 / weighted by I-state	3D-MoRSE descriptors
Mor14s	signal 14 / weighted by I-state	3D-MoRSE descriptors
Mor22s	signal 22 / weighted by I-state	3D-MoRSE descriptors
Mor24s	signal 24 / weighted by I-state	3D-MoRSE descriptors
Mor31s	signal 31 / weighted by I-state	3D-MoRSE descriptors

Name	Description	Block
L2u	2nd component size directional WHIM index / unweighted	WHIM descriptors
G3v	3rd component symmetry directional WHIM index / weighted by van der Waals volume	WHIM descriptors
L1p	1st component size directional WHIM index / weighted by polarizability	WHIM descriptors
G3p	3rd component symmetry directional WHIM index / weighted by polarizability	WHIM descriptors
Km	K global shape index / weighted by mass	WHIM descriptors
De	D total accessibility index / weighted by Sanderson electronegativity	WHIM descriptors
Dp	D total accessibility index / weighted by polarizability	WHIM descriptors
HTm	H total index / weighted by mass	GETAWAY descriptors
HATS6m	leverage-weighted autocorrelation of lag 6 / weighted by mass	GETAWAY descriptors
H0v	H autocorrelation of lag 0 / weighted by van der Waals volume	GETAWAY descriptors
HTv	H total index / weighted by van der Waals volume	GETAWAY descriptors
HATS4e	leverage-weighted autocorrelation of lag 4 / weighted by Sanderson electronegativity	GETAWAY descriptors
HTp	H total index / weighted by polarizability	GETAWAY descriptors
HTs	H total index / weighted by I-state	GETAWAY descriptors
R7u+	R maximal autocorrelation of lag 7 / unweighted	GETAWAY descriptors
R2e	R autocorrelation of lag 2 / weighted by Sanderson electronegativity	GETAWAY descriptors
R3e+	R maximal autocorrelation of lag 3 / weighted by Sanderson electronegativity	GETAWAY descriptors
RTp	R total index / weighted by polarizability	GETAWAY descriptors
R1i+	R maximal autocorrelation of lag 1 / weighted by ionization potential	GETAWAY descriptors
R3i+	R maximal autocorrelation of lag 3 / weighted by ionization potential	GETAWAY descriptors
R1s+	R maximal autocorrelation of lag 1 / weighted by I-state	GETAWAY descriptors
R3s+	R maximal autocorrelation of lag 3 / weighted by I-state	GETAWAY descriptors

Name	Description	Block
DP11	molecular profile no. 11	Randic molecular profiles
DP19	molecular profile no. 19	Randic molecular profiles
nR=Cp	number of terminal primary C(sp2)	Functional group counts
nRCONH2	number of primary amides (aliphatic)	Functional group counts
nArCHO	number of aldehydes (aromatic)	Functional group counts
nArCN	number of nitriles (aromatic)	Functional group counts
nN+	number of positively charged N	Functional group counts
nArOH	number of aromatic hydroxyls	Functional group counts
nOHp	number of primary alcohols	Functional group counts
nArOR	number of ethers (aromatic)	Functional group counts
nRSR	number of sulfides	Functional group counts
nOxiranes	number of Oxiranes	Functional group counts
nPyrrolidines	number of Pyrrolidines	Functional group counts
nOxolanes	number of Oxolanes	Functional group counts
nPyrazoles	number of Pyrazoles	Functional group counts
nPyridines	number of Pyridines	Functional group counts
nHBonds	number of intramolecular H-bonds (with N,O,F)	Functional group counts
C-002	CH2R2	Atom-centred fragments
C-003	CHR3	Atom-centred fragments
C-004	CR4	Atom-centred fragments

Name	Description	Block
C-005	CH3X	Atom-centred fragments
C-007	CH2X2	Atom-centred fragments
C-012	CR2X2	Atom-centred fragments
C-026	R--CX--R	Atom-centred fragments
C-027	R--CH--X	Atom-centred fragments
H-051	H attached to alpha-C	Atom-centred fragments
O-057	phenol / enol / carboxyl OH	Atom-centred fragments
N-066	Al-NH2	Atom-centred fragments
N-072	RCO-N< / >N-X=X	Atom-centred fragments
N-074	R#N / R=N-	Atom-centred fragments
F-081	F attached to C1(sp3)	Atom-centred fragments
SsCH3	Sum of sCH3 E-states	Atom-type E-state indices
SaaCH	Sum of aaCH E-states	Atom-type E-state indices
SsNH2	Sum of sNH2 E-states	Atom-type E-state indices
SdsN	Sum of dsN E-states	Atom-type E-state indices
StN	Sum of tN E-states	Atom-type E-state indices
SaaO	Sum of aaO E-states	Atom-type E-state indices
NssNH	Number of atoms of type ssNH	Atom-type E-state

Name	Description	Block
		indices
NtN	Number of atoms of type tN	Atom-type E-state indices
NsssP	Number of atoms of type sssP	Atom-type E-state indices
NssS	Number of atoms of type ssS	Atom-type E-state indices
CATS2D_03_DA	CATS2D Donor-Acceptor at lag 03	CATS 2D
CATS2D_07_DN	CATS2D Donor-Negative at lag 07	CATS 2D
CATS2D_08_DN	CATS2D Donor-Negative at lag 08	CATS 2D
CATS2D_05_DL	CATS2D Donor-Lipophilic at lag 05	CATS 2D
CATS2D_07_AA	CATS2D Acceptor-Acceptor at lag 07	CATS 2D
CATS2D_06_AP	CATS2D Acceptor-Positive at lag 06	CATS 2D
CATS2D_07_AP	CATS2D Acceptor-Positive at lag 07	CATS 2D
CATS2D_00_AN	CATS2D Acceptor-Negative at lag 00	CATS 2D
CATS2D_02_AN	CATS2D Acceptor-Negative at lag 02	CATS 2D
CATS2D_09_AN	CATS2D Acceptor-Negative at lag 09	CATS 2D
CATS2D_02_AL	CATS2D Acceptor-Lipophilic at lag 02	CATS 2D
CATS2D_03_PP	CATS2D Positive-Positive at lag 03	CATS 2D
CATS2D_09_PP	CATS2D Positive-Positive at lag 09	CATS 2D
CATS2D_06_PN	CATS2D Positive-Negative at lag 06	CATS 2D
CATS2D_09_PL	CATS2D Positive-Lipophilic at lag 09	CATS 2D
CATS2D_04_NN	CATS2D Negative-Negative at lag 04	CATS 2D

Name	Description	Block
CATS2D_05_NL	CATS2D Negative-Lipophilic at lag 05	CATS 2D
CATS2D_09_NL	CATS2D Negative-Lipophilic at lag 09	CATS 2D
CATS2D_01_LL	CATS2D Lipophilic-Lipophilic at lag 01	CATS 2D
CATS2D_06_LL	CATS2D Lipophilic-Lipophilic at lag 06	CATS 2D
T(N..Cl)	sum of topological distances between N..Cl	2D Atom Pairs
B01[C-C]	Presence/absence of C - C at topological distance 1	2D Atom Pairs
B01[C-N]	Presence/absence of C - N at topological distance 1	2D Atom Pairs
B01[C-O]	Presence/absence of C - O at topological distance 1	2D Atom Pairs
B01[C-P]	Presence/absence of C - P at topological distance 1	2D Atom Pairs
B01[O-P]	Presence/absence of O - P at topological distance 1	2D Atom Pairs
B02[C-P]	Presence/absence of C - P at topological distance 2	2D Atom Pairs
B02[N-O]	Presence/absence of N - O at topological distance 2	2D Atom Pairs
B02[N-S]	Presence/absence of N - S at topological distance 2	2D Atom Pairs
B02[O-O]	Presence/absence of O - O at topological distance 2	2D Atom Pairs
B02[O-F]	Presence/absence of O - F at topological distance 2	2D Atom Pairs
B03[C-F]	Presence/absence of C - F at topological distance 3	2D Atom Pairs
B03[N-O]	Presence/absence of N - O at topological distance 3	2D Atom Pairs
B03[N-S]	Presence/absence of N - S at topological distance 3	2D Atom Pairs
B03[N-Cl]	Presence/absence of N - Cl at topological distance 3	2D Atom Pairs
B03[O-Br]	Presence/absence of O - Br at topological distance 3	2D Atom Pairs
B04[C-N]	Presence/absence of C - N at topological distance 4	2D Atom Pairs
B04[C-F]	Presence/absence of C - F at topological distance 4	2D Atom Pairs

Name	Description	Block
B04[C-Cl]	Presence/absence of C - Cl at topological distance 4	2D Atom Pairs
B04[P-P]	Presence/absence of P - P at topological distance 4	2D Atom Pairs
B05[C-C]	Presence/absence of C - C at topological distance 5	2D Atom Pairs
B05[C-N]	Presence/absence of C - N at topological distance 5	2D Atom Pairs
B05[C-P]	Presence/absence of C - P at topological distance 5	2D Atom Pairs
B05[O-P]	Presence/absence of O - P at topological distance 5	2D Atom Pairs
B06[N-O]	Presence/absence of N - O at topological distance 6	2D Atom Pairs
B06[N-P]	Presence/absence of N - P at topological distance 6	2D Atom Pairs
B06[N-Cl]	Presence/absence of N - Cl at topological distance 6	2D Atom Pairs
B06[O-P]	Presence/absence of O - P at topological distance 6	2D Atom Pairs
B07[C-F]	Presence/absence of C - F at topological distance 7	2D Atom Pairs
B07[O-P]	Presence/absence of O - P at topological distance 7	2D Atom Pairs
B08[C-F]	Presence/absence of C - F at topological distance 8	2D Atom Pairs
B08[O-O]	Presence/absence of O - O at topological distance 8	2D Atom Pairs
B09[C-P]	Presence/absence of C - P at topological distance 9	2D Atom Pairs
B09[C-I]	Presence/absence of C - I at topological distance 9	2D Atom Pairs
B09[N-N]	Presence/absence of N - N at topological distance 9	2D Atom Pairs
B09[N-O]	Presence/absence of N - O at topological distance 9	2D Atom Pairs
B10[N-N]	Presence/absence of N - N at topological distance 10	2D Atom Pairs
B10[O-O]	Presence/absence of O - O at topological distance 10	2D Atom Pairs
F02[C-S]	Frequency of C - S at topological distance 2	2D Atom Pairs
F02[C-Br]	Frequency of C - Br at topological distance 2	2D Atom Pairs

Name	Description	Block
F02[F-F]	Frequency of F - F at topological distance 2	2D Atom Pairs
F03[C-C]	Frequency of C - C at topological distance 3	2D Atom Pairs
F03[N-O]	Frequency of N - O at topological distance 3	2D Atom Pairs
F03[Cl-Cl]	Frequency of Cl - Cl at topological distance 3	2D Atom Pairs
F04[C-X]	Frequency of C - X at topological distance 4	2D Atom Pairs
F04[N-O]	Frequency of N - O at topological distance 4	2D Atom Pairs
F04[N-S]	Frequency of N - S at topological distance 4	2D Atom Pairs
F04[Br-Br]	Frequency of Br - Br at topological distance 4	2D Atom Pairs
F05[C-N]	Frequency of C - N at topological distance 5	2D Atom Pairs
F05[C-S]	Frequency of C - S at topological distance 5	2D Atom Pairs
F05[N-S]	Frequency of N - S at topological distance 5	2D Atom Pairs
F05[O-Cl]	Frequency of O - Cl at topological distance 5	2D Atom Pairs
F05[S-P]	Frequency of S - P at topological distance 5	2D Atom Pairs
F06[C-S]	Frequency of C - S at topological distance 6	2D Atom Pairs
F06[N-O]	Frequency of N - O at topological distance 6	2D Atom Pairs
F06[N-Br]	Frequency of N - Br at topological distance 6	2D Atom Pairs
F07[C-Cl]	Frequency of C - Cl at topological distance 7	2D Atom Pairs
F08[C-F]	Frequency of C - F at topological distance 8	2D Atom Pairs
F09[N-N]	Frequency of N - N at topological distance 9	2D Atom Pairs
F10[C-C]	Frequency of C - C at topological distance 10	2D Atom Pairs
F10[C-O]	Frequency of C - O at topological distance 10	2D Atom Pairs
F10[N-P]	Frequency of N - P at topological distance 10	2D Atom Pairs

Name	Description	Block
F10[O-P]	Frequency of O - P at topological distance 10	2D Atom Pairs
G(N..P)	sum of geometrical distances between N..P	3D Atom Pairs
G(S..S)	sum of geometrical distances between S..S	3D Atom Pairs
MLOGP	Moriguchi octanol-water partition coeff. (logP)	Molecular properties
ALOGP	Ghose-Crippen octanol-water partition coeff. (logP)	Molecular properties
ALOGP2	squared Ghose-Crippen octanol-water partition coeff. (logP^2)	Molecular properties
Ro5	Lipinski Rule of 5	Drug-like indices
DLS_04	modified drug-like score from Chen et al. (7 rules)	Drug-like indices
DLS_05	modified drug-like score from Zheng et al. (2 rules)	Drug-like indices
DLS_07	modified drug-like score from Veber et al. (2 rules)	Drug-like indices
LLS_02	modified lead-like score from Monge et al. (8 rules)	Drug-like indices
Psychotic-80	Ghose-Viswanadhan-Wendoloski antipsychotic-like index at 80%	Drug-like indices
Hypertens-50	Ghose-Viswanadhan-Wendoloski antihypertensive-like index at 50%	Drug-like indices
Hypnotic-80	Ghose-Viswanadhan-Wendoloski hypnotic-like index at 80%	Drug-like indices
Neoplastic-50	Ghose-Viswanadhan-Wendoloski antineoplastic-like index at 50%	Drug-like indices
CATS3D_12_DD	CATS3D Donor-Donor BIN 12 (12.000 - 13.000 Å)	CATS 3D
CATS3D_15_DD	CATS3D Donor-Donor BIN 15 (15.000 - 16.000 Å)	CATS 3D
CATS3D_07_DA	CATS3D Donor-Acceptor BIN 07 (7.000 - 8.000 Å)	CATS 3D
CATS3D_14_DA	CATS3D Donor-Acceptor BIN 14 (14.000 - 15.000 Å)	CATS 3D
CATS3D_11_DP	CATS3D Donor-Positive BIN 11 (11.000 - 12.000 Å)	CATS 3D
CATS3D_13_DP	CATS3D Donor-Positive BIN 13 (13.000 - 14.000 Å)	CATS 3D
CATS3D_17_DP	CATS3D Donor-Positive BIN 17 (17.000 - 18.000 Å)	CATS 3D

Name	Description	Block
CATS3D_04_DN	CATS3D Donor-Negative BIN 04 (4.000 - 5.000 Å)	CATS 3D
CATS3D_06_DN	CATS3D Donor-Negative BIN 06 (6.000 - 7.000 Å)	CATS 3D
CATS3D_07_DN	CATS3D Donor-Negative BIN 07 (7.000 - 8.000 Å)	CATS 3D
CATS3D_12_DN	CATS3D Donor-Negative BIN 12 (12.000 - 13.000 Å)	CATS 3D
CATS3D_05_DL	CATS3D Donor-Lipophilic BIN 05 (5.000 - 6.000 Å)	CATS 3D
CATS3D_07_DL	CATS3D Donor-Lipophilic BIN 07 (7.000 - 8.000 Å)	CATS 3D
CATS3D_13_DL	CATS3D Donor-Lipophilic BIN 13 (13.000 - 14.000 Å)	CATS 3D
CATS3D_00_AA	CATS3D Acceptor-Acceptor BIN 00 (0.000 - 1.000 Å)	CATS 3D
CATS3D_05_AP	CATS3D Acceptor-Positive BIN 05 (5.000 - 6.000 Å)	CATS 3D
CATS3D_02_AN	CATS3D Acceptor-Negative BIN 02 (2.000 - 3.000 Å)	CATS 3D
CATS3D_06_AN	CATS3D Acceptor-Negative BIN 06 (6.000 - 7.000 Å)	CATS 3D
CATS3D_06_AL	CATS3D Acceptor-Lipophilic BIN 06 (6.000 - 7.000 Å)	CATS 3D
CATS3D_09_AL	CATS3D Acceptor-Lipophilic BIN 09 (9.000 - 10.000 Å)	CATS 3D
CATS3D_07_PP	CATS3D Positive-Positive BIN 07 (7.000 - 8.000 Å)	CATS 3D
CATS3D_09_PN	CATS3D Positive-Negative BIN 09 (9.000 - 10.000 Å)	CATS 3D
CATS3D_15_PN	CATS3D Positive-Negative BIN 15 (15.000 - 16.000 Å)	CATS 3D
CATS3D_09_PL	CATS3D Positive-Lipophilic BIN 09 (9.000 - 10.000 Å)	CATS 3D
CATS3D_15_PL	CATS3D Positive-Lipophilic BIN 15 (15.000 - 16.000 Å)	CATS 3D
CATS3D_01_NN	CATS3D Negative-Negative BIN 01 (1.000 - 2.000 Å)	CATS 3D
CATS3D_02_LL	CATS3D Lipophilic-Lipophilic BIN 02 (2.000 - 3.000 Å)	CATS 3D
CATS3D_03_LL	CATS3D Lipophilic-Lipophilic BIN 03 (3.000 - 4.000 Å)	CATS 3D