Molecular Descriptors Guide

Description of the Molecular Descriptors Appearing in the ChemoPy Software Package

Version1.0

Table of Contents

Descriptors of drugs

A small or drug molecule could be represented by its chemical structure. In the ChemoPy software, we calculate ten types of molecular descriptors to represent drug molecules, including constitutional descriptors, topological descriptors, connectivity indices, E-state indices, autocorrelation descriptors, charge descriptors, molecular properties, kappa shape indices, MOE-type descriptors, and molecular fingerprints. These descriptors capture and magnify distinct aspects of chemical structures.

1. Molecular constitutional descriptors

1. Molecular weight (*Weight*)
2. Count of hydrogen atoms (*nhyd*)
3. Count of halogen atoms (*nhal*)
4. Count of hetero atoms (*nhet*)
5. Count of heavy atoms (*nhev*)
6. Count of F atoms (*ncof*)
7. Count of Cl atoms (*ncocl*)
8. Count of Br atoms (*ncobr*)
9. Count of I atoms (*ncoi*)
10. Count of C atoms (*ncarb*)
11. Count of P atoms (*nphos*)
12. Count of S atoms (*nsulph*)
13. Count of O atoms (*noxy*)
14. Count of N atoms (*nnitro*)
15. Number of rings (*nring*)
16. Number of rotatable bonds (*nrot*)
17. Number of H-bond donors (*ndonr*)
18. Number of H-bond acceptors (*naccr*)
19. Number of single bonds (*nsb*)
20. Number of double bonds (*ndb*)
21. Number of triple bonds (*ntb*)
22. Number of aromatic bonds (*naro*)
23. Number of all atoms (*nta*)
24. Average molecular weight (*AWeight*)
25. Molecular path counts of length 1 (*PC1*)
26. Molecular path counts of length 2 (*PC2*)
27. Molecular path counts of length 3 (*PC3*)
28. Molecular path counts of length 4 (*PC4*)
29. Molecular path counts of length 5 (*PC5*)
30. Molecular path counts of length 6 (*PC6*)

Introduction:

(1) The molecular weight (MW) is the sum of molecular weights of the individual atoms, defined as:

$$MW = \sum_{i=1}^{A} MW_i$$

And the average molecular weight (AWeight) is given as follows:

AWeight=MW/nAT

where nAT is the number of atoms

(2) The number of hydrogen (*nhyd*), carbon (*ncarb*), nitrogen (*nnitro*), oxygen (*noxy*), phosphorus (*nphos*), sulfur (*nsulph*), fluorine (*ncof*), chlorine (*ncocl*), bromine (*ncobr*), and iodine (*ncoi*) atoms are simply the total number of each of these types of atoms in the molecule.

The number of halogen atoms (*nhal*) is simply the sum of the counts of the halogen atoms; the number of heavy atoms (*nhev*) and hetero atoms (*nhet*) are defined the similar way.

(3) From descriptor 15 to 22, they are simply the number of ring, single bond, double bond, aromatic bond and H-acceptor, etc, in the molecule.

(4) From descriptor 25 to 30, they represent the number of path of length 1-6. The path of length *n* indicates the shortest distance equal *n* between two atoms in a topological molecular graph.

2. Topological descriptors

1. Weiner index (*W*)
2. Average Weiner index (*AW*)
3. Balaban's J index (*J*)
4. Harary number (T_{hara})
5. Schiultz index (T_{sch})
6. Graph distance index (*Tigdi*)
7. Platt number (*Platt*)
8. Xu index (*Xu*)
9. Polarity number (*Pol*)
10. Pogliani index (*Dz*)
11. Ipc index (*Ipc*)
12. BertzCT (*BertzCT*)

13. Gutman molecular topological index based on simple vertex degree (*GMTI*)
14. Zagreb index with order 1 (*ZM1*)
15. Zagreb index with order 2 (*ZM2*)
16. Modified Zagreb index with order 1 (*MZM1*)
17. Modified Zagreb index with order 2 (*MZM2*)
18. Quadratic index (*Qindex*)
19. Largest value in the distance matrix (*diametert*)
20. Radius based on topology (*radiust*)
21. Petitjean based on topology (*petitjeant*)
22. The logarithm of the simple topological index by Narumi (*Sito*)
23. Harmonic topological index proposed by Narnumi (*Hato*)
24. Geometric topological index by Narumi (*Geto*)
25. Arithmetic topological index by Narumi (*Arto*)
26. Total information index on molecular size (*ISIZ*)
27. Total information index on atomic composition (*TIAC*)
28. Total information index on distance equality (*IDET*)
29. Mean information index on distance equality (*IDE*)
30. Total information index on vertex equality (*IVDE*)
31. Logarithm of the simple topological index by Narumi (*Sitov*)
32. Harmonic topological index proposed by Narnumi (*Hatov*)
33. Geometric topological index by Narumi (*Getov*)
34. Gravitational topological index based on topological distance (*Gravto*)
35. Gutman molecular topological index based on valence vertex degree (*GMTIV*)

Introduction:

(1) Weiner index (*W*)

$$W = (\sum d_{ij}) / 2$$

d_{ij} is the entries of distance matrix D from H-depleted molecular graph.

(2) Average Weiner index (*AW*)

The average Weiner index is given by

$$WA = \frac{2W}{A(A-1)}$$

where A is the total number of atoms in the molecule, W and AW are described in more detail

on pa 497 of the Handbook of Molecular Descriptors

(3) Balaban's J index (*J*)

$$J = \frac{B}{C+1}\sum_b (\sigma_i \sigma_j)_b^{-1/2}$$

where σ_i and σ_j are the vertex distance degree of adjacent atoms, and the sum run over all the molecular bond b, B is the number of bonds in the molecular graph and C is the number of rings. *J* are described in more detail on pa 21 of the Handbook of Molecular Descriptors

(4) Harary number (T_{hara})

$$H = \frac{1}{2}\sum_i \sum_j d_{ij}^{-1}$$

The Harary index is a molecular topological index derived from the reciprocal distance matrix D^{-1}

(5) Schiultz index (T_{sch})

$$MTI = \sum_{i=1}^{A} [(A+D)v]_i$$

It is a topological index derived from the adjacency matrix A, the distance matrix D and the A dimensional column vector v constituted by the vertex degree of the A atoms.

(6) Graph distance index (*Tigdi*)

The graph distance index is defined as the squared sum of all graph distance counts:

$$GDI = \sum_{k=1}^{D} ({}^k f)^2$$

where D is the topological diameter, ${}^k f$ is the total number of distances in the graph equal to k.

(7) Platt number (*Platt*)

Platt number is also known as the total edge adjacency index A_E, it is the sum over all entries of

the edge adjacency matrix:

$$A_E = \sum_{i=1}^{B}\sum_{j=1}^{B} E_{ij}$$

where B is the number of edges in molecular graph

(8) Xu index (*Xu*)

It is a topological molecular descriptor based on the adjacency matrix and distance matrix; it is defined as:

$$Xu = \sqrt{A}\log\frac{\sum_{i=1}^{A}\delta_i\sigma_i^2}{\sum_{i=1}^{A}\delta_i\sigma_i}$$

where A is the number of atoms, δ is vertex degree and σ is distance degree of all the atoms.

(9) Polarity number (*Pol*)

It is usually assumed that the polarity number accounts for the flexibility of acyclic structure; it is usually calculated on the distance matrix as the number of pairs of vertices at a topological distance equal to three. Some other polarity number also been defined based on different rules.

(10) Pogliani index (*Dz*)

$$D^Z = \sum_{i=1}^{A}\frac{Z_i^v}{L_i}$$

where A is the number of atoms, Z is the number of valence electrons and L the principal quantum number.

(11) Ipc index (*Ipc*)

Ipc index is the information for polynomial coefficients based information theory.

(12) BertzCT (*BertzCT*)

It is the most popular complexity index, taking into account both the variety of kinds of bond connectivities and atom types. It is defined as:

$$I_{CPX} = I_{CPB} + I_{CPA}$$

where I_{CPB} and I_{CPA} are the information contents related to the bond connectivity and atom type diversity

(13) Gutman molecular topological index based on simple vertex degree (*GMTI*)

$$S_G = \sum_{i=1}^{A} \sum_{j=1}^{A} \delta_i \delta_j d_{ij}$$

where $\delta_i \delta_j d_{ij}$ is the topological distance between vertex i and vertex j weighted by the product of the endpoint vertex degrees.

(14) Zagreb index with order 1 (*ZM1*)

The first Zagreb index (Weighted by vertex degrees) is given by

$$M1 = \sum_a \delta_a^2$$

where *a* runs over the A atoms of the molecule and δ is the vertex degree.

(15) Zagreb index with order 2 (*ZM2*)

$$M2 = \sum_b (\delta_i \delta_j)_b$$

where b runs over all the bonds in the molecule

The Zagreb indices are described on pg 509 of Handbook of Molecular Descriptors

(16) Modified Zagreb index with order 1 (*MZM1*)

(17) Modified Zagreb index with order 2 (*MZM2*)

(18) Quadratic index (*Qindex*)

$$Q = \frac{\sum_g (g^2 - 2g)^g F + 2}{2}$$

Quadratic index also called normalized quadratic index, where *g* are the different vertex degree values and $^g F$ is the vertex degree count.

(19) Largest value in the distance matrix (*diametert*)

$$D = \max_i (\eta_i)$$

$$\eta_i = \max_j (d_{ij})$$

η_i called atom eccentricity is the maximum distance from the *ith* vertex to the other vertices.

(20) Radius based on topology (*radiust*)

$$R = \min_i(\eta_i)$$

(21) Petitjean based on topology (*petitjeant*)

$$I_2 = \frac{D-R}{R}$$

(22) The logarithm of the simple topological index by Narumi (*Sito*)

$$S = \prod_{i=1}^{A} \delta_i$$

where A is the number of atoms, *Sito* is a molecular descriptor related to molecular branching proposed as the product of the vertex degrees.

(23) Harmonic topological index proposed by Narumi (*Hato*)

$$H = \frac{A}{\sum_{i=1}^{A} 1/\delta_i}$$

(24) Geometric topological index by Narumi (*Geto*)

$$G = \left(\prod_{i=1}^{A} \delta_i\right)^{1/A}$$

(25) Arithmetic topological index by Narumi (*Arto*)

$$A = \frac{\sum_{i=1}^{A} \delta_i}{A}$$

(26) Total information index on molecular size (ISIZ)

(27) Total information index on atomic composition (TIAC)

(28) Total information index on distance equality (IDET)

(29) Mean information index on distance equality (IDE)

(30) Total information index on vertex equality (IVDE)

(31) Logarithm of the simple topological index by Narumi (Sitov)

(32) Harmonic topological index proposed by Narnumi (Hatov)

(33) Geometric topological index by Narumi (Getov)

(34) Gravitational topological index based on topological distance (Gravto)

(35) Gutman molecular topological index based on valence vertex degree (GMTIV)

3. Molecular connectivity indices

1. Valence molecular connectivity Chi index for path order 0 (${}^{0}\chi^{v}$)
2. Valence molecular connectivity Chi index for path order 1(${}^{1}\chi^{v}$)
3. Valence molecular connectivity Chi index for path order 2(${}^{3}\chi^{v}$)
4. Valence molecular connectivity Chi index for path order 3(${}^{4}\chi^{v}$)
5. Valence molecular connectivity Chi index for path order 4(${}^{5}\chi^{v}$)
6. Valence molecular connectivity Chi index for path order 5(${}^{6}\chi^{v}$)
7. Valence molecular connectivity Chi index for path order 6(${}^{7}\chi^{v}$)
8. Valence molecular connectivity Chi index for path order 7 (${}^{8}\chi^{v}$)
9. Valence molecular connectivity Chi index for path order 8(${}^{9}\chi^{v}$)
10. Valence molecular connectivity Chi index for path order 9(${}^{10}\chi^{v}$)
11. Valence molecular connectivity Chi index for path order 10(${}^{11}\chi^{v}$)
12. Valence molecular connectivity Chi index for three cluster (${}^{3}\chi^{v}{}_{c}$)
13. Valence molecular connectivity Chi index for four cluster (${}^{4}\chi^{v}{}_{c}$)
14. Valence molecular connectivity Chi index for path/cluster (${}^{4}\chi^{v}{}_{pc}$)
15. Valence molecular connectivity Chi index for cycles of 3 (${}^{3}\chi^{v}{}_{CH}$)
16. Valence molecular connectivity Chi index for cycles of 4 (${}^{4}\chi^{v}{}_{CH}$)
17. Valence molecular connectivity Chi index for cycles of 5 (${}^{5}\chi^{v}{}_{CH}$)
18. Valence molecular connectivity Chi index for cycles of 6 (${}^{6}\chi^{v}{}_{CH}$)
19. Simple molecular connectivity Chi indices for path order 0 (${}^{0}\chi$)
20. Simple molecular connectivity Chi indices for path order 1 (${}^{1}\chi$)
21. Simple molecular connectivity Chi indices for path order 2 (${}^{2}\chi$)
22. Simple molecular connectivity Chi indices for path order 3 (${}^{3}\chi_{p}$)
23. Simple molecular connectivity Chi indices for path order 4 (${}^{4}\chi_{p}$)
24. Simple molecular connectivity Chi indices for path order 5 (${}^{5}\chi_{p}$)
25. Simple molecular connectivity Chi indices for path order 6 (${}^{6}\chi_{p}$)
26. Simple molecular connectivity Chi indices for path order 7 (${}^{7}\chi_{p}$)
27. Simple molecular connectivity Chi indices for path order 8 (${}^{8}\chi_{p}$)
28. Simple molecular connectivity Chi indices for path order 9 (${}^{9}\chi_{p}$)

29. Simple molecular connectivity Chi indices for path order 10 ($^{10}\chi_p$)
30. Simple molecular connectivity Chi indices for three cluster ($^{3}\chi_c$)
31. Simple molecular connectivity Chi indices for four cluster ($^{4}\chi_c$)
32. Simple molecular connectivity Chi indices for path/cluster ($^{4}\chi_{pc}$)
33. Simple molecular connectivity Chi indices for cycles of 3 ($^{3}\chi_{CH}$)
34. Simple molecular connectivity Chi indices for cycles of 4 ($^{4}\chi_{CH}$)
35. Simple molecular connectivity Chi indices for cycles of 5 ($^{5}\chi_{CH}$)
36. Simple molecular connectivity Chi indices for cycles of 6 ($^{6}\chi_{CH}$)
37. mean chi1 (Randic) connectivity index (*mChi1*)
38. the difference between chi3c and chi4pc (*knotp*)
39. the difference between chi0v and chi0 (*dchi0*)
40. the difference between chi1v and chi1 (*dchi1*)
41. the difference between chi2v and chi2 (*dchi0*)
42. the difference between chi3v and chi3 (*dchi3*)
43. the difference between chi4v and chi4 (*dchi4*)
44. the difference between chiv3c and chiv4pc (*knotpv*)

Introduction:

1. Simple molecular connectivity index (No.19~36)

   The general formula for the molecular connectivity indices ($^{m}\chi_t$) is as follows:

$$^{m}\chi_q = \sum_{k=1}^{K}\left(\prod_{a=1}^{n} \delta_a\right)_k^{-1/2}$$

   where *k* runs over all of the *mth* order sub-graphs constituted by *n* atoms; *K* is the total number of *mth* order sub-graphs present in the molecular graph and in the case of the path sub-graphs equals the *mth* order path count ^{m}P. The product is over the simple vertex degrees of all the vertices involved in each sub-graph. The subscript "*q*" for the connectivity indices refers to the type of molecular sub-graph and *ch* for chain or ring, *pc* for path-cluster, *c* for cluster, and *p* for path. For the first three path indices ($^{0}\chi$, $^{1}\chi$, $^{2}\chi$), the calculation type, *p*, is often omitted from the variable name in the software.

2. Valence molecular connectivity indices (No.1~18)

   The valence connectivity indices ($^{m}\chi^{v}_{t}$) are calculated in the same fashion as the simple connectivity indices except that the vertex degree are replaced by the valence vertex degree, and the valence

degree is given by: $\delta^v=Z^v-h=\sigma+\pi+n-h$. Where Z^v is the number of valence electrons, π is the number of electrons in *pi* orbital and *n* is the number of electrons in lone-pair orbitals.

The valence connectivity indices are described on page 86 of the Handbook of Molecular Descriptors. The connectivity indices are described in detail in the literature.

3. The remains connectivity indices are simple combination of the above simple connectivity indices and valence connectivity indices.

4. Kappa shape descriptors

1. Kappa alpha index for 1 bonded fragment ($^1\kappa_\alpha$)
2. Kappa alpha index for 2 bonded fragment ($^2\kappa_\alpha$)
3. Kappa alpha index for 3 bonded fragment ($^3\kappa_\alpha$)
4. Kier molecular flexibility index (*phi*)
5. Molecular shape Kappa index for 1 bonded fragment ($^1\kappa$)
6. Molecular shape Kappa index for 2 bonded fragment ($^1\kappa$)
7. Molecular shape Kappa index for 3 bonded fragment ($^1\kappa$)

Introduction:

(1) Kappa alpha index

The first order kappa shape index ($^1\kappa$) is given by

$$^1k = 2\,{}^1P_{\max}\,{}^1P_{\min} / ({}^1P_i)^2 = A(A-1)^2 / ({}^1P_i)^2$$

where P_i=# of paths of bond length *i* in the hydrogen suppressed molecule and A is the number of non hydrogen atoms in the molecule.

The second order kappa shape index ($^2\kappa$) is given by

$$^2k = 2\,{}^2P_{\max}\,{}^2P_{\min} / ({}^2P_i)^2 = (A-1)(A-2)^2 / ({}^2P_i)^2$$

The kappa shape indices are described on pg 248 of the Handbook of Molecular Descriptors.

The first order kappa alpha shape index ($^1\kappa_\alpha$) is given by

$$^1k_a = \frac{(A+a)(A+a-1)^2}{({}^1P+a)^2}$$

where

$$a = 1 - \frac{r_x}{r_{x(sp^3)}}$$

where r_x is the covalent radius of the atom being evaluated and $r_{x(sp^3)}$ is the covalent radius of a carbon sp^3 atom (0.77Å).

The second order kappa alpha shape index (${}^2\kappa_\alpha$) is given by

$${}^2k_a = \frac{(A+a-1)(A+a-2)^2}{({}^2P+a)^2}$$

The third order kappa alpha shape index (${}^3\kappa_\alpha$) is given by

$${}^3k_a = \frac{(A+a-1)(A+a-3)^2}{({}^3P+a)^2} \quad \text{if A is odd}$$

$${}^3k_a = \frac{(A+a-3)(A+a-2)^2}{({}^3P+a)^2} \quad \text{if A is even}$$

The kappa shape indices are described on page 250 of the Handbook of Molecular Descriptors.

The kappa flexibility index (*phi*) is given by

$$phi = \frac{{}^1k_a \, {}^2k_a}{A}$$

The kappa flexibility index is described on page 178 of the Handbook of Molecular Descriptors.

5. Basak descriptors

(1) The information content with order 0 proposed by Basak (IC0)
(2) The information content with order 1 proposed by Basak(IC1)
(3) Tthe information content with order 2 proposed by Basak(IC2)
(4) The information content with order 3 proposed by Basak(IC3)
(5) The information content with order 4 proposed by Basak(IC4)
(6) The information content with order 5 proposed by Basak(IC5)
(7) The information content with order 6 proposed by Basak(IC6)

(8) The structural information content with order 0 proposed by Basak (SIC0)
(9) The structural information content with order 1 proposed by Basak(SIC1)
(10) The structural information content with order 2 proposed by Basak(SIC2)
(11) The structural information content with order 3 proposed by Basak(SIC3)
(12) The structural information content with order 4 proposed by Basak(SIC4)
(13) The structural information content with order 5 proposed by Basak(SIC5)
(14) The structural information content with order 6 proposed by Basak(SIC6)
(15) The complementary information content with order 0 proposed by Basak(CIC0)
(16) The complementary information content with order 1 proposed by Basak(CIC1)
(17) The complementary information content with order 2 proposed by Basak(CIC2)
(18) The complementary information content with order 3 proposed by Basak(CIC3)
(19) The complementary information content with order 4 proposed by Basak(CIC4)
(20) The complementary information content with order 5 proposed by Basak(CIC5)
(21) The complementary information content with order 6 proposed by Basak(CIC6)

6. Electrotopological State Indices

1. Sum of E-State of atom type: sLi (*S1*)
2. Sum of E-State of atom type: ssBe (*S2*)
3. Sum of E-State of atom type: ssssBe (*S3*)
4. Sum of E-State of atom type: ssBH (*S4*)
5. Sum of E-State of atom type: sssB (*S5*)
6. Sum of E-State of atom type: ssssB (*S6*)
7. Sum of E-State of atom type: sCH3 (*S7*)
8. Sum of E-State of atom type: dCH2 (*S8*)
9. Sum of E-State of atom type: ssCH2 (*S9*)
10. Sum of E-State of atom type: tCH (*S10*)
11. Sum of E-State of atom type: dsCH (*S11*)
12. Sum of E-State of atom type: aaCH (*S12*)
13. Sum of E-State of atom type: sssCH (*S13*)
14. Sum of E-State of atom type: ddC (*S14*)
15. Sum of E-State of atom type: tsC (*S15*)
16. Sum of E-State of atom type: dssC (*S16*)

17. Sum of E-State of atom type: aasC (*S17*)
18. Sum of E-State of atom type: aaaC (*S18*)
19. Sum of E-State of atom type: ssssC (*S19*)
20. Sum of E-State of atom type: sNH3 (*S20*)
21. Sum of E-State of atom type: sNH2 (*S21*)
22. Sum of E-State of atom type: ssNH2 (*S22*)
23. Sum of E-State of atom type: dNH (*S23*)
24. Sum of E-State of atom type: ssNH (*S24*)
25. Sum of E-State of atom type: aaNH (*S25*)
26. Sum of E-State of atom type: tN (*S26*)
27. Sum of E-State of atom type: sssNH (*S27*)
28. Sum of E-State of atom type: dsN (*S28*)
29. Sum of E-State of atom type: aaN (*S29*)
30. Sum of E-State of atom type: sssN (*S30*)
31. Sum of E-State of atom type: ddsN (*S31*)
32. Sum of E-State of atom type: aasN (*S32*)
33. Sum of E-State of atom type: ssssN (*S33*)
34. Sum of E-State of atom type: sOH (*S34*)
35. Sum of E-State of atom type: dO (*S35*)
36. Sum of E-State of atom type: ssO (*S36*)
37. Sum of E-State of atom type: aaO (*S37*)
38. Sum of E-State of atom type: sF (*S38*)
39. Sum of E-State of atom type: sSiH3 (*S39*)
40. Sum of E-State of atom type: ssSiH2 (*S40*)
41. Sum of E-State of atom type: sssSiH (*S41*)
42. Sum of E-State of atom type: ssssSi (*S42*)
43. Sum of E-State of atom type: sPH2 (*S43*)
44. Sum of E-State of atom type: ssPH (*S44*)
45. Sum of E-State of atom type: sssP (*S45*)
46. Sum of E-State of atom type: dsssP (*S46*)
47. Sum of E-State of atom type: sssssP (*S47*)
48. Sum of E-State of atom type: sSH (*S48*)
49. Sum of E-State of atom type: dS (*S49*)

50. Sum of E-State of atom type: ssS (*S50*)
51. Sum of E-State of atom type: aaS (*S51*)
52. Sum of E-State of atom type: dssS (*S52*)
53. Sum of E-State of atom type: ddssS (*S53*)
54. Sum of E-State of atom type: sCl (*S54*)
55. Sum of E-State of atom type: sGeH3 (*S55*)
56. Sum of E-State of atom type: ssGeH2 (*S56*)
57. Sum of E-State of atom type: sssGeH (*S57*)
58. Sum of E-State of atom type: ssssGe (*S58*)
59. Sum of E-State of atom type: sAsH2 (*S59*)
60. Sum of E-State of atom type: ssAsH (*S60*)
61. Sum of E-State of atom type: sssAs (*S61*)
62. Sum of E-State of atom type: sssdAs (*S62*)
63. Sum of E-State of atom type: ssssAs (*S63*)
64. Sum of E-State of atom type: sSeH (*S64*)
65. Sum of E-State of atom type: dSe (*S65*)
66. Sum of E-State of atom type: ssSe (*S66*)
67. Sum of E-State of atom type: aaSe (*S67*)
68. Sum of E-State of atom type: dssSe (*S68*)
69. Sum of E-State of atom type: ddssSe (*S69*)
70. Sum of E-State of atom type: sBr (*S70*)
71. Sum of E-State of atom type: sSnH3 (*S71*)
72. Sum of E-State of atom type: ssSnH2 (*S72*)
73. Sum of E-State of atom type: sssSnH (*S73*)
74. Sum of E-State of atom type: ssssSn (*S74*)
75. Sum of E-State of atom type: sI (*S75*)
76. Sum of E-State of atom type: sPbH3 (*S76*)
77. Sum of E-State of atom type: ssPbH2 (*S77*)
78. Sum of E-State of atom type: sssPbH (*S78*)
79. Sum of E-State of atom type: ssssPb (*S79*)
80-158. maximum of E-State value of specified atom type (*Smax1~Smax79*)
159-237. minimum of E-State value of specified atom type (*Smin1~Smin79*)
238. EState indices over all non-hydrogen atoms (*Shev*)

239. The sum of the EState indices over all C atoms (*Scar*)

240. The sum of the EState indices over all Halogen atoms (*Shal*)

241. The sum of the EState indices over all hetero atoms (*Shet*)

242. The sum of the EState indices over all non-hydrogen atoms divided by the number of non-hydrogen atoms (*Save*)

243. The maximal Estate value in all atoms (*Smax*)

244. The minimal Estate value in all atoms (*Smin*)

245. The difference between Smax and Smin (*DS*)

Introduction:

The E-State value for a given non hydrogen atom *i* in a molecule is given by its intrinsic state (I_i) plus the sum of the perturbations on that atom from all the other atoms in the molecule:

$$S_k = I_k + \sum_{i=1}^{A} \Delta I_{ki}$$

where the intrinsic state (I_k) is given by

$$I_k = \frac{(2/N)^2 \delta_k^v + 1}{\delta_k}$$

where N=principle quantum number (which is equal to the element's period or row in the element table).

The perturbation of atom *k* due to atom *i* is given by

$$\Delta I_{ki} = \frac{(I_i - I_k)}{r_{ki}^2}$$

where

$$r_{ki} = d_{ki} + 1$$

d_{ki} is the number of bonds that separate atom *k* from atom *i*.

The atom type non hydrogen indices (SX) are obtained by summing the E-State values for all the atoms of a given type *t* that are present in the molecule.

$$SX = \sum S(t)$$

In addition, the symbol present in molecular descriptors, *s*, *d*, *t* and *a* indicate single bond, double bond, triple bond and aromatic bond, respectively.

7. Burden descriptors

1. Highest eigenvaluen.1 of Burden matrix/weighted by atomic masses (*bcutm1*)
2. Highest eigenvaluen.2 of Burden matrix/weighted by atomic masses (*bcutm2*)
3. Highest eigenvaluen.3 of Burden matrix/weighted by atomic masses (*bcutm 3*)
4. Highest eigenvaluen.4ofBurden matrix/weighted by atomic masses (*bcutm 4*)
5. Highest eigenvaluen.5 of Burden matrix/weighted by atomic masses (*bcutm 5*)
6. Highest eigenvaluen.6 of Burden matrix/weighted by atomic masses (*bcutm 6*)
7. Highest eigenvaluen.7 of Burden matrix/weighted by atomic masses (*bcutm7*)
8. Highest eigenvaluen.8 of Burden matrix/weighted by atomic masses (*bcutm8*)
9. Lowest eigenvaluen.1 of Burden matrix/weighted by atomic masses (*bcutm1*)
10. Lowest eigenvaluen.2 of Burden matrix/weighted by atomic masses (*bcutm2*)
11. Lowest eigenvaluen.3 of Burden matrix/weighted by atomic masses (*bcutm3*)
12. Lowest eigenvaluen.4 of Burden matrix/weighted by atomic masses (*bcutm4*)
13. Lowest eigenvaluen.5 of Burden matrix/weighted by atomic masses (*bcutm5*)
14. Lowest eigenvaluen.6 of Burden matrix/weighted by atomic masses (*bcutm6*)
15. Lowest eigenvaluen.7 of Burden matrix/weighted by atomic masses (*bcutm7*)
16. Lowest eigenvaluen.8 of Burden matrix/weighted by atomic masses (*bcutm8*)
17. Highest eigenvaluen.1 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv1*)
18. Highest eigenvaluen.2 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv2*)
19. Highest eigenvaluen.3 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv3*)
20. Highest eigenvaluen.4 of Burden matrix/weighted by atomic vander Waals volumes(*bcutv4*)
21. Highest eigenvaluen.5 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv5*)
22. Highest eigenvaluen.6 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv6*)
23. Highest eigenvaluen.7 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv7*)
24. Highest eigenvaluen.8 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv8*)
25. Lowest eigenvaluen.1of Burden matrix/weighted by atomic vander Waals volumes (*bcutv1*)

26. Lowest eigenvaluen.2 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv2*)
27. Lowest eigenvaluen.3 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv3*)
28. Lowest eigenvaluen.4 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv4*)
29. Lowest eigenvaluen.5 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv5*)
30. Lowest eigenvaluen.6 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv6*)
31. Lowest eigenvaluen.7of Burden matrix/weighted by atomic vander Waals volumes (*bcutv7*)
32. Lowest eigenvaluen.8 of Burden matrix/weighted by atomic vander Waals volumes (*bcutv8*)
33. Highest eigenvaluen.1 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute1*)
34. Highest eigenvaluen.2 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute2*)
35. Highest eigenvaluen.3 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute3*)
36. Highest eigenvaluen.4 of Burden matrix/weighted by atomic Sandersonel ectronegativities (*bcute4*)
37. Highest eigenvaluen.5 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute5*)
38. Highest eigenvaluen.6 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute6*)
39. Highest eigenvaluen.7of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute7*)
40. Highest eigenvaluen.8 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute8*)
41. Lowest eigenvaluen.1 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute1*)
42. Lowes teigenvaluen.2 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute2*)
43. Lowest eigenvaluen.3 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute3*)
44. Lowest eigenvaluen.4 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute4*)
45. Lowest eigenvaluen.5 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute5*)
46. Lowest eigenvaluen.6 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute6*)
47. Lowesteigenvaluen.7 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute7*)
48. Lowest eigenvaluen.8 of Burden matrix/weighted by atomic Sanderson electronegativities (*bcute8*)
49. Highest eigenvaluen.1 of Burden matrix/weighted by atomic polarizabilities (*bcutp1*)
50. Highest eigenvaluen.2 of Burden matrix/weighted by atomic polarizabilities (*bcutp2*)
51. Highesteigenvaluen.3 of Burden matrix/weighted by atomic polarizabilities (*bcutp3*)
52. Highest eigenvaluen.4 of Burden matrix/weighted by atomic polarizabilities (*bcutp4*)
53. Highest eigenvaluen.5 of Burden matrix/weighted by atomic polarizabilities (*bcutp5*)
54. Highesteigenvaluen.6 of Burden matrix/weighted by atomic polarizabilities (*bcutp6*)
55. Highesteigenvaluen.7 of Burden matrix/weighted by atomic polarizabilities (*bcutp7*)
56. Highest eigenvaluen.8 of Burden matrix/weighted by atomic polarizabilities (*bcutp8*)
57. Lowes teigenvaluen.1 of Burden matrix/weighted by atomic polarizabilities (*bcutp1*)
58. Lowest eigenvaluen.2 of Burden matrix/weighted by atomic polarizabilities (*bcutp2*)

59. Lowest eigenvaluen.3 of Burden matrix/weighted by atomic polarizabilities (*bcutp3*)

60. Lowest eigenvaluen.4 of Burden matrix/weighted by atomic polarizabilities (*bcutp4*)

61. Lowest eigenvaluen.5 of Burden matrix/weighted by atomic polarizabilities (*bcutp5*)

62. Lowest eigenvaluen.6 of Burden matrix/weighted by atomic polarizabilities (*bcutp6*)

63. Lowest eigenvaluen.7of Burden matrix/weighted by atomic polarizabilities (*bcutp7*)

64. Lowest eigenvaluen.8 of Burden matrix/weighted by atomic polarizabilities (*bcutp8*)

Introduction:

The Burden eigenvalue descriptors are determined by solving the following general eigenvalue equation:

where B is a real connectivity matrix to be defined, V is a matrix of eigenvectors, and e is a diagonal matrix of eigenvalues. The rules defining B are as follows:

a. Hydrogen atoms are included.

b. The diagonal elements of B, Bii, are either given by the carbon normalized atomic mass, vander Waals volume, Sanderson electronegativity, and polarizability of atom i.

c. The element of B connecting atoms i and j, Bij, is equal to the square root of the bond order between atoms i and j.

d. All other elements of B (corresponding non bonded atom pairs) are set to 0.001.

The carbon normalized weights are as follows:

Symbol	Mass	VdWVolume	Electronegativity	Polarizability
C	1.000	1.000	1.000	1.000
H	0.084	0.299	0.942	0.381
N	1.167	0.695	1.160	0.625
O	1.332	0.512	1.327	0.455
P	2.579	1.181	0.916	2.063
S	2.670	1.088	1.076	1.648
F	1.582	0.410	1.455	0.318
Cl	2.952	1.000	1.265	1.239
Br	6.653	1.384	1.171	1.733
I	10.566	1.728	1.011	3.040
Si	2.339	1.424	0.778	3.057
As	6.238	1.181	1.025	2.449
Sn	9.884	2.041	0.836	4.375
Hg	16.702	0.695	0.800	3.239
Pb	17.252	1.538	0.833	3.864

The lowest eigenvalues are the absolute values of the negative eigenvalues. The highest eigenvalues are

the eight largest positive eigenvalues. The Burden eigenvalues descriptors are described on the Handbook of Molecular Descriptors (Todeschini and Consonni 2000)

8. Autocorrelation descriptors

The Broto-Moreau autocorrelation descriptors (ATSdw) are given by

$$ATSdw = \sum_{i=1}^{A}\sum_{j=1}^{A}\delta_{ij}\omega_i\omega_j$$

where d is the considered topological distance (i.e. the lag in the autocorrelation terms), δ_{ij} is the Kronecker delta function ($\delta_{ij}=1$ if $d_{ij}=d$, zero otherwise), and w_i and w_j are the weights (normalized atomic properties) for atoms i and j respectively. The normalized atomic mass, van der Waals volume, electronegativity, or polarizability can be used for the weights. To match Dragon, the Broto-Moreau autocorrelation descriptors are calculated in the Software as follows:

$$ATSdw = \ln\left(1+\sum_{i=1}^{A}\sum_{j=1}^{A}\delta_{ij}\cdot w_i\cdot w_j\right)$$

The Moran autocorrelation descriptors (MATSdw) are given by

$$MATSdw = \frac{\frac{1}{\Delta}\cdot\sum_{i=1}^{A}\sum_{j=1}^{A}\delta ij\cdot\left(w_i-\overline{w}\right)\left(w_j-\overline{w}\right)}{\frac{1}{A}\sum_{i=1}^{A}\left(w_i-\overline{w}\right)^2}$$

where $\overline{w}$ is the average value of the property for the molecule and $\triangle$ is the number of vertex pairs at distance equal to d .

The Geary autocorrelation descriptors are given by

$$GATSdw = \frac{\frac{1}{2\Delta}\cdot\sum_{i=1}^{A}\sum_{j=1}^{A}\delta ij\cdot\left(w_i-w_j\right)^2}{\frac{1}{A-1}\sum_{i=1}^{A}\left(w_i-\overline{w}\right)^2}$$

The 2D autocorrelation descriptors are described on page17-19 of the Handbook of Molecular Descriptors.

8.1 Moreau-Broto autocorrelation descriptors

1. Broto-Moreau autocorrelation of a topological structure-lag1/weighted by atomic masses (*ATSm1*)
2. Broto-Moreau autocorrelation of a topological structure-lag2/weighted by atomic masses (*ATSm2*)
3. Broto-Moreau autocorrelation of a topological structure-lag3/weighted by atomic masses (*ATSm3*)
4. Broto-Moreau autocorrelation of a topologicalstructure-lag4/weighted by atomic masses (*ATSm4*)
5. Broto-Moreau autocorrelation of a topological structure-lag5/weighted by atomic masses (*ATSm5*)
6. Broto-Moreau autocorrelation of a topological structure-lag6/weighted by atomic masses (*ATSm6*)
7. Broto-Moreau autocorrelation of a topological structure-lag7/weighted by atomic masses (*ATSm7*)
8. Broto-Moreau autocorrelation of a topological structure-lag8/weighted by atomic masses (*ATSm8*)
9. Broto-Moreau autocorrelation of a topological structure-lag1/weighted by atomic van der Waals volumes (*ATSv1*)
10. Broto-Moreau autocorrelation of a topological structure-lag2/weighted by atomic van der Waals volumes (*ATSv2*)
11. Broto-Moreau autocorrelation of a topological structure-lag3/weighted by atomic van der Waals volumes (*ATSv3*)
12. Broto-Moreau autocorrelation of a topological structure-lag4/weighted by atomic van der Waals volumes (*ATSv4*)
13. Broto-Moreau autocorrelation of a topological structure-lag5/weighted by atomic van der Waals volumes (*ATSv5*)
14. Broto-Moreau autocorrelation of a topological structure-lag6/weighted by atomi van der Waals volumes (*ATSv6*)
15. Broto-Moreau autocorrelation of a topological structure-lag7/weighted by atomic van der Waals volumes (*ATSv7*)
16. Broto-Moreau autocorrelation of a topological structure-lag8/weighted by atomic van der Waals volumes (*ATSv8*)
17. Broto-Moreau autocorrelation of a topological structure-lag1/weighted by atomic Sanderson electronegativities (*ATSe1*)
18. Broto-Moreau autocorrelation of a topological structure-lag2/weighted by atomic Sanderson electronegativities (*ATSe2*)
19. Broto-Moreau autocorrelation of a topological structure-lag3/weighted by atomic Sanderson electronegativities (*ATSe3*)

20. Broto-Moreau autocorrelation of a topological structure-lag4/weighted by atomic Sanderson electronegativities (*ATSe4*)
21. Broto-Moreau autocorrelation of a topological structure-lag5/weighted by atomic Sanderson electronegativities (*ATSe5*)
22. Broto-Moreau autocorrelation of a topological structure-lag6/weighted by atomic Sanderson electronegativities (*ATSe6*)
23. Broto-Moreau autocorrelation of a topological structure-lag7/weighted by atomic Sanderson electronegativities (*ATSe7*)
24. Broto-Moreau autocorrelation of a topological structure-lag8/weighted by atomic Sanderson electronegativities (*ATSe8*)
25. Broto-Moreau autocorrelation of a topological structure-lag1/weighted by atomic polarizabilities (*ATSp1*)
26. Broto-Moreau autocorrelation of a topological structure-lag2/weighted by atomic polarizabilities (*ATSp2*)
27. Broto-Moreau autocorrelation of a topological structure-lag3/weighted by atomic polarizabilities (*ATSp3*)
28. Broto-Moreau autocorrelation of a topological structure-lag4/weighted by atomic polarizabilities (*ATSp4*)
29. Broto-Moreau autocorrelation of a topological structure-lag5/weighted by atomic polarizabilities (*ATSp5*)
30. Broto-Moreau autocorrelation of a topological structure-lag6/weighted by atomic polarizabilities (*ATSp6*)
31. Broto-Moreau autocorrelation of a topological structure-lag7/weighted by atomic polarizabilities (*ATSp7*)
32. Broto-Moreau autocorrelation of a topological structure-lag8/weightedbyatomic polarizabilities (*ATSp8*)

8.2 Moran autocorrelation descriptors

33. Moran autocorrelation-lag1/weighted by atomic masses (*MATSm1*)
34. Moran autocorrelation-lag2/weighted by atomic masses (*MATSm2*)
35. Moran autocorrelation-lag3/weighted by atomic masses (*MATSm3*)

36. Moran autocorrelation-lag4/weighted by atomic masses (*MATSm4*)
37. Moran autocorrelation-lag5/weighted by atomic masses (*MATSm5*)
38. Moran autocorrelation-lag6/weighted by atomic masses (*MATSm6*)
39. Moran autocorrelation-lag7/weighted by atomic masses (*MATSm7*)
40. Moran autocorrelation-lag 8/weighted by atomic masses (*MATSm8*)
41. Moran autocorrelation-lag1/weighted by atomic van der Waals volumes (*MATSv1*)
42. Moran autocorrelation-lag2/weighted by atomic van der Waals volumes (*MATSv2*)
43. Moran autocorrelation-lag3/weighted by atomic van der Waals volumes (*MATSv3*)
44. Moran autocorrelation-lag4/weighted by atomic van der Waals volumes (*MATSv4*)
45. Moran autocorrelation-lag5/weighted by atomic van der Waals volumes (*MATSv5*)
46. Moran autocorrelation-lag6/weighted by atomic van der Waals volumes (*MATSv6*)
47. Moran autocorrelation-lag7/weighted by atomic van der Waals volumes (*MATSv7*)
48. Moran autocorrelation-lag8/weighted by atomic van der Waals volumes (*MATSv8*)
49. Moran autocorrelation-lag1/weighted by atomic Sanderson electronegativities (*MATSe1*)
50. Moran autocorrelation-lag2/weighted by atomic Sanderson electronegativities (*MATSe2*)
51. Moran autocorrelation-lag3/weighted by atomic Sanderson electronegativities (*MATSe3*)
52. Moran autocorrelation-lag4/weighted by atomic Sanderson electronegativities (*MATSe4*)
53. Moran autocorrelation-lag5/weighted by atomic Sanderson electronegativities (*MATSe5*)
54. Moran autocorrelation-lag6/weighted by atomic Sanderson electronegativities (*MATSe6*)
55. Moran autocorrelation-lag7/weighted by atomic Sanderson electronegativities (*MATSe7*)
56. Moran autocorrelation-lag8/weighted by atomic Sanderson electronegativities (*MATSe8*)
57. Moran autocorrelation-lag1/weighted by atomic polarizabilities (*MATSp1*)
58. Moran autocorrelation-lag2/weighted by atomic polarizabilities (*MATSp2*)
59. Moran autocorrelation-lag3/weighted by atomic polarizabilities (*MATSp3*)
60. Moran autocorrelation-lag4/weighted by atomic polarizabilities (*MATSp4*)
61. Moran autocorrelation-lag5/weighted by atomic polarizabilities (*MATSp5*)
62. Moran autocorrelation-lag6/weighted by atomic polarizabilities (*MATSp6*)
63. Moran autocorrelation-lag7/weighted by atomic polarizabilities (*MATSp7*)
64. Moran autocorrelation-lag8/weighted by atomic polarizabilities (*MATSp8*)

8.3 Geary autocorrelation descriptors

65. Geary autocorrelation-lag1/weighted by atomic masses (*GATSm1*)

66. Geary autocorrelation-lag2/weighted by atomic masses (*GATSm2*)
67. Geary autocorrelation-lag3/weighted by atomic masses (*GATSm3*)
68. Geary autocorrelation-lag4/weighted by atomic masses (*GATSm4*)
69. Geary autocorrelation-lag5/weighted by atomic masses (*GATSm5*)
70. Geary autocorrelation-lag6/weighted by atomic masses (*GATSm6*)
71. Geary autocorrelation-lag7/weighted by atomic masses (*GATSm7*)
72. Geary autocorrelation-lag8/weighted by atomic masses (*GATSm8*)
73. Geary autocorrelation-lag1/weighted by atomic van der Waals volumes (*GATSv1*)
74. Geary autocorrelation-lag2/weighted by atomic van der Waals volumes (*GATSv2*)
75. Geary autocorrelation-lag3/weighted by atomic van der Waals volumes (*GATSv3*)
76. Geary autocorrelation-lag4/weighted by atomic van der Waals volumes (*GATSv4*)
77. Geary autocorrelation-lag5/weighted by atomic van der Waals volumes (*GATSv5*)
78. Geary autocorrelation-lag6/weighted by atomic van der Waals volumes (*GATSv6*)
79. Geary autocorrelation-lag7/weighted by atomic van der Waals volumes (*GATSv7*)
80. Geary autocorrelation-lag8/weighted by atomic van der Waals volumes (*GATSv8*)
81. Geary autocorrelation-lag1/weighted by atomic Sanderson electronegativities (*GATSe1*)
82. Geary autocorrelation-lag2/weighted by atomic Sanderson electronegativities (*GATSe2*)
83. Gearyautocorrelation-lag3/weighted by atomic Sanderson electronegativities (*GATSe3*)
84. Geary autocorrelation-lag4/weighted by atomic Sanderson electronegativities (*GATSe4*)
85. Geary autocorrelation-lag5/weighted by atomic Sanderson electronegativities (*GATSe5*)
86. Geary autocorrelation-lag6/weighted by atomic Sanderson electronegativities (*GATSe6*)
87. Geary autocorrelation-lag7/weighted by atomic Sanderson electronegativities (*GATSe7*)
88. Geary autocorrelation-lag8/weighted by atomic Sanderson electronegativities (*GATSe8*)
89. Geary autocorrelation-lag1/weighted by atomic polarizabilities (*GATSp1*)
90. Geary autocorrelation-lag2/weighted by atomic polarizabilities (*GATSp2*)
91. Geary autocorrelation-lag3/weighted by atomic polarizabilities (*GATSp3*)
92. Geary autocorrelation-lag4/weighted by atomic polarizabilities (*GATSp4*)
93. Geary autocorrelation-lag5/weighted by atomic polarizabilities (*GATSp5*)
94. Geary autocorrelation-lag6/weighted by atomic polarizabilities (*GATSp6*)
95. Geary autocorrelation-lag7/weighted by atomic polarizabilities (*GATSp7*)
96. Geary autocorrelation-lag8/weighted by atomic polarizabilities (*GATSp8*)

9. Charge descriptors

1. Most positive charge on H atoms (Q_{Hmax})
2. Most positive charge on C atoms (Q_{Cmax})
3. Most positive charge on N atoms (Q_{Nmax})
4. Most positive charge on O atoms (Q_{Omax})
5. Most negative charge on H atoms (Q_{Hmin})
6. Most negative charge on C atoms (Q_{Cmin})
7. Most negative charge on N atoms (Q_{Nmin})
8. Most negative charge on O atoms (Q_{Omin})
9. Most positive charge in a molecule (Q_{max})
10. Most negative charge in a molecule (Q_{min})
11. Sum of squares of charges on H atoms (Q_{HSS})
12. Sum of squares of charges on C atoms (Q_{CSS})
13. Sum of squares of charges on N atoms (Q_{NSS})
14. Sum of squares of charges on O atoms (Q_{OSS})
15. Sum of squares of charges on all atoms (Q_{aSS})
16. Mean of positive charges (*Mpc*)
17. Total of positive charges (*Tpc*)
18. Mean of negative charges (*Mnc*)
19. Total of negative charges (*Tnc*)
20. Mean of absolute charges (*Mac*)
21. Total of absolute charges (*Tac*)
22. Relative positive charge (*Rpc*)
23. Relative negative charge (*Rnc*)
24. Submolecular polarity parameter (*SPP*)
25. Local dipole index (*LDI*)

Introduction:

These are electronic descriptors defined in terms of atomic charges and used to describe electronic aspects of the whole molecule and of particular regions, such as atoms, bonds and molecular fragments. Charge descriptors are calculated by computational chemistry and therefore can be considered among quantum chemical descriptors.

Electrical charges in the molecule are the driving force of electrostatic interactions, and it is well known the local electron density or charge plays a fundamental role in many chemical reactions and physic-chemical properties.

Some most used charge descriptors are displayed here as followed:

(1) Most positive charge in a molecule (Q_{max})

The maximum positive charge of the atoms in a molecule:

$$Q_{\max} = \max_a (q_a^+)$$

where q^+ are net atom positive charges

(2) Most negative charge in a molecule (Q_{min})

The maximum negative charge of the atoms in a molecule:

$$Q_{\min} = \max_a (q_a^-)$$

where q^- are net atom negative charges

(3) Total of positive charges (*Tpc*)

The sum of all of the positive charges of the atoms in a molecule:

$$Tpc = \sum_a (q_a^+)$$

where q^+ are net atom positive charges

(4) Total of negative charges (*Tnc*)

The sum of all of the negative charges of the atoms in a molecule:

$$Tnc = \sum_a (q_a^-)$$

where q^- are net atom negative charges

10. Molecular properties

1. Molar refractivity (*MREF*)
2. LogP value based on the Crippen method (*logP*)
3. Square of LogP value based on the Crippen method (*logP2*)
4. Topological polarity surface area (*TPSA*)
5. Unsaturation index (*UI*)
6. Hydrophilic index (*Hy*)

Introduction:

(1) Molar refractivity (*MREF*)

Molecular descriptor of a liquid which contains both information about molecular volume and polarizability, usually defined by the Lorenz-Lorentz equation:

$$MR = \frac{n^2 - 1}{n^2 + 2} \frac{MW}{\rho}$$

where MW is the molecular weight, ρ is the liquid density, and n the refractive index of the liquid.

（2） LogP value based on the Crippen method (*logP*)

The Ghose-Crippen contribution method is based on hydrophobic atomic constants a_k measuring the lipophilic contributions of atoms in the molecule, each described by its neighbouring atoms.

$$LogP = \sum_k a_k N_k$$

where N_k is the occurrence of the *kth* atom type

（3） Topological polarity surface area (*TPSA*)

It is the sum of solvent-accessible surface areas of atoms with absolute value of partial charges greater than or equal to 0.2.

$$TPSA = \sum_a SA_a$$
$$|q_a| \geq 0.2$$

（4） Unsaturation index (*UI*)

The unsaturation index (*UI*) is defined as

$$UI = \log_2(1 + nDB + nTB + nAB)$$

where nDB=the number of double bonds, nTB=the number of triple bonds and nAB=the number of aromatic bonds. The unsaturation index is described in the user manual for Dragon.

(5) Hydrophilic index (*Hy*)

The hydrophilic index is given by

$$Hy = \frac{(1+N_{Hy})\log_2(1+N_{Hy}) + N_c(\frac{1}{A}\log_2\frac{1}{A}) + \sqrt{\frac{N_{Hy}}{A^2}}}{\log_2(1+A)}$$

where N_{Hy} is the number of hydrophilic groups (or the total number of hydrogen attached to oxygen, sulfur and nitrogen atoms), N_c is the number of carbon atoms, and A is the number of non hydrogen atoms. The hydrophilic index is described in more detail on page 225 of the Handbook of Molecular Descriptors (Todeschini and Consonni 2000).

11. MOE-type descriptors

1. topological polar surface area based on fragments (*TPSA*)
2. Labute's Approximate Surface Area (*LabuteASA*)
3. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA1*)
4. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA2*)
5. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA3*)
6. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA4*)
7. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA5*)
8. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA6*)
9. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA7*)
10. MOE-type descriptors using SLogP contributions and surface area contributions (*SLOGPVSA8*)
11. MOE-type descriptors using SLogP contributions and surface area contributions(*SLOGPVSA9*)
12. MOE-type descriptors using SLogP contributions and surface area contributions(*SLOGPVSA10*)
13. MOE-type descriptors using SLogP contributions and surface area contributions(*SLOGPVSA11*)
14. MOE-type descriptors using SLogP contributions and surface area contributions(*SLOGPVSA12*)
15. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA1*)
16. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA2*)
17. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA3*)
18. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA4*)
19. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA5*)
20. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA6*)
21. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA7*)

22. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA8*)
23. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA9*)
24. MOE-type descriptors using MR contributions and surface area contributions (*SMRVSA10*)
25. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA1*)
26. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA2*)
27. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA3*)
28. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA4*)
29. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA5*)
30. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA6*)
31. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA7*)
32. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA8*)
33. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA9*)
34. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA10*)
35. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA11*)
36. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA12*)
37. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA13*)
38. MOE-type descriptors using partial charges and surface area contributions (*PEOEVSA14*)
39. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA1*)
40. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA2*)
41. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA3*)
42. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA4*)
43. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA5*)
44. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA6*)
45. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA7*)
46. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA8*)
47. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA9*)
48. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA10*)
49. MOE-type descriptors using Estate indices and surface area contributions (*EstateVSA11*)
50. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate1*)
51. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate2*)
52. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate3*)
53. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate4*)
54. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate5*)

55. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate6*)
56. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate7*)
57. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate8*)
58. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate9*)
59. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate10*)
60. MOE-type descriptors using surface area contributions and Estate indices (*VSAEstate11*)

12. Geometric descriptors

1. 3-D Wiener index based gemetrical distance matrix (including Hs) (*W3DH*)
2. 3-D Wiener index based geometrical distance matrix (Not including Hs) (*W3D*)
3. Petitjean Index based on molecular geometrical distance matrix (*Petitj3D*)
4. The longest distance between two atoms (geometrical diameter) (*GeDi*)
5. Gravitational 3D index (*grav1*)
6. Radius of gyration (*rygr*)
7. 3D-Harary index (*Harary3D*)
8. The average geometric distance degree (*AGDD*)
9. The absolute eigenvalue sum on geometry matrix (*SEig*)
10. The span R (*SPAN*)
11. The average span R (*ASPAN*)
12. The molecular eccentricity (*MEcc*)

13. CPSA descriptors

1. partial negative surface area (*PNSA1*)
2. partial positive surface area (*PPSA1*)
3. total charge weighted negative surface area (*PNSA2*)
4. total charge weighted positive surface area (*PPSA2*)
5. atomic charge weighted negative surface area (*PNSA3*)
6. atomic charge weighted positive surface area (*PPSA3*)
7. difference in charged partial surface area (*DPSA1*)
8. difference in total charge weighted surface area (*DPSA2*)
9. difference in atomic charge weighted surface area (*DPSA3*)

10-12. fractional charged partial negative surface areas (*FNSA1-3*)

13-15. fractional charged partial positive surface areas (*FPSA1-3*)

16-18. surface weighted charged partial negative surface areas (*WNSA1-3*)

19-21. surface weighted charged partial positive surface areas (*WPSA1-3*)

22. relative negative charge surface area (*RNCS*)

23. relative positive charge surface area (*RPCS*)

24. total hydrophobic surface area (*TASA*)

25. total polar surface area (*PSA*)

26. relative hydrophobic surface area (*RASA*)

27. relative polar surface area (*RPSA*)

28. The fraction between TASA and TPSA (*FrTATP*)

29. Solvent-accessible surface areas (*ASA*)

30. Molecular surface areas (*MSA*)

Introduction:

1. partial negative surface area

It is the sum of the solvent-accessible surface areas of all negatively charged atoms, that is,

$$PNSA_1 = \sum_{a-} SA_a^-$$

where the sum is restricted to negatively charged atoms a_.

2. partial positive surface area

It is the sum of the solvent-accessible surface areas of all positively charged atoms, that is,

$$PPSA_1 = \sum_{a+} SA_a^+$$

where the sum is restricted to positively charged atoms at.

3. total charge weighted negative surface area

It is the partial negative solvent-accessible surface area multiplied by the ! total negative charge Q_, that is,

$$PNSA_2 = Q^- \cdot \sum_{a-} SA_a^-$$

4. total charge weighted positive surface area

It is the partial positive solvent-accessible surface area multiplied by the ! total positive charge Qt, that is,

$$PPSA_2 = Q^+ \cdot \sum_{a+} SA_a^+$$

5. atomic charge weighted negative surface area

It is the sum of the product of atomic solvent-accessible surface area by the partial charge q_ a over all negatively charged atoms, that is,

$$PNSA_3 = \sum_{a-} q_a^- SA_a^-$$

6. atomic charge weighted positive surface area

It is the sum of the product of atomic solvent-accessible surface area by the partial charge qt a over all positively charged atoms, that is,

$$PPSA_3 = \sum_{a+} q_a^+ SA_a^+$$

7. difference in charged partial surface area

It is the partial positive solvent-accessible surface area minus the partial negative solventaccessible surface area, that is,

$$DPSA_1 = PPSA_1 - PNSA_1$$

8. difference in total charge weighted surface area

It is the total charge weighted positive solvent-accessible surface area minus the total charge weighted negative solvent-accessible surface area, that is,

$$DPSA_2 = PPSA_2 - PNSA_2$$

9. difference in atomic charge weighted surface area

It is the atomic charge weighted positive solvent-accessible surface area minus the atomic charge weighted negative solvent-accessible surface area, that is,

$$DPSA_3 = PPSA_3 - PNSA_3$$

10-12. fractional charged partial negative surface areas

They are the partial negative surface area (PNSA1), the total charge weighted negative surface area (PNSA2), and the atomic charge weighted negative surface area (PNSA3), divided by the

total molecular solvent-accessible surface area (SASA), that is,

$$FNSA_1 = \frac{PNSA_1}{SASA} \qquad FNSA_2 = \frac{PNSA_2}{SASA} \qquad FNSA_3 = \frac{PNSA_3}{SASA}$$

13-15. fractional charged partial positive surface areas

They are the partial positive surface area (PPSA1), the total charge weighted positive surface area (PPSA2), and the atomic charge weighted positive surface area (PPSA3), divided by the total molecular solvent-accessible surface area (SASA), that is,

$$FPSA_1 = \frac{PPSA_1}{SASA} \qquad FPSA_2 = \frac{PPSA_2}{SASA} \qquad FPSA_3 = \frac{PPSA_3}{SASA}$$

16-18. surface weighted charged partial negative surface areas

They are the partial negative surface area (PNSA1), the total charge weighted negative surface area (PNSA2), and the atomic charge weighted negative surface area (PNSA3), multiplied by the total molecular solvent-accessible surface area (SASA) and divided by 1000, that is,

$$WNSA_1 = \frac{PNSA_1 \cdot SASA}{1000} \qquad WNSA_2 = \frac{PNSA_2 \cdot SASA}{1000} \qquad WNSA_3 = \frac{PNSA_3 \cdot SASA}{1000}$$

19-21. surface weighted charged partial positive surface areas

They are the partial positive surface area (PPSA1), the total charge weighted positive surface area (PPSA2), and the atomic charge weighted positive surface area (PPSA3), multiplied by the total molecular solvent-accessible surface area (SASA) and divided by 1000, that is,

$$WPSA_1 = \frac{PPSA_1 \cdot SASA}{1000} \qquad WPSA_2 = \frac{PPSA_2 \cdot SASA}{1000} \qquad WPSA_3 = \frac{PPSA_3 \cdot SASA}{1000}$$

22. relative negative charge surface area

It is the solvent-accessible surface area of the most negative atom divided by the relative negative charge (RNCG), that is,

$$RNCS = \frac{SA^{-}_{\max}}{RNCG}$$

23. relative positive charge surface area

It is the solvent-accessible surface area of the most positive atom divided by the relative positive

charge (RPCG), that is,

$$RPCS = \frac{SA^{+}_{\max}}{RPCG}$$

24. total hydrophobic surface area

It is the sum of solvent-accessible surface areas of atoms with absolute value of partial charges less than 0.2, that is,

$$TASA = \sum_a SA_a \quad \forall a : |q_a| < 0.2$$

25. total polar surface area

It is the sum of solvent-accessible surface areas of atoms with absolute value of partial charges greater than or equal to 0.2.

$$TPSA = \sum_a SA_a \quad \forall a : |q_a| \geq 0.2$$

26. relative hydrophobic surface area

It is the total hydrophobic surface area (TASA) divided by the total molecular solvent-accessible surface area (SASA), that is,

$$RASA = \frac{TASA}{SASA}$$

27. relative polar surface area

It is the total polar surface area (TPSA) divided by the total molecular solvent-accessible surface area (SASA), that is,

$$RPSA = \frac{TPSA}{SASA}$$

28. The fraction between TASA and TPSA (FrTATP)

29. solvent-accessible surface areas (ASA)

30. molecular surface areas (MSA)

14. WHIM descriptors

1. 1st component size directional WHIM index / unweighted (*L1u*)
2. 2nd component size directional WHIM index / unweighted (*L2u*)
3. 3rd component size directional WHIM index / unweighted (*L3u*)
4. 1st component shape directional WHIM index / unweighted (*P1u*)
5. 2nd component shape directional WHIM index / unweighted (*P2u*)
6. 1st component symmetry directional WHIM index / unweighted (*G1u*)
7. 2nd component symmetry directional WHIM index / unweighted (*G2u*)
8. 3rd component symmetry directional WHIM index / unweighted (*G3u*)
9. 1st component accessibility directional WHIM index / unweighted (*E1u*)
10. 2nd component accessibility directional WHIM index / unweighted (*E2u*)
11. 3rd component accessibility directional WHIM index / unweighted (*E3u*)
12. T total size index / unweighted (*Tu*)
13. A total size index / unweighted (*Au*)
14. G total symmetry index / unweighted (*Gu*)
15. 1st component size directional WHIM index / weighted by atomic masses (*L1m*)
16. 2nd component size directional WHIM index / weighted by atomic masses (*L2m*)
17. 3rd component size directional WHIM index / weighted by atomic masses (*L3m*)
18. 1st component shape directional WHIM index / weighted by atomic masses (*P1m*)
19. 2nd component shape directional WHIM index / weighted by atomic masses (*P2m*)
20. 1st component symmetry directional WHIM index / weighted by atomic masses (*G1m*)
21. 2nd component symmetry directional WHIM index / weighted by atomic masses (*G2m*)
22. 3rd component symmetry directional WHIM index / weighted by atomic masses (*G3m*)
23. 1st component accessibility directional WHIM index / weighted by atomic masses (*E1m*)
24. 2nd component accessibility directional WHIM index / weighted by atomic masses (*E2m*)
25. 3rd component accessibility directional WHIM index / weighted by atomic masses (*E3m*)
26. T total size index / weighted by atomic masses (*Tm*)
27. A total size index / weighted by atomic masses (*Am*)
28. G total symmetry index / weighted by atomic masses (*Gm*)
29. 1st component size directional WHIM index / weighted by atomic van der Waals volumes (*L1v*)
30. 2nd component size directional WHIM index / weighted by atomic van der Waals volumes (*L2v*)
31. 3rd component size directional WHIM index / weighted by atomic van der Waals volumes (*L3v*)

32. 1st component shape directional WHIM index / weighted by atomic van der Waals volumes (*P1v*)
33. 2nd component shape directional WHIM index / weighted by atomic van der Waals volumes (*P2v*)
34. 1st component symmetry directional WHIM index / weighted by atomic van der Waals volumes (*G1v*)
35. 2nd component symmetry directional WHIM index / weighted by atomic van der Waals volumes (*G2v*)
36. 3rd component symmetry directional WHIM index / weighted by atomic van der Waals volumes (*G3v*)
37. 1st component accessibility directional WHIM index / weighted by atomic van der Waals volumes (*E1v*)
38. 2nd component accessibility directional WHIM index / weighted by atomic van der Waals volumes (*E2v*)
39. 3rd component accessibility directional WHIM index / weighted by atomic van der Waals volumes (*E3v*)
40. T total size index / weighted by atomic van der Waals volumes (*Tv*)
41. A total size index / weighted by atomic van der Waals volumes (*Av*)
42. G total symmetry index / weighted by atomic van der Waals volumes (*Gv*)
43. 1st component size directional WHIM index / weighted by atomic polarizabilities (*L1p*)
44. 2nd component size directional WHIM index / weighted by atomic polarizabilities (*L2p*)
45. 3rd component size directional WHIM index / weighted by atomic polarizabilities (*L3p*)
46. 1st component shape directional WHIM index / weighted by atomic polarizabilities (*P1p*)
47. 2nd component shape directional WHIM index / weighted by atomic polarizabilities (*P2p*)
48. 1st component symmetry directional WHIM index / weighted by atomic polarizabilities (*G1p*)
49. 2nd component symmetry directional WHIM index / weighted by atomic polarizabilities (*G2p*)
50. 3rd component symmetry directional WHIM index / weighted by atomic polarizabilities (*G3p*)
51. 1st component accessibility directional WHIM index / weighted by atomic polarizabilities (*E1p*)
52. 2nd component accessibility directional WHIM index / weighted by atomic polarizabilities (*E2p*)
53. 3rd component accessibility directional WHIM index / weighted by atomic polarizabilities (*E3p*)
54. T total size index / weighted by atomic polarizabilities (*Tp*)
55. A total size index / weighted by atomic polarizabilities (*Ap*)
56. G total symmetry index / weighted by atomic polarizabilities (*Gp*)
57. 1st component size directional WHIM index / weighted by atomic electrotopological states (*L1e*)
58. 2nd component size directional WHIM index / weighted by atomic electrotopological states (*L2e*)

59. 3rd component size directional WHIM index / weighted by atomic electrotopological states (*L3e*)
60. 1st component shape directional WHIM index / weighted by atomic electrotopological states (*P1e*)
61. 2nd component shape directional WHIM index / weighted by atomic electrotopological states (*P2e*)
62. 1st component symmetry directional WHIM index / weighted by atomic electrotopological states (*G1e*)
63. 2nd component symmetry directional WHIM index / weighted by atomic electrotopological states (*G2e*)
64. 3rd component symmetry directional WHIM index / weighted by atomic electrotopological states (*G3e*)
65. 1st component accessibility directional WHIM index / weighted by atomic electrotopological states (*E1e*)
66. 2nd component accessibility directional WHIM index / weighted by atomic electrotopological states (*E2e*)
67. 3rd component accessibility directional WHIM index / weighted by atomic electrotopological states (*E3e*)
68. T total size index / weighted by atomic electrotopological states (*Te*)
69. A total size index / weighted by atomic electrotopological states (*Ae*)
70. G total symmetry index / weighted by atomic electrotopological states (*Ge*)

15. MoRSE descriptors

1. 3D-MoRSE - signal 01 / unweighted (*MoRSEU1*)
2. 3D-MoRSE - signal 02 / unweighted (*MoRSEU2*)
3. 3D-MoRSE - signal 03 / unweighted (*MoRSEU3*)
4. 3D-MoRSE - signal 04 / unweighted (*MoRSEU4*)
5. 3D-MoRSE - signal 05 / unweighted (*MoRSEU5*)
6. 3D-MoRSE - signal 06 / unweighted (*MoRSEU6*)
7. 3D-MoRSE - signal 07 / unweighted (*MoRSEU7*)
8. 3D-MoRSE - signal 08 / unweighted (*MoRSEU8*)
9. 3D-MoRSE - signal 09 / unweighted (*MoRSEU9*)
10. 3D-MoRSE - signal 10 / unweighted (*MoRSEU10*)
11. 3D-MoRSE - signal 11 / unweighted (*MoRSEU11*)

12. 3D-MoRSE - signal 12 / unweighted (*MoRSEU12*)
13. 3D-MoRSE - signal 13 / unweighted (*MoRSEU13*)
14. 3D-MoRSE - signal 14 / unweighted (*MoRSEU14*)
15. 3D-MoRSE - signal 15 / unweighted (*MoRSEU15*)
16. 3D-MoRSE - signal 16 / unweighted (*MoRSEU16*)
17. 3D-MoRSE - signal 17/ unweighted(*MoRSEU17*)
18. 3D-MoRSE - signal 18 / unweighted (*MoRSEU18*)
19. 3D-MoRSE - signal19 / unweighted (*MoRSEU19*)
20. 3D-MoRSE - signal 20 / unweighted (*MoRSEU20*)
21. 3D-MoRSE - signal 21 / unweighted (*MoRSEU21*)
22. 3D-MoRSE - signal 22 / unweighted (*MoRSEU22*)
23. 3D-MoRSE - signal 23 / unweighted (*MoRSEU23*)
24. 3D-MoRSE - signal 24 / unweighted (*MoRSEU24*)
25. 3D-MoRSE - signal 25 / unweighted (*MoRSEU25*)
26. 3D-MoRSE - signal 26 / unweighted (*MoRSEU26*)
27. 3D-MoRSE - signal 27 / unweighted (*MoRSEU27*)
28. 3D-MoRSE - signal 28 / unweighted (*MoRSEU28*)
29. 3D-MoRSE - signal 29 / unweighted (*MoRSEU29*)
30. 3D-MoRSE - signal 30 / unweighted (*MoRSEU30*)
31. 3D-MoRSE - signal 01 / weighted by atomic masses (*MoRSEM1*)
32. 3D-MoRSE - signal 02 / weighted by atomic masses (*MoRSEM2*)
33. 3D-MoRSE - signal 03 / weighted by atomic masses (*MoRSEM3*)
34. 3D-MoRSE - signal 04 / weighted by atomic masses (*MoRSEM4*)
35. 3D-MoRSE - signal 05 / weighted by atomic masses (*MoRSEM5*)
36. 3D-MoRSE - signal 06 / weighted by atomic masses (*MoRSEM6*)
37. 3D-MoRSE - signal 07 / weighted by atomic masses (*MoRSEM7*)
38. 3D-MoRSE - signal 08 / weighted by atomic masses (*MoRSEM8*)
39. 3D-MoRSE - signal 09 / weighted by atomic masses (*MoRSEM9*)
40. 3D-MoRSE - signal 10 / weighted by atomic masses (*MoRSEM10*)
41. 3D-MoRSE - signal 11 / weighted by atomic masses (*MoRSEM11*)
42. 3D-MoRSE - signal 12 / weighted by atomic masses (*MoRSEM12*)
43. 3D-MoRSE - signal 13 / weighted by atomic masses (*MoRSEM13*)
44. 3D-MoRSE - signal 14 / weighted by atomic masses (*MoRSEM14*)

45. 3D-MoRSE - signal 15 / weighted by atomic masses (*MoRSEM15*)
46. 3D-MoRSE - signal 16 / weighted by atomic masses (*MoRSEM16*)
47. 3D-MoRSE - signal 17/ weighted by atomic masses (*MoRSEM17*)
48. 3D-MoRSE - signal 18 / weighted by atomic masses (*MoRSEM18*)
49. 3D-MoRSE - signal19 / weighted by atomic masses (*MoRSEM19*)
50. 3D-MoRSE - signal 20 / weighted by atomic masses (*MoRSEM20*)
51. 3D-MoRSE - signal 21 / weighted by atomic masses (*MoRSEM21*)
52. 3D-MoRSE - signal 22 / weighted by atomic masses (*MoRSEM22*)
53. 3D-MoRSE - signal 23 / weighted by atomic masses (*MoRSEM23*)
54. 3D-MoRSE - signal 24 / weighted by atomic masses (*MoRSEM24*)
55. 3D-MoRSE - signal 25 / weighted by atomic masses (*MoRSEM25*)
56. 3D-MoRSE - signal 26 / weighted by atomic masses (*MoRSEM26*)
57. 3D-MoRSE - signal 27 / weighted by atomic masses (*MoRSEM27*)
58. 3D-MoRSE - signal 28 / weighted by atomic masses (*MoRSEM28*)
59. 3D-MoRSE - signal 29 / weighted by atomic masses (*MoRSEM29*)
60. 3D-MoRSE - signal 30 / weighted by atomic masses (*MoRSEM30*)
61. 3D-MoRSE - signal 01 / weighted by atomic number (*MoRSEN1*)
62. 3D-MoRSE - signal 02 / weighted by atomic number (*MoRSEN2*)
63. 3D-MoRSE - signal 03 / weighted by atomic number (*MoRSEN3*)
64. 3D-MoRSE - signal 04 / weighted by atomic number (*MoRSEN4*)
65. 3D-MoRSE - signal 05 / weighted by atomic number (*MoRSEN5*)
66. 3D-MoRSE - signal 06 / weighted by atomic number (*MoRSEN6*)
67. 3D-MoRSE - signal 07 / weighted by atomic number (*MoRSEN7*)
68. 3D-MoRSE - signal 08 / weighted by atomic number (*MoRSEN8*)
69. 3D-MoRSE - signal 09 / weighted by atomic number (*MoRSEN9*)
70. 3D-MoRSE - signal 10 / weighted by atomic number (*MoRSEN10*)
71. 3D-MoRSE - signal 11 / weighted by atomic number (*MoRSEN11*)
72. 3D-MoRSE - signal 12 / weighted by atomic number (*MoRSEN12*)
73. 3D-MoRSE - signal 13 / weighted by atomic number (*MoRSEN13*)
74. 3D-MoRSE - signal 14 / weighted by atomic number (*MoRSEN14*)
75. 3D-MoRSE - signal 15 / weighted by atomic number (*MoRSEN15*)
76. 3D-MoRSE - signal 16 / weighted by atomic number (*MoRSEN16*)
77. 3D-MoRSE - signal 17/ weighted by atomic number (*MoRSEN17*)

78. 3D-MoRSE - signal 18 / weighted by atomic number (*MoRSEN18*)
79. 3D-MoRSE - signal19 / weighted by atomic number (*MoRSEN19*)
80. 3D-MoRSE - signal 20 / weighted by atomic number (*MoRSEN20*)
81. 3D-MoRSE - signal 21 / weighted by atomic number (*MoRSEN21*)
82. 3D-MoRSE - signal 22 / weighted by atomic number (*MoRSEN22*)
83. 3D-MoRSE - signal 23 / weighted by atomic number (*MoRSEN23*)
84. 3D-MoRSE - signal 24 / weighted by atomic number (*MoRSEN24*)
85. 3D-MoRSE - signal 25 / weighted by atomic number (*MoRSEN25*)
86. 3D-MoRSE - signal 26 / weighted by atomic number (*MoRSEN26*)
87. 3D-MoRSE - signal 27 / weighted by atomic number (*MoRSEN27*)
88. 3D-MoRSE - signal 28 / weighted by atomic number (*MoRSEN28*)
89. 3D-MoRSE - signal 29 / weighted by atomic number (*MoRSEN29*)
90. 3D-MoRSE - signal 30 / weighted by atomic number (*MoRSEN30*)
91. 3D-MoRSE - signal 01 / weighted by atomic van der Waals volumes (*MoRSEV1*)
92. 3D-MoRSE - signal 02 / weighted by atomic van der Waals volumes (*MoRSEV2*)
93. 3D-MoRSE - signal 03 / weighted by atomic van der Waals volumes (*MoRSEV3*)
94. 3D-MoRSE - signal 04 / weighted by atomic van der Waals volumes (*MoRSEV4*)
95. 3D-MoRSE - signal 05 / weighted by atomic van der Waals volumes (*MoRSEV5*)
96. 3D-MoRSE - signal 06 / weighted by atomic van der Waals volumes (*MoRSEV6*)
97. 3D-MoRSE - signal 07 / weighted by atomic van der Waals volumes (*MoRSEV7*)
98. 3D-MoRSE - signal 08 / weighted by atomic van der Waals volumes (*MoRSEV8*)
99. 3D-MoRSE - signal 09 / weighted by atomic van der Waals volumes (*MoRSEV9*)
100. 3D-MoRSE - signal 10 / weighted by atomic van der Waals volumes (*MoRSEV10*)
101. 3D-MoRSE - signal 11 / weighted by atomic van der Waals volumes (*MoRSEV11*)
102. 3D-MoRSE - signal 12 / weighted by atomic van der Waals volumes (*MoRSEV12*)
103. 3D-MoRSE - signal 13 / weighted by atomic van der Waals volumes (*MoRSEV13*)
104. 3D-MoRSE - signal 14 / weighted by atomic van der Waals volumes (*MoRSEV14*)
105. 3D-MoRSE - signal 15 / weighted by atomic van der Waals volumes (*MoRSEV15*)
106. 3D-MoRSE - signal 16 / weighted by atomic van der Waals volumes (*MoRSEV16*)
107. 3D-MoRSE - signal 17/ weighted by atomic van der Waals volumes (*MoRSEV17*)
108. 3D-MoRSE - signal 18 / weighted by atomic van der Waals volumes (*MoRSEV18*)
109. 3D-MoRSE - signal19 / weighted by atomic van der Waals volumes (*MoRSEV19*)
110. 3D-MoRSE - signal 20 / weighted by atomic van der Waals volumes (*MoRSEV20*)

111. 3D-MoRSE - signal 21 / weighted by atomic van der Waals volumes (*MoRSEV21*)
112. 3D-MoRSE - signal 22 / weighted by atomic van der Waals volumes (*MoRSEV22*)
113. 3D-MoRSE - signal 23 / weighted by atomic van der Waals volumes (*MoRSEV23*)
114. 3D-MoRSE - signal 24 / weighted by atomic van der Waals volumes (*MoRSEV24*)
115. 3D-MoRSE - signal 25 / weighted by atomic van der Waals volumes (*MoRSEV25*)
116. 3D-MoRSE - signal 26 / weighted by atomic van der Waals volumes (*MoRSEV26*)
117. 3D-MoRSE - signal 27 / weighted by atomic van der Waals volumes (*MoRSEV27*)
118. 3D-MoRSE - signal 28 / weighted by atomic van der Waals volumes (*MoRSEV28*)
119. 3D-MoRSE - signal 29 / weighted by atomic van der Waals volumes (*MoRSEV29*)
120. 3D-MoRSE - signal 30 / weighted by atomic van der Waals volumes (*MoRSEV30*)
121. 3D-MoRSE - signal 01 / weighted by atomic Sanderson electronegativities (*MoRSEE1*)
122. 3D-MoRSE - signal 02 / weighted by atomic Sanderson electronegativities (*MoRSEE2*)
123. 3D-MoRSE - signal 03 / weighted by atomic Sanderson electronegativities (*MoRSEE3*)
124. 3D-MoRSE - signal 04 / weighted by atomic Sanderson electronegativities (*MoRSEE4*)
125. 3D-MoRSE - signal 05 / weighted by atomic Sanderson electronegativities (*MoRSEE5*)
126. 3D-MoRSE - signal 06 / weighted by atomic Sanderson electronegativities (*MoRSEE6*)
127. 3D-MoRSE - signal 07 / weighted by atomic Sanderson electronegativities (*MoRSEE7*)
128. 3D-MoRSE - signal 08 / weighted by atomic Sanderson electronegativities (*MoRSEE8*)
129. 3D-MoRSE - signal 09 / weighted by atomic Sanderson electronegativities (*MoRSEE9*)
130. 3D-MoRSE - signal 10 / weighted by atomic Sanderson electronegativities (*MoRSEE10*)
131. 3D-MoRSE - signal 11 / weighted by atomic Sanderson electronegativities (*MoRSEE11*)
132. 3D-MoRSE - signal 12 / weighted by atomic Sanderson electronegativities (*MoRSEE12*)
133. 3D-MoRSE - signal 13 / weighted by atomic Sanderson electronegativities (*MoRSEE13*)
134. 3D-MoRSE - signal 14 / weighted by atomic Sanderson electronegativities (*MoRSEE14*)
135. 3D-MoRSE - signal 15 / weighted by atomic Sanderson electronegativities (*MoRSEE15*)
136. 3D-MoRSE - signal 16 / weighted by atomic Sanderson electronegativities (*MoRSEE16*)
137. 3D-MoRSE - signal 17/ weighted by atomic Sanderson electronegativities (*MoRSEE17*)
138. 3D-MoRSE - signal 18 / weighted by atomic Sanderson electronegativities (*MoRSEE18*)
139. 3D-MoRSE - signal19 / weighted by atomic Sanderson electronegativities (*MoRSEU19*)
140. 3D-MoRSE - signal 20 / weighted by atomic Sanderson electronegativities (*MoRSEE20*)
141. 3D-MoRSE - signal 21 / weighted by atomic Sanderson electronegativities (*MoRSEE21*)
142. 3D-MoRSE - signal 22 / weighted by atomic Sanderson electronegativities (*MoRSEE22*)
143. 3D-MoRSE - signal 23 / weighted by atomic Sanderson electronegativities (*MoRSEE23*)

144. 3D-MoRSE - signal 24 / weighted by atomic Sanderson electronegativities (*MoRSEE24*)
145. 3D-MoRSE - signal 25 / weighted by atomic Sanderson electronegativities (*MoRSEE25*)
146. 3D-MoRSE - signal 26 / weighted by atomic Sanderson electronegativities (*MoRSEE26*)
147. 3D-MoRSE - signal 27 / weighted by atomic Sanderson electronegativities (*MoRSEE27*)
148. 3D-MoRSE - signal 28 / weighted by atomic Sanderson electronegativities (*MoRSEE28*)
149. 3D-MoRSE - signal 29 / weighted by atomic Sanderson electronegativities (*MoRSEE29*)
150. 3D-MoRSE - signal 30 / weighted by atomic Sanderson electronegativities (*MoRSEE30*)
151. 3D-MoRSE - signal 01 / weighted by atomic polarizabilities (*MoRSEP1*)
152. 3D-MoRSE - signal 02 / weighted by atomic polarizabilities (*MoRSEP2*)
153. 3D-MoRSE - signal 03 / weighted by atomic polarizabilities (*MoRSEP3*)
154. 3D-MoRSE - signal 04 / weighted by atomic polarizabilities (*MoRSEP4*)
155. 3D-MoRSE - signal 05 / weighted by atomic polarizabilities (*MoRSEP5*)
156. 3D-MoRSE - signal 06 / weighted by atomic polarizabilities (*MoRSEP6*)
157. 3D-MoRSE - signal 07 / weighted by atomic polarizabilities (*MoRSEP7*)
158. 3D-MoRSE - signal 08 / weighted by atomic polarizabilities (*MoRSEP8*)
159. 3D-MoRSE - signal 09 / weighted by atomic polarizabilities (*MoRSEP9*)
160. 3D-MoRSE - signal 10 / weighted by atomic polarizabilities (*MoRSEP10*)
161. 3D-MoRSE - signal 11 / weighted by atomic polarizabilities (*MoRSEP11*)
162. 3D-MoRSE - signal 12 / weighted by atomic polarizabilities (*MoRSEP12*)
163. 3D-MoRSE - signal 13 / weighted by atomic polarizabilities (*MoRSEP13*)
164. 3D-MoRSE - signal 14 / weighted by atomic polarizabilities (*MoRSEP14*)
165. 3D-MoRSE - signal 15 / weighted by atomic polarizabilities (*MoRSEP15*)
166. 3D-MoRSE - signal 16 / weighted by atomic polarizabilities (*MoRSEP16*)
167. 3D-MoRSE - signal 17/ weighted by atomic polarizabilities (*MoRSEP17*)
168. 3D-MoRSE - signal 18 / weighted by atomic polarizabilities (*MoRSEP18*)
169. 3D-MoRSE - signal19 / weighted by atomic polarizabilities (*MoRSEP19*)
170. 3D-MoRSE - signal 20 / weighted by atomic polarizabilities (*MoRSEP20*)
171. 3D-MoRSE - signal 21 / weighted by atomic polarizabilities (*MoRSEP21*)
172. 3D-MoRSE - signal 22 / weighted by atomic polarizabilities (*MoRSEP22*)
173. 3D-MoRSE - signal 23 / weighted by atomic polarizabilities (*MoRSEP23*)
174. 3D-MoRSE - signal 24 / weighted by atomic polarizabilities (*MoRSEP24*)
175. 3D-MoRSE - signal 25 / weighted by atomic polarizabilities (*MoRSEP25*)
176. 3D-MoRSE - signal 26 / weighted by atomic polarizabilities (*MoRSEP26*)

177. 3D-MoRSE - signal 27 / weighted by atomic polarizabilities (*MoRSEP27*)
178. 3D-MoRSE - signal 28 / weighted by atomic polarizabilities (*MoRSEP28*)
179. 3D-MoRSE - signal 29 / weighted by atomic polarizabilities (*MoRSEP29*)
180. 3D-MoRSE - signal 30 / weighted by atomic polarizabilities (*MoRSEP30*)
181. 3D-MoRSE - signal 01 / weighted by atomic charge (*MoRSEC1*)
182. 3D-MoRSE - signal 02 / weighted by atomic charge (*MoRSEC2*)
183. 3D-MoRSE - signal 03 / weighted by atomic charge (*MoRSEC3*)
184. 3D-MoRSE - signal 04 / weighted by atomic charge (*MoRSEC4*)
185. 3D-MoRSE - signal 05 / weighted by atomic charge (*MoRSEC5*)
186. 3D-MoRSE - signal 06 / weighted by atomic charge (*MoRSEC6*)
187. 3D-MoRSE - signal 07 / weighted by atomic charge (*MoRSEC7*)
188. 3D-MoRSE - signal 08 / weighted by atomic charge (*MoRSEC8*)
189. 3D-MoRSE - signal 09 / weighted by atomic charge (*MoRSEC9*)
190. 3D-MoRSE - signal 10 / weighted by atomic charge (*MoRSEC10*)
191. 3D-MoRSE - signal 11 / weighted by atomic charge (*MoRSEC11*)
192. 3D-MoRSE - signal 12 / weighted by atomic charge (*MoRSEC12*)
193. 3D-MoRSE - signal 13 / weighted by atomic charge (*MoRSEC13*)
194. 3D-MoRSE - signal 14 / weighted by atomic charge (*MoRSEC14*)
195. 3D-MoRSE - signal 15 / weighted by atomic charge (*MoRSEC15*)
196. 3D-MoRSE - signal 16 / weighted by atomic charge (*MoRSEC16*)
197. 3D-MoRSE - signal 17/ weighted by atomic charge (*MoRSEC17*)
198. 3D-MoRSE - signal 18 / weighted by atomic charge (*MoRSEC18*)
199. 3D-MoRSE - signal19 / weighted by atomic charge (*MoRSEC19*)
200. 3D-MoRSE - signal 20 / weighted by atomic charge (*MoRSEC20*)
201. 3D-MoRSE - signal 21 / weighted by atomic charge (*MoRSEC21*)
202. 3D-MoRSE - signal 22 / weighted by atomic charge (*MoRSEC22*)
203. 3D-MoRSE - signal 23 / weighted by atomic charge (*MoRSEC23*)
204. 3D-MoRSE - signal 24 / weighted by atomic charge (*MoRSEC24*)
205. 3D-MoRSE - signal 25 / weighted by atomic charge (*MoRSEC25*)
206. 3D-MoRSE - signal 26 / weighted by atomic charge (*MoRSEC26*)
207. 3D-MoRSE - signal 27 / weighted by atomic charge (*MoRSEC27*)
208. 3D-MoRSE - signal 28 / weighted by atomic charge (*MoRSEC28*)
209. 3D-MoRSE - signal 29 / weighted by atomic charge (*MoRSEC29*)

210. 3D-MoRSE - signal 30 / weighted by atomic charge (*MoRSEC30*)

16 RDF descriptors

1. 3D-RDF - signal 01 / unweighted (*RDFU1*)
2. 3D-RDF - signal 02 / unweighted (*RDFU2*)
3. 3D-RDF - signal 03 / unweighted (*RDFU3*)
4. 3D-RDF - signal 04 / unweighted (*RDFU4*)
5. 3D-RDF - signal 05 / unweighted (*RDFU5*)
6. 3D-RDF - signal 06 / unweighted (*RDFU6*)
7. 3D-RDF - signal 07 / unweighted (*RDFU7*)
8. 3D-RDF - signal 08 / unweighted (*RDFU8*)
9. 3D-RDF - signal 09 / unweighted (*RDFU9*)
10. 3D-RDF - signal 10 / unweighted (*RDFU10*)
11. 3D-RDF - signal 11 / unweighted (*RDFU11*)
12. 3D-RDF - signal 12 / unweighted (*RDFU12*)
13. 3D-RDF - signal 13 / unweighted (*RDFU13*)
14. 3D-RDF - signal 14 / unweighted (*RDFU14*)
15. 3D-RDF - signal 15 / unweighted (*RDFU15*)
16. 3D-RDF - signal 16 / unweighted (*RDFU16*)
17. 3D-RDF - signal 17/ unweighted(*RDFU17*)
18. 3D-RDF - signal 18 / unweighted (*RDFU18*)
19. 3D-RDF - signal19 / unweighted (*RDFU19*)
20. 3D-RDF - signal 20 / unweighted (*RDFU20*)
21. 3D-RDF - signal 21 / unweighted (*RDFU21*)
22. 3D-RDF - signal 22 / unweighted (*RDFU22*)
23. 3D-RDF - signal 23 / unweighted (*RDFU23*)
24. 3D-RDF - signal 24 / unweighted (*RDFU24*)
25. 3D-RDF - signal 25 / unweighted (*RDFU25*)
26. 3D-RDF - signal 26 / unweighted (*RDFU26*)

27. 3D-RDF - signal 27 / unweighted (*RDFU27*)
28. 3D-RDF - signal 28 / unweighted (*RDFU28*)
29. 3D-RDF - signal 29 / unweighted (*RDFU29*)
30. 3D-RDF - signal 30 / unweighted (*RDFU30*)
31. 3D-RDF - signal 01 / weighted by atomic masses (*RDFM1*)
32. 3D-RDF - signal 02 / weighted by atomic masses (*RDFM2*)
33. 3D-RDF - signal 03 / weighted by atomic masses (*RDFM3*)
34. 3D-RDF - signal 04 / weighted by atomic masses (*RDFM4*)
35. 3D-RDF - signal 05 / weighted by atomic masses (*RDFM5*)
36. 3D-RDF - signal 06 / weighted by atomic masses (*RDFM6*)
37. 3D-RDF - signal 07 / weighted by atomic masses (*RDFM7*)
38. 3D-RDF - signal 08 / weighted by atomic masses (*RDFM8*)
39. 3D-RDF - signal 09 / weighted by atomic masses (*RDFM9*)
40. 3D-RDF - signal 10 / weighted by atomic masses (*RDFM10*)
41. 3D-RDF - signal 11 / weighted by atomic masses (*RDFM11*)
42. 3D-RDF - signal 12 / weighted by atomic masses (*RDFM12*)
43. 3D-RDF - signal 13 / weighted by atomic masses (*RDFM13*)
44. 3D-RDF - signal 14 / weighted by atomic masses (*RDFM14*)
45. 3D-RDF - signal 15 / weighted by atomic masses (*RDFM15*)
46. 3D-RDF - signal 16 / weighted by atomic masses (*RDFM16*)
47. 3D-RDF - signal 17/ weighted by atomic masses (*RDFM17*)
48. 3D-RDF - signal 18 / weighted by atomic masses (*RDFM18*)
49. 3D-RDF - signal19 / weighted by atomic masses (*RDFM19*)
50. 3D-RDF - signal 20 / weighted by atomic masses (*RDFM20*)
51. 3D-RDF - signal 21 / weighted by atomic masses (*RDFM21*)
52. 3D-RDF - signal 22 / weighted by atomic masses (*RDFM22*)
53. 3D-RDF - signal 23 / weighted by atomic masses (*RDFM23*)
54. 3D-RDF - signal 24 / weighted by atomic masses (*RDFM24*)
55. 3D-RDF - signal 25 / weighted by atomic masses (*RDFM25*)
56. 3D-RDF - signal 26 / weighted by atomic masses (*RDFM26*)
57. 3D-RDF - signal 27 / weighted by atomic masses (*RDFM27*)
58. 3D-RDF - signal 28 / weighted by atomic masses (*RDFM28*)
59. 3D-RDF - signal 29 / weighted by atomic masses (*RDFM29*)

60. 3D-RDF - signal 30 / weighted by atomic masses (*RDFM30*)
61. 3D-RDF - signal 01 / weighted by atomic van der Waals volumes (*RDFV1*)
62. 3D-RDF - signal 02 / weighted by atomic van der Waals volumes (*RDFV2*)
63. 3D-RDF - signal 03 / weighted by atomic van der Waals volumes (*RDFV3*)
64. 3D-RDF - signal 04 / weighted by atomic van der Waals volumes (*RDFV4*)
65. 3D-RDF - signal 05 / weighted by atomic van der Waals volumes (*RDFV5*)
66. 3D-RDF - signal 06 / weighted by atomic van der Waals volumes (*RDFV6*)
67. 3D-RDF - signal 07 / weighted by atomic van der Waals volumes (*RDFV7*)
68. 3D-RDF - signal 08 / weighted by atomic van der Waals volumes (*RDFV8*)
69. 3D-RDF - signal 09 / weighted by atomic van der Waals volumes (*RDFV9*)
70. 3D-RDF - signal 10 / weighted by atomic van der Waals volumes (*RDFV10*)
71. 3D-RDF - signal 11 / weighted by atomic van der Waals volumes (*RDFV11*)
72. 3D-RDF - signal 12 / weighted by atomic van der Waals volumes (*RDFV12*)
73. 3D-RDF - signal 13 / weighted by atomic van der Waals volumes (*RDFV13*)
74. 3D-RDF - signal 14 / weighted by atomic van der Waals volumes (*RDFV14*)
75. 3D-RDF - signal 15 / weighted by atomic van der Waals volumes (*RDFV15*)
76. 3D-RDF - signal 16 / weighted by atomic van der Waals volumes (*RDFV16*)
77. 3D-RDF - signal 17/ weighted by atomic van der Waals volumes (*RDFV17*)
78. 3D-RDF - signal 18 / weighted by atomic van der Waals volumes (*RDFV18*)
79. 3D-RDF - signal19 / weighted by atomic van der Waals volumes (*RDFV19*)
80. 3D-RDF - signal 20 / weighted by atomic van der Waals volumes (*RDFV20*)
81. 3D-RDF - signal 21 / weighted by atomic van der Waals volumes (*RDFV21*)
82. 3D-RDF - signal 22 / weighted by atomic van der Waals volumes (*RDFV22*)
83. 3D-RDF - signal 23 / weighted by atomic van der Waals volumes (*RDFV23*)
84. 3D-RDF - signal 24 / weighted by atomic van der Waals volumes (*RDFV24*)
85. 3D-RDF - signal 25 / weighted by atomic van der Waals volumes (*RDFV25*)
86. 3D-RDF - signal 26 / weighted by atomic van der Waals volumes (*RDFV26*)
87. 3D-RDF - signal 27 / weighted by atomic van der Waals volumes (*RDFV27*)
88. 3D-RDF - signal 28 / weighted by atomic van der Waals volumes (*RDFV28*)
89. 3D-RDF - signal 29 / weighted by atomic van der Waals volumes (*RDFV29*)
90. 3D-RDF - signal 30 / weighted by atomic van der Waals volumes (*RDFV30*)
91. 3D-RDF - signal 01 / weighted by atomic Sanderson electronegativities (*RDFE1*)
92. 3D-RDF - signal 02 / weighted by atomic Sanderson electronegativities (*RDFE2*)

93. 3D-RDF - signal 03 / weighted by atomic Sanderson electronegativities (*RDFE3*)
94. 3D-RDF - signal 04 / weighted by atomic Sanderson electronegativities (*RDFE4*)
95. 3D-RDF - signal 05 / weighted by atomic Sanderson electronegativities (*RDFE5*)
96. 3D-RDF - signal 06 / weighted by atomic Sanderson electronegativities (*RDFE6*)
97. 3D-RDF - signal 07 / weighted by atomic Sanderson electronegativities (*RDFE7*)
98. 3D-RDF - signal 08 / weighted by atomic Sanderson electronegativities (*RDFE8*)
99. 3D-RDF - signal 09 / weighted by atomic Sanderson electronegativities (*RDFE9*)
100. 3D-RDF - signal 10 / weighted by atomic Sanderson electronegativities (*RDFE10*)
101. 3D-RDF - signal 11 / weighted by atomic Sanderson electronegativities (*RDFE11*)
102. 3D-RDF - signal 12 / weighted by atomic Sanderson electronegativities (*RDFE12*)
103. 3D-RDF - signal 13 / weighted by atomic Sanderson electronegativities (*RDFE13*)
104. 3D-RDF - signal 14 / weighted by atomic Sanderson electronegativities (*RDFE14*)
105. 3D-RDF - signal 15 / weighted by atomic Sanderson electronegativities (*RDFE15*)
106. 3D-RDF - signal 16 / weighted by atomic Sanderson electronegativities (*RDFE16*)
107. 3D-RDF - signal 17/ weighted by atomic Sanderson electronegativities (*RDFE17*)
108. 3D-RDF - signal 18 / weighted by atomic Sanderson electronegativities (*RDFE18*)
109. 3D-RDF - signal19 / weighted by atomic Sanderson electronegativities (*RDFU19*)
110. 3D-RDF - signal 20 / weighted by atomic Sanderson electronegativities (*RDFE20*)
111. 3D-RDF - signal 21 / weighted by atomic Sanderson electronegativities (*RDFE21*)
112. 3D-RDF - signal 22 / weighted by atomic Sanderson electronegativities (*RDFE22*)
113. 3D-RDF - signal 23 / weighted by atomic Sanderson electronegativities (*RDFE23*)
114. 3D-RDF - signal 24 / weighted by atomic Sanderson electronegativities (*RDFE24*)
115. 3D-RDF - signal 25 / weighted by atomic Sanderson electronegativities (*RDFE25*)
116. 3D-RDF - signal 26 / weighted by atomic Sanderson electronegativities (*RDFE26*)
117. 3D-RDF - signal 27 / weighted by atomic Sanderson electronegativities (*RDFE27*)
118. 3D-RDF - signal 28 / weighted by atomic Sanderson electronegativities (*RDFE28*)
119. 3D-RDF - signal 29 / weighted by atomic Sanderson electronegativities (*RDFE29*)
120. 3D-RDF - signal 30 / weighted by atomic Sanderson electronegativities (*RDFE30*)
121. 3D-RDF - signal 01 / weighted by atomic polarizabilities (*RDFP1*)
122. 3D-RDF - signal 02 / weighted by atomic polarizabilities (*RDFP2*)
123. 3D-RDF - signal 03 / weighted by atomic polarizabilities (*RDFP3*)
124. 3D-RDF - signal 04 / weighted by atomic polarizabilities (*RDFP4*)
125. 3D-RDF - signal 05 / weighted by atomic polarizabilities (*RDFP5*)

126. 3D-RDF - signal 06 / weighted by atomic polarizabilities (*RDFP6*)
127. 3D-RDF - signal 07 / weighted by atomic polarizabilities (*RDFP7*)
128. 3D-RDF - signal 08 / weighted by atomic polarizabilities (*RDFP8*)
129. 3D-RDF - signal 09 / weighted by atomic polarizabilities (*RDFP9*)
130. 3D-RDF - signal 10 / weighted by atomic polarizabilities (*RDFP10*)
131. 3D-RDF - signal 11 / weighted by atomic polarizabilities (*RDFP11*)
132. 3D-RDF - signal 12 / weighted by atomic polarizabilities (*RDFP12*)
133. 3D-RDF - signal 13 / weighted by atomic polarizabilities (*RDFP13*)
134. 3D-RDF - signal 14 / weighted by atomic polarizabilities (*RDFP14*)
135. 3D-RDF - signal 15 / weighted by atomic polarizabilities (*RDFP15*)
136. 3D-RDF - signal 16 / weighted by atomic polarizabilities (*RDFP16*)
137. 3D-RDF - signal 17/ weighted by atomic polarizabilities (*RDFP17*)
138. 3D-RDF - signal 18 / weighted by atomic polarizabilities (*RDFP18*)
139. 3D-RDF - signal19 / weighted by atomic polarizabilities (*RDFP19*)
140. 3D-RDF - signal 20 / weighted by atomic polarizabilities (*RDFP20*)
141. 3D-RDF - signal 21 / weighted by atomic polarizabilities (*RDFP21*)
142. 3D-RDF - signal 22 / weighted by atomic polarizabilities (*RDFP22*)
143. 3D-RDF - signal 23 / weighted by atomic polarizabilities (*RDFP23*)
144. 3D-RDF - signal 24 / weighted by atomic polarizabilities (*RDFP24*)
145. 3D-RDF - signal 25 / weighted by atomic polarizabilities (*RDFP25*)
146. 3D-RDF - signal 26 / weighted by atomic polarizabilities (*RDFP26*)
147. 3D-RDF - signal 27 / weighted by atomic polarizabilities (*RDFP27*)
148. 3D-RDF - signal 28 / weighted by atomic polarizabilities (*RDFP28*)
149. 3D-RDF - signal 29 / weighted by atomic polarizabilities (*RDFP29*)
150. 3D-RDF - signal 30 / weighted by atomic polarizabilities (*RDFP30*)
151. 3D-RDF - signal 01 / weighted by atomic charge (*RDFC1*)
152. 3D-RDF - signal 02 / weighted by atomic charge (*RDFC2*)
153. 3D-RDF - signal 03 / weighted by atomic charge (*RDFC3*)
154. 3D-RDF - signal 04 / weighted by atomic charge (*RDFC4*)
155. 3D-RDF - signal 05 / weighted by atomic charge (*RDFC5*)
156. 3D-RDF - signal 06 / weighted by atomic charge (*RDFC6*)
157. 3D-RDF - signal 07 / weighted by atomic charge (*RDFC7*)
158. 3D-RDF - signal 08 / weighted by atomic charge (*RDFC8*)

159. 3D-RDF - signal 09 / weighted by atomic charge (*RDFC9*)
160. 3D-RDF - signal 10 / weighted by atomic charge (*RDFC10*)
161. 3D-RDF - signal 11 / weighted by atomic charge (*RDFC11*)
162. 3D-RDF - signal 12 / weighted by atomic charge (*RDFC12*)
163. 3D-RDF - signal 13 / weighted by atomic charge (*RDFC13*)
194. 3D-RDF - signal 14 / weighted by atomic charge (*RDFC14*)
165. 3D-RDF - signal 15 / weighted by atomic charge (*RDFC15*)
166. 3D-RDF - signal 16 / weighted by atomic charge (*RDFC16*)
167. 3D-RDF - signal 17/ weighted by atomic charge (*RDFC17*)
168. 3D-RDF - signal 18 / weighted by atomic charge (*RDFC18*)
169. 3D-RDF - signal19 / weighted by atomic charge (*RDFC19*)
170. 3D-RDF - signal 20 / weighted by atomic charge (*RDFC20*)
171. 3D-RDF - signal 21 / weighted by atomic charge (*RDFC21*)
172. 3D-RDF - signal 22 / weighted by atomic charge (*RDFC22*)
173. 3D-RDF - signal 23 / weighted by atomic charge (*RDFC23*)
174. 3D-RDF - signal 24 / weighted by atomic charge (*RDFC24*)
175. 3D-RDF - signal 25 / weighted by atomic charge (*RDFC25*)
176. 3D-RDF - signal 26 / weighted by atomic charge (*RDFC26*)
177. 3D-RDF - signal 27 / weighted by atomic charge (*RDFC27*)
178. 3D-RDF - signal 28 / weighted by atomic charge (*RDFC28*)
179. 3D-RDF - signal 29 / weighted by atomic charge (*RDFC29*)
180. 3D-RDF - signal 30 / weighted by atomic charge (*RDFC30*)

17. Molecular fingerprint

Molecular fingerprints are string representations of chemical structures designed to enhance the efficiency of chemical database searching and analysis. They can encode the 2D and/or 3D features of molecules as an array of binary values or counts. Therefore, molecular fingerprints consist of bins, each bin being a substructure descriptor associated with a specific molecular feature.

Molecular fingerprints directly encode molecular structure in a series of binary bits that represent the presence or absence of particular substructures in the molecule. Although it divides the whole molecule into a large number of fragments, it has the potential to keep overall complexity of drug

molecules. Additionally, it does not need reasonable three-dimensional conformation of drug molecules and thereby does not lead to error accumulation from the description of molecular structures. Thus by means of such descriptors, each molecule can be described based on a set of fingerprints of structural keys, which is represented as a Boolean array. A SMARTS list of substructure patterns is first determined as a predefined dictionary. There is a one-to-one correspondence between each SMARTS pattern and bit in the fingerprint. For each SMARTS pattern, if its corresponding substructure is present in the given molecule, the corresponding bit in the fingerprint is set to 1; conversely, it is set to 0 if the substructure is absent in the molecule (see Figure 1). Note that different molecular fingerprint systems abstract and magnify different aspects of molecular topology.

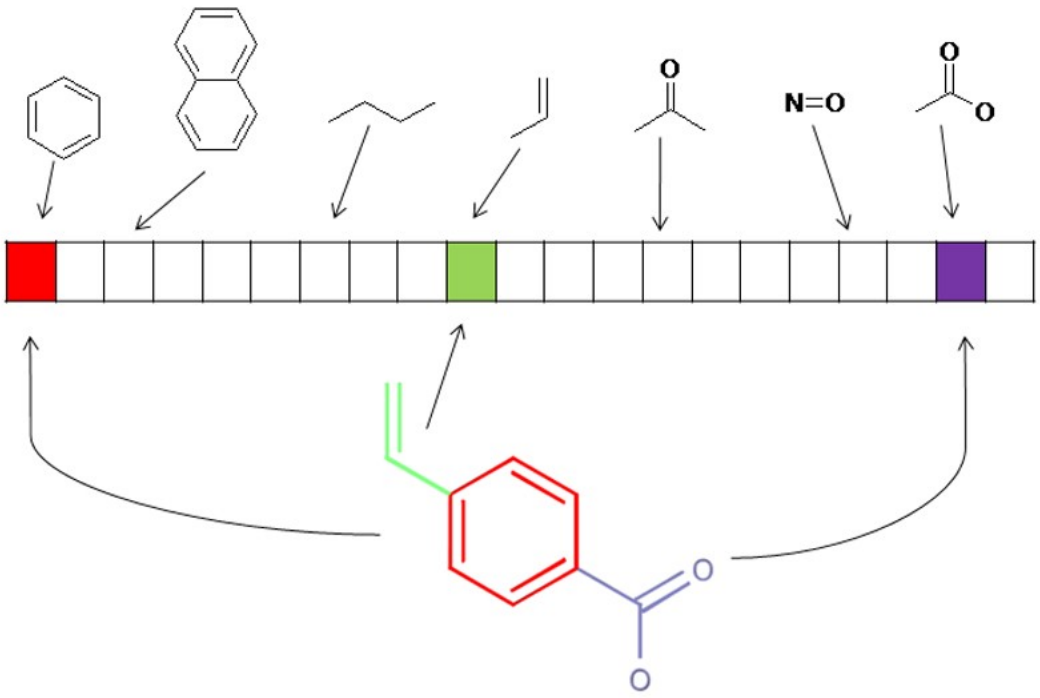

Figure 1 Representation of a molecular substructure fingerprint with a substructure fingerprint dictionary of given substructure patterns. This molecule is represented in a series of binary bits that represent the presence or absence of particular substructures in the molecules. This Figure is from Ref. 2 in section 3 and 4.

17.1 Daylight-type fingerprint

The Daylight fingerprints (DFP) are hashed fingerprints encoding each atom type, all Augmented

Atoms and all paths of length 2–7 atoms, giving a total string of 1024 bits [Daylight-James, Weininger et al., 1997].

17.2 MACCS keys and FP4 fingerprint

The FP4 and MACCS fingerprints are used to construct the substructure dictionaries, respectively. The dictionary of FP4 fingerprint contains 307 mostly common substructure patterns. It is originally written in an attempt to represent the classification of organic compounds from the viewpoint of an organic chemist. The MACCS fingerprint uses a dictionary of MDL keys, which contains a set of 166 mostly common substructure features. These are referred to as the MDL public MACCS keys. Both the definitions of FP4 and MACCS fingerprints are available from OpenBabel (version 2.3.0, http://openbabel.org/, accessed October, 2010). All calculations for these substructure fingerprints are performed in PyDPI package, developed by our group.

17.3 E-state fingerprint

Electrotopological State (E-state) fingerprints represent the presence/absence of 79 E-state substructures defined Kier and Hall in a molecule. The definition of 79 atom types can be found in section 1.5.

17.4 Atom pairs and topological torsions fingerprints

Atom pairs fingerprint:

Atom pairs are substructure descriptors defined in terms of any pair of atoms and bond types connecting them. An atom pair is composed of two non-hydrogen atoms and an interatomic separation:

$$\mathrm{AP} = \{[i\text{th atom description}][\text{separation}][j\text{th atom description}]\}$$

The two considered atoms need not be directly connected and the separation can be the topological distance between them [Carhart, Smith et al., 1985]; these descriptors are usually called topological atom pairs being based on the topological representation of the molecules. Atom type is defined by the element itself, the number of heavy-atom connections and number of p electron pairs on each atom.

Unlike topological torsions, atom pairs are sensitive to long-range correlations between the atoms in molecules and therefore to small changes in one part of even large molecules. Atom pair descriptors usually are Boolean variables encoding the presence or absence of a particular atom pair in each

molecule.

Topological torsion fingerprint:

The topological torsion descriptor (TT) is related to the 4-atom linear subfragment descriptor of Klopman because it is defined as a Boolean variable for the presence/absence of a linear sequence of four consecutively bonded non-hydrogen atoms *k–i–j–l*, each described by its atom type (TYPE), the number of *p* electrons (NPI) on each atom, and the number of non-hydrogen atoms (NBR) bonded to it [Nilakantan, Bauman et al., 1987]. Usually NBR does not include *k–i–j–l* atoms that go to make the torsion itself; therefore, it is -1 for k and l atoms and -2 for the two central atoms *i* and *j*. The torsion around the *i-j* bond and defined by the four indices *k–i–j–l* is represented by the following TT descriptor:

$$\mathrm{TT} = \{[\mathrm{NPI}-\mathrm{TYPE}-\mathrm{NBR}]_k[\mathrm{NPI}-\mathrm{TYPE}-\mathrm{NBR}]_i[\mathrm{NPI}-\mathrm{TYPE}-\mathrm{NBR}]_j[\mathrm{NPI}-\mathrm{TYPE}-\mathrm{NBR}]_l\}$$

The TT descriptor is a topological analogue of the 3D torsion angle, defined by four consecutively bonded atoms. The topological torsion is a short-range descriptor, that is, it is sensitive only to local changes in the molecule and is independent of the total number of atoms in the molecule.

The use of atom-centered fragments and related descriptors greatly increases the specific chemical information concerning different functional groups, but cannot discriminate between different arrangements of functional groups within a molecule.

17.5 Morgan fingerprint

This family of fingerprints, better known as circular fingerprints, is built by applying the Morgan algorithm to a set of user-supplied atom invariants. When generating Morgan fingerprints, the radius of the fingerprint need be provided. For detailed information about Morgan fingerprint, please refer to Ref. [19]. Note The default atom invariants use connectivity information similar to those used for the well known ECFP family of fingerprints. When comparing the ECFP/FCFP fingerprints and the Morgan fingerprints generated by the PyDPI, remember that the 4 in ECFP4 corresponds to the diameter of the atom environments considered, while the Morgan fingerprints take a radius parameter. So the examples above, with radius=2, are roughly equivalent to ECFP4 and FCFP4.

References:

[1] Aguiara, P.F.d., Bourguignon, B., Khotsa, M.S., Massarta, D.L., and Phan-Than-Luub, R. 1995.D-

optimal designs. Chemometrics and Intelligent Laboratory Systems 30:199-210.

[2] Daylight Chemical Information Systems ,Inc. Simplified Molecular Input Line Entry System. 2006, http://www.daylight.com/smiles/index.html.

[3] Elsevier MDL. MDL QSAR Version 2.2. 2006, http://www.mdl.com/products/predictive/qsar/index.jsp.

[4] Ghose, A.K., Viswanadhan,V. N., and Wendoloski, J.J. 1998. Prediction of Hydrophilic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods: An analysis of ALOG an CLOGP Methods. J. Phys. Chem.102:3762-3772.

[5] Gramatica, P., Corradi, M., and Consonni, V. 2000. Model ligand Prediction of Soil Sorption Coefficients of Non-ionic Organic Pesticides by Molecular Descriptors. Chemosphere 41:763-777.

[6] Hall, L.H., and Kier, L.B. 1991. The Molecular Connectivity Chi Indices and Kappa Shape Indices in Structure-Property Relations. In Reviews of Computational Chemistry, edited by D. Boyd and K. Lipkowitz. New York: VCH Publishers,Inc.,367-422.

[7] Hall, L.H., and Kier,L.B.1999. Molecular Connectivity Chi Indices for Database Analysis and Structure-Property Modeling. In Methods for QSAR Modelling, edited by J. Devillers.

[8] Kier,L.B.1987.Inclusion of symmetry as a shape attribute in Kappa index analysis. Quantit. Struct.-Act. Relat.6: 8-12.

[9] Kier, L.B., and Hall, L.H.1976. Molecular Connectivity in Chemistry and Drug Research. New York: Academic Press Inc.

[10] Kier, L.B.,and Hall, L.H. 1986. Molecular Connectivity in Structure-Activity Analysis. New York: John Wiley and Sons.

[11] Kier,L.B., and Hall, L.H. 1999. Molecule Structure Description: The Electrotopological State. New York: Academic Press.

[12] Martin, T.M., Harten, P., Venkatapathy, R., Das, S., and Young, D.M. 2008. A Hierarchical Clustering Methodology for the Estimation of Toxicity. Toxicology Mechanisms and Methods 18:251-266.

[13] Talete. Dragon Version 5.4. 2006, http://www.talete.mi.it/dragon_net.htm.

Todeschini, R., and Consonni, V. 2000. Handbook of Molecular Descriptors. Weinheim, Germany: Wiley-VCH.

[14] Viswanadhan, V.N., Ghose, A.K., Revankar, G. R., and Robins, R.K. 1989. Atomic Physicochemical Parameters for Three Dimensional Structure Directed Quantitative Structure-Activity Relationships. 4. Additional Parameters for Hydrophobic and Dispersive Interactions and Their Application for an Automated Superposition of Certain Naturally Occurring Nucleoside Antibiotics. J.

Chem. Inf. Comput. Sci. 29:163-172.

[15] Wang, R., Gao, Y., and Lai, L. 2000. Calculating partition coefficient by atom-additive method. Perspectives in Drug Discovery and Design19:47-66.

[16] R. E. Carhart, D.H. Smith, R. Venkataraghavan. Atom Pairs as Molecular Features in Structure-Activity Studies: Definition and Applications. J. Chem. Inf. Comput. Sci. 1985, 265 64-73.

[17] R. Nilakantan, N. Bauman, J.S. Dixon, R. Venkataraghavan. Topological Torsions: A New Molecular Descriptor for SAR Applications. Comparison with Other Descriptors. J. Chem. Inf. Comput. Sci. 1987, 27, 82-85.

[18] David Rogers, Mather Hahn. Extended-Connectivity Fingerprints. 2010, 50, 742-754.

[19] Paul Labute. A widely applicable set of descriptors. Journal of Molecular Graphics and Modeling. 2000, 18, 464-477.

[20] C. A. James, *D.* Weininger, J. Delany, Daylight Theory Manual 1997, http://www.daylight.com/dayhtml/doc/theory/theory.toc.html.

[21] Burden, F.R. (1997) A chemically intuitive molecular index based on the eigenvalues of a modified adjacency matrix. Quant. Struct. -Act. Relat., 16, 309–314

[22] Jurs, P.C., Dixon, J.S. and Egolf, L.M. (1995) Representations of molecules, in Chemometrics Methods in Molecular Design, Vol. 2 (ed. H. van de Waterbeemd), VCH Publishers, New York, pp. 15–38.

[23] *Basak, S.C. (1999) Information theoretic indices of neighborhood complexity and their applications, in Bibliography Topological Indices and Related Descriptors in QSAR and QSPR (eds J. Devillers and A.T. Balaban), Gordon and Breach Science Publishers, Amsterdam, The Netherlands, pp. 563–593.*

Table S1 List of ChemoPy computed descriptors for small molecules

Molecular descriptors		
	Constitutional descriptors (30)	
1	Weight	Molecular weight
2	nhyd	Count of hydrogen atoms
3	nhal	Count of halogen atoms
4	nhet	Count of hetero atoms
5	nhev	Count of heavy atoms
6	ncof	Count of F atoms
7	ncocl	Count of Cl atoms
8	ncobr	Count of Br atoms
9	ncoi	Count of I atoms
10	ncarb	Count of C atoms
11	nphos	Count of P atoms
12	nsulph	Count of S atoms
13	noxy	Count of O atoms
14	nnitro	Count of N atoms
15	nring	Number of rings
16	nrot	Number of rotatable bonds
17	ndonr	Number of H-bond donors
18	naccr	Number of H-bond acceptors

19	nsb	Number of single bonds
20	ndb	Number of double bonds
21	ntb	Number of triple bonds
22	naro	Number of aromatic bonds
23	nta	Number of all atoms
24	AWeight	Average molecular weight
25-30	PC1 PC2 PC3 PC4 PC5 PC6	Molecular path counts of length 1-6
	Topological descriptors (35)	
1	W	Weiner index
2	AW	Average Wiener index
3	J	Balaban's J index
4	T_{hara}	Harary number
5	T_{sch}	Schiultz index
6	Tigdi	Graph distance index
7	Platt	Platt number
8	Xu	Xu index
9	Pol	Polarity number
10	Dz	Pogliani index

11	Ipc	Ipc index
12	BertzCT	BertzCT
13	GMTI	Gutman molecular topological index based on simple vertex degree
14-15	ZM1 ZM2	Zagreb index with order 1-2
16-17	MZM1 MZM2	Modified Zagreb index with order 1-2
18	Qindex	Quadratic index
19	diametert	Largest value in the distance matrix
20	radiust	radius based on topology
21	petitjeant	Petitjean based on topology
22	Sito	the logarithm of the simple topological index by Narumi
23	Hato	harmonic topological index proposed by Narnumi
24	Geto	Geometric topological index by Narumi
25	Arto	Arithmetic topological index by Narumi
26	ISIZ	Total information index on molecular size
27	TIAC	Total information index on atomic composition
28	DET	Total information index on distance equality
29	IDE	Mean information index on distance equality
30	IVDE	Total information index on vertex equality
31	Sitov	Logarithm of the simple topological index by Narumi
32	Hatov	Harmonic topological index proposed by Narnumi
33	Getov	Geometric topological index by Narumi

34	Gravto	Gravitational topological index based on topological distance
35	GMTIV	Gutman molecular topological index based on valence vertex degree(log10)
	Connectivity descriptors (44)	
1-11	$^{0}\chi^{v}$ $^{1}\chi^{v}$ $^{2}\chi^{v}$ $^{3}\chi_{p}^{v}$ $^{4}\chi_{p}^{v}$ $^{5}\chi_{p}^{v}$ $^{6}\chi_{p}^{v}$ $^{7}\chi_{p}^{v}$ $^{8}\chi_{p}^{v}$ $^{9}\chi_{p}^{v}$ $^{10}\chi_{p}^{v}$	Valence molecular connectivity Chi index for path order 0-10
12	$^{3}\chi^{v}_{c}$	Valence molecular connectivity Chi index for three cluster
13	$^{4}\chi^{v}_{c}$	Valence molecular connectivity Chi index for four cluster
14	$^{4}\chi^{v}_{pc}$	Valence molecular connectivity Chi index for path/cluster
15-18	$^{3}\chi^{v}_{CH}$ $^{4}\chi^{v}_{CH}$ $^{5}\chi^{v}_{CH}$ $^{6}\chi^{v}_{CH}$	Valence molecular connectivity Chi index for cycles of 3-6
19-29	$^{0}\chi$ $^{1}\chi$ $^{2}\chi$ $^{3}\chi_{p}$ $^{4}\chi_{p}$ $^{5}\chi_{p}$	Simple molecular connectivity Chi indices for path order 0-10

	$^{6}\chi_{p}$ $^{7}\chi_{p}$ $^{8}\chi_{p}$ $^{9}\chi_{p}$ $^{10}\chi_{p}$	
30	$^{3}\chi_{c}$	Simple molecular connectivity Chi indices for three cluster
31	$^{4}\chi_{c}$	Simple molecular connectivity Chi indices for four cluster
32	$^{4}\chi_{pc}$	Simple molecular connectivity Chi indices for path/cluster
33-36	$^{3}\chi_{CH}$ $^{4}\chi_{CH}$ $^{5}\chi_{CH}$ $^{6}\chi_{CH}$	Simple molecular connectivity Chi indices for cycles of 3-6
37	mChi1	mean chi1 (Randic) connectivity index
38	knotp	the difference between chi3c and chi4pc
39	dchi0	the difference between chi0v and chi0
40	dchi1	the difference between chi1v and chi1
41	dchi2	the difference between chi2v and chi2
42	dchi3	the difference between chi3v and chi3
43	dchi4	the difference between chi4v and chi4
44	knotpv	the difference between chiv3c and chiv4pc
	Kappa descriptors (7)	
1	$^{1}\kappa_{\alpha}$	Kappa alpha index for 1 bonded fragment
2	$^{2}\kappa_{\alpha}$	Kappa alpha index for 2 bonded fragment
3	$^{3}\kappa_{\alpha}$	Kappa alpha index for 3 bonded fragment

4	phi	Kier molecular flexibility index
5	$^{1}\kappa$	Molecular shape Kappa index for 1 bonded fragment
6	$^{2}\kappa$	Molecular shape Kappa index for 2 bonded fragment
7	$^{3}\kappa$	Molecular shape Kappa index for 3 bonded fragment
	Basak descriptors (21)	
1	IC0	Information content with order 0 proposed by Basak
2	IC1	Information content with order 1 proposed by Basak
3	IC2	Information content with order 2 proposed by Basak
4	IC3	Information content with order 3 proposed by Basak
5	IC4	Information content with order 4 proposed by Basak
6	IC5	Information content with order 5 proposed by Basak
7	IC6	Information content with order 6 proposed by Basak
8	SIC0	Complementary information content with order 0 proposed by Basak
9	SIC1	Structural information content with order 1 proposed by Basak
10	SIC2	Structural information content with order 2 proposed by Basak
11	SIC3	Structural information content with order 3 proposed by Basak
12	SIC4	Structural information content with order 4 proposed by Basak
13	SIC5	Structural information content with order 5 proposed by Basak
14	SIC6	Structural information content with order 6 proposed by Basak
15	CIC0	Complementary information content with order 0

		proposed by Basak
16	CIC1	Complementary information content with order 1 proposed by Basak
17	CIC2	Complementary information content with order 2 proposed by Basak
18	CIC3	Complementary information content with order 3 proposed by Basak
19	CIC4	Complementary information content with order 4 proposed by Basak
20	CIC5	Complementary information content with order 5 proposed by Basak
21	CIC6	Complementary information content with order 6 proposed by Basak
	E-state descriptors (245)	
1	S1	Sum of E-State of atom type: sLi
2	S2	Sum of E-State of atom type: ssBe
3	S3	Sum of E-State of atom type: ssssBe
4	S4	Sum of E-State of atom type: ssBH
5	S5	Sum of E-State of atom type: sssB
6	S6	Sum of E-State of atom type: ssssB
7	S7	Sum of E-State of atom type: sCH3
8	S8	Sum of E-State of atom type: dCH2
9	S9	Sum of E-State of atom type: ssCH2
10	S10	Sum of E-State of atom type: tCH
11	S11	Sum of E-State of atom type: dsCH
12	S12	Sum of E-State of atom type: aaCH

13	S13	Sum of E-State of atom type: sssCH
14	S14	Sum of E-State of atom type: ddC
15	S15	Sum of E-State of atom type: tsC
16	S16	Sum of E-State of atom type: dssC
17	S17	Sum of E-State of atom type: aasC
18	S18	Sum of E-State of atom type: aaaC
19	S19	Sum of E-State of atom type: ssssC
20	S20	Sum of E-State of atom type: sNH3
21	S(21)	Sum of E-State of atom type: sNH2
22	S22	Sum of E-State of atom type: ssNH2
23	S23	Sum of E-State of atom type: dNH
24	S24	Sum of E-State of atom type: ssNH
25	S25	Sum of E-State of atom type: aaNH
26	S26	Sum of E-State of atom type: tN
27	S27	Sum of E-State of atom type: sssNH
28	S28	Sum of E-State of atom type: dsN
29	S29	Sum of E-State of atom type: aaN
30	S30	Sum of E-State of atom type: sssN
31	S31	Sum of E-State of atom type: ddsN
32	S32	Sum of E-State of atom type: aasN
33	S33	Sum of E-State of atom type: ssssN

34	S34	Sum of E-State of atom type: sOH
35	S35	Sum of E-State of atom type: dO
36	S36	Sum of E-State of atom type: ssO
37	S37	Sum of E-State of atom type: aaO
38	S38	Sum of E-State of atom type: sF
39	S39	Sum of E-State of atom type: sSiH3
40	S40	Sum of E-State of atom type: ssSiH2
41	S41	Sum of E-State of atom type: sssSiH
42	S42	Sum of E-State of atom type: ssssSi
43	S43	Sum of E-State of atom type: sPH2
44	S44	Sum of E-State of atom type: ssPH
45	S45	Sum of E-State of atom type: sssP
46	S46	Sum of E-State of atom type: dsssP
47	S47	Sum of E-State of atom type: ssssssP
48	S48	Sum of E-State of atom type: sSH
49	S49	Sum of E-State of atom type: dS
50	S50	Sum of E-State of atom type: ssS
51	S51	Sum of E-State of atom type: aaS
52	S52	Sum of E-State of atom type: dssS
53	S53	Sum of E-State of atom type: ddssS
54	S54	Sum of E-State of atom type: sCl

55	S55	Sum of E-State of atom type: sGeH3
56	S56	Sum of E-State of atom type: ssGeH2
57	S57	Sum of E-State of atom type: sssGeH
58	S58	Sum of E-State of atom type: ssssGe
59	S59	Sum of E-State of atom type: sAsH2
60	S60	Sum of E-State of atom type: ssAsH
61	S61	Sum of E-State of atom type: sssAs
62	S62	Sum of E-State of atom type: sssdAs
63	S63	Sum of E-State of atom type: sssssAs
64	S64	Sum of E-State of atom type: sSeH
65	S65	Sum of E-State of atom type: dSe
66	S66	Sum of E-State of atom type: ssSe
67	S67	Sum of E-State of atom type: aaSe
68	S68	Sum of E-State of atom type: dssSe
69	S69	Sum of E-State of atom type: ddssSe
70	S70	Sum of E-State of atom type: sBr
71	S71	Sum of E-State of atom type: sSnH3
72	S72	Sum of E-State of atom type: ssSnH2
73	S73	Sum of E-State of atom type: sssSnH
74	S74	Sum of E-State of atom type: ssssSn
75	S75	Sum of E-State of atom type: sI

76	S76	Sum of E-State of atom type: sPbH3
77	S77	Sum of E-State of atom type: ssPbH2
78	S78	Sum of E-State of atom type: sssPbH
79	S79	Sum of E-State of atom type: ssssPb
80-158	Smax1-Smax79	Maxmum of E-State value of specified atom type
159-237	Smin1-Smin79	Minimum of E-State value of specified atom type
238	Shev	The sum of the EState indices over all non-hydrogen atoms
239	Scar	The sum of the EState indices over all C atoms
240	Shal	The sum of the EState indices over all Halogen atoms
241	Shet	The sum of the EState indices over all hetero atoms
242	Save	The sum of the EState indices over all non-hydrogen atoms divided by the number of non-hydrogen atoms
243	Smax	The maximal Estate value in all atoms
244	Smin	The minimal Estate value in all atoms
245	DS	The difference between Smax and Smin
	Burden descriptors (64)	
1-16	bcutm1-16	Burden descriptors based on atomic mass
17-32	bcutv1-16	Burden descriptors based on atomic vloumes
33-48	bcute1-16	Burden descriptors based on atomic electronegativity
49-64	bcutp1-16	Burden descriptors based on polarizability
	Autocorrelation descriptors (96)	
1-8	ATSm1-ATSm8	Moreau-Broto autocorrelation descriptors based on

		atom mass
9-16	ATSv1-ATSv8	Moreau-Broto autocorrelation descriptors based on atomic van der Waals volume
17-24	ATSe1-ATSe8	Moreau-Broto autocorrelation descriptors based on atomic Sanderson electronegativity
25-32	ATSp1-ATSp8	Moreau-Broto autocorrelation descriptors based on atomic polarizability
33-40	MATSm1-MATSm8	Moran autocorrelation descriptors based on atom mass
41-48	MATSv1-MATSv8	Moran autocorrelation descriptors based on atomic van der Waals volume
49-56	MATSe1-MATSe8	Moran autocorrelation descriptors based on atomic Sanderson electronegativity
57-64	MATSp1-MATSp8	Moran autocorrelation descriptors based on atomic polarizability
65-72	GATSm1-GATSm8	Geary autocorrelation descriptors based on atom mass
73-80	GATSv1-GATSv8	Geary autocorrelation descriptors based on atomic van der Waals volume
81-88	GATSe1-GATSe8	Geary autocorrelation descriptors based on atomic Sanderson electronegativity
89-96	GATSp1-GATSp8	Geary autocorrelation descriptors based on atomic polarizability
	Charge descriptors (25)	
1-4	Q_{Hmax} Q_{Cmax} Q_{Nmax} Q_{Omax}	Most positive charge on H,C,N,O atoms
5-8	Q_{Hmin} Q_{Cmin} Q_{Nmin} Q_{Omin}	Most negative charge on H,C,N,O atoms

9-10	Q_{max} Q_{min}	Most positive and negative charge in a molecule
11-15	Q_{HSS} Q_{CSS} Q_{NSS} Q_{OSS} Qass	Sum of squares of charges on H,C,N,O and all toms
16-17	Mpc Tpc	Mean and total of positive charges
18-19	Mnc Tnc	Mean and total of negative charges
20-21	Mac Tac	Mean and total of absolute charges
22	Rpc	Relative positive charge
23	Rnc	Relative negative charge
24	SPP	Submolecular polarity parameter
25	LDI	Local dipole index
	Molecular property descriptors (6)	
1	MREF	Molar refractivity
2	logP	LogP value based on the Crippen method
3	logP2	Square of LogP value based on the Crippen method
4	TPSA	Topological polarity surface area
5	UI	Unsaturation index
6	Hy	Hydrophilic index
	MOE-type descriptors (60)	
1	TPSA	Topological polar surface area based on fragments

2	LabuteASA	Labute's Approximate Surface Area
3-14	SLOGPVSA	MOE-type descriptors using SLogP contributions and surface area contributions
15-24	SMRVSA	MOE-type descriptors using MR contributions and surface area contributions
25-38	PEOEVSA	MOE-type descriptors using partial charges and surface area contributions
39-49	EstateVSA	MOE-type descriptors using Estate indices and surface area contributions
50-60	VSAEstate	MOE-type descriptors using surface area contributions and Estate indices
	Geometric descriptors (12)	
1	W3DH	3-D Wiener index based geometrical distance matrix (including Hs)
2	W3D	3-D Wiener index based geometrical distance matrix (Not including Hs)
3	Petitj3D	Petitjean Index based on molecular geometrical distance matrix
4	GeDi	The longest distance between two atoms (geometrical diameter)
5	grav1	Gravitational 3D index
6	rygr	Radius of gyration
7	Harary3D	The 3D-Harary index
8	AGDD	The average geometric distance degree
9	SEig	The absolute eigenvalue sum on geometry matrix
10	SPAN	The span R
11	ASPAN	The average span R
12	MEcc	The molecular eccentricity

	CPSA descriptors (30)	
1	ASA	Solvent-accessible surface areas
2	MSA	Molecular surface areas
3	PNSA1	Partial negative area
4	PPSA1	Partial negative area
5	PNSA2	Total charge weighted negative surface area
6	PPSA2	Total charge weighted negative surface area
7	PNSA3	Atom charge weighted negative surface areas
8	PPSA3	Atom charge weighted positive surface areas
9	DPSA1	Difference in charged partial surface area
10	DPSA2	Difference in total charge weighted partial surface area
11	DPSA3	Difference in atomic charge weighted surface area
12	FNSA1	Fractional charged partial negative surface areas
13	FNSA2	Fractional charged partial negative surface areas
14	FNSA3	Fractional charged partial negative surface areas
15	FPSA1	Fractional charged partial negative surface areas
16	FPSA2	Fractional charged partial negative surface areas
17	FPSA3	Fractional charged partial negative surface areas
18	WNSA1	Surface weighted charged partial negative surface areas
19	WNSA2	Surface weighted charged partial negative surface areas
20	WNSA3	Surface weighted charged partial negative surface areas

21	WPSA1	Surface weighted charged partial negative surface areas
22	WPSA2	Surface weighted charged partial negative surface areas
23	WPSA3	Surface weighted charged partial negative surface areas
24	TASA	Total hydrophobic surface area
25	PSA	Total polar surface area
26	FrTATP	The fraction between TASA and TPSA
27	RASA	Relative hydrophobic surface area
28	RPSA	Relative polar surface area
29	RNCS	Relative negative charge surface area
30	RPCS	Relative positive charge surface area
	WHIM descriptors (70)	
1-14	---	Unweighted WHIM descriptors
15-28	---	WHIM descriptors based on atomic mass
29-42	---	WHIM descriptors based on Sanderson Electronegativity
43-56	---	WHIM descriptors based on VDW Volume
57-70	---	WHIM descriptors based on Polarizability
	MoRSE descriptors (210)	
1-30	MoRSEU1-30	Unweighted 3-D MoRse descriptors
31-60	MoRSEC1-30	3-D MoRse descriptors based on atomic charge
61-90	MoRSEM1-30	3-D MoRse descriptors based on atomic mass
91-120	MoRSEN1-30	3-D MoRse descriptors based on atomic number

121-150	MoRSEP1-30	3-D MoRse descriptors based on atomic polarizablity
151-180	MoRSEE1-30	3-D MoRse descriptors based on atomic Sanderson electronegativity
181-210	MoRSEV1-30	3-D MoRse descriptors based on atomic van der Waals volume
	RDF descriptors (180)	
1-30	RDFU1-30	Unweighted radial distribution function (RDF) descriptors
31-60	RDFC1-30	Radial distribution function (RDF) descriptors based on atomic charge.
61-90	RDFM1-30	Radial distribution function (RDF) descriptors based on atomic mass
91-120	RDFP1-30	Radial distribution function (RDF) descriptors based on atomic polarizability
121-150	RDFE1-30	Radial distribution function (RDF) descriptors based on atomic electronegativity
151-180	RDFV1-30	Radial distribution function (RDF) descriptors based on atomic van der Waals volume
	Fragment/Fingerprint-based descriptors	
1	FP2	(Topological fingerprint) A Daylight-like fingerprint based on hashing molecular subgraphs
2	MACCS	(MACCS keys)Using the 166 public keys implemented as SMARTS
3	E-state	79 E-state fingerprints or fragments
4	FP4	307 FP4 fingerprints
5	Atom Paris	Atom Paris fingerprints
6	Torsions	Topological torsion fingerprints
7	Morgan/Circular	Fingerprints based on the Morgan algorithm