Properties codebook for structuprint version 1.001

The following properties were recorded for the 20 common amino acids, using the **Molecular Operating Environment** (MOE) software package (version 2012.10), developed by the Chemical Computing Group (Montreal, Canada).

Property	Description
a_acc	Number of hydrogen bond acceptor atoms (not counting acidic atoms but counting atoms that are both hydrogen bond donors and acceptors such as -OH).
a_acid	Number of acidic atoms.
a_aro	Number of aromatic atoms.
a_base	Number of basic atoms.
a_count	Number of atoms (including implicit hydrogens). This is calculated as the sum of $(1 + h_i)$ over all non-trivial atoms i.
a_don	Number of hydrogen bond donor atoms (not counting basic atoms but counting atoms that are both hydrogen bond donors and acceptors such as -OH).
a_heavy	Number of heavy atoms $\#\{Z_i \mid Z_i > 1\}$.
a_hyd	Number of hydrophobic atoms.
a_IC	Atom information content (total). This is calculated to be a_ICM times n.
a_ICM	Atom information content (mean). This is the entropy of the element distribution in the molecule (including implicit hydrogens but not lone pair pseudo-atoms). Let n_i be the number of occurrences of atomic number i in the molecule. Let $pi = n_i/n$ where n is the sum of the n_i. The value of a_ICM is the negative of the sum over all i of $p_i \log p_i$.
a_nB	Number of boron atoms $\#\{Z_i \mid Z_i = 5\}$.
a_nBr	Number of bromine atoms $\#\{Z_i \mid Z_i = 35\}$.
a_nC	Number of carbon atoms $\#\{Z_i \mid Z_i = 6\}$.
a_nCl	Number of chlorine atoms $\#\{Z_i \mid Z_i = 17\}$.
a_nF	Number of fluorine atoms $\#\{Z_i \mid Z_i = 9\}$.
a_nH	Number of hydrogen atoms (including implicit hydrogens). This is calculated as the sum of h_i over all non-trivial atoms i plus the number of non-trivial hydrogen atoms.
a_nI	Number of iodine atoms $\#\{Z_i \mid Z_i = 53\}$.
a_nN	Number of nitrogen atoms $\#\{Z_i \mid Z_i = 7\}$.
a_nO	Number of oxygen atoms $\#\{Z_i \mid Z_i = 8\}$.
a_nP	Number of phosphorus atoms $\#\{Z_i \mid Z_i = 15\}$.
a_nS	Number of sulfur atoms $\#\{Z_i \mid Z_i = 16\}$.
AM1_dipole	The dipole moment calculated using the AM1 Hamiltonian [MOPAC1993].
AM1_E	The total SCF energy (kcal/mol) calculated using the AM1 Hamiltonian [MOPAC1993].
AM1_Eele	The electronic energy (kcal/mol) calculated using the AM1 Hamiltonian [MOPAC1993].

Property	Description
AM1_HF	The heat of formation (kcal/mol) calculated using the AM1 Hamiltonian [MOPAC1993].
AM1_HOMO	The energy (eV) of the Highest Occupied Molecular Orbital calculated using the AM1 Hamiltonian [MOPAC1993].
AM1_IP	The ionization potential (kcal/mol) calculated using the AM1 Hamiltonian [MOPAC1993].
AM1_LUMO	The energy (eV) of the Lowest Unoccupied Molecular Orbital calculated using the AM1 Hamiltonian [MOPAC1993].
apol	Sum of the atomic polarizabilities (including implicit hydrogens) with polarizabilities taken from [CRC1994].
ASA	Water accessible surface area calculated using a radius of 1.4 Å for the water molecule. A polyhedral representation is used for each atom in calculating the surface area.
ASA_H	Water accessible surface area of all hydrophobic ($\|q_i\| < 0.2$) atoms.
ASA_minus	Water accessible surface area of all atoms with negative partial charge (strictly less than 0).
ASA_P	Water accessible surface area of all polar ($\|q_i\| \geq 0.2$) atoms.
ASA_plus	Water accessible surface area of all atoms with positive partial charge (strictly greater than 0).
b_1rotN	Number of rotatable single bonds. Conjugated single bonds are not included (e.g. ester and peptide bonds).
b_1rotR	Number of rotatable single bonds: b_1rotN divided by b_heavy.
b_ar	Number of aromatic bonds.
b_count	Number of bonds (including implicit hydrogens). This is calculated as the sum of $(d_i/2 + h_i)$ over all non-trivial atoms i.
b_double	Number of double bonds. Aromatic bonds are not considered to be double bonds.
b_heavy	Number of bonds between heavy atoms.
b_rotN	Number of rotatable bonds. A bond is rotatable if it has order 1, is not in a ring, and has at least two heavy neighbors.
b_rotR	Fraction of rotatable bonds: b_rotN divided by b_heavy.
b_single	Number of single bonds (including implicit hydrogens). Aromatic bonds are not considered to be single bonds.
b_triple	Number of triple bonds (including implicit hydrogens). Aromatic bonds are not considered to be triple bonds.
balabanJ	Balaban's connectivity topological index [Balaban1982].
BCUT_PEOE_0	The BCUT descriptors [Pearlman1998] are calculated from the eigenvalues of a modified adjacency matrix. Each ij entry of the adjacency matrix takes the value 1/sqrt(b_{ij}) where b_{ij} is the formal bond order between bonded atoms i and j. The diagonal takes the value of the PEOE partial charges. The resulting eigenvalues are sorted and the smallest, 1/3-ile, 2/3-ile and largest eigenvalues are reported.
BCUT_PEOE_1	See above.
BCUT_PEOE_2	See above.
BCUT_PEOE_3	See above.
BCUT_SLOGP_0	The BCUT descriptors using atomic contribution to logP (using the Wildman and Crippen SlogP method) instead of partial charge.

Property	Description
BCUT_SLOGP_1	See above.
BCUT_SLOGP_2	See above.
BCUT_SLOGP_3	See above.
BCUT_SMR_0	The BCUT descriptors using atomic contribution to molar refractivity (using the Wildman and Crippen SlogP method) instead of partial charge.
BCUT_SMR_1	See above.
BCUT_SMR_2	See above.
BCUT_SMR_3	See above.
bpol	Sum of the absolute value of the difference between atomic polarizabilities of all bonded atoms in the molecule (including implicit hydrogens) with polarizabilities taken from [CRC1994].
CASA_minus	Negative charge weighted surface area, ASA_minus times max $\{ q_i < 0 \}$ [Stanton1990].
CASA_plus	Positive charge weighted surface area, ASA_plus times max $\{ q_i > 0 \}$ [Stanton1990].
chi0	Atomic connectivity index (order 0) from [Hall1991] and [Kier1977]. This is calculated as the sum of 1/sqrt(d_i) over all heavy atoms i with $d_i > 0$.
chi0_C	Carbon connectivity index (order 0). This is calculated as the sum of 1/sqrt(v_i) over all carbon atoms i with $d_i > 0$.
chi0v	Atomic valence connectivity index (order 0) from [Hall1991] and [Kier1977]. This is calculated as the sum of 1/sqrt(v_i) over all heavy atoms i with $v_i > 0$.
chi0v_C	Carbon valence connectivity index (order 0). This is calculated as the sum of 1/sqrt(v_i) over all carbon atoms i with $v_i > 0$.
chi1	Atomic connectivity index (order 1) from [Hall1991] and [Kier1977]. This is calculated as the sum of 1/sqrt($d_i d_j$) over all bonds between heavy atoms i and j where $i < j$.
chi1_C	Carbon connectivity index (order 1). This is calculated as the sum of 1/sqrt($d_i d_j$) over all bonds between carbon atoms i and j where $i < j$.
chi1v	Atomic valence connectivity index (order 1) from [Hall1991] and [Kier1977]. This is calculated as the sum of 1/sqrt($v_i v_j$) over all bonds between heavy atoms i and j where $i < j$.
chi1v_C	Carbon valence connectivity index (order 1). This is calculated as the sum of 1/sqrt($v_i v_j$) over all bonds between carbon atoms i and j where $i < j$.
chiral	The number of chiral centers.
chiral_u	The number of unconstrained chiral centers.
DASA	Absolute value of the difference between ASA_plus and ASA_minus.
DCASA	Absolute value of the difference between CASA_plus and CASA_minus [Stanton1990].
dens	Mass density: molecular weight divided by van der Waals volume as calculated in the vol descriptor.
density	Molecular mass density: Weight divided by vdw_vol (amu/Å^3).
diameter	Largest value in the distance matrix [Petitjean1992].
dipole	Dipole moment calculated from the partial charges of the molecule.

Property	Description
dipoleX	The x component of the dipole moment (external coordinates).
dipoleY	The y component of the dipole moment (external coordinates).
dipoleZ	The z component of the dipole moment (external coordinates).
E	Value of the potential energy.
E_ang	Angle bend potential energy.
E_ele	Electrostatic component of the potential energy.
E_nb	Value of the potential energy with all bonded terms disabled.
E_oop	Out-of-plane potential energy.
E_sol	Solvation energy.
E_stb	Bond stretch-bend cross-term potential energy.
E_str	Bond stretch potential energy.
E_strain	Local strain energy: the current energy minus the value of the energy at a near local minimum. The current energy is calculated as for the E descriptor. The local minimum energy is the value of the E descriptor after first performing an energy minimization.
E_tor	Torsion (proper and improper) potential energy.
E_vdw	Van der Waals component of the potential energy.
FASA_H	Fractional ASA_H calculated as ASA_H/ASA.
FASA_minus	Fractional ASA_minus calculated as ASA_minus/ASA.
FASA_P	Fractional ASA_P calculated as ASA_P/ASA.
FASA_plus	Fractional ASA_plus calculated as ASA_plus/ASA.
FCASA_minus	Fractional CASA_minus calculated as CASA_minus/ASA.
FCASA_plus	Fractional CASA_plus calculated as CASA_plus/ASA.
FCharge	Total charge of the molecule (sum of formal charges).
GCUT_PEOE_0	The GCUT descriptors are calculated from the eigenvalues of a modified graph distance adjacency matrix. Each ij entry of the adjacency matrix takes the value 1/sqrt(d_{ij}) where d_{ij} is the (modified) graph distance between atoms i and j. The diagonal takes the value of the PEOE partial charges. The resulting eigenvalues are sorted and the smallest, 1/3-ile, 2/3-ile and largest eigenvalues are reported.
GCUT_PEOE_1	See above.
GCUT_PEOE_2	See above.
GCUT_PEOE_3	See above.
GCUT_SLOGP_0	The GCUT descriptors using atomic contribution to logP (using the Wildman and Crippen SlogP method) instead of partial charge.
GCUT_SLOGP_1	See above.
GCUT_SLOGP_2	See above.
GCUT_SLOGP_3	See above.
GCUT_SMR_0	The GCUT descriptors using atomic contribution to molar refractivity (using the Wildman and Crippen SMR method) instead of partial charge.
GCUT_SMR_1	See above.
GCUT_SMR_2	See above.
GCUT_SMR_3	See above.

Property	Description
glob	Globularity, or inverse condition number (smallest eigenvalue divided by the largest eigenvalue) of the covariance matrix of atomic coordinates. A value of 1 indicates a perfect sphere while a value of 0 indicates a two- or one-dimensional object.
Kier1	First kappa shape index: $(n-1)^2/m^2$ [Hall1991].
Kier2	Second kappa shape index: $(n-1)^2/m^2$ [Hall1991].
Kier3	Third kappa shape index: $(n-1)(n-3)^2/p_3^2$ for odd n, and $(n-3)(n-2)^2/p_3^2$ for even n [Hall1991].
KierA1	First alpha modified shape index: $s(s-1)^2/m^2$ where $s = n + a$ [Hall1991].
KierA2	Second alpha modified shape index: $s(s-1)^2/m^2$ where $s = n + a$ [Hall1991].
KierA3	Third alpha modified shape index: $(s-1)(s-3)^2/p_3^2$ for odd n, and $(s-3)(s-2)^2/p_3^2$ for even n where $s = n + a$ [Hall1991].
KierFlex	Kier molecular flexibility index: (KierA1)(KierA2)/n [Hall1991].
lip_acc	The number of O and N atoms.
lip_don	The number of OH and NH atoms.
lip_druglike	One if and only if lip_violation < 2 otherwise zero.
lip_violation	The number of violations of Lipinski's Rule of Five [Lipinski2012].
logP_par_o_div_w_par_	Log of the octanol/water partition coefficient (including implicit hydrogens). This property is calculated from a linear atom type model [LOGP1998] with $r^2 = 0.931$, RMSE=0.393 on 1,827 molecules.
logS	Log of the aqueous solubility (mol/L). This property is calculated from an atom contribution linear atom type model [Hou2004] with $r^2 = 0.90$, ~1,200 molecules.
MNDO_dipole	The dipole moment calculated using the MNDO Hamiltonian [MOPAC1993].
MNDO_E	The total SCF energy (kcal/mol) calculated using the MNDO Hamiltonian [MOPAC1993].
MNDO_Eele	The electronic energy (kcal/mol) calculated using the MNDO Hamiltonian [MOPAC1993].
MNDO_HF	The heat of formation (kcal/mol) calculated using the MNDO Hamiltonian [MOPAC1993].
MNDO_HOMO	The energy (eV) of the Highest Occupied Molecular Orbital calculated using the MNDO Hamiltonian [MOPAC1993].
MNDO_IP	The ionization potential (kcal/mol) calculated using the MNDO Hamiltonian [MOPAC1993].
MNDO_LUMO	The energy (eV) of the Lowest Unoccupied Molecular Orbital calculated using the MNDO Hamiltonian [MOPAC1993].
mr	Molecular refractivity (including implicit hydrogens). This property is calculated from an 11 descriptor linear model [MREF1998] with $r^2 = 0.997$, RMSE = 0.168 on 1,947 small molecules.
mutagenic	Indicator of the presence of potentially toxic groups. A non-zero value indicates that the molecule contains a mutagenic group. The table of mutagenic groups is based on the Kazius set [Kazius2005].
nmol	The number of molecules (connected components).
npr1	Normalized PMI ratio pmi1/pmi3.
npr2	Normalized PMI ratio pmi2/pmi3.

Property	Description
opr_brigid	The number of rigid bonds from [Oprea2000].
opr_leadlike	One if and only if opr_violation < 2 otherwise zero.
opr_nring	The number of ring bonds from [Oprea2000].
opr_nrot	The number of rotatable bonds from [Oprea2000].
opr_violation	The number of violations of Oprea's lead-like test [Oprea2000].
PEOE_PC_minus	Total negative partial charge: the sum of the negative q_i.
PEOE_PC_plus	Total positive partial charge: the sum of the positive q_i.
PEOE_RPC_minus	Relative negative partial charge: the smallest negative q_i divided by the sum of the negative q_i.
PEOE_RPC_plus	Relative positive partial charge: the largest positive q_i divided by the sum of the positive q_i.
PEOE_VSA_FHYD	Fractional hydrophobic van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is less than or equal to 0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
PEOE_VSA_FNEG	Fractional negative van der Waals surface area. This is the sum of the v_i such that q_i is negative divided by the total surface area. The v_i are calculated using a connection table approximation.
PEOE_VSA_FPNEG	Fractional negative polar van der Waals surface area. This is the sum of the v_i such that q_i is less than -0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
PEOE_VSA_FPOL	Fractional polar van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is greater than 0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
PEOE_VSA_FPOS	Fractional positive van der Waals surface area. This is the sum of the v_i such that q_i is non-negative divided by the total surface area. The v_i are calculated using a connection table approximation.
PEOE_VSA_FPPOS	Fractional positive polar van der Waals surface area. This is the sum of the v_i such that q_i is greater than 0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
PEOE_VSA_HYD	Total hydrophobic van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is less than or equal to 0.2. The v_i are calculated using a connection table approximation.
PEOE_VSA_minus0	Sum of v_i where q_i is in the range [-0.05,0.00).
PEOE_VSA_minus1	Sum of v_i where q_i is in the range [-0.10,-0.05).
PEOE_VSA_minus2	Sum of v_i where q_i is in the range [-0.15,-0.10).
PEOE_VSA_minus3	Sum of v_i where q_i is in the range [-0.20,-0.15).
PEOE_VSA_minus4	Sum of v_i where q_i is in the range [-0.25,-0.20).
PEOE_VSA_minus5	Sum of v_i where q_i is in the range [-0.30,-0.25).
PEOE_VSA_minus6	Sum of v_i where q_i is less than -0.3.
PEOE_VSA_NEG	Total negative van der Waals surface area. This is the sum of the v_i such that q_i is negative. The v_i are calculated using a connection table approximation.
PEOE_VSA_plus0	Sum of v_i where q_i is in the range [0.00,0.05).
PEOE_VSA_plus1	Sum of v_i where q_i is in the range [0.05,0.10).
PEOE_VSA_plus2	Sum of v_i where q_i is in the range [0.10,0.15).
PEOE_VSA_plus3	Sum of v_i where q_i is in the range [0.15,0.20).
PEOE_VSA_plus4	Sum of v_i where q_i is in the range [0.20,0.25).

Property	Description
PEOE_VSA_plus5	Sum of v_i where q_i is in the range [0.25,0.30).
PEOE_VSA_plus6	Sum of v_i where q_i is greater than 0.3.
PEOE_VSA_PNEG	Total negative polar van der Waals surface area. This is the sum of the v_i such that q_i is less than -0.2. The v_i are calculated using a connection table approximation.
PEOE_VSA_POL	Total polar van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is greater than 0.2. The v_i are calculated using a connection table approximation.
PEOE_VSA_POS	Total positive van der Waals surface area. This is the sum of the v_i such that q_i is non-negative. The v_i are calculated using a connection table approximation.
PEOE_VSA_PPOS	Total positive polar van der Waals surface area. This is the sum of the v_i such that q_i is greater than 0.2. The v_i are calculated using a connection table approximation.
petitjean	Value of (diameter - radius)/diameter.
petitjeanSC	Petitjean graph Shape Coefficient as defined in [Petitjean1992]: (diameter - radius)/radius.
PM3_dipole	The dipole moment calculated using the PM3 Hamiltonian [MOPAC1993].
PM3_E	The total SCF energy (kcal/mol) calculated using the PM3 Hamiltonian [MOPAC1993].
PM3_Eele	The electronic energy (kcal/mol) calculated using the PM3 Hamiltonian [MOPAC1993].
PM3_HF	The heat of formation (kcal/mol) calculated using the PM3 Hamiltonian [MOPAC1993].
PM3_HOMO	The energy (eV) of the Highest Occupied Molecular Orbital calculated using the PM3 Hamiltonian [MOPAC1993].
PM3_IP	The ionization potential (kcal/mol) calculated using the PM3 Hamiltonian [MOPAC1993].
PM3_LUMO	The energy (eV) of the Lowest Unoccupied Molecular Orbital calculated using the PM3 Hamiltonian [MOPAC1993].
pmi	Principal moment of inertia.
pmi1	First diagonal element of diagonalized moment of inertia tensor.
pmi2	Second diagonal element of diagonalized moment of inertia tensor.
pmi3	Third diagonal element of diagonalized moment of inertia tensor.
pmiX	x component of the principal moment of inertia (external coordinates).
pmiY	y component of the principal moment of inertia (external coordinates).
pmiZ	z component of the principal moment of inertia (external coordinates).
Q_PC_minus	Total negative partial charge: the sum of the negative q_i.
Q_PC_plus	Total positive partial charge: the sum of the positive q_i.
Q_RPC_minus	Relative negative partial charge: the smallest negative q_i divided by the sum of the negative q_i.
Q_RPC_plus	Relative positive partial charge: the largest positive q_i divided by the sum of the positive q_i.

Property	Description
Q_VSA_FHYD	Fractional hydrophobic van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is less than or equal to 0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
Q_VSA_FNEG	Fractional negative van der Waals surface area. This is the sum of the v_i such that q_i is negative divided by the total surface area. The v_i are calculated using a connection table approximation.
Q_VSA_FPNEG	Fractional negative polar van der Waals surface area. This is the sum of the v_i such that q_i is less than -0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
Q_VSA_FPOL	Fractional polar van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is greater than 0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
Q_VSA_FPOS	Fractional positive van der Waals surface area. This is the sum of the v_i such that q_i is non-negative divided by the total surface area. The v_i are calculated using a connection table approximation.
Q_VSA_FPPOS	Fractional positive polar van der Waals surface area. This is the sum of the v_i such that q_i is greater than 0.2 divided by the total surface area. The v_i are calculated using a connection table approximation.
Q_VSA_HYD	Total hydrophobic van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is less than or equal to 0.2. The v_i are calculated using a connection table approximation.
Q_VSA_NEG	Total negative van der Waals surface area. This is the sum of the v_i such that q_i is negative. The v_i are calculated using a connection table approximation.
Q_VSA_PNEG	Total negative polar van der Waals surface area. This is the sum of the v_i such that q_i is less than -0.2. The v_i are calculated using a connection table approximation.
Q_VSA_POL	Total polar van der Waals surface area. This is the sum of the v_i such that $\|q_i\|$ is greater than 0.2. The v_i are calculated using a connection table approximation.
Q_VSA_POS	Total positive van der Waals surface area. This is the sum of the v_i such that q_i is non-negative. The v_i are calculated using a connection table approximation.
Q_VSA_PPOS	Total positive polar van der Waals surface area. This is the sum of the v_i such that q_i is greater than 0.2. The v_i are calculated using a connection table approximation.
radius	If r_i is the largest matrix entry in row i of the distance matrix D, then the radius is defined as the smallest of the r_i [Petitjean1992].
reactive	Indicator of the presence of reactive groups. A non-zero value indicates that the molecule contains a reactive group. The table of reactive groups is based on the [Oprea2000] set and includes metals, phospho-, N/O/S-N/O/S single bonds, thiols, acyl halides, Michael Acceptors, azides, esters, etc.
rgyr	Radius of gyration.
rings	The number of rings.
RPC_minus	Relative negative partial charge: the smallest negative q_i divided by the sum of the negative q_i.

Property	Description
RPC_plus	Relative positive partial charge: the largest positive q_i divided by the sum of the positive q_i.
rsynth	A value in [0,1] indicating the synthetic reasonableness, or feasibility, of the chemical structure. A value of 0 means it is unlikely that the molecule can be synthesized while a value of 1 means that it is likely that the molecule can be synthesized. The value reflects the fraction of heavy atoms in the molecule that can be traced back to starting materials fragments resulting from retrosynthetic disconnection rules.
SlogP	Log of the octanol/water partition coefficient (including implicit hydrogens). This property is an atomic contribution model [Wildman1999] that calculates logP from the given structure; i.e. the correct protonation state (washed structures). Results may vary from the logP_par_o_div_w_par_ descriptor. The training set for SlogP was ~7000 structures.
SlogP_VSA0	Sum of v_i such that $L_i \leq -0.4$.
SlogP_VSA1	Sum of v_i such that L_i is in (-0.4,-0.2].
SlogP_VSA2	Sum of v_i such that L_i is in (-0.2,0].
SlogP_VSA3	Sum of v_i such that L_i is in (0,0.1].
SlogP_VSA4	Sum of v_i such that L_i is in (0.1,0.15].
SlogP_VSA5	Sum of v_i such that L_i is in (0.15,0.20].
SlogP_VSA6	Sum of v_i such that L_i is in (0.20,0.25].
SlogP_VSA7	Sum of v_i such that L_i is in (0.25,0.30].
SlogP_VSA8	Sum of v_i such that L_i is in (0.30,0.40].
SlogP_VSA9	Sum of v_i such that $L_i > 0.40$.
SMR	Molecular refractivity (including implicit hydrogens). This property is an atomic contribution model [Wildman1999] that assumes the correct protonation state (washed structures). The model was trained on ~7000 structures and results may vary from the mr descriptor.
SMR_VSA0	Sum of v_i such that R_i is in [0,0.11].
SMR_VSA1	Sum of v_i such that R_i is in (0.11,0.26].
SMR_VSA2	Sum of v_i such that R_i is in (0.26,0.35].
SMR_VSA3	Sum of v_i such that R_i is in (0.35,0.39].
SMR_VSA4	Sum of v_i such that R_i is in (0.39,0.44].
SMR_VSA5	Sum of v_i such that R_i is in (0.44,0.485].
SMR_VSA6	Sum of v_i such that R_i is in (0.485,0.56].
SMR_VSA7	Sum of v_i such that $R_i > 0.56$.
std_dim1	Standard dimension 1: the square root of the largest eigenvalue of the covariance matrix of the atomic coordinates. A standard dimension is equivalent to the standard deviation along a principal component axis.
std_dim2	Standard dimension 2: the square root of the second largest eigenvalue of the covariance matrix of the atomic coordinates. A standard dimension is equivalent to the standard deviation along a principal component axis.

Property	Description
std_dim3	Standard dimension 3: the square root of the third largest eigenvalue of the covariance matrix of the atomic coordinates. A standard dimension is equivalent to the standard deviation along a principal component axis.
TPSA	Polar surface area ($Å^2$) calculated using group contributions to approximate the polar surface area from connection table information only. The parameterization is that of [Ertl2000].
VAdjEq	Vertex adjacency information (equality): $-(1-f)log_2(1-f) - flog_2f$ where $f = (n^2-m)/n^2$, n is the number of heavy atoms and m is the number of heavy-heavy bonds. If f is not in the open interval (0,1), then 0 is returned.
VAdjMa	Vertex adjacency information (magnitude): $1+log_2m$ where m is the number of heavy-heavy bonds. If m is zero, then zero is returned.
VDistEq	If m is the sum of the distance matrix entries then VdistEq is defined to be the sum of $log_2m - p_ilog_2p_i/m$ where p_i is the number of distance matrix entries equal to i.
VDistMa	If m is the sum of the distance matrix entries then VDistMa is defined to be the sum of $log_2m - D_{ij}log_2D_{ij}/m$ over all i and j.
vdw_area	Area of van der Waals surface ($Å^2$) calculated using a connection table approximation.
vdw_vol	Van der Waals volume ($Å^3$) calculated using a connection table approximation.
vol	Van der Waals volume calculated using a grid approximation (spacing 0.75 Å).
VSA	Van der Waals surface area. A polyhedral representation is used for each atom in calculating the surface area.
vsa_acc	Approximation to the sum of VDW surface areas ($Å^2$) of pure hydrogen bond acceptors (not counting acidic atoms and atoms that are both hydrogen bond donors and acceptors such as -OH).
vsa_acid	Approximation to the sum of VDW surface areas of acidic atoms ($Å^2$).
vsa_base	Approximation to the sum of VDW surface areas of basic atoms ($Å^2$).
vsa_don	Approximation to the sum of VDW surface areas of pure hydrogen bond donors (not counting basic atoms and atoms that are both hydrogen bond donors and acceptors such as -OH) ($Å^2$).
vsa_hyd	Approximation to the sum of VDW surface areas of hydrophobic atoms ($Å^2$).
vsa_other	Approximation to the sum of VDW surface areas ($Å^2$) of atoms typed as "other".
vsa_pol	Approximation to the sum of VDW surface areas ($Å^2$) of polar atoms (atoms that are both hydrogen bond donors and acceptors), such as -OH.
vsurf_A	Amphiphilic moment.
vsurf_CP	Critical packing parameter.
vsurf_CW1	Capacity factor (1).
vsurf_CW2	Capacity factor (2).

Property	Description
vsurf_CW3	Capacity factor (3).
vsurf_CW4	Capacity factor (4).
vsurf_CW5	Capacity factor (5).
vsurf_CW6	Capacity factor (6).
vsurf_CW7	Capacity factor (7).
vsurf_CW8	Capacity factor (8).
vsurf_D1	Hydrophobic volume (1).
vsurf_D2	Hydrophobic volume (2).
vsurf_D3	Hydrophobic volume (3).
vsurf_D4	Hydrophobic volume (4).
vsurf_D5	Hydrophobic volume (5).
vsurf_D6	Hydrophobic volume (6).
vsurf_D7	Hydrophobic volume (7).
vsurf_D8	Hydrophobic volume (8).
vsurf_DD12	Contact distances of vsurf_DDmin (12).
vsurf_DD13	Contact distances of vsurf_DDmin (13).
vsurf_DD23	Contact distances of vsurf_DDmin (23).
vsurf_DW12	Contact distances of vsurf_EWmin (12).
vsurf_DW13	Contact distances of vsurf_EWmin (13).
vsurf_DW23	Contact distances of vsurf_EWmin (23).
vsurf_EDmin1	Lowest hydrophobic energy (1).
vsurf_EDmin2	Lowest hydrophobic energy (2).
vsurf_EDmin3	Lowest hydrophobic energy (3).
vsurf_EWmin1	Lowest hydrophilic energy (1).
vsurf_EWmin2	Lowest hydrophilic energy (2).
vsurf_EWmin3	Lowest hydrophilic energy (3).
vsurf_G	Surface globularity.
vsurf_HB1	H-bond donor capacity (1).
vsurf_HB2	H-bond donor capacity (2).
vsurf_HB3	H-bond donor capacity (3).
vsurf_HB4	H-bond donor capacity (4).
vsurf_HB5	H-bond donor capacity (5).
vsurf_HB6	H-bond donor capacity (6).
vsurf_HB7	H-bond donor capacity (7).
vsurf_HB8	H-bond donor capacity (8).
vsurf_HL1	Hydrophilic-Lipophilic (1).
vsurf_HL2	Hydrophilic-Lipophilic (2).
vsurf_ID1	Hydrophobic integy moment (1).
vsurf_ID2	Hydrophobic integy moment (2).
vsurf_ID3	Hydrophobic integy moment (3).
vsurf_ID4	Hydrophobic integy moment (4).
vsurf_ID5	Hydrophobic integy moment (5).
vsurf_ID6	Hydrophobic integy moment (6).
vsurf_ID7	Hydrophobic integy moment (7).
vsurf_ID8	Hydrophobic integy moment (8).
vsurf_IW1	Hydrophilic integy moment (1).
vsurf_IW2	Hydrophilic integy moment (2).
vsurf_IW3	Hydrophilic integy moment (3).

Property	Description
vsurf_IW4	Hydrophilic integy moment (4).
vsurf_IW5	Hydrophilic integy moment (5).
vsurf_IW6	Hydrophilic integy moment (6).
vsurf_IW7	Hydrophilic integy moment (7).
vsurf_IW8	Hydrophilic integy moment (8).
vsurf_R	Surface rugosity.
vsurf_S	Interaction field surface area.
vsurf_V	Interaction field volume.
vsurf_W1	Hydrophilic volume (1).
vsurf_W2	Hydrophilic volume (2).
vsurf_W3	Hydrophilic volume (3).
vsurf_W4	Hydrophilic volume (4).
vsurf_W5	Hydrophilic volume (5).
vsurf_W6	Hydrophilic volume (6).
vsurf_W7	Hydrophilic volume (7).
vsurf_W8	Hydrophilic volume (8).
vsurf_Wp1	Polar volume (1).
vsurf_Wp2	Polar volume (2).
vsurf_Wp3	Polar volume (3).
vsurf_Wp4	Polar volume (4).
vsurf_Wp5	Polar volume (5).
vsurf_Wp6	Polar volume (6).
vsurf_Wp7	Polar volume (7).
vsurf_Wp8	Polar volume (8).
Weight	Molecular weight (including implicit hydrogens) in atomic mass units with atomic weights taken from [CRC1994].
wienerPath	Wiener path number: half the sum of all the distance matrix entries as defined in [Balaban1979] and [Wiener1947].
wienerPol	Wiener polarity number: half the sum of all the distance matrix entries with a value of 3 as defined in [Balaban1979].
zagreb	Zagreb index: the sum of d_i^2 over all heavy atoms i.

References

[Balaban1979] Alexandra T Balaban. Five new topological indices for the branching of tree-like graphs. *Theoretica Chimica Acta*, 53:355–375, 1979.

[Balaban1982] Alexandru T Balaban. Highly discriminating distance-based topological index. *Chemical Physics Letters*, 89(5):399–404, 1982.

[CRC1994] David R Lide, editor. *CRC Handbook of Chemistry and Physics*. CRC Press, 1994.

[Ertl2000] Peter Ertl, Bernhard Rohde, and Paul Selzer. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, 43(20):3714–3717, 2000.

[Hall1991] Lowell H Hall and Lemont B Kier. The molecular connectivity chi indexes and kappa shape indexes in structure-property modeling. In *Reviews in Computational Chemistry, Volume 2*. Wiley Online Library, 1991.

[Hou2004] TJ Hou, Ke Xia, Wei Zhang, and XJ Xu. Adme evaluation in drug discovery. 4. prediction of aqueous solubility based on atom contribution approach. *Journal of Chemical Information and Computer Sciences*, 44(1):266–275, 2004.

[Kazius2005] Jeroen Kazius, Ross McGuire, and Roberta Bursi. Derivation and validation of toxicophores for mutagenicity prediction. *Journal of Medicinal Chemistry*, 48(1):312–320, 2005.

[Kier1977] Lemont B Kier and Lowell H Hall. The nature of structure-activity-relationships and their relation to molecular connectivity. *European Journal of Medicinal Chemistry*, 12(4):307–312, 1977.

[LOGP1998] P Labute. MOE LogP(Octanol/Water) Model *unpublished*, 1998.

[Lipinski2012] Christopher A Lipinski, Franco Lombardo, Beryl W Dominy, and Paul J Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 64:4–17, 2012.

[MOPAC1993] JJP Stewart. MOPAC Manual (Seventh Edition), 1993.

[MREF1998] P Labute. MOE Molar Refractivity Model *unpublished*, 1998.

[Oprea2000] Tudor I Oprea. Property distribution of drug-related chemical databases. *Journal of Computer-Aided Molecular Design*, 14(3):251–264, 2000.

[Pearlman1998] Robert S Pearlman and KM Smith. Novel software tools for chemical diversity. In *3D QSAR in drug design*. Springer, 1998.

[Petitjean1992] Michel Petitjean. Applications of the radius-diameter diagram to the classification of topological and geometrical shapes of chemical compounds. *Journal of Chemical Information and Computer Sciences*, 32(4):331–337, 1992.

[Stanton1990] David T Stanton and Peter C Jurs. Development and use of charged partial surface area structural descriptors in computer-assisted quantitative structure-property relationship studies. *Analytical Chemistry*, 62(21):2323–2329, 1990.

[Wiener1947] Harry Wiener. Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1):17–20, 1947.

[Wildman1999] Scott A Wildman and Gordon M Crippen. Prediction of physicochemical parameters by atomic contributions. *Journal of Chemical Information and Computer Sciences*, 39(5):868–873, 1999.