Subject: Molecular Flexibility Descriptor
Posted by paulc on Thu, 19 Dec 2019 16:24:38 GMT
View Forum Message <> Reply to Message

Can anyone elucidate how the molecular flexibility descriptor is defined ?

Much Appreciated,
Paul

Subject: Re: Molecular Flexibility Descriptor
Posted by thomas on Sun, 29 Dec 2019 15:28:03 GMT
View Forum Message <> Reply to Message

I will try to give at least a short explanation that summarizes the calculation:

Molecular flexibility values range from 0.0 (rigid) to 1.0 (complately flexible).
Four steps are needed for its calculation:

- All rotatable bonds are determined. These are the only considered contributors to flexibility.
  Bending or stretching the molecule or individual bonds is not considered. Double or triple bonds
  and bonds within aromatic rings or within 3-, 4- or 5-membered rings are not considered flexible.
  All remaining single bonds are considered rotatable, if both atoms are sp2 or sp3 and carry at least
  one more non-hydrogen neighbor. For chains of conjugated triple bonds the following applies:
  If at least one terminal sp2/sp3 atom has no external neighbor, then no single bond is considered rotatable.
  Otherwise that terminal single bond connecting the smaller substituent is considered the only rotatable bond
  of the linear atom strand.

- Dihedral angle changes of a bond in the center of a molecule have a much larger effect on the
  overall molecule shape that the torsion change of e outer substituent bond. Therefore an influence
  factor is calculated for every rotatable bond. It also considers the limited degree of freedom of
  ring bond rotation:
  - for ring bonds: factor=0.33, since ring bonds cannot be changed without typically affecting two other rings bonds
  - other bonds: factor=sqrt(2.0 * smallerSideNonHydrogenAtomCount / moleculeNonHydrogenAtomCount)

- The flexibility contribution of every rotatable bond is determined from statistical torsion data of similar
  bonds in the Crystallography Open Database (COD). For instance C-O-C-C (C-O bond) is considered much more flexible
  than the central bond in a C-C-C-C chain, which has a strong preference towards the anti-conformation.
  For every rotatable bond a characterizing substructure is generated from its local environment.

This substructure consists of the two bond atoms, their direct neighbour atoms and the next shell of neighbour atoms.
It includes all connecting bond orders and various properties, e.g. whether atoms are in a ring, are aromatic,
the stereo configuration, etc. Then canonical representation of the bond's characterizing substructure is built.
The same mechanism was used before to characterize all rotatable bonds of all organic structures of the COD.
Now using the canonical key DataWarrior gets a torsion angle distribution curve of equal bond environments
from all organic structures of the COD. The number of peaks, their heights and their widths are now used to
generate the bond specific flexibility value. Frequency distributions with wide and multiple distribution maxima
of similar heights receive local flexibility values close to 1.0 while histograms with one narrow single
peak are close to 0.0.

- Finally, from the number of all molecule bonds, the number of rotatable bonds, their specific flexibility values,
and their geometry weighting factor an overall flexibility value is calculated with a non-linear incremental approach.

For more details I suggest to check the source code in
com.actelion.research.chem.conf.MolecularFlexibilityCalculat or

---

Subject: Re: Molecular Flexibility Descriptor
Posted by molcreator on Tue, 14 Jan 2020 02:59:34 GMT
View Forum Message <> Reply to Message

Good!