Clean3D is a project at the Department of Organic Chemistry, Eötvös Loránd University, Budapest which aims to automate 3D coordinates determination and store information on transmittable structures.

Clean3D is integrated into the software package Marvin from ChemAxon as a part of API and as a stand-alone component. It is also called within other applications through API.

At the beginning of the Clean3D development a good idea arisen which was worth to develop. The described method was aimed to determine multiple energetically favorable structures (conformers) for organic molecules. The multiple conformations can be stored in a database. During the building process database access can be implemented by querying structural fragments.

Development status: A direct method was developed which allows the off-line 3D coordinate process is implemented and validated. This method can be easily modeled from custom applications. The tuning of the cleaning process can be done by passing parameters.

Build steps can be classified as follow:

- Single atom fusing
- Fragment-fragment fusing
- Neighbor of an sp2 atom
- Atom-wise bond closing
- Neighbor of an sp3 atom
- Ring atom
- Other atoms

Fragment-fusing is currently under development. In many cases, the stored conformer count can be reduced dramatically (for the same quality in content with the parent atom fuse build process) if parts are built separated and joined together.

This problem can be solved by means of the following tools:

- Determine possible values of internal coordinates (bond lengths, angles, dihedral angles) by local heuristics
- Try to satisfy most important coordinate variations via triangulation
- Try to satisfy most important coordinate variations via triangulation
- Do geometry optimization if required
- Threw identical conformers
- Solved conformers to survive for the next step

The coordinate generation failed with the single closing atom or, at fragment-fragment fusing, align single close heuristic.

Multiple conformations can be stored in a database. During the building process database access can be implemented by querying structural fragments.

Multiple atom fusing

The connected multiple atom fragments will be joined together. The connecting bond defines an orientation for the fragment-fragment fusing. Overlapping directions can help to align the fragments.

Fragment database query

The frequency document fragments with multiple conformations can be derived. In database-driven access database access can be implemented by querying structural fragments.

Component hierarchy

- Interfaces
- API
- Command line

2D structures and conformers can be processed in Marvin like line and curve, Hidden lines can be removed.

Interfaces

- Topology analysis and querying
- Local heuristics
- Equivalency perception
- Selection and placement
- Direct molecule placement
- Molecular-mechanical force field (Dreiding)
- Command line

With the functionality given through the public API (Application Programming Interface) the 3D coordinate generation can be easily modeled from custom applications. The tuning of the cleaning process can be done by passing parameters.

Acknowledgements

The authors wish to express their thanks to the Hungarian National Science Fund (OTKA) for financial support (Grant No. K60108). Further thanks to the Hungarian Research Foundation (OTKA) for financial support (Grant No. 81642). Further thanks to the European Commission for the support of the COST Action C19, which is also acknowledged with gratitude.

References


