Conformator Command Line Documentation 2.3

Sascha Jung and Florian Flachsenberg

December 18, 2024

Contents

1 Introduction

All links, references, table of contents lines etc. in this pdf are clickable.

Please note that this package is a command line package.

Conformator is a command line tool to generate a single 3D conformation or ensembles of conformers for an input molecule.

Conformator at the command line features:

- fast generation of a single 3D conformation for the input molecules
- generation of ensembles of conformers

Please note: Generation of conformers for molecules containing rings with more than 9 atoms ("macrocycles") is currently not supported!

Conformator is a knowledge-based method for generating conformer ensembles. It is based on the same-named method[1] developed at the ZBH - Center for Bioinformatics (Universität Hamburg) and concepts of the CONFECT algorithm[3] and the torsion angle library of Guba et al.[2].

The first step is to generate initial 3D coordinates. This involves assigning ideal bond lengths and angles according to the VSEPR model to all acyclic bonds in the molecule and generating coordinates for the non-macrocyclic rings. The torsion angles of the acyclic bonds in the molecule are then adjusted discretely to select preferred torsion angles[2] and to minimize intramolecular clashes. If only one conformer is requested, that conformer is reported.

If more than one conformer is requested, Conformator will continue with the sampling of torsion angles and conformers of non-macrocyclic rings. Candidate conformers are generated that are as clash-free as possible by discretely adjusting the torsion angles of acyclic bonds and selecting conformers for non-macrocyclic rings. Here, preferred torsion angles from the torsion angle library[2] are used and acyclic bonds located near the center of the molecule are sampled more extensively. The resulting set of candidate conformers is clustered to the user-specified maximum number of conformers using an RMSD-based adaptive sphere exclusion clustering algorithm to generate a diverse conformer ensemble.

2 Jump Start: Swiftly Generate Conformers

Your license is all set? You are familiar with command line usage? Then here is a typical call to generate 3D conformers for your input molecules:

```
./conformator -i <path/to/my_2D_molecules.sdf> -o <path/to/my_3D_molecules.sdf>
```

The input molecules can be given in SD, smiles (.smi or .smiles file, containing line-separated SMILES), MOL and MOL2 format and the file can have multiple entries. To quickly generate a 3D conformation for a single molecule, the input can just as well be a SMILES string in quotation marks, for example:

```
./conformator -i "CC(C)C(=O)N" -o isobutyramide.sdf
```

With the above calls, a single conformer is generated for every input molecule. To generate multiple conformers (ensembles) for each input molecule, specify the desired number with the `--max-nof-conf` option:

```
./conformator -i "CC(C)C(=O)N" -o isobutyramide.sdf --max-nof-conf 20
```

3 Technical Prerequisites

3.1 Required Software

Conformator is a command line application.

Technically, you will need:

- The Conformator **application package** (from `https://biosolveit.de/download/?product=conformator`). Depending on your operating system, some libraries may have to be installed. Get in touch with us if that is the case:
  `mailto:support@biosolveit.com`. Please mention any errors/warnings that you see in your mail.
- A **shell** (Linux/Unix) or a terminal (macOS), or a command line environment (Windows; e.g.: cmd.exe or PowerShell)
- A valid **license** (from `mailto:license@biosolveit.de`), see below.

3.2 Licensing

Conformator needs a license to operate which is available from us. There are various sorts of licenses, but in most of the cases, your early testing will use a license file that simply needs to be placed next to the executable, see below.

The license setup instructions will come with the license that we will send out — or that has already been sent out to you. In case you do not have a license yet, please get in touch with us at `mailto:contact@biosolveit.de`, and provide us with the necessary information (see Obtaining a License File).

Please note: A valid SeeSAR license will also be accepted by Conformator.

License File Locations A "test license" that you can request online and that is sent to you instantaneously can simply be placed next to the executable (`conformator.exe`, `Conformator` or `conformator` — depending on your operating system). For macOS please read on...

---

macOS Specialties On macOS, the executable will typically reside inside the *.app package:

```
/Applications/Conformator.app/Contents/MacOS/Conformator
```

To place the short term test license there, you will have to go into the *.app package using a right mouse click (or CTRL-click) on Conformator.app in the Finder, and click on "Show package contents". In there, you will see the Contents/ subfolder, in there the MacOS subfolder, and in there, the Conformator executable. If you are about to use the **test license**, place is right there, next to the executable. A longer term license will be handled separately, we will tell you how when we send that very license.

When you call Conformator for the first time, go to the Finder, and navigate to the Applications folder. Do a right(!) click on Conformator.app, and — if applicable — confirm that you want to open the program. It will flash up once, and you are good to go at the terminal prompt from there on.

Obtaining a License File Using `--license-info` you can obtain information about the specification of your license server machine, the searched directories, and the validity of the currently used license files. This may also be useful when Conformator is not starting up as you would expect it to.

Call Conformator with the `--license-info` option, to see an output like this:

```
./conformator --license-info
================================================================================
 License Information:
================================================================================
Host-ID: "28f0763fa408 38c98636b9cb"
BIOSOLVE_LICENSE_FILE: /Applications/BioSolveIT/License
LM_LICENSE_FILE:

Currently used key/path after environment variables:
/software/conformator-2.3.0/conformator

Conformator:
 >> Valid key, expires on Tuesday, 31 December 2024

Request a license:
***
*** https://www.biosolveit.de/license/?product=conformator&operating_system=linux...
***
```

Request an evaluation or longer-term license using the link that is provided at the very bottom of the output. Also, this output may help us to find out if there are any problems with your license or its setup.

4 Command Line Options

4.1 Overview

An overview of all command line options is available by calling Conformator with `--help`.

```
./conformator -h

Calculates conformations for molecules in the input file

Program options:
-i [ --input ] arg               Input molecule file or single input molecule as SMILES. Supported file
                                 types are *.smi, *.smiles, *.mol2, *.mol and *.sdf
-o [ --output ] arg              Output file, supported file types are *.mol2 and *.sdf

Configuration:
--keep-input-sd-tags [=arg(=1)]  Keep input sd tags for the conformations.
--max-nof-conf arg (=1)          Set maximum number of conformations to be generated.
--rmsd-ensemble [=arg(=1)]       Calculate the minimum RMSD within the ensemble. Only available if more
                                 than 1 conformer is requested with --max-nof-conf. Implies
--write-log-file.
--rmsd-input [=arg(=1)]          Calculate the minimum RMSD of the closest ensemble member to the input
                                 conformation. Implies --write-log-file.
--write-log-file [=arg(=1)]      Write a logfile with additional information about the conformer
                                 calculation results (e.g., number of conformers per molecule, RMSD if
                                 requested). The logfile will be placed next to the output file (given by
                                 --output) with the same base name and a ".log" extension.

General options:
-h [ --help ]                    Print this help message.
--license-info                   Print license info.
--thread-count arg               Maximum number of threads used for calculations. The default is to use all
                                 available cores.
--version                        Print version info
-v [ --verbosity ] arg (=2)      Set verbosity level
                                      0 [quiet]
                                      1 [error]
                                      2 [warning]
                                      3 [info]
                                      4 [steps]
```

Please note that the abbreviated, one-letter options are preceded with one dash `-` whereas the longer, named options are preceded with two dashes: `--`.

In addition, we are available to support your endeavors with Conformator by email (`mailto:support@biosolveit.de`). We try to answer within a day, during business hours.

4.2 Program Options

–i [––input] arg Specify a file containing the input molecules. Supported file formats are SDF, MOL and MOL2. You can also provide a text file containing multiple line-separated SMILES (file extension must be .smi or .smiles). Instead of a molecule file, you can also specify a single SMILES string enclosed by quotation marks, e.g. ”CNCCC1=CC=CC(=C1)C(O)=O”.
NOTE: The `-i` option is required.

Examples:

```
conformator -i my_molecules.sdf
```

```
conformator -i "CC1=CC=CN=C1"
```

–o [––output] arg Specify a path and name for the output file. The file will contain all conformers generated for the molecules in the input file. Supported file formats are SDF and MOL2.
NOTE: The `-o` option is required.

Example:

```
conformator -o my_conformers.sdf
```

4.3 Configuration

––keep-input-sd-tags For every molecule in the input SD file the associated SD tags are read and added to every conformer generated for that molecule in the output SD file.
NOTE: Requires an input SD file and the specification of an output SD file.

Example:

```
conformator --keep-input-sd-tags
```

––max-nof-conf arg(=1) Specify the maximum number of conformers generated for each input molecule. Default value is 1 (a single conformer).

Example:

```
conformator --max-nof-conf 50
```

––rmsd-ensemble Calculates the minimum symmetry corrected heavy atom RMSD between all generated conformers in the ensemble. The minimum ensemble RMSD will be annotated in the log file which is generated with the `--write-log-file` option (RMSD_Ensemble entry, see above).
NOTE: This option can only be used if you generate ensembles of conformers, i.e., if `--max-nof-conf` is set to > 1.
NOTE: If this option is used, a log file is implicitly generated (see `--write-log-file` option above).

Example:

```
conformator --rmsd-ensemble
```

--rmsd-input Calculates the symmetry corrected heavy atom RMSD of the generated conformers to the input conformation. The RMSD can only be calculated if the input molecule has 3D coordinates. The RMSD of the conformer which has the closest conformation to the input structure will be annotated in the log file which is generated with the `--write-log-file` option (RMSD_Input entry, see above). NOTE: If this option is used, a log file is implicitly generated (see `--write-log-file` option above).

Example:

```
conformator --rmsd-input
```

--write-log-file Write a log file with additional information about the conformer results (see below). The log file will be placed next to the output file (given by `--output`) with the same base name and a `.log` extension.

Example:

```
conformator -i my_molecules.smi -o my_conformers.sdf --write-log-file
```

The above call will generate a log file (in CSV format) named `my_conformers.log` with the following information:

- **Entry:** Position of molecule in the input file (consecutive numbers starting at 1)
- **Molecule:** Name of the molecule (if annotated in the input file)
- **NofConfs:** Number of effectively generated conformers for the input molecule
- **Stereo:** Information about stereo chemistry generated for the conformers:
  - *NONE*: input molecule is achiral (no stereo centers)
  - *INPUT*: stereo chemistry of generated conformers fits to the stereo chemistry of input molecule
  - *PURE*: stereo chemistry was not annotated/defined for all stereo centers or bonds in the input molecule, but all conformers generated have the same defined stereo chemistry
- **RMSD_Input:** RMSD of the conformer which has the closest conformation to the input molecule. This can only be calculated if the input molecule has 3D coordinates, otherwise "inf" is reported here. This entry is only written if requested via the `--rmsd-input` option (see below).
- **RMSD_Ensemble:** Smallest RMSD difference between two conformers of the generated ensemble. This entry is only written if requested via the `--rmsd-ensemble` option (see below).
- **Status:** status of the conformer generation for this input molecule
  - *OK*: generation of the conformer (ensemble) was successful
  - *WARNING*: input molecule had too many rotatable bonds, which means that some bonds were only partially sampled. The resulting ensemble of conformers may not cover the whole conformational space.
  - *ERROR*: errors occurred during coordinate generation, no conformers could be generated for the respective input molecule

4.4 General Options

–h [––help] Displays the command line help with short descriptions for every argument option. For more information see Section 4.1.

Example:

```
conformator --help
```

––license-info Shows command line information about the license setup you currently use. If you have any problems with your license, send an email to `mailto:support@biosolveit.com` and include this information. For more information see Section 3.2.

Example:

```
conformator --license-info
```

––thread-count arg Specify the maximum number of threads used for the conformer generation. By default, all available logical cores of your computer are used. You may want to reduce the number of threads used if you want to run other computations on your computer at the same time, or if you share the compute resource.

Example:

```
conformator --thread-count 4
```

––version Displays information on the version of Conformator on the command line. In quoting Conformator, please mention this version number.

Example:

```
conformator --version
```

–v [––verbosity] arg(=2) Set the verbosity level, e.g., the level of console output, with an integer argument. The default value is 2. The following options are available:

- `0` Quiet. No messages will be displayed on the console. Errors will be ignored whenever possible.
- `1` Error. Only error messages will be displayed.
- `2` Warning. The default setting, warnings and error messages will be displayed.
- `3` Info. Additional information beyond errors and warnings are displayed.
- `4` Steps. In addition to the 'Info' option, the progress is displayed in detail.

Example:

```
conformator -v 0
```

5 Further Reading, References

The original ideas behind the Conformator method are covered in the original publication by Nils-Ole Friedrich et al.[1]. Further details can also be found in the publication by Schärfer et al.[3] and the publication on the torsion angle library by Guba et al.[2].

Complementary tools, especially also the graphical platform SeeSAR, can be obtained from the BioSolveIT website (`https://biosolveit.com/download/?product=seesar`).

References

[1] Nils-Ole Friedrich, Florian Flachsenberg, Agnes Meyder, Kai Sommer, Johannes Kirchmair, and Matthias Rarey. Conformator: A novel method for the generation of conformer ensembles. *Journal of Chemical Information and Modelling*, 59(2):731–742, 2019.

[2] Wolfgang Guba, Agnes Meyder, Matthias Rarey, and Jérôme Hert. Torsion library reloaded: A new version of expert-derived smarts rules for assessing conformations of small molecules. *Journal of Chemical Information and Modeling*, 56(1):1–5, 2016.

[3] Christin Schärfer, Tanja Schulz-Gasch, Jérôme Hert, Lennart Heinzerling, Benjamin Schulz, Therese Inhester, Martin Stahl, and Matthias Rarey. Confect: Conformations from an expert collection of torsion patterns. *ChemMedChem*, 8(10):1690–1700, 2013.

We wish you great success and much joy with Conformator!