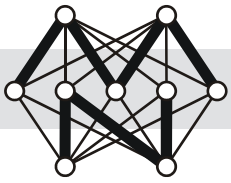

Molecular Networks

Inspiring Chemical Discovery

August 2006

New Features and Improvements of CORINA Version 3.4

Since August 2006, version 3.4 of the 3D Structure Generator CORINA is available. Due to almost 20 years of continual development, CORINA matured over a series of versions. With CORINA version 3.4, its unrivaled and outstanding performance and quality could be further improved. A brief summary of all major new features and improvements of CORINA version 3.4 are described in the following.

New Features

(1) The new input option `-i sdfi2c=<value>` ("sdf item-to-comment") can be used to copy the contents of the SDF data field named `<value>` (data header) in an SDF input file to a comment line in the output file, e.g., to the 3rd line in the header block of an SDFile. Thus, a piece of information from a data field in an SDFile can be transported to any file format which supports a comment (such as SDF, SYBYL MOL2 or PDB).

For example, the option `-i sdfi2c=MySDFdataField` copies the data entry "`My entry in my SDF data field`" of the SDF data field

```
>  <MySDFdataField>
My entry in my SDF data field
```

to the comment line of the SDF output file (3rd line in the header block):

```
ethane
CHStkserve06270620222D 0   0.00000     0.000004
My entry in my SDF data field
  8  7  0  0  0  0  0  0  0  0  1 V2000
```

Note. The original SDF data field of the input SDFile is preserved and also copied to the output SDFile. The size of the comment line may not exceed 80 characters.

(2) The new output option `-o sdfc2i=<value>` ("sdf comment-to-item") can be used to copy the comment line of an input file, e.g., 3rd line in the headerblock of an SDFile, to a newly generated SDF data field named `<value>` (data header) in the SDF output file. Thus, a comment line from file formats which support comments (such as SDF or SYBYL MOL2) can be transported to a data field of an SDF output file.

For example, the option `-o sdfc2i=MyNewSDFdataField` copies the comment line of the SDF input file

```
ethane
CHStkserve06270620222D 0   0.00000     0.000004
My input SDFile comment line
  8  7  0  0  0  0  0  0  0  0  1 V2000
```

to the data field MyNewSDFdataField in the SDF output file:

```
>  <MyNewSDFdataField>
My input SDFile comment line
```

Note. The original SDFile comment line is overwritten by the comment produced by CORINA. The size of the comment line may not exceed 80 characters.

(3) The new output option `-o pascom` ("pass comment") passes comment lines between file formats which support comments (e.g., SDF, SYBYL MOL2 or PDB). In the case of SDFiles as input and output this option preserves the comment line in the header block (3rd line) of the input SDFile and writes it to the comment line of the SDF output file. By default, CORINA overwrites the comment line of the input SDFile with information about the program version. This option prevents overwriting and passes the comment line of the input to the output file.

(4) The handling of large molecules has been changed. In general, CORINA does not have any limitations regarding the number of atoms or bonds of an input structure that should be converted into 3D. However, CORINA has been designed to process small to medium sized organic (typically "drug-like") molecules. The larger a molecules gets the more the intra-molecular interactions gain in importance influencing the secondary structure of a molecule. CORINA can model these interactions only to a limited extend and, therefore, is not able to correctly predict 3D structures of polymers and biopolymers such as proteins, enzymes or nucleic acids. For this reason, the following changes and improvements have been implemented.

- By default, the maximum number of atoms and bonds of an input structure is now limited to 999. This default limitation can be extended by the new driver option `-d maxat=<value>`. For example, the option `-d maxat=1001` extends this limit to a maximum of 1001 atoms and bonds. **Note.** Some file formats are limited by definition to a certain number of atom and bonds, e.g., SDFile is limited to 999. This driver option will not circumvent any of these limitations.

- The functionality of file format conversion (driver option `-d no3d`) is not affected by this limitation. In addition, all file format interfaces that are commonly used for macromolecular structures (e.g., PDB, MacroModel, Maestro, CIF and SYBYL MOL2) have been carefully reviewed to guarantee a proper file format conversion.

Improvements

(1) The stereoisomer module (driver option `-d stergen`) has been improved.

- Input 3D coordinates which may define the configuration at a stereo center are now ignored if no stereo descriptors (wedges, parity flags) are set. This is also the case if the additional driver option `-d preserve` is set in order to preserve defined stereo centers.
- The handling of spiro compounds and the hashcoding algorithm have been further improved for avoiding the generation of duplicate stereoisomers.

(2) New structural parameters (standard bond lengths and angles) have been added for the following atom types and substance classes according to some analysis of X-ray structures.

- New standard bond angles for S(sp3) (103°), P(sp3) (102°) and Se(sp3) (101°) have been added.
- The standard value for the CSC bond angle in thiophenes (s1cccc1) has been set to 91°.

(3) Several improvements in the file format interfaces to MDL SDF/RDF (read/write), SMILES (read), SYBYL MOL/MOL2 (read/write), MacroModel (write) and Maestro (write) file formats made the read and write routines more stable and reliable.

Availability

CORINA version 3.4 is available for the following hardware platforms and operating systems.

- Silicon Graphics, MIPS, IRIX 6.5
- Sun SPARC, SunOS 5.8 (Solaris 8) or higher
- x86 Linux kernel 2.4 or higher (distributions by SUSE and RedHat)
- Microsoft Windows (win32)