

# ConfGen 2.9

---

## User Manual

ConfGen User Manual Copyright © 2014 Schrödinger, LLC. All rights reserved.

While care has been taken in the preparation of this publication, Schrödinger assumes no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

Canvas, CombiGlide, ConfGen, Epik, Glide, Impact, Jaguar, Liaison, LigPrep, Maestro, Phase, Prime, PrimeX, QikProp, QikFit, QikSim, QSite, SiteMap, Strike, and WaterMap are trademarks of Schrödinger, LLC. Schrödinger, BioLuminate, and MacroModel are registered trademarks of Schrödinger, LLC. MCPRO is a trademark of William L. Jorgensen. DESMOND is a trademark of D. E. Shaw Research, LLC. Desmond is used with the permission of D. E. Shaw Research. All rights reserved. This publication may contain the trademarks of other companies.

Schrödinger software includes software and libraries provided by third parties. For details of the copyrights, and terms and conditions associated with such included third party software, use your browser to open [third\\_party\\_legal.html](#), which is in the docs folder of your Schrödinger software installation.

This publication may refer to other third party software not included in or with Schrödinger software ("such other third party software"), and provide links to third party Web sites ("linked sites"). References to such other third party software or linked sites do not constitute an endorsement by Schrödinger, LLC or its affiliates. Use of such other third party software and linked sites may be subject to third party license agreements and fees. Schrödinger, LLC and its affiliates have no responsibility or liability, directly or indirectly, for such other third party software and linked sites, or for damage resulting from the use thereof. Any warranties that we make regarding Schrödinger products and services do not apply to such other third party software or linked sites, or to the interaction between, or interoperability of, Schrödinger products and services and such other third party software.

August 2014

# Contents

---

Document Conventions .....	v
Chapter 1: Introduction .....	1
1.1 ConfGen Overview .....	1
1.2 Limitations .....	4
1.3 Running Schrödinger Software .....	4
1.4 Starting Jobs from the Maestro Interface .....	6
1.5 Citing ConfGen in Publications .....	8
Chapter 2: Running ConfGen from Maestro .....	9
2.1 Running a ConfGen Standard Search .....	9
2.1.1 Setting Up and Running the Job .....	9
2.1.2 Comparison of Methods .....	11
2.2 Running a Customized ConfGen Search .....	12
2.2.1 The Potential Tab .....	13
2.2.2 The Mini Tab .....	15
2.2.3 The ConfGen Tab .....	16
2.3 Starting the Job .....	18
Chapter 3: Running ConfGen from the Command Line .....	21
3.1 Using the confgen Command .....	21
3.2 Using MacroModel .....	23
3.2.1 Operation Codes (Opcodes) for ConfGen .....	23
3.2.2 Command File Example .....	34
Getting Help .....	37



---

# Document Conventions

In addition to the use of italics for names of documents, the font conventions that are used in this document are summarized in the table below.

Font	Example	Use
Sans serif	Project Table	Names of GUI features, such as panels, menus, menu items, buttons, and labels
Monospace	<code>\$SCHRODINGER/maestro</code>	File names, directory names, commands, environment variables, command input and output
Italic	<i>filename</i>	Text that the user must replace with a value
Sans serif uppercase	CTRL+H	Keyboard keys

Links to other locations in the current document or to other PDF documents are colored like this: [Document Conventions](#).

In descriptions of command syntax, the following UNIX conventions are used: braces { } enclose a choice of required items, square brackets [ ] enclose optional items, and the bar symbol | separates items in a list from which one item must be chosen. Lines of command syntax that wrap should be interpreted as a single command.

File name, path, and environment variable syntax is generally given with the UNIX conventions. To obtain the Windows conventions, replace the forward slash / with the backslash \ in path or directory names, and replace the \$ at the beginning of an environment variable with a % at each end. For example, `$SCHRODINGER/maestro` becomes `%SCHRODINGER%\maestro`.

Keyboard references are given in the Windows convention by default, with Mac equivalents in parentheses, for example CTRL+H (⌘H). Where Mac equivalents are not given, COMMAND should be read in place of CTRL. The convention CTRL-H is not used.

In this document, to *type* text means to type the required text in the specified location, and to *enter* text means to type the required text, then press the ENTER key.

References to literature sources are given in square brackets, like this: [10].



# Introduction

When generating collections of conformations for ligands, it is often sufficient to generate broad rather than exhaustive coverage of conformational space. With growing interest in generating conformations for many ligands, speed is becoming an important consideration. ConfGen, which was originally developed for rapidly and systematically exploring ligand conformations in Glide, has been enhanced and is available as a rapid conformational generation product. For a complete technical description of ConfGen along with validation studies, see the reference given in [Section 1.5](#).

ConfGen carefully examines the structure of the ligand to understand where to expect local minima as a function of rotations about rotatable bonds. It then systematically generates the conformations that arise from various combinations of these local minima. Thus it provides a broad and fairly uniform coverage of the available conformational space. This systematic approach used in ConfGen avoids the enormous amount of resampling of conformations that occurs in most conformational searching methods designed for exhaustive sampling. In addition to sampling rotatable bonds ConfGen also samples ring conformations, nitrogen atom inversions, and amide bond conformations. For a more detailed description of the conformational search, see [Section 2.1](#) of the *Glide User Manual*.

ConfGen's speed and ability to generate compact collections of quality candidate conformations forms a powerful combination with the force fields available with Schrödinger software, the GB/SA solvation model, minimization procedures and the redundant conformer elimination facilities of MacroModel, to provide very good coverage of conformational space. This facility is focused on the efficient generation of ligand conformations and may not be applicable or suitable to other types of problems.

ConfGen functions differently from the conformational searching methods supported in MacroModel: it is specifically focused on extremely efficient generation of ligand conformations similar to those found in protein-ligand complexes. For information on conformational searching procedures in MacroModel, see [Chapter 8](#) of the *MacroModel User Manual*.

## 1.1 ConfGen Overview

ConfGen technology was originally developed for rapid and effective systematic ligand conformation generation in Glide. A number of enhancements have been introduced, including the use of a larger ring conformation library, more control over amide bond geometries, taking into account molecular symmetry when generating conformers, and better invertible N atom

detection. The combination of ConfGen's speed and ability to generate compact yet diverse collections of candidate conformations with MacroModel's proven force fields, GB/SA solvation model, redundant conformer elimination, and minimization procedures provides a versatile and effective conformation generation facility.

What makes ConfGen so useful is that it carefully and systematically selects which conformations to produce, based upon an examination of the structure of the ligand being processed. Many other methods used for conformation generation construct new candidate conformers by random variation of internal coordinates or by systematic variation of internal coordinates without examining the geometric preferences of the molecule in question. These strategies are designed to eventually find all conformations for the molecule in question, but often end up resampling the same structures many times. Since the candidate structure generation is not intended to produce structures close to a local minimum in the potential energy surface, it is often necessary to carry out a time-consuming minimization in order to obtain a viable structure. For many problem-solving efforts, particularly those involving the rapid examination of the conformers of many ligands, complete enumeration of the conformations is not strictly necessary. Instead, rapid and broad coverage of conformation space by systematically generating structures close to most of the local minima on the potential energy surface, particularly the low-energy minima, often suffices. This is exactly what ConfGen is designed to do. To better simulate the structures found in protein-ligand complexes, ConfGen also tends to eliminate compact ligand conformations.

The flow of ConfGen conformational searches can be broken down into the following stages:

1. Read in the structure to be searched.
2. MacroModel minimizes the input structure (optional).
3. ConfGen generates and saves the collection of candidate conformations.
4. Cycle through MacroModel's conformer generation procedure:
  - a. Request a new candidate conformer from ConfGen.
  - b. Minimize the energy of the conformer (optional).
  - c. Examine the structure for acceptability (energy window (optional), appropriate chiralities, etc.).
  - d. Carry out the redundant conformer elimination process by comparing the current conformer with previously retained conformers to see if a conformer may be eliminated (optional).
5. Record results.

Before generating conformations, ConfGen analyzes the ligand structure to detect pyramidal nitrogen atoms with at most 3 non-hydrogen attachments, flexible rings, and rotatable bonds.



Terminal  $-\text{CH}_3$ ,  $-\text{NH}_2$ ,  $-\text{NH}_3^+$  groups are not considered to be rotatable because rotating such groups does not result in significant conformational changes. Other groups may also be optionally excluded from consideration as rotatable. For each rotatable bond, the potential energy as a function of the dihedral angle (using a truncated version of the OPLS\_2005 force field) is examined for local minima, the energy and location of which are recorded for later use. Some types of symmetry within the molecule are recognized during this process and the corresponding redundant minima eliminated from further consideration. During conformational sampling, only dihedral angles corresponding to these minima are sampled.

During conformation generation, the ligand is first divided into a core region and rotamer groups. A rotamer group is defined as a terminal group that does not contain a rotatable bond but is attached to the rest of the molecule by a rotatable bond (the *terminal rotatable bond*). The core is any part of the molecule that is not within a rotamer group. Internally, ConfGen estimates the energy of a conformation using the sum of the energy of the ring conformations and the energy for each rotatable bond (the *Cen* or *ConfGen energy*). Within ConfGen, relative energies are calculated as the Cen of the current conformer minus the Cen of the lowest energy conformer.

A collection of conformations is first built up based on all combinations of ring conformations for the various rings present, nitrogen atom inversions, and all minima for each rotatable bond in the core with all terminal rotatable bonds in their lowest energy minima. Conformers whose Cen is larger than a predefined cut-off value are discarded. The geometry of these conformers is then cleaned up by adjusting the torsional angles to minimize the sum of the Cen, and introducing a short-range excluded-volume energy term, to reduce atom collisions. The maximum number of core conformations to retain (*maxcore*) is set, based upon the number of ring systems, the number of invertible nitrogens atoms, and the number of rotatable bonds in the core. If more conformers than this are generated for the core, their number is reduced to *maxcore* by eliminating the more compact conformers.

After the core conformations are collected, the peripheral groups are sampled. ConfGen either samples all combinations of all minima for the terminal rotatable bonds, or samples the terminal rotatable bonds one group at a time, while all other terminal rotatable bonds remain in their lowest energy minima depending on the value of *arg4* of the *CGEN* command. Conformations whose relative energy exceeds a preset cut-off are eliminated. The remaining structures are minimized in the same manner as the core conformations.

ConfGen has an internal limit of 20,000 conformations. When this value is reached, ConfGen stops looking for more conformations. Typical ligands have, on average, a few hundred distinct conformers, but the actual number can vary dramatically from ligand to ligand.

At the end of the conformation generation process, ConfGen saves the conformations generated for future use and prints a summary message to the *.log* file:

Number of rotatable bonds	4	
Total ring and/or		
N inversion conformations	2	
Core rotatable bonds, maxkeep	3	1000
Maximum excitation level	11	
Energy cutoff	25.017	
Total core conformations	54	
Total conformations minimized	100	
Effective number of internal		
degrees of freedom	5.00	

The conformations are written to the output file, along with the conformational energies, given in kJ/mol.

## 1.2 Limitations

- Molecules containing more than 200 atoms or more than 35 rotatable bonds are skipped.
- ConfGen is designed to function optimally for ligand-like molecules. Other classes of molecules may not be well sampled by it.
- ConfGen searches are not exhaustive. However, the sampling is usually quite representative and the collection of conformers generated often includes conformers quite similar to those missing.
- ConfGen only samples up to 20,000 conformations
- In testing, between a small percentage of typical ligands contain ring systems that lack specific templates. Such ring systems are treated as rigid. You can generate your own templates with the `ring_templating` script—see [Section 17.8](#) of the *MacroModel User Manual*.
- ConfGen can generate conformers in which atoms that are distant in terms of connectivity may lie very close to each other. Minimization by MacroModel of the generated structures generally removes such problematic conformations.

## 1.3 Running Schrödinger Software

Schrödinger applications can be run from a graphical interface or from the command line. The software writes input and output files to a directory (folder) which is termed the *working directory*. If you run applications from the command line, the directory from which you run the application is the working directory for the job.

**Linux:**

To run any Schrödinger program on a Linux platform, or start a Schrödinger job on a remote host from a Linux platform, you must first set the `SCHRODINGER` environment variable to the installation directory for your Schrödinger software. To set this variable, enter the following command at a shell prompt:

```
csh/tcsh:      setenv SCHRODINGER installation-directory
bash/ksh:     export SCHRODINGER=installation-directory
```

Once you have set the `SCHRODINGER` environment variable, you can run programs and utilities with the following commands:

```
$SCHRODINGER/program &
$SCHRODINGER/utilities/utility &
```

You can start the Maestro interface with the following command:

```
$SCHRODINGER/maestro &
```

It is usually a good idea to change to the desired working directory before starting the Maestro interface. This directory then becomes the working directory.

**Windows:**

The primary way of running Schrödinger applications on a Windows platform is from a graphical interface. To start the Maestro interface, double-click on the Maestro icon, on a Maestro project, or on a structure file; or choose `Start` → `All Programs` → `Schrodinger-2014-3` → `Maestro`. You do not need to make any settings before starting Maestro or running programs. The default working directory is the Schrodinger folder in your Documents folder.

If you want to run applications from the command line, you can do so in one of the shells that are provided with the installation and have the Schrödinger environment set up:

- Schrödinger Command Prompt—DOS shell.
- Schrödinger Power Shell—Windows Power Shell (if available).

You can open these shells from `Start` → `All Programs` → `Schrodinger-2014-3`. You do not need to include the path to a program or utility when you type the command to run it. If you want access to Unix-style utilities (such as `awk`, `grep`, and `sed`), preface the commands with `sh`, or type `sh` in either of these shells to start a Unix-style shell.

**Mac:**

The primary way of running Schrödinger software on a Mac is from a graphical interface. To start the Maestro interface, click its icon on the dock. If there is no Maestro icon on the dock,

you can put one there by dragging it from the SchrodingerSuite2014-3 folder in your Applications folder. This folder contains icons for all the available interfaces. The default working directory is the Schrodinger folder in your Documents folder ( $\$HOME/Documents/Schrodinger$ ).

Running software from the command line is similar to Linux—open a terminal window and run the program. You can also start Maestro from the command line in the same way as on Linux. The default working directory is then the directory from which you start Maestro. You do not need to set the `SCHRODINGER` environment variable, as this is set in your default environment on installation. To set other variables, on OS X 10.7 use the command

```
defaults write ~/.MacOSX/environment variable "value"
```

and on OS X 10.8 and 10.9 use the command

```
launchctl setenv variable "value"
```

## 1.4 Starting Jobs from the Maestro Interface

To run a job from the Maestro interface, you open a panel from one of the menus (e.g. Tasks), make settings, and then submit the job to a host or a queueing system for execution. The panel settings are described in the help topics and in the user manuals. When you have finished making settings, you can use the Job toolbar to start the job.



You can start a job immediately by clicking Run. The job is run on the currently selected host with the current job settings and the job name in the Job name text box. If you want to change the job name, you can edit it in the text box before starting the job. Details of the job settings are reported in the status bar, which is below the Job toolbar.

If you want to change the job settings, such as the host on which to run the job and the number of processors to use, click the Settings button. (You can also click and hold, and choose Job Settings from the menu that is displayed.)



You can then make the settings in the Job Settings dialog box, and choose to just save the settings by clicking OK, or save the settings and start the job by clicking Run. These settings apply only to jobs that are started from the current panel.

If you want to save the input files for the job but not run it, click the Settings button and choose Write. A dialog box opens in which you can provide the job name, which is used to name the files. The files are written to the current working directory.

The **Settings** button also allows you to change the panel settings. You can choose **Read**, to read settings from an input file for the job and apply them to the panel, or you can choose **Reset Panel** to reset all the panel settings to their default values.

You can also set preferences for all jobs and how the interface interacts with the job at various stages. This is done in the **Preferences** panel, which you can open at the **Jobs** section by choosing **Preferences** from the **Settings** button menu.

**Note:** The items present on the **Settings** menu can vary with the application. The descriptions above cover all of the items.

The icon on the **Job Status** button shows the status of jobs for the application that belong to the current project. It starts spinning when the first job is successfully launched, and stops spinning when the last job finishes. It changes to an exclamation point if a job is not launched successfully.



Clicking the button shows a small job status window that lists the job name and status for all active jobs submitted for the application from the current project, and a summary message at the bottom. The rows are colored according to the status: yellow for submitted, green for launched, running, or finished, red for incorporated, died, or killed. You can double-click on a row to open the **Monitor** panel and monitor the job, or click the **Monitor** button to open the **Monitor** panel and close the job status window. The job status is updated while the window is open. If a job finishes while the window is open, the job remains displayed but with the new status. Click anywhere outside the window to close it.

Jobs are run under the **Job Control** facility, which manages the details of starting the job, transferring files, checking on status, and so on. For more information about this facility and how it operates, as well as details of the **Job Settings** dialog box, see the *Job Control Guide*.

## 1.5 Citing ConfGen in Publications

The use of this product should be acknowledged in publications as:

ConfGen, version 2.9, Schrödinger, LLC, New York, NY, 2014.

Please also cite the following paper:

Watts, K. S.; Dalal, P.; Murphy, R. B.; Sherman, W.; Friesner, R. A.; Shelley, J. C. ConfGen: A Conformational Search Method for Efficient Generation of Bioactive Conformers. *J. Chem. Inf. Model.* **2010**, *50*, 534.



# Running ConfGen from Maestro

This chapter describes how to run ConfGen from Maestro. There are two panels that you can use: the ConfGen Standard panel, which provides a selection of preset strategies for conformational searching, and the ConfGen Advanced panel, which provides the flexibility needed to design your own conformational search. These two panels are described in the sections below.

The input structures for all ConfGen calculations should be 3D, all-atom representations compatible with OPLS all-atom force fields. 2D structures should not be used as input. If you have 2D structures, you should convert them to 3D using LigPrep first.

## 2.1 Running a ConfGen Standard Search

In the ConfGen Standard panel you can run standard ConfGen searches by selecting one of four predefined and well-characterized conformational search methods. The ConfGen facility itself uses the OPLS\_2005 force field. Minimizations and energy filtering are performed using a distance dependent dielectric constant with a prefactor of 4.

To open the ConfGen Standard panel, choose Applications > ConfGen > Standard.

### 2.1.1 Setting Up and Running the Job

The source of input structures for the conformational search can be specified with the Use structures from option menu, which offers the following choices:

- Workspace (included entries)—Use the structures that are displayed in the Workspace
- Project Table (selected entries)—Use the entries that are selected in the Project Table.
- File—Use the structures from the specified file. To specify the file, enter the path to the file in the File name text box, or click Browse and navigate to the structure file. The file must be a Maestro file.

There are several search strategies available in the Search strategy section of the panel:

- Very fast (no energy filtering)—Generate up to 5 conformers per degree of freedom. As conformers are processed, eliminate conformers if they have an RMSD value of less than 1.25 Å and all of the dihedral angles involving polar hydrogen atoms are within 60 degrees relative to a previously accepted conformer.

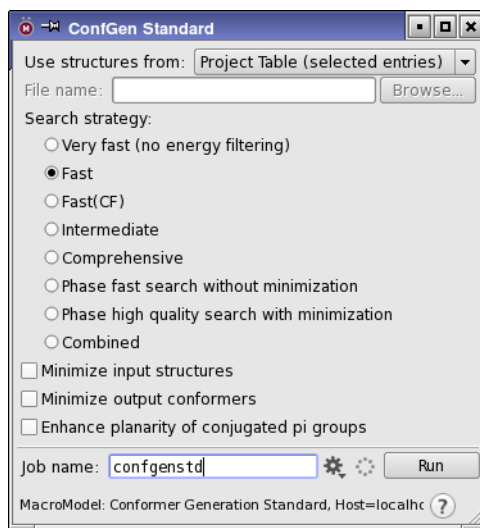


Figure 2.1. The ConfGen Standard panel.

- **Fast**—Similar to Very fast except that an RMSD value of 1.0 Å is used to detect redundant conformers, and conformers whose energy is more than 25 kcal/mol (104.67 kJ/mol) higher than the lowest energy conformer are eliminated.
- **Fast(CF)**—An alternate fast protocol, corresponding to the optimized protocol from Chen, I.-J.; Foloppe, N. *J. Chem. Inf. Model.*, **2010**, *50*, 822-839.
- **Intermediate**—Similar to Fast except that 75 conformers are generated per degree of freedom and the quality of ConfGen sampling is enhanced.
- **Comprehensive**—Similar to Intermediate except that an RMSD value of 0.5 Å is used to detect redundant conformations and the energy window for eliminating candidate structures is increased to roughly 120 kcal/mol (500 kJ/mol).
- **Phase fast search without minimization**—Search for conformers with the same settings used by Phase for database generation. No minimization is done; if you select this option the minimization options and the Enhance planarity of conjugated pi groups option are not available.
- **Phase high quality search with minimization**—Search for conformers with the same settings used by Phase in the Develop Pharmacophore Model workflow. Minimization is done on both the input and output structures. If you select this option the minimization options and the Enhance planarity of conjugated pi groups option are not available, because the settings are determined by the selection of this option.



- **Combined**—Run a Comprehensive search together with a Very fast search (if no output minimization is requested) or a Fast search (if output minimization is requested). The results are an improvement on the comprehensive search without taking much extra time.

The descriptions of the strategies just given focus on the gross changes in search strategies between the first five types of calculations. More settings are varied than are mentioned here.

For all except the Very fast strategy and the Phase strategies, you can choose to minimize the input structures before performing the conformational search, and you can choose to minimize the output conformers, by selecting Minimize input structures and Minimize output conformers. The minimization is done with the truncated Newton (TCNG) method, using 100 iterations for the input structures, and 50 iterations for the output structures. Minimization can help to correct poor geometries, but is not a substitute for 2D–3D conversion if the structures are 2D structures. For this purpose, you should use LigPrep (see the *LigPrep User Manual*).

If you are concerned about obtaining planar geometry for  $sp^2$  atoms, you can select Enhance planarity of conjugated pi groups, which increases the torsional potential around bonds between atoms whose geometry are formally planar. This option should make aromatic rings, amides, esters, and so on, less likely to adopt a nonplanar geometry. Nonplanar geometries are not in themselves unphysical, because the potential well at the planar geometry is not infinite, and some deviations from planarity can be induced by other nearby atoms.

When you have finished making settings, click the Settings button to open the Job Settings dialog box, in which you can make job settings and run the job, or click Run to start the job immediately with the current job settings. For more information on this panel, see [Section 2.2](#) of the *Job Control Guide*. See [Section 1.4 on page 6](#) for more information on starting jobs from Maestro.

### 2.1.2 Comparison of Methods

[Table 2.1](#) provides information on the trade-off of speed against quality for the four standard search methods with and without minimization of input and output structures (where applicable).

In general the progression from Very fast, Fast, Intermediate and Comprehensive resulted in more bioactive conformations of ligands being reproduced (particularly for smaller resolutions), more output conformations, and more CPU time being used.

For these structures minimization of the input structures had little effect on the quality or CPU time in part because the structures were produced by LigPrep and thus were low energy structures already. Minimization of the output structures generally resulted in slightly more ligands matching the bioactive conformations at all resolutions and for all the three methods compatible with minimization while increasing the CPU time required by a factor between 4 and 13.

Table 2.1. Results for standard ConfGen search methods.

Method	Percentage of ligands with at least one conformation within the given cutoff <sup>a</sup>				Conformers per ligand	CPU time per ligand (sec) <sup>b</sup>
	< 0.5Å	< 1.0Å	< 1.5Å	< 2.0 Å		
Very fast	16	52	84	96	14.6	1.0
Fast	19	55	82	95	13.5	2.3
Fast + mini <sup>c</sup>	20	60	85	96	17.0	9.1
Intermediate	21	64	85	97	38.4	6.5
" + mini <sup>c</sup>	24	69	89	97	46.6	80.2
Comprehensive	34	70	90	98	150.2	16.3
" + mini <sup>c</sup>	37	74	91	97	122.2	97.0

a. RMSD relative to the bioactive conformation.

b. Run on a single core of a 2.2 GHz dual-core Opteron 275 processor.

c. Minimization of input and output structures.

Which search method is best depends on how the conformations will be used in subsequent studies. For instance, for pharmacophore studies, such as those conducted using Phase a resolution of 2 Å is usually sufficient and conformer energies are relatively unimportant, so the Very fast method may be the most appropriate for your needs. For follow-on studies that require low energy conformers that closely reproduce the bioactive conformation, minimization in combination with intermediate or comprehensive searches may be most appropriate.

## 2.2 Running a Customized ConfGen Search

To run ConfGen conformational searches from Maestro with your own selection of settings, you can use the ConfGen Advanced panel. To open this panel, choose Applications > Confgen > Advanced. The panel contains three tabs: Potential, Mini, and CSearch, in which you can make the required settings. These tabs are described in detail in the following sections.

### To set up a ConfGen Advanced calculation:

1. Choose Advanced from the ConfGen submenu of the Applications menu.
2. Choose Workspace (included entries) or Project Table (selected entries) from the Use structures from option menu.
3. Select the appropriate settings in the Potential, Mini, and ConfGen tabs.

It is strongly recommended that you use either a distance-dependent dielectric or GB/SA solvation.

4. Click the Settings button to set up and launch the job, or choose Write from the Settings button menu to write the job files so they can be run from the command line or edited.

### 2.2.1 The Potential Tab

This tab is a modified version of the standard Potential tab described in [Section 4.2](#) of the *MacroModel User Manual*.

It is strongly recommended that you use either a distance-dependent dielectric or GB/SA solvation for ConfGen searches.

**Selecting a Force Field:** To specify a force field, select the desired force field from the Force field option menu. The force field list is limited to the MMFF, MMFFs, and OPLS force fields from OPLS\_2005 on. The default force field is OPLS\_2005. For information about implementation of the supported force fields, see [Section 2.1](#) of the *MacroModel User Manual*.

**Solvation and Electrostatic Treatment:** The settings for the Solvent, Electrostatic treatment, and Dielectric constant are interrelated. The default values of None, Distance dependent and 4.0 for these controls should work well for most ligands. If Solvent is set to any value other than None, a GB/SA solvation model is used and the recommended settings are Constant dielectric for the Electrostatic treatment and 1.0 for the Dielectric constant. The Force field defined setting for the Electrostatic treatment is equivalent to choosing a constant dielectric

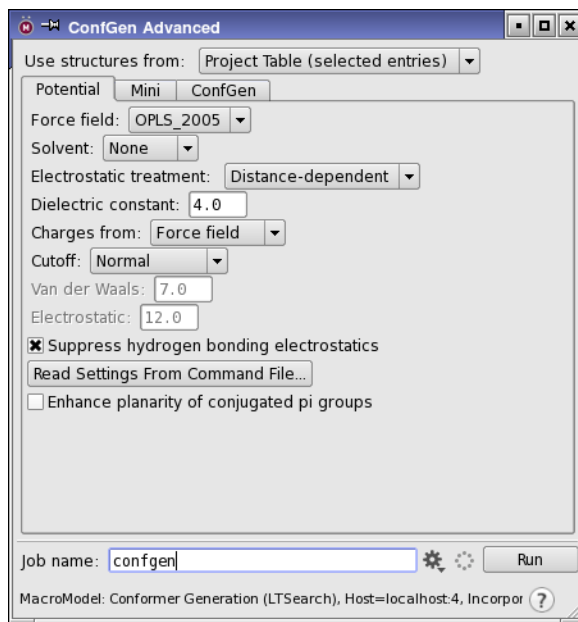


Figure 2.2. The Potential tab of the ConfGen Advanced panel.

with a value of 1.0. The MMFF and MMFFs force fields use a “buffered” constant dielectric treatment.

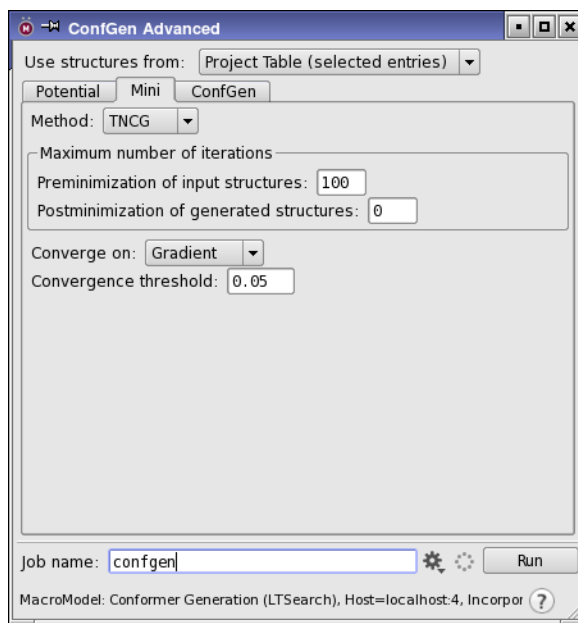
**Using Charges From a Structure File:** The charges used in the electrostatic portion of an energy calculation can either be assigned by the force field or obtained from the structure. Regardless of the source, charges are written to the structure file when a job is started. By default, the Force field option is used. To use charge information from the structure, select Structure file from the Charges from option menu.

**Choosing a Nonbonded Cutoff Setting:** The maximum distances over which van der Waals and electrostatic contributions to the molecular potential energy are computed may be controlled using the Cutoff option menu. Default cutoff distances are 7 Å for van der Waals and 12 Å for electrostatics. The van der Waals and electrostatic cutoff distances are the center of a soft cutoff that starts at 1 Å smaller than the specified distance and ends at 1 Å larger than the specified distance.

**Suppressing Hydrogen-Bonding Electrostatics:** Suppression of hydrogen-bonding electrostatics can aid in generating conformations closer to those found in protein-ligand complexes, where ligand donor and acceptor groups are likely to be hydrogen-bonded to the protein rather than to each other. To do so, select Suppress hydrogen bonding electrostatics. Deselect this option to model the interactions between ligand donor and acceptor groups according to the force field and solvation model. For more information, see the [CHYD](#) opcode in [Chapter 3](#).

**Using an Existing Command File:** To read a MacroModel command file and have the settings on the energetics panel update to reflect all potential energy settings, click Read Settings From Command File. Setting up MacroModel calculations in this way is helpful because you can easily reproduce your potential energy settings for multiple calculations.

**Treatment of formally planar functional groups:** If you are concerned about obtaining planar geometry for  $sp^2$  atoms, you can select Enhance planarity of conjugated pi groups, which increases the torsional potential around bonds between atoms whose geometries are formally planar. This option should make aromatic rings, amides, esters, and so on, less likely to adopt a nonplanar geometry. Nonplanar geometries are not in themselves unphysical, because the potential well at the planar geometry is not infinite, and some deviations from planarity can be induced by other nearby atoms.



**Figure 2.3.** The Mini tab of the ConfGen Advanced panel.

## 2.2.2 The Mini Tab

This tab provides options for preminimization of the input structures and postminimization of the output conformers. It is very similar to the standard Mini tab described in [Section](#) of the *MacroModel User Manual*.

**Choosing a Minimization Method:** From the Method option menu, select a minimization method. The following methods are available, and are described in more detail in [Section 6.2.1](#) of the *MacroModel User Manual*:

- PRCG (Polak-Ribiere Conjugate Gradient)
- TNCG (Truncated Newton Conjugate Gradient)
- OSVM (Oren-Spedicato Variable Metric)
- SD (Steepest Descent)
- FMNR (Full Matrix Newton Raphson)
- LBFGS (Low-memory Broyden-Fletcher-Goldfarb-Shanno)
- Optimal

**Setting the Number of Iterations:** You can set the number of iterations for preminimization of input structures and postminimization of generated structures independently, using the two text boxes in the Maximum number of iterations section:

- Preminimization of input structures—specify the maximum number of iterations for the input structures.
- Postminimization of generated structures—specify the maximum number of minimization steps for the generated structures (conformers).

Minimizing the input structures is important for obtaining useful collections of conformers since the quality of torsional potentials within ConfGen, and thus the estimates for the locations of the local minima in torsional space, is quite sensitive to the bond lengths and angles provided to ConfGen. If the input conformers are already minimized this step may be skipped.

Unlike other search methods, ConfGen attempts to generate conformations that are close to a local potential minimum. As a result, for many applications, it is usually not necessary to carefully minimize these generated structures. Omitting the postminimization can result in substantial reduction in the overall processing time.

**Setting Convergence Parameters:** You can select the convergence criterion from the Converge on option menu, and enter the convergence threshold in the Convergence threshold text box. The criterion choices are:

- Gradient—converge the gradient to the specified threshold in  $\text{kJ mol}^{-1} \text{Å}^{-1}$ .
- Energy—converge the energy to the specified threshold in  $\text{kJ mol}^{-1}$ .
- Movement—converge the maximum displacement to the specified threshold in Å.
- Nothing—run to the full number of iterations specified.

### 2.2.3 The ConfGen Tab

This tab controls how the ConfGen search is conducted in terms of the number and types of conformations generated, and the elimination of high-energy or redundant conformers.

**Specifying the search parameters:** There are several ways of limiting the number of conformations generated and the number of conformers saved:

- Set a limit on the total number of conformations returned from ConfGen in the Total text box under Maximum number of conformations returned. If ConfGen finds more conformations than this for a given ligand, it uniformly selects conformations from this collection to ensure broad coverage of conformational space. If ConfGen finds fewer conformations than this for a given ligand, only that many conformations are returned.
- Set a limit on the number of conformations returned as a multiple of the number of rotatable bonds in the molecule, by selecting Number of rotatable bonds times and entering the multiple in the text box. The number of conformations is limited to the specified multiple of the number of rotatable bonds, if this is less than the total number set.

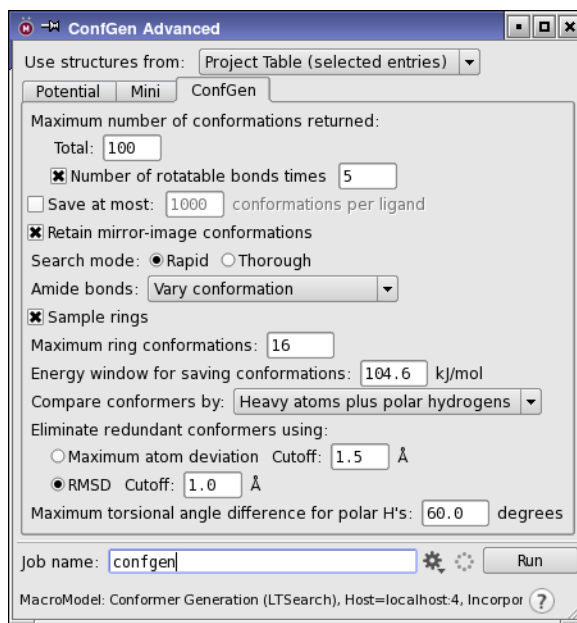


Figure 2.4. The ConfGen tab of the ConfGen Advanced panel.

- Set a limit on the total number of conformations saved per ligand, by selecting Save at most  $N$  conformations per ligand and entering the limit in the text box. The  $N$  lowest-energy conformers are saved, after redundant conformer elimination and application of other criteria. If this option is deselected, all generated conformations are saved that meet any of the other criteria set.
- Select a search mode. The search mode controls what combinations of rotations about terminal rotatable bonds (TRB) are sampled. In Rapid mode, the minima for each TRB are sampled separately with all other TRBs in their lowest energy orientation. In Thorough mode, all combinations of TRB minima are sampled.

In addition, you can choose to treat mirror-image structures as separate conformers, by selecting Retain mirror-image conformations. Distinguishing mirror images is important for ligand–protein interactions, where it is usually only one enantiomer that is active.

**Specifying structure options:** You can select one of three options from the Amide bonds option menu to control how amide bonds that do not lie in a ring are sampled by ConfGen:

- Vary conformation
- Retain original conformation
- Set to trans conformation

You can also select the Sample rings option to enable ConfGen to seek specific templates for rings systems present in each ligand. The number of ring conformations returned can be set in the Maximum ring conformations text box. Deselect this option to prevent ConfGen from sampling ring systems.

**Filtering and eliminating conformers:** Conformers are returned in order of the force-field energy, relative to the lowest-energy conformer. To set the energy cutoff for eliminating high-energy conformers, enter a value in the Energy window for saving conformations text box. Any conformer with a force-field energy greater than this cutoff above the minimum energy conformer is not retained in the output structure file.

Redundant conformers may be eliminated using geometric criteria. Redundant conformer elimination involves comparing each pair of conformers to see if they are different enough to justify retaining both of them. The choice of atoms for comparison is made from the Compare conformers by option menu, which has the following choices:

- None: no redundant conformers are eliminated
- Heavy atoms—conformational comparisons are preformed using the positions of non-hydrogen atoms (activates Distance cutoff for redundant conformers).
- Heavy atoms plus polar hydrogens: conformational comparisons are performed using the positions of non-hydrogen atoms and the torsional angles for polar hydrogens.

You can choose the criterion used for assessing whether conformers are the same by selecting either Maximum atom deviation or RMSD under Eliminate redundant conformers using, and you can set the threshold in the corresponding Cutoff text box. The default cutoff is 1.5 Å for maximum atom deviation and 1.0 Å for RMSD. If you chose Heavy atoms plus polar hydrogens, you can also set a threshold in the Max torsional angle difference for polar H's text box. Conformers are considered distinct if a dihedral angle involving a polar hydrogen in the two conformers differs by more than the value in this text box. The default is 60°.

## 2.3 Starting the Job

Once you have made all the desired settings, click the Settings button to open the Job Settings dialog box. This dialog box allows you to decide what to do with the output, set the job name, and also to distribute the job over multiple hosts and processors. A detailed description of this dialog box is given in [Section 2.2](#) of the *Job Control Guide*. If you want to reuse the settings from the last ConfGen job, click Run. See [Section 1.4 on page 6](#) for more information on starting jobs.

You can also save the input files so that you can run the job from the command line. This can be useful if you want to customize the job. To do this, choose Write from the Settings button



menu, and provide a job name in the dialog box that opens. The files are written to the current working directory, using the job name as the stem of the file name: so, for example, the structure file is written in Maestro format as *jobname.maegz*. You can then use the `confgen` command to run the job—see [Section 3.1 on page 21](#) for details.



# Running ConfGen from the Command Line

ConfGen searches can be run from the command line in two ways. One is to use the `confgen` command, which is described in [Section 3.1](#). The other is to run the underlying MacroModel command. This approach is described in [Section 3.2](#).

You can run ConfGen from the command line on Linux or Mac in a terminal window, after setting the `SCHRODINGER` environment variable. The command must be prepended with `$$SCHRODINGER/`. To run ConfGen from the command line on Windows, you can open a Schrodinger Command Prompt window from the Start menu. This window is a DOS shell, but you can use the same syntax as running from Linux, except that you do not need to prepend the command with `$$SCHRODINGER/`. See [Section 1.3 on page 4](#) for details for all platforms.

## 3.1 Using the `confgen` Command

The primary way of running ConfGen is with the `confgen` command. This command allows you to run ConfGen with the simplest options. The command syntax is as follows:

```
confgen [options] inputfile
```

The input file can be a Maestro file or a ConfGen input file, or both. The ConfGen input file is the same as the MacroModel simplified input file—see [Section 3](#) of the *MacroModel Reference Manual* for details. A Maestro file specified on the command line supersedes the structure file specified in the ConfGen input file, so you can easily reuse the ConfGen input file with many Maestro files. If only a Maestro file is given, the default calculation is run.

The standard Job Control options, described in [Table 2.1](#) of the *Job Control Guide*, are accepted, and in addition, the `-LOCAL`, `-NOJOBID`, and `-WAIT` options, described in [Table 2.2](#) of the *Job Control Guide*, are accepted. The remaining options are described in [Table 3.1](#).

You can run jobs with more than one mode, by quoting the options for each mode. The options can be separated by commas or spaces. For example,

```
$$SCHRODINGER/confgen '-FAST, -MIN_INP' '-VERY_FAST' inputfile
```

runs two jobs, one with the options `-FAST` and `-MIN_INP`, and another with the option `-VERY_FAST`. The results of the two jobs are collated into a single output file, without eliminating any redundant conformers. Running with more than one mode and combining the results often gives better coverage of conformational space. These jobs can be distributed, and each mode is split into the number of subjobs specified by `-NJOBS`.

Table 3.1. Options for the confgen command.

Option	Description
-h[e p]	Print usage message and exit.
-v	Display the version number and exit.
-NJOBS <i>n</i>	Divide the computational work into <i>n</i> subjobs.
-JOBCTS <i>nstr</i>	Ensure that each subjob has no more than this many structures to process. Default: 100.
-MIN_INP	Minimize input structures.
-MIN_OUT	Minimize output structures.
-FORCE_PLANAR	Use force-field parameters that increase the penalty for nonplanarity of conjugated pi groups.
-DRIVERHOST <i>host</i>	Host on which to run the driver (master) job. Default is the first host specified by -HOST. If used with -HOST, -HOST specifies the subjob hosts.
-SUBHOST <i>subhosts</i>	Hosts on which to run the subjobs, when the job is distributed. Default: the hosts specified by -HOST, except the first.
<i>Mode options, only applicable with a Maestro input file. See Section 2.1.1 on page 9 for details.</i>	
-VERY_FAST	Run in very fast mode.
-FAST	Run in fast mode (Default).
-FAST_CF	Run with the alternate fast protocol of Chen and Foloppe.
-PHASE_FAST	Run with the settings used in Phase for database generation.
-PHASE_QUALITY	Run with the settings used in the Phase Develop Common Pharmacophore Hypotheses workflow.
-INTERMEDIATE	Run in intermediate mode.
-COMPREHENSIVE	Run in comprehensive mode.

If you want to use custom ring templates, you can generate them with the `ring_templating` script. See Section 17.8 of the *MacroModel User Manual* for details, including information on where to place the templates for personal and general use.

## 3.2 Using MacroModel

Since ConfGen's code is supported by MacroModel, you can use `bmin` to run the searches. The input files, both the structure file and the command file, are written when you click Write in either of the ConfGen panels.

The syntax for command-line operation is as follows:

```
$SCHRODINGER/bmin [options] jobname
```

where the command file is named `jobname.com`. See [Section 2.2.2](#) and [Section 2.4.3.2](#) of the *MacroModel Reference Manual* for more information on this command, including tables of options. The results are returned in a compressed Maestro file (`.maegz`).

You can also run ConfGen searches with the simplified MacroModel input file and the following command syntax.

```
$SCHRODINGER/macromodel [options] jobname
```

See [Chapter 3](#) of the *MacroModel Reference Manual* for more information on the simplified input file and the `macromodel` command.

### 3.2.1 Operation Codes (Opcodes) for ConfGen

This section lists the operation codes that are used in the MacroModel input (`.com`) file for ConfGen. In addition to these opcodes, `DEBG 200` turns on greater verbosity for ConfGen, and `DEBG 202` overrides the force field setting and uses `OPLS_2001`.

ConfGen searches are inherently serial in that each input structure is used to seed a separate conformation generation calculation. The `.com` files should be constructed for serial searches (see [Section 3.2.2 on page 34](#) for an example `.com` file) even when processing only one structure. ConfGen has many options that can affect its behavior. Additional options for ConfGen can be selected using the `CGOP`, `CGO2`, `CGO3`, `CGO4`, `CGO5`, and `CGO6` opcodes. It can also be advantageous to suppress hydrogen bonding electrostatics using the `CHYD` opcode during CGEN conformational searches.

#### CGEN — ConfGen

*arg1* Maximum number of structures to request from ConfGen

*arg2* Maximum number of structures to retain while running

- 0 The maximum number of conformations to retain while running is the value specified in *arg1* or 10,000, whichever is greater

> 0 This is the maximum number of conformations to retain while running

*arg4 Peripheral sampling option*

- 1 Rapid Sampling: only generate conformers in which at most one peripheral group is rotated away from its lowest internal energy conformation
- 2 Thorough Sampling: sample all combinations of rotations of peripheral groups

*arg6 Allowable interatomic approach distance*

Fraction of sum of van der Waals radii which is used as a closest atomic approach limit (default: 0.25)

*arg7 Enhance planarity of groups with bonds between  $sp^2$  atoms*

Increase the torsional potential around bonds between  $sp^2$  atoms, such as aromatic rings, amides, and esters, whose geometry should ideally be planar.

- 0 Default. Do not enhance planarity of groups with bonds between  $sp^2$  atoms.
- 1 Enhance planarity of groups with bonds between  $sp^2$  atoms.

*arg8 Verbose reporting on CGEN processing*

Related DEBG flag: 200.

### CGOP — ConfGen Options

CGOP is one of a number of opcodes (CGOP, CGO2, CGO3, CGO4, CGO5, CGO6) that provide access to additional options for the ConfGen facility.

*arg1 Sampling symmetric terminal groups*

Terminal  $-CH_3$ ,  $-NH_2$  and  $-NH_3^+$  groups are never sampled by ConfGen. Sampling the conformations of other types of symmetric terminal groups is sometimes not useful. The identification of such groups is complex but boils down to an atom with identical groups attached to it. These groups can be monoatomic or composed of atoms with only 2 or 3 hydrogen atoms bonded to them. Checks for rotational symmetry are also imposed (e.g. if there are two groups but they do not lie  $180^\circ$  apart from each other, they would need to be sampled). Typical examples include  $-SO_3^-$ ,  $-N(CH_3)_3^+$ , and  $-NO_2$ .

0 do not sample these symmetric terminal groups

1 sample such terminal groups

*arg2 Ring conformation sampling*

Ring sampling is done by matching the rings in the molecules with templates for which the low-energy conformers have been previously identified. Two such template matching systems are supported:

- one using general forms for flexible 5 and 6 atom rings.
- one employing an extensive collection of specific templates for a large variety of ring systems (`ring_conf` utility).

The specific templating system has broader coverage (several hundred templates) and the template conformations were generated using MacroModel searches employing the MMFFs force field.

0 or 2 use the specific templating system (`ring_conf`)

1 use the general templating system

3 do not sample ring conformations

*arg3 Minimization of generated structures*

0 minimize generated conformations using at most the number of iterations in the MINI command

-1 do not minimize the generated conformations, but estimate the current energy, superimpose the generated conformations, and eliminate redundant conformations

-2 do not post process the conformations provided by ConfGen

-3 do not estimate the energies for the conformations produced but superimpose the generated conformations, and eliminate redundant conformations

> 0 minimize generated conformations using at most this number of iterations

*arg4 Non-ring amide bond conformation sampling*

0 or 1 default: vary the geometry of amide bonds

2 retain the original amide bond geometry

3 make amide bond geometry trans

*arg5* Maximum relative ring conformation energy in kJ/mol

0 use default value of 48 kJ/mol

> 0 use this value (kJ/mol)

*arg6* Upper limit on the number of combinations of ring conformations sampled

0 use default value of 16

> 0 use this value

*arg7* Upper limit on the number of ring conformations sampled per ring system

0 use default value of 8

>0 use this value

*arg8* Maximum relative ConfGen internal energy

0 use default value of 50.0 kJ/mol

> 0 use this value (kJ/mol)

#### **CGO2 — additional ConfGen Options**

CGO2 is one of a number of opcodes (CGOP, CGO2, CGO3, CGO4, CGO5, CGO6) that provide access to additional options for the ConfGen facility.

*arg1* Carboxylic acid bond conformation sampling

0, 1 default: vary the geometry of carboxylic acid bonds

2 retain the original carboxylic acid bond geometry

3 make carboxylic acid bond geometry trans

*arg3* Method for estimating the number of degrees of freedom

When [AUOP arg5](#) is greater than zero, the number of conformations to sample is given by the product of this value and the number of degrees of freedom. This argument controls which method is used to estimate the number of degrees of freedom.

0, 2 Use ConfGen's internal estimate for the number of degrees of freedom, which is given by the number of rotatable bonds +  $\log_2$ (number of ring conformation-nitrogen atom inversion combinations)



1 Use the number of rotatable bonds given by MacroModel.

*arg4 Include the most extended conformers*

When selecting a subset of conformers from those produced by ConfGen include the *arg4* most extended conformers first. In each conformer the largest distance between any two heavy atoms is calculated. The ratio of this value to the largest such distance amongst all of the conformations for this molecule is used as a measure for how extended the conformer is.

*arg5 Van derWaals radius scaling factor for close atomic approaches*

ConfGen rejects conformations with atoms closer than this factor times the sum of their van der Waals radii.

$\leq 0$  use default value of 0.6

$> 0$  use this value

*arg6 Scaling factor for close atom gradients*

ConfGen internally minimizes structures on a simple potential function that includes a penalty term for close approaches of atoms. This argument specifies the scaling factor for this penalty term.

$\leq 0$  use default value of 1.0

$> 0$  use this value

*arg7 Minimum van derWaals atom radius used for rejecting structures*

If the van der Waals radius for an atom is less than this value, use this value instead to determine when atoms are too close within a candidate structure.

0 use default value of 1.0 Å

$> 0$  use this value

### **CGO3 — Additional ConfGen Options for enhanced sampling of weak torsional potentials**

CGO3 allows the user to activate and adjust enhanced ConfGen capabilities for sampling weak torsional potentials. CGO3 is one of a number of opcodes (CGOP, CGO2, CGO3, CGO4, CGO5, CGO6) that provide access to additional options for the ConfGen facility.

sp<sup>2</sup>-sp<sup>3</sup> bonds often have only two potential minima and usually have low barriers to rotation. As a result, environmental factors can force such bonds to adopt conformations that do not closely resemble either minimum. This is a concern when trying to reproduce the bioactive conformations of ligands. CGO3 activates functionality that detects rotatable bonds with low barriers to rotation and replaces the original torsional potential with a simple cosine function with a user-specified periodicity. The cosine function is shifted so that one of its minima coincides with the deepest minimum in the original torsional potential. ConfGen then systematically samples all combinations of minima in all torsional potentials including the new artificial minima for the cosine functions.

If CGO3 is not present then the original weak potential and its minima are used. The presence of CGO3 activates replacement of weak potentials by cosine functions.

*arg1 Cosine function frequency*

Number of times the cosine function repeats when rotating by 360 degrees. This also is the number of minima spaced 360/arg1 degrees apart that will be sampled.

0 use the default frequency: 6 (spacing of 60°).

> 0 use this frequency (maximum 36).

*arg5 Energy range*

The difference in energy between the minima and the maxima in the cosine potential. This energy range is used to determine which torsional potentials are weak. If the extremes of the original torsional potential all lie within arg5 of each other then the original potential is deemed weak and is replaced by a cosine function.

0 use the default energy range 12 kJ/mol

> 0 use this value (kJ/mol)

*arg6 Maximal original energy for new minima*

All arg1 minima of the cosine potential are normally sampled by ConfGen. Some of these minima may correspond to relatively high potential regions of the original potential. The value in arg6 sets an upper bound on the energy difference between the lowest energy in the original potential and the energy in the original potential corresponding to a minimum in the cosine potential. If that bound is exceeded then this minimum in the cosine function is not sampled.

0 do not eliminate minima

>0 use this value as the upper bound on the energy difference (kJ/mol)

*arg7* *restraining potential prefactor*

If the conformations are being minimized inside MacroModel itself, the dihedrals with weak torsional potentials have a periodic flat-bottomed cosine restraining potential applied. This value is the prefactor for the cosine potential.

0 use 1000 kJ/mol

> 0 use this value (kJ/mol)

*arg8* *Half-width of the flat-bottom potential*

If the conformations are being minimized inside MacroModel itself, the dihedrals with weak torsional potentials have a periodic flat-bottomed cosine restraining potential applied. This value is the half width of the flat portion of the potential.

0 use 10 degrees

> 0 use this value (degrees)

**CGO4 — Additional ConfGen Options for rejecting conformations with close contacts between charged functional groups**

CGO4 allows the user to activate and adjust enhanced ConfGen capabilities for rejecting conformations that have formally charged functional groups that lie too close to each other. CGO4 is one of a number of opcodes (CGOP, CGO2, CGO3, CGO4, CGO5, CGO6) that provide access to additional options related to this facility.

ConfGen inherently generates conformations without considering electrostatics. This can occasionally lead to conformations that have functional groups with a formal charge of the same sign lying near each other. If minimization or energy filtering of conformations is not done inside MacroModel itself (CGEN arg3 = -1 or -2, respectively) these conformations may end up being saved in the output structure file. CGO4 turns on a penalty mechanism that can exclude such conformations within ConfGen itself. The penalty for conformation  $k$  is calculated from the fractional formal charges  $q_i$  on the atoms using the expression

$$P_k = \sum_{ij} p_{ij}$$

where the sum runs overall all pairs of atoms  $i$  and  $j$  that are separated by at least 3 bonds, and

$$\begin{aligned} p_{ij} &= 0, & q_i q_j &\leq 0 \\ &= A q_i q_j \exp(-r_{ij}^2 / 2s^2), & q_i q_j &> 0 \end{aligned}$$

where  $A$  and  $s$  are constants and  $r_{ij}$  is the distance between atoms  $i$  and  $j$ .

If  $P_k - P_{\min} > P_{\text{cutoff}}$  where  $P_{\min}$  is the smallest penalty for any of the conformations then the conformation is eliminated within ConfGen.

*arg5* Standard deviation of the Gaussian penalty,  $s$

0 use 5 Ang

> 0 use this value (Ang)

*arg6* Maximum value for the Gaussian,  $A$

0 use 2.5

> 0 use this value

*arg7* Relative penalty cutoff,  $P_{\text{cutoff}}$

0 use 0.75

> 0 use this value

### **CGO5 — Additional ConfGen Options for functional groups that are close to polar hydrogens**

CGO5 allows the user to activate and adjust enhanced ConfGen capabilities for rejecting conformations that have polar hydrogen atoms too close to groups with positive formal charges, and also to reward intramolecular hydrogen bonds. CGO5 is one of a number of opcodes (CGOP, CGO2, CGO3, CGO4, CGO5, CGO6) that provide access to additional options for the ConfGen facility.

ConfGen inherently generates conformations without considering electrostatics. This can occasionally lead to conformations that have functional groups with a bond to a hydrogen close to and pointing towards a functional group with a positive formal charge. These conformations may end up being saved in the output structure file particularly if minimization or energy filtering of conformations is not done inside MacroModel itself (CGEN *arg3* = -1 or -2, respectively). CGO5 turns on a penalty mechanism that can exclude such conformations within ConfGen itself. The penalty for conformation  $k$  is calculated from the fractional formal charges  $q_i$  on the atoms using the expression

$$P_k = \sum_{ij} p_{ij}$$

where the sum runs over all all pairs of polar bonds to hydrogens  $i$  and positively charged atoms  $j$ . Polar bonds to hydrogens involve a hydrogen atom (H) bonded to a O or N atom. The O or N atom must be at least three bonds away from atom  $j$ . The penalty term is defined by

$$P_{ij} = P_{Hj} - P_{ONj}$$

where  $p_{Hj}$  and  $p_{ONj}$  have the same functional form,

$$p_{Xj} = Aq_j \exp(-r_{Xj}^2/2s^2)$$

$A$  and  $s$  are constants and  $r_{Xj}$  is the distance between atoms  $X$  and  $j$ .

If  $P_k - P_{\min} > P_{\text{cutoff}}$  where  $P_{\min}$  is the smallest penalty for any of the conformations then the conformation is eliminated within ConfGen.

Another situation that can be encountered is when a polar hydrogen from a cation is too close to an aromatic or pi ring system. In this case, such conformations are rejected if the polar hydrogen is closer to the ring than a specified cutoff.

Similarly, ConfGen might not find conformations in which there is a hydrogen bond: a polar bond to hydrogen pointing towards a hydrogen-bond acceptor (an atom with a partial negative charge). For such situations, a reward is applied that has the same form as the penalties, and the sum is over negatively-charged atoms  $j$ . No cutoffs are applied to the reward term.

*arg2* Add rewards for intra molecular hydrogen bonds

- 0 do not reward intramolecular hydrogen bonds.
- 1 reward intramolecular hydrogen bonds. *arg5* and *arg6* are used to specify the reward.

*arg5* Standard deviation of the Gaussian,  $s$

This value is used as specified for penalties, and multiplied by 0.2 for rewards.

- 0 use 5.0 Å
- > 0 use this value (Å)

*arg6* Maximum value for the Gaussian,  $A$

This value is used as specified for penalties, and multiplied by 10.0 for rewards.

0 use 1

> 0 use this value

*arg7* Relative penalty cutoff,  $P_{cutoff}$

0 use 0.25

> 0 use this value

*arg8* Distance cutoff between cation-H and pi ring system, in angstroms

0 use 1.0

>0 use this value

#### **CGO6 — Additional ConfGen Options for rejecting conformations with close contacts between nonhydrogen atoms**

CGO6 allows the user to activate and adjust enhanced ConfGen capabilities for rejecting conformations that have too many nonhydrogen atoms in close proximity. CGO6 is one of a number of opcodes (CGOP, CGO2, CGO3, CGO4, CGO5, CGO6) that provide access to additional options for the ConfGen facility.

Bioactive conformations tend to be relatively extended. ConfGen can generate conformations in which topologically distant parts of the ligand lie close to each other. These conformations may end up being saved in the output structure file. CGO6 turns on a penalty mechanism that can exclude such conformations within ConfGen itself particularly if the groups approaching each other are rings. The penalty for conformation  $k$  is calculated using the expression:

$$P_k = \sum_{ij} p_{ij}$$

where the sum runs over all all pairs of nonhydrogen atoms  $i$  and  $j$  that are at least three bonds away from each other. The penalty term is defined by

$$p_{ij} = A \exp(-r_{ij}^2/2s^2)$$

$A$  and  $s$  are constants and  $r_{Xj}$  is the distance between atoms  $X$  and  $j$ .

If  $P_k - P_{\min} > P_{\text{cutoff}}$  where  $P_{\min}$  is the smallest penalty for any of the conformations then the conformation is eliminated within ConfGen.

*arg5* Standard deviation of the Gaussian penalty,  $s$

0 use 2.5 Å

> 0 use this value (Å)

*arg6* Maximum value for the Gaussian,  $A$

0 use 1

> 0 use this value

*arg7* Relative penalty cutoff,  $P_{\text{cutoff}}$

0 use 1.00

> 0 use this value

### CHYD — Suppress Hydrogen Bond Electrostatics

The presence of this opcode suppresses electrostatic interactions between the charges arising from the bond dipole between the H and donor atom (O, N or S atom) and acceptor atom (N, O, S or F) within a ligand. The exception is that  $sp^2$  nitrogen atoms with 2 or 3 non-hydrogen attachments or nitrogen atoms that carry a positive formal charge cannot be considered acceptors.

Intramolecular hydrogen bonds in ligand-sized molecules can have a strong effect on the relative energies of conformations and yield compact conformations as the global minimum energy state. This effect is partially mitigated by using solvation treatments such as a distant dependent dielectric or GB/SA. However, intramolecular hydrogen bonds still can have a significant influence. In a protein environment, intraligand hydrogen-bonding competes with ligand-protein hydrogen bonding. In addition, ligand conformations in ligand protein complexes are usually extended rather than compact. Thus when modeling just the ligand it can be advantageous to suppress intra-ligand hydrogen bonding interactions to better imitate ligand conformations within ligand-protein complexes. Since the electrostatic contributions dominate hydrogen bond energetics and eliminating excluded volume effects can lead to problematic behavior, the CHYD opcode simply turns off the electrostatic portion of the hydrogen bonding interaction. CHYD does not affect the electrostatics of donor and acceptor atoms if they interact via a dihedral angle potential.

*arg1* Control hydrogen-bond electrostatics

-1 Use normal hydrogen-bond electrostatics.

0,1 Suppress hydrogen-bond electrostatics.

### 3.2.2 Command File Example

To run MacroModel calculations, a molecular structure file and a command file are required. The molecular structure file contains the structures to be used as input in the calculation. The command file contains the name of the input structure file, the name of the output structure file, and an ordered list of operation codes (*opcodes*) for the calculations.

Once you set up a job in Maestro and click either **Start** or **Write**, Maestro writes out a molecular structure file and a command file. For many types of jobs, command files written this way are complete and adequate, but for some types of jobs, you may need to adjust the Maestro-generated command file.

The command files and the log files for the examples given in this section can be found in \$SCHRODINGER/macromodel-vversion/samples/Examples.

Below is an example of a .com file for a ConfGen calculation. Descriptions of the opcodes in the file follow. More information on these opcodes can be found in [Section 3.2 on page 23](#).

```
mmod_confgen.mae
mmod_confgen-out.mae
MMOD      0      1      0      0      0.0000      0.0000      0.0000      0.0000
FFLD      10     1      0      1      1.0000      0.0000      0.0000      0.0000
SOLV      3      1      0      0      0.0000      0.0000      0.0000      0.0000
BDCC      0      0      0      0      41.5692 99999.0000      0.0000      0.0000
READ      0      0      0      0      0.0000      0.0000      0.0000      0.0000
CRMS      0      0      0      0      0.0000      0.5000      60.0000      0.0000
CGEN     1000      0      0      2      0.0000      0.0000      0.0000      0.0000
CGOP      0      2      50     1      48.0000      16.0000      8.0000      0.0000
CHYD      1      0      0      0      0.0000      0.0000      0.0000      0.0000
MCOP      1      0      0      0      0.0000      0.0000      0.0000      0.0000
DEMX      0      0      0      0      25.0000      0.0000      0.0000      0.0000
MSYM      0      0      0      0      0.0000      0.0000      0.0000      0.0000
NANT      0      0      0      0      0.0000      0.0000      0.0000      0.0000
AUTO      0      4      1      2      -1.0000      1.0000      0.0000      0.0000
CONV      2      0      0      0      0.0500      0.0000      0.0000      0.0000
MINI      9      0     100     0      0.0000      0.0000      0.0000      0.0000
```

**MMOD:** Creates and updates an intermediate structure file so that structures can be displayed in Maestro as the job progresses.



**FFLD:** Specifies the force field and aspects of the electrostatic treatment to use in MacroModel. Arg1 defines the force field used in the calculations. 10 means MMFF or MMFFs will be used. ConfGen calculations are compatible with the MMFF, MMFFs, OPLS\_2001 and OPLS\_2005 force fields. Arg2 defines the electrostatic treatment for the calculation. In this case a constant dielectric is used due to the use of solvation model 3 (see SOLV below). Arg4 selects either MMFF (0) and MMFFs (1). Arg5 is used to indicate that the dielectric constant should be set to 1 and is appropriate for solvation model 3.

**SOLV:** Controls which solvation model to use in MacroModel. Arg1 specifies the type of solvation treatment to use. 1 means that GB/SA is used. Arg2 selects the solvent and a value of 3 corresponds to water.

**BDCO:** Use the Bond Dipole CutOff (BDCO) method for truncating electrostatic interactions. Arg5 and arg6 are used to specify the cutoffs used for charge-dipole and charge-charge interactions, respectively.

**READ:** Read the input file.

**CRMS:** Sets parameters for redundant conformer elimination. Arg6 controls the maximum distance between atoms in equal conformers and is set to 0.5 Å. Arg7 controls the maximum difference in dihedral angles for polar hydrogens and is set to 60°.

**CGEN:** Use ConfGen to generate conformers. Arg1 defines the maximum number of conformers MacroModel should process during the search. Arg4 controls what combinations of rotations of terminal rotatable bonds are sampled. A value of 2 means that all combinations of minima for terminal rotatable bonds are sampled.

**CGOP:** Selects options for ConfGen. Arg1 controls sampling of certain types of compact symmetric terminal groups e.g.  $-\text{CF}_3$  or  $-\text{SO}_3^-$ . A value of 0 means “do not rotate these groups”. Arg2 controls how rings are sampled. Arg2=2 means use specific ring system templates when sampling ring flexibility. Arg3 controls minimization of conformers produced by ConfGen. If it is a negative value, the structures are not minimized. If it is a positive value, the structures are minimized using this number of steps. Arg3=0 means the number of steps listed in MINI Arg3 are used. Arg4 controls the sampling of non-ring amide bonds. Arg4=1 specifies that both cis and trans conformations should be generated. Arg5 controls the maximum relative energy between ring conformations inside ConfGen. Arg6 controls the maximum number of ring conformations to sample for the ligand as a whole. Arg7 controls the maximum number of ring conformations to sample for each ring system.

**CHYD:** Suppress hydrogen bond electrostatic interactions.

**MCOP:** Options to control what and how often data is written to the .log file during a conformation generation calculation. Arg1=1 updates the log file for every search step.

**DEMEX:** This command is used to prevent saving of high-energy conformers during the search. Arg5 defines the allowed energy window above the currently found global minimum. New conformers that are not within arg5 kJ/mol will be discarded. Additionally, a preliminary energy test can be performed during the energy minimization, to ensure that a reasonable structure has been found.

**MSYM:** Invokes the numbering symmetry library `mmsym`, which automatically and more generally identifies a suitable numbering order for use in comparing molecular conformations.

**NANT:** Regards enantiomers as distinct and retains both forms in the output structure file.

**AUTO:** Run MacroModel's automatic setup for conformational generation. Arg2 controls what atoms and types of comparisons amongst the atoms are used to identify redundant conformers. A value of 4 indicates that the positions of non-hydrogen atoms and dihedral angles for polar hydrogens should be used. Arg3 controls chiral atom checking during the calculation. Arg3=1 means identify and enforce chirality conservation for chiral atoms during the search. Arg4 controls the identification of torsional constraints during conformational searches. Arg4=2 means that checks are carried out to ensure that C=C bond geometries (i.e. cis or trans) match those in the input structure file. Arg5 controls the identification of torsions to sample within MacroModel and a value of -1 turns this off since this is handled within ConfGen itself. Arg6 is used to indicate whether this is a serial calculation (i.e. perform a separate conformational search on each structure in the input structure file) or not. Arg6=1 means that this is a serial calculation. All ConfGen jobs should be run as serial calculations.

**CONV:** Defines convergence criteria. Arg1=2 specifies derivative convergence (default criterion is 0.05 kJ/mol-Å, and this value is set in arg5).

**MINI:** Starts the minimization. Arg1 defines the type of minimization algorithm to be used. Arg1=9 means that Truncated Newton-Raphson Conjugate Gradient will be used. Arg3 defines the number of minimization steps used in the structures read in from the input structure file prior to passing them to ConfGen. Arg3 also defines the number of minimization steps used on structures generated by ConfGen if `CGOP arg3=0`.

---

# Getting Help

Information about Schrödinger software is available in two main places:

- The `docs` folder (directory) of your software installation, which contains HTML and PDF documentation. Index pages are available in this folder.
- The Schrödinger web site, <http://www.schrodinger.com/>, In particular, you can use the Knowledge Base, <http://www.schrodinger.com/kb>, to find current information on a range of topics, and the Known Issues page, <http://www.schrodinger.com/knownissues>, to find information on software issues.

## Finding Information in Maestro

Maestro provides access to nearly all the information available on Schrödinger software.

### To get information:

- Pause the pointer over a GUI feature (button, menu item, menu, ...). In the main window, information is displayed in the Auto-Help text box, which is located at the foot of the main window, or in a tooltip. In other panels, information is displayed in a tooltip.

If the tooltip does not appear within a second, check that Show tooltips is selected under General → Appearance in the Preferences panel, which you can open with CTRL+, (⌘,). Not all features have tooltips.

- Click the Help button in the lower right corner of a panel or press F1, for information about a panel or the tab that is displayed in a panel. The help topic is displayed in the Help panel. The button may have text or an icon:



- Choose Help → Online Help or press CTRL+H (⌘H) to open the default help topic.
- When help is displayed in the Help panel, use the navigation links in the help topic or search the help.
- Choose Help → Manuals Index, to open a page that has links to all the PDF documents. Click a link to open the document.

- Choose Help → Search Manuals to search the manuals. The search tab in Adobe Reader opens, and you can search across all the PDF documents. You must have Adobe Reader installed to use this feature.

### For information on:

- Problems and solutions: choose Help → Knowledge Base or Help → Known Issues → *product*.
- New software features: choose Help → New Features.
- Scripts available for download: choose Scripts → Update.
- Python scripting: choose Help → Python Module Overview.
- Utility programs: choose Help → About Utilities.
- Keyboard shortcuts: choose Help → Keyboard Shortcuts.
- Installation and licensing: see the *Installation Guide*.
- Running and managing jobs: see the *Job Control Guide*.
- Using Maestro: see the *Maestro User Manual*.
- Maestro commands: see the *Maestro Command Reference Manual*.

## Contacting Technical Support

If you have questions that are not answered from any of the above sources, contact Schrödinger using the information below.

Web: <http://www.schrodinger.com/supportcenter>  
E-mail: [help@schrodinger.com](mailto:help@schrodinger.com)  
Mail: Schrödinger, 101 SW Main Street, Suite 1300, Portland, OR 97204  
Phone: +1 503 299-1150 (USA, 9am – 5pm Pacific Time)  
+49 621 438-55173 (Europe, 9am – 5pm Central European Time)  
Fax: +1 503 299-4532 (USA, Portland office)  
FTP: <ftp://ftp.schrodinger.com>

Generally, using the web form is best because you can add machine output and upload files, if necessary. You will need to include the following information:

- All relevant user input and machine output
- ConfGen purchaser (company, research institution, or individual)
- Primary ConfGen user
- Installation, licensing, and machine information as described below.

## Gathering Information for Technical Support

The instructions below describe how to gather the required machine, licensing, and installation information, and any other job-related or failure-related information, to send to technical support. Where the instructions depend on the profile used for Maestro, the profile is indicated.

### For general enquiries or problems:

1. Open the Diagnostics panel.
  - **Maestro:** Help → Diagnostics
  - **Windows:** Start → All Programs → Schrodinger-2014-3 → Diagnostics
  - **Mac:** Applications → Schrodinger2014-3 → Diagnostics
  - **Command line:** `$SCHRODINGER/diagnostics`
2. When the diagnostics have run, click Technical Support.

A dialog box opens, with instructions. You can highlight and copy the name of the file.

3. Attach the file specified in the dialog box to the web form.

If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can attach the file when you fill in the form.

### If your job failed:

1. Open the Monitor panel, using the instructions for your profile as given below:
  - **Maestro/Jaguar/Elements:** Tasks → Monitor Jobs
  - **BioLuminate/MaterialsScience:** Tasks → Job Monitor
2. Select the failed job in the table, and click Postmortem.

The Postmortem panel opens.

3. If your data is not sensitive and you can send it, select Include structures and deselect Automatically obfuscate path names.
4. Click Create.

An archive file is created in your working directory, and an information dialog box with the name of the file opens. You can highlight and copy the name of the file.

5. Attach the file specified in the dialog box to the web form.

If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can attach the file when you fill in the form.

6. Copy and paste any log messages from the window used to start the interface or the job into the web form (or an e-mail message), or attach them as a file.
  - **Windows:** Right-click in the window and choose **Select All**, then press **ENTER** to copy the text.
  - **Mac:** Start the **Console** application (**Applications** → **Utilities**), filter on the application that you used to start the job (**Maestro**, **BioLuminate**, **Elements**), copy the text.

### If Maestro failed:

1. Open the **Diagnostics** panel.
  - **Windows:** **Start** → **All Programs** → **Schrodinger-2014-3** → **Diagnostics**
  - **Mac:** **Applications** → **SchrodingerSuite2014-3** → **Diagnostics**
  - **Linux/command line:** `$SCHRODINGER/diagnostics`

2. When the diagnostics have run, click **Technical Support**.

A dialog box opens, with instructions. You can highlight and copy the name of the file.

3. Attach the file specified in the dialog box to the web form.

If you have already submitted a support request, use the upload link in the e-mail response from Schrödinger to upload the file. If you need to submit a new request, you can attach the file when you fill in the form.

4. Attach the error files to the web form.

The files should be in the following location:

- **Windows:** `%LOCALAPPDATA%\Schrodinger\appcrash`  
(Choose **Start** → **Run** and paste this location into the **Open** text box.)  
Attach `maestro_error_pid.txt` and `maestro.exe_pid_timestamp.dmp`.
- **Mac:** `$HOME/Library/Logs/CrashReporter`  
(Go → **Home** → **Library** → **Logs** → **CrashReporter**)  
Attach `maestro_error_pid.txt` and `maestro_timestamp_machinename.crash`.
- **Linux:** `$HOME/.schrodinger/appcrash`  
Attach `maestro_error_pid.txt` and `crash_report_timestamp_pid.txt`.



120 West 45th Street  
17th Floor  
New York, NY 10036

155 Gibbs St  
Suite 430  
Rockville, MD 20850-0353

Quatro House  
Frimley Road  
Camberley GU16 7ER  
United Kingdom

101 SW Main Street  
Suite 1300  
Portland, OR 97204

Dynamostraße 13  
D-68165 Mannheim  
Germany

8F Pacific Century Place  
1-11-1 Marunouchi  
Chiyoda-ku, Tokyo 100-6208  
Japan

245 First Street  
Riverview II, 18th Floor  
Cambridge, MA 02142

Zeppelinstraße 73  
D-81669 München  
Germany

No. 102, 4th Block  
3rd Main Road, 3rd Stage  
Sharada Colony  
Basaveshwaranagar  
Bangalore 560079, India

8910 University Center Lane  
Suite 270  
San Diego, CA 92122

Potsdamer Platz 11  
D-10785 Berlin  
Germany

**SCHRÖDINGER**