

# **Silico**

Computational Chemistry Management

*Quick Reference Guide*

Version 0.13.6

31 July 2020

## Table of Contents

<b>How to Use This Document.....</b>	<b>1</b>
<b>About Silico.....</b>	<b>2</b>
<b>Installation.....</b>	<b>3</b>
<b>Submitting Calculations.....</b>	<b>4</b>
Prepare Files.....	4
Upload Files.....	4
Submit Files.....	5
Calculation Browser.....	5
Select a Method.....	6
Select a Program.....	7
Select a Calculation.....	8
Confirm Selection.....	9
Using Calculation Codes.....	10
Common Codes.....	11
Advanced: Modifying Calculations.....	12
<b>Folder Structure.....</b>	<b>15</b>
Input.....	15
Output.....	15
Flags.....	15
Results.....	15
Report.....	15
<b>Monitoring Calculations.....</b>	<b>16</b>
The Flags Folder.....	16
List of Flags.....	17
Multiple Calculations.....	18
<b>Post Analysis.....</b>	<b>19</b>
Results.....	19
Available Result Files.....	20
Report.....	21
<b>Analysis of Results.....</b>	<b>22</b>
Generating Summaries.....	22
Analysis of Multiple Results.....	22
<b>Generating Reports.....</b>	<b>24</b>
Normal Usage.....	24
Including Emission Energy.....	25

# How to Use This Document

This document is designed to act as rapid introduction and quick-reference to the Silico Computational Chemistry Management system.

Main instructions are written as this line.

Images are included for comparison and reference.

Supplementary information and helpful tips are displayed separately, as follows:

**NOTE:** This is a tip.

Commands that should be typed by the user are displayed in the following format:

```
> echo Hello world
```

The arrow character (>) should not be typed; it is used to differentiate commands that should be typed from the resulting output (which will be displayed without the arrow character):

```
> echo Hello world  
Hello world
```

In the above example, the user is being instructed to type: echo Hello world. The computer, in response, gives the output: Hello world.

Ellipses (...) indicate that the real, full output has been truncated:

```
> cat /etc/fstab  
# /etc/fstab: static file system information.  
#  
...
```

## About Silico

Silico is a set of programs designed to automate and simplify all aspects of computational chemistry, without which can be: confusing, frustrating, time-consuming, bizarre and/or upsetting. Silico is under development and suggestions are welcome, please contact [osl@st-andrews.ac.uk](mailto:osl@st-andrews.ac.uk).

A non-exhaustive list of features is as follows:

- Submission to computational programs through a simple and unified interface.
- Simultaneous submission of multiple molecules/systems.
- Automatic in series submission of results from completed calculations to subsequent calculations.
- Automatic conversion of input files (including ChemDraw) to types appropriate for the selected computational program.
- Automatic and manual analysis of computation results, including tabulation to csv format.
- Automatic and manual generation of PDF reports from computation results, including rendered 3D structures, orbital images and graphs.

# Installation

To install Silico, log in to the Kennedy server as normal and use the following command:

```
> /gpfs1/apps/EZC-tools/install-silico
```

**NOTE:** Installation only ever needs to be performed once (for each user). Updates to Silico will be become available automatically without the need to reinstall.

Installation should be near instantaneous. Once complete, a message will be printed:

```
> /gpfs1/apps/EZC-tools/install-silico  
Installed successfully. Please log out and in again to complete
```

Use the 'exit' command to log out:

```
> /gpfs1/apps/EZC-tools/install-silico  
Installed successfully. Please log out and in again to complete  
> exit  
logout
```

Once logged back in again, Silico will now be available for use. The installation can be tested by using the 'silico' command with the '-v' (version) option. If successful, the version will be printed:

```
> silico -v  
0.13.6
```

If the installation was not successfully, the following error message will be printed:

```
> silico -v  
-bash: silico: command not found
```

If unsuccessful, seek help from another group member.

# Submitting Calculations

## Prepare Files

To submit a calculation, the system(s) of interest first need to be prepared by drawing the relevant structures electronically (on a program). Silico supports a wide variety of input file types, including both 2D and 3D formats. Notable entries include:

- G09 output (.log)
- GaussView (.com, .gjf, .gjc and .gau)
- ChemDraw (.cdx and .cdxml)
- Avogadro (.cml)
- MarvinSketch (.mrv)

**NOTE:** MarvinSketch (.cml) and Avogadro (.mrv) are free alternatives to ChemDraw and GaussView, respectively. You have probably already used MarvinSketch; it is the program used to draw structures on Reaxys.

Please see <https://open-babel.readthedocs.io/en/latest/FileFormats/Overview.html> for a full list of supported formats.

**NOTE:** When using ChemDraw files (.cdx), remember to save each molecule individually to its own file. If a single file contains multiple structures, then all these structures will be submitted together to a single calculation.

## Upload Files

Once the structures have been drawn, upload the input files to the server as normal.

## Silico – Computational Chemistry Management – Quick Reference Guide

### Submit Files

To submit files, use the 'silico' command, followed by the list of filenames. For example, to submit a file named 'Benzene.com':

```
> silico Benzene.com
```

If the file name contains spaces, remember to use speech marks:

```
> silico "Methyl benzene.com"
```

To submit multiple files at once, separate each with a space:

```
> silico Benzene.com "Methyl benzene.com"
```

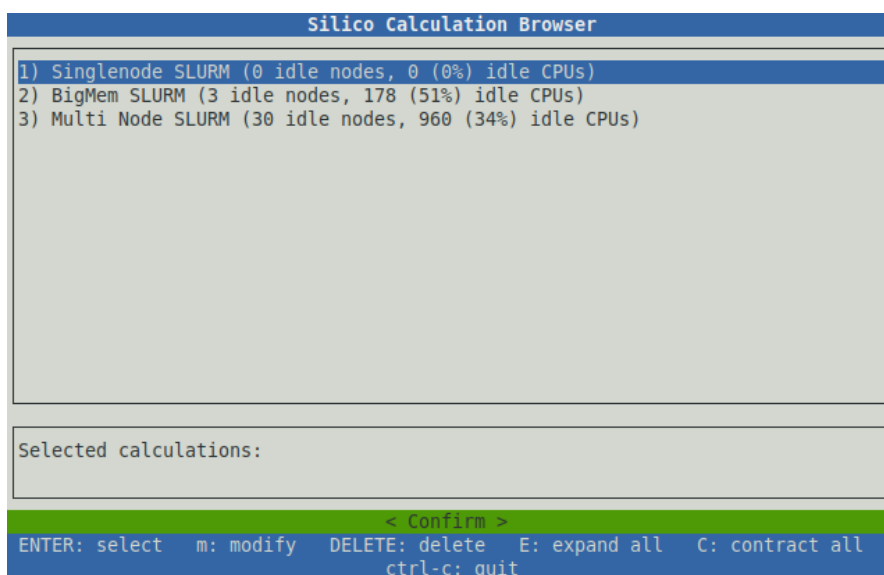
Finally, groups of files can be submitted by using the wildcard character (\*). For example, to submit all files ending in '.com':

```
> silico *.com
```

**NOTE:** All files submitted with the same silico command will be submitted to the same calculation. To submit to different calculations, use the silico command multiple times.

### Calculation Browser

The Silico calculation browser will then open, from which a calculation to submit to can be chosen.



## Silico – Computational Chemistry Management – Quick Reference Guide

This browser is interactive; you can navigate up and down using the arrow keys. 'Enter' or 'space' can be used to expand or contract each item. 'Esc' or ctrl-c (the 'ctrl' key and the 'c' key pressed together) can be used to quit at any time.

### Select a Method

The first items to choose from are the calculation methods; these represent different groups (called partitions) of computing nodes (individual computers). Three are available at time of writing, they are:

- **Singlenode SLURM:** The common DFT partition, this should be chosen if in doubt.
- **BigMem SLURM:** A partition containing nodes with additional memory capacity.
- **Multi Node SLURM:** A partition intended for calculations that require multiple nodes. Because G09 does not readily support parallelisation across multiple nodes, this partition should be avoided for DFT.

**NOTE:** Although BigMem SLURM does have more memory available than Singlenode SLURM, the latter already offers > 100 GB per node. This is likely to be more than sufficient for most cases.

Next to the name of each method, the status of the relevant partition is printed. This gives an indication of how busy each is at the current moment, which may help to choose:

```
Silico Calculation Browser
1) Singlenode SLURM (0 idle nodes, 0 (0%) idle CPUs)
2) BigMem SLURM (3 idle nodes, 178 (51%) idle CPUs)
3) Multi Node SLURM (30 idle nodes, 960 (34%) idle CPUs)

Selected calculations:

< Confirm >
ENTER: select  m: modify  DELETE: delete  E: expand all  C: contract all
                        ctrl-c: quit
```



## Silico – Computational Chemistry Management – Quick Reference Guide

To select a method, navigate up and/or down using the arrow keys and use 'enter' to make your selection. For example, to select the second method, BigMem SLURM:

```
Silico Calculation Browser

1) Singlenode SLURM (0 idle nodes, 0 (0%) idle CPUs)
2) BigMem SLURM (3 idle nodes, 178 (51%) idle CPUs)
  L[+]>1) Gaussian 09
3) Multi Node SLURM (30 idle nodes, 960 (34%) idle CPUs)

Selected calculations:

< Confirm >
ENTER: select  m: modify  DELETE: delete  E: expand all  C: contract all
                        ctrl-c: quit
```

### Select a Program

The next item to choose from the browser is the calculation program. However, because the Kennedy server only has one supported program, G09, there is only one item to 'choose' from. Other programs (ORCA) will become available once they are supported.

To choose Gaussian 09, navigate down with the arrow keys and use 'enter' to make your selection:

```
Silico Calculation Browser

1) Singlenode SLURM (0 idle nodes, 0 (0%) idle CPUs)
2) BigMem SLURM (3 idle nodes, 178 (51%) idle CPUs)
  L[-]>1) Gaussian 09
      L[+]>Auto Characterisation
      L[+]>Single Point Singlet
      L[+]>Opt Freq
      L[+]>Opt Freq Unrestricted Triplet
      L[+]>TDA
      L[+]>TD-DFT
3) Multi Node SLURM (30 idle nodes, 960 (34%) idle CPUs)

Selected calculations:

< Confirm >
ENTER: select  m: modify  DELETE: delete  E: expand all  C: contract all
                        ctrl-c: quit
```

## Silico – Computational Chemistry Management – Quick Reference Guide

### Select a Calculation

The final item to choose is the calculation itself. The available calculations are sorted into categories with self descriptive names, each category can be expanded in the same manner as before; with the arrow keys and enter key. The categories are as follows:

- Calculation type (Opt, Freq, SP, TD-DFT etc)
- Solvent (or gas phase)
- Functional
- Basis set

**NOTE:** There are 3 functionals available. PBE1PBE and B3LYP are popular and cheap (fast). M062X is more accurate but also more expensive (slow).

Once the desired calculation has been chosen, pressing enter will add it to the calculation list, represented by the box below the browser:

```
Silico Calculation Browser
1) Singlenode SLURM (0 idle nodes, 0 (0%) idle CPUs)
2) BigMem SLURM (3 idle nodes, 178 (51%) idle CPUs)
  [-] 1) Gaussian 09
    [+>] Auto Characterisation
    [->] Single Point Singlet
    [->] Opt Freq
      [->] Gas Phase
        [->] PBE1PBE
          239) 6-31G(d,p)
          240) 6-31G+(d,p)
          241) 6-31G++(d,p)
          242) 6-311G(d,p)
          243) 6-311G+(d,p)
          244) 6-311G++(d,p)

Selected calculations: 2/1/239

< Confirm >
ENTER: select  m: modify  DELETE: delete  E: expand all  C: contract all
ctrl-c: quit
```

The calculation chosen is represented by 3 numbers corresponding to the selected method, program and calculation, which together make up a unique calculation code. In the above example, the choices were:

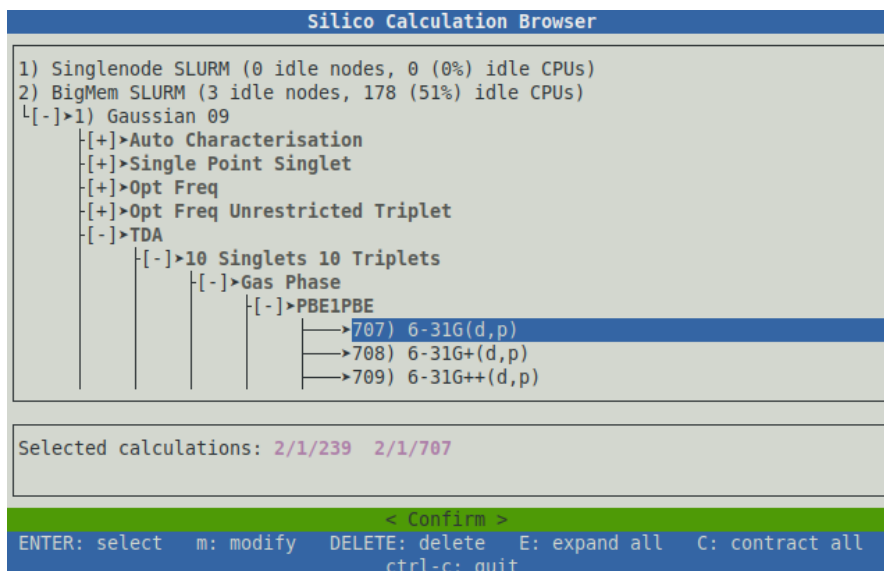
- Method 2 (BigMem SLURM)
- Program 1 (Gaussian 09)
- Calculation 239 (Opt Freq PBE1PBE/6-31G(d,p) in the gas phase)

Resulting in the calculation code '2/1/239'.

**NOTE:** Pressing the 'Delete' key will delete the last calculation from the calculations list.

## Silico – Computational Chemistry Management – Quick Reference Guide

Multiple calculations can be chosen in the same way and will each be added to the calculations list:



Multiple calculations selected in this way will be run *in series*. That is, the first calculation (2/1/239) will be performed by itself. Once complete, the output of this calculation will be automatically submitted to the next in the list (2/1/707), and so on until all calculations requested have been performed.

**NOTE:** The calculations list is freely editable. Press the 'tab' key to first switch to the edit box before typing. Use 'shift-tab' to return to the browser.

### Confirm Selection

Once the desired calculations have been queued, press the 'tab' key to select the 'Confirm' button and then press 'enter' to confirm. If no error messages are printed, then the files have been submitted successfully.

If one of the files selected was a 2D format (ChemDraw for example), a warning similar to the following will be printed:

```
> silico Benzene.cdx
silico: WARNING: Generating 3D coordinates from file 'Benzene.cdx';
this will scramble atom coordinates
```

This warning indicates that the original 2D coordinates have been modified to produce 3D coordinates. In nearly all cases this is actually desirable and so this warning can be safely ignored.

## Silico – Computational Chemistry Management – Quick Reference Guide

### Using Calculation Codes

The codes that identify each calculation (such as 2/1/239 and 2/1/707 used in the above example) are constant, unique and will always refer to the same calculation. This property allows computations to be submitted more quickly than by using the browser, if the calculation codes of interest are already known.

**NOTE:** Calculation codes are only constant for the same Silico version. Updates may add, remove or rearrange the available calculations, resulting in changes to the relevant codes. You will be notified if this occurs.

To submit a calculation without using the browser, use the '-c' (calculation) option to the silico command, followed by the relevant codes:

```
> silico Benzene.com -c 2/1/239
```

Calculations can be queued by separating each by a space, so to recreate the example above, the full command would be:

```
> silico Benzene.com -c 2/1/239 2/1/707
```

## Silico – Computational Chemistry Management – Quick Reference Guide

### Common Codes

For reference, some of the more common calculation codes are printed here:

Code	Description	Comment
1/1/1	Performs the following at PBE1PBE/6-31G(d,p): <ul style="list-style-type: none"><li>• Ground optimisation (with frequency)</li><li>• TDA excited states (50/50)</li></ul>	Used for characterising organic fluorescent and TADF emitters
1/1/4	Performs the following at B3LYP/6-31G(d,p), using SBKJC-VDZ-ECP for metals, in MeCN: <ul style="list-style-type: none"><li>• Ground optimisation (with frequency)</li><li>• TD-DFT excited states (50/50)</li><li>• Optimised triplet (with frequency)</li><li>• Single point singlet at the same geometry</li></ul>	Used for characterising organometallic phosphorescent emitters
1/1/239	Optimisation (with frequency) at PBE1PBE/6-31G(d,p)	EZC recommended for organic emitters
1/1/252	Optimisation (with frequency) at B3LYP/6-31G(d,p)	Popular method & basis set in the literature
1/1/302	Optimisation (with frequency) at B3LYP/6-31G(d,p), using SBKJC-VDZ-ECP for metals, in MeCN	EZC recommended for organometallic emitters
1/1/707	TDA excited states at PBE1PBE/6-31G(d,p)	EZC recommended for organic emitters
1/1/1940	TD-DFT excited states at B3LYP/6-31G(d,p), using SBKJC-VDZ-ECP for metals, in MeCN	EZC recommended for organometallic emitters
1/1/1175	TDA excited states (100 singlets) at PBE1PBE/6-31G(d,p)	Useful for simulating absorption spectra of organic emitters
1/1/2408	TD-DFT excited states (100 singlets) at B3LYP/6-31G(d,p), using SBKJC-VDZ-ECP for metals, in MeCN	Useful for simulating absorption spectra of organometallic emitters

## Silico – Computational Chemistry Management – Quick Reference Guide

### Advanced: Modifying Calculations

The calculations available in Silico can be modified prior to submission. This allows altering of such properties as: memory, CPUs, basis set, functional and more. To modify a calculation, follow the submission process until the calculation browser is displayed:

```
Silico Calculation Browser

1) Singlenode SLURM (0 idle nodes, 0 (0%) idle CPUs)
2) BigMem SLURM (3 idle nodes, 178 (51%) idle CPUs)
  [-]>1) Gaussian 09
    [-]>1) Auto Characterisation
    [-]>1) Single Point Singlet
    [-]>1) Opt Freq
    [-]>1) Opt Freq Unrestricted Triplet
    [-]>1) TDA
      [-]>10 Singlets 10 Triplets
        [-]>Gas Phase
          [-]>PBE1PBE
            707) 6-31G(d,p)
            708) 6-31G+(d,p)
            709) 6-31G++(d,p)

Selected calculations: 2/1/239 2/1/707

< Confirm >
ENTER: select  m: modify  DELETE: delete  E: expand all  C: contract all
ctrl-c: quit
```

Once a calculation has been selected with the arrow keys, press 'm' (**m**odify) or 'e' (**e**dit) to open the edit window:

```
Editing calculation 'Gaussian-DFT' with ID: 707

CPU list:
1) 
A list of integers specifying specific CPUs to use for the calculation,
starting at 0. CPU_list and num_CPUs are mutually exclusive

GROUP:
1) TDA
2) 10 Singlets 10 Triplets
3) Gas Phase
4) PBE1PBE
5)
An ordered list of group names, used for categorisation

GROUP_NAME: 6-31G(d,p)
An alternative to NAME; GROUP_NAME is used to set a shorter name that appears
as part of a GROUP

NAME:
The unique name of this Configurable and the name of the folder under which
this calculation will run. If left blank and a GROUP_NAME is set, a name will

< Cancel > < Confirm >
```

Here, numerous properties of the calculation can be changed. The up and down arrow keys are used to navigate between the different fields. Each comes with a help message, displayed below the field, explaining the property and how it impacts the calculation. There are a great many more fields than can be displayed at one time, navigating to the bottom of the screen will scroll down so more fields can be viewed:

## Silico – Computational Chemistry Management – Quick Reference Guide

Editing calculation 'Gaussian-DFT' with ID: 5

**NAME:**  
The unique name of this Configurable and the name of the folder under which this calculation will run. If left blank and a GROUP\_NAME is set, a name will be generated automatically

**charge: auto**  
Forcibly set the system charge. Use 'auto' to use the charge given in the input file

**external\_ECPs:**  
1)   
A list of external ECPs (effective core potentials) to use. Possible values are: 'LANL2DZ for >= K', 'LANL2DZ', 'SBKJC-VDZ for >= K', 'SBKJC\_VDZ'

**external\_basis\_sets:**  
1)   
A list of external basis sets to use. The order given here is the order the basis sets will be appended to the input file. Possible values are: '6-31+G(d,p) for H to Ar', '6\_31\_plus\_starstar\_H\_Ar', 'LANL2DZ for >= K', 'LANL2DZ', 'SBKJC-VDZ for >= K', 'SBKJC\_VDZ'

< Cancel > < Confirm >

Most fields are freely editable (can take any value). Some can only accept values from a list; in which case the list of possible values is given in the help. Some can only accept a 'True' or 'False' value, these fields appear as follows:

Editing calculation 'Gaussian-DFT' with ID: 5

**programs:**  
1) g09  
2) g16  
3)   
A list of programs that this calculation is compatible with

**scratch\_options**  
Options that control the use of the scratch directory

☒ use\_scratch  
Whether to use a scratch directory. False will disable the scratch directory, and is not recommended

**path: /scratch**  
Path to the top of the scratch directory. For each calculation, a sub-directory will be created under this path

☒ use\_username  
Whether to create a subdirectory for each user

☐ keep

< Cancel > < Confirm >

In such fields, the large 'X' indicates the option is set to 'True' or 'Yes', its absence indicates the opposite. Space or enter can be used to toggle the value.

**NOTE:** The 'Cancel' button will discard any changes and return to the calculation browser. The 'Esc' key also has the same effect.

Once the necessary changes have been made, press 'tab' to select the 'Confirm' button, then press 'enter' to confirm. The screen will then return to the calculation browser. You can now add the modified calculation to the calculation list as normal. It is important to note that any changes

## Silico – Computational Chemistry Management – Quick Reference Guide

made will effect all calculations with the same calculation code, regardless of whether the calculation was added to the list before or after changes were made.

**NOTE:** All modifications made in this way are temporary. The next time the calculation browser is opened, settings will revert to their default. If an edit is made in error, and the original value cannot be remembered, keep pressing 'Esc' until the browser is closed, then begin the submission process again. Future versions of Silico will offer to permanently save modifications.



### Folder Structure

Silico will create a series of folders for each file and each calculation submitted. The topmost folder is named after the molecule submitted. Within this molecule directory, a separate folder is then created for each calculation, using the name of that calculation. If the same input file is submitted to multiple of the same calculation, a number will be appended to the calculation folder name to ensure each calculation is performed in a unique directory.

**NOTE:** The name of the 'molecule' directory is simply taken from the name of the input file; Silico does not try to automatically name structures.

Within each calculation directory 5 sub folders are created. They are as follows.

#### Input

Contains the input file for the calculation. The file within will be fully prepared for Gaussian, meaning it can be inspected to determine the specific parameters of the calculation. If an external basis set was requested (SBKJC-VDZ-ECP, for example) then this will be appended to the input file. If a non-native format was submitted (ChemDraw, for example), then this will have been converted to a Gaussian input file.

**NOTE:** The file originally submitted is not stored in the Input directory; only the prepared input file is saved.

#### Output

Contains output files written by the calculation program, which for Gaussian is the .log and .fchk files.

**NOTE:** By default, Silico will automatically convert .chk files to .fchk (before deleting the original .chk) in order to save file space. Occasionally, the original .chk file is required for post-analysis, in which case this option can be disabled (see **Submitting Calculations/Advanced: Modifying Calculations**).

#### Flags

Contains file flags, text files that convey information about the status of the calculation. See the section on **Monitoring Calculations (The Flags Folder)** for more information about file flags.

#### Results

Contains text result files that are automatically written by Silico during post analysis.

#### Report

Contains a PDF report summary of the completed calculation, along with the rendered images used in the report. In addition, a smaller, mini report is also generated containing only atom coordinates, intended for easy inclusion into ESI.

# Monitoring Calculations

## The Flags Folder

Within each calculation directory, Silico manages a special folder with the name 'Flags'. This folder contains a number of empty text files where the name of each file conveys status about the calculation. These files are created and destroyed at key points in the calculation submission process, so the calculation can be monitored by observing which files are present (or absent) from the 'Flags' folder at any given moment.

To check the file flags, use the 'ls' command, followed by the path to the relevant Flags folder. For example:

```
> ls "Benzene/Opt Freq Gas Phase PBE1PBE 6-31G(d,p)/Flags"
RUNNING  STARTED
```

Here, two flags are present: 'RUNNING' and 'STARTED', which indicate that the calculation is currently ongoing.

Typing the full path to the flags folder is time consuming, so it is recommended to make use of autocompletion by using the 'tab' key. Begin typing the path, for example:

```
> ls Benzene/Opt
```

Then press the 'tab' key. Linux will automatically guess the rest of the filename (based on the available files that match what has already been typed):

```
> ls Benzene/Opt
> ls Benzene/Opt\ Freq\ Gas\ Phase\ PBE1PBE\ 6-31G\ (d\,p\)/
```

Then type 'Flags' to complete the path:

```
> ls Benzene/Opt
> ls Benzene/Opt\ Freq\ Gas\ Phase\ PBE1PBE\ 6-31G\ (d\,p\)/
> ls Benzene/Opt\ Freq\ Gas\ Phase\ PBE1PBE\ 6-31G\ (d\,p\)/Flags
```

**NOTE:** Autocompletion is a feature of Linux, not Silico, meaning you can use it with any Linux command.

If the autocompletion function does not guess the filename, press 'tab' a second time to see a list of options:

## Silico – Computational Chemistry Management – Quick Reference Guide

```
> ls Benzene/Opt\ Freq\ Gas\ Phase\  
Opt Freq Gas Phase B3LYP 6-31G(d,p)/      Opt Freq Gas Phase PBE1PBE 6-  
31G(d,p)/  
> ls Benzene/Opt\ Freq\ Gas\ Phase\
```

Choose from the options listed by typing:

```
> ls Benzene/Opt\ Freq\ Gas\ Phase\  
Opt Freq Gas Phase B3LYP 6-31G(d,p)/      Opt Freq Gas Phase PBE1PBE 6-  
31G(d,p)/  
> ls Benzene/Opt\ Freq\ Gas\ Phase\ B
```

Then use 'tab' again to autocomplete the rest of the path:

```
> ls Benzene/Opt\ Freq\ Gas\ Phase\  
Opt Freq Gas Phase B3LYP 6-31G(d,p)/      Opt Freq Gas Phase PBE1PBE 6-  
31G(d,p)/  
> ls Benzene/Opt\ Freq\ Gas\ Phase\ B3LYP\ 6-31G\ (d\,p\)/
```

### List of Flags

The currently available file flags are as follows:

Name	Description
PENDING	The calculation has been submitted but has not yet begun; most likely because it is waiting in the queue.
STARTED	The calculation has begun. This flag is never deleted, so it is useful for confirming that the calculation at least started, even if it did not finish.
RUNNING	The calculation is currently ongoing.
SUCCESS	The calculation finished successfully.
CONVERGED	The optimisation converged successfully. This flag is only used for optimisation calculations.
NOT_CONVERGED	The optimisation did not converge successfully. This flag is only used for optimisation calculations.
CLEANUP	The calculation has finished (successfully or otherwise) and Silico is currently cleaning up (saving files etc).
ERROR	The calculation has stopped because an error occurred.
POST	The calculation has finished and Silico is currently performing post analysis (writing result and report files).
DONE	All work (including post-analysis) has been completed; Silico will not make any changes after this flag. It is safe to move, download or delete the calculation folder.

## Silico – Computational Chemistry Management – Quick Reference Guide

### Multiple Calculations

Multiple calculations can be checked simultaneously by specifying multiple directories to the 'ls' command:

```
> ls "Benzene/Opt Freq Gas Phase PBE1PBE 6-31G(d,p)/Flags" "Benzene/TDA
10 Singlets 10 Triplets Gas Phase PBE1PBE 6-31G(d,p)/Flags"
Benzene/Opt Freq Gas Phase PBE1PBE 6-31G(d,p)/Flags:
CONVERGED  DONE  STARTED  SUCCESS

Benzene/TDA 10 Singlets 10 Triplets Gas Phase PBE1PBE 6-31G(d,p)/Flags:
DONE  STARTED  SUCCESS
```

This is more conveniently achieved using the Linux wildcard feature, using the asterisk character (\*). For example, to check all calculations in the Benzene directory:

```
> ls Benzene/*/Flags
Benzene/Opt Freq Gas Phase B3LYP 6-31G(d,p)/Flags:
CONVERGED  DONE  STARTED  SUCCESS

Benzene/Opt Freq Gas Phase PBE1PBE 6-31G(d,p)/Flags:
CONVERGED  DONE  STARTED  SUCCESS

Benzene/Single Point Singlet Gas Phase PBE1PBE 6-31G(d,p)/Flags:
DONE  STARTED  SUCCESS

Benzene/TDA 10 Singlets 10 Triplets Gas Phase PBE1PBE 6-31G(d,p)/Flags:
DONE  STARTED  SUCCESS
```

Similarly, to check the status of all calculations of all molecules in the current directory, two wildcards can be used:

```
> ls */*/Flags
246tCzPPC/TDA Optimised S(1) B3LYP 6-31G(d,p)/Flags:
PENDING

26tCzPPC/TDA Optimised S(1) B3LYP 6-31G(d,p)/Flags:
RUNNING  STARTED

2CzIPN/TDA Optimised S(1) B3LYP 6-31G(d,p)/Flags:
CONVERGED  DONE  STARTED  SUCCESS

2CzPN/TDA Optimised S(1) B3LYP 6-31G(d,p)/Flags:
CONVERGED  DONE  STARTED  SUCCESS

2CzTPN/TDA Optimised S(1) B3LYP 6-31G(d,p)/Flags:
CONVERGED  DONE  STARTED  SUCCESS
...
```

### Post Analysis

#### Results

Once a calculation is complete (the DONE flag has been set), result files will be available for analysis in the 'Results' folder. Perhaps the most important of these is the 'summary' file, which will be named something like 'Benzene.summary'. This file can be viewed on the server using the 'less' or 'cat' commands:

```
> less "Benzene/Opt Freq Gas Phase B3LYP 6-31G(d,p)/Results/Benzene.summary"
```

**NOTE:** When using the 'less' command, use the 'q' key to quit.

The summary file, as its name would suggest, contains a summary of the most important calculation results. An example file is display below:

```
> cat "Benzene/Opt Freq Gas Phase B3LYP 6-31G(d,p)/Results/Benzene.summary"
-----
|      Metadata      |
-----
Name: Benzene/Opt Freq Gas Phase B3LYP 6-31G(d,p)/Output/Benzene.log
Date: 30/07/2020 at 10:27:17
Duration: 0 days, 0 hours, 1 minutes
Computational package: Gaussian (2009+D.01)
Calculations: Optimisation, Frequencies
Methods: DFT
Functional: B3LYP
Basis set: 6-31G(d,p)
Multiplicity: Singlet
Charge: 0
Orbital spin: restricted
Success: True
Converged: True
Calculation temperature /K: 298.15
Calculation pressure /atm: 1.00

-----
|  Vibrational Frequencies  |
-----
No. vibrations: 30
No. negative frequency: 0

-----
|      Geometry      |
-----
Formula: C6H6
Exact mass /gmol-1: 78.0469
Molar mass /gmol-1: 78.1118
No. atoms: 12
Alignment method: Minimal
```

## Silico – Computational Chemistry Management – Quick Reference Guide

```
X extension /Å: 5.00
Y extension /Å: 3.58
Z extension /Å: 2.85
Linearity ratio: 0.28
Planarity ratio: 0.20

-----
|      SCF Energy      |
-----
No. of steps: 31
Energy /eV: -6315.9237
Energy /kJmol-1: -609393.9974

-----
|      Orbitals      |
-----
HOMO eV: -5.73
LUMO eV: -1.88
HOMO/LUMO energy /eV: 3.85
No. virtual: 99
No. occupied: 21

-----
|      Permanent Dipole Moment      |
-----
Total /D: 1.23
Origin X /D: 0.00
Origin Y /D: 0.00
Origin Z /D: 0.00
Vector X /D: -0.25
Vector Y /D: 1.12
Vector Z /D: -0.43
X axis angle /°: 78.40
XY plane angle /°: 20.61
```

The remaining files in the 'Results' directory contain specific calculation results in CSV format. These too can be read in the same way as the summary file, but are more easily manipulated once downloaded to a personal computer.

**NOTE:** CSV format can be easily imported into a spreadsheet (such as Microsoft Excel, Libre Office or Google Sheets) and graphed, if desired.

### Available Result Files

The currently available result files are listed below (*FILE* is used as a stand-in for the actual filename):

Name	Description
<i>FILE</i> .summary	Human readable summary of all calculation results
<i>FILE</i> .atoms.csv	Optimised (if applicable) atom coordinates
<i>FILE</i> .orbitals.csv	Orbital numbers, labels and energies (restricted calculations only)
<i>FILE</i> .alpha.csv	Alpha orbital numbers, labels and energies (unrestricted calculations only)

## Silico – Computational Chemistry Management – Quick Reference Guide

<i>FILE.beta.csv</i>	Beta orbital numbers, labels and energies (unrestricted calculations only)
<i>FILE.SCF.csv</i>	SCF (self consistent field) energy. For optimisations, includes the energy at each step which can be used to graph the calculation's convergence. The SCF energy is the HF (Hartree–Fock) or DFT energy.
<i>FILE.MP.csv</i>	Same as <i>FILE.SCF.csv</i> , but for Møller–Plesset energies. The energies are total (including HF and MP correction) and are from the highest MP level of the calculation.
<i>FILE.CC.csv</i>	Same as <i>FILE.SCF.csv</i> , but for coupled-cluster energies. The energies are total (including HF and CC correction, if applicable).
<i>FILE.ES.csv</i>	Excited state results (from TD-DFT, TDA etc). Includes energy, symmetry, multiplicity and orbital contributions.
<i>FILE.transitions.csv</i>	Excited state transitions. This information is also included in <i>FILE.ES.csv</i> , but the format here may be more convenient for directly comparing transitions.
<i>FILE.TDM.csv</i>	Excited state transition dipole moments.
<i>FILE.UV-Vis.csv</i>	A simulated UV-Vis spectrum in corrected units of wavelength (nm).
<i>FILE.absorptions.csv</i>	A simulated UV-Vis spectrum in uncorrected units of energy (eV).
<i>FILE.vibrations.csv</i>	Vibrational frequencies.
<i>FILE.IR.csv</i>	A simulated IR spectrum.

### Report

Silico will also generate a PDF report file at the end of each calculation, this is available under the 'Report' folder. The report contains much the same data as is available in the result files, but in a more convenient and easily transferable format. In addition, the PDF report contains graphical results that can not be represented in text, including simulated spectra, convergence graphs, energy diagrams and orbital density images.

In addition to the report itself, all the rendered and generated image files are available under the 'Report/image' directory. For each 3D rendered structure image two images are available, a lower quality but smaller sized .jpg, and a higher quality, larger .png file.

**NOTE:** The PDF report itself uses the smaller .jpg files to ensure the file size is small enough to be sent by email. The higher quality .png files are included for publication purposes.

Finally, a smaller mini report is also available, named something like *FILE.atoms.pdf*, which only contains final atom coordinates. The intention of this mini-report is for it to be easily included in the ESI of journal papers.

# Analysis of Results

In addition to aiding calculation submission, Silico also includes a number of tools for analysing calculation result files. Analysis is performed using the 'cresult' program.

## Generating Summaries

Used without other arguments, cresult will generate a text summary of the given calculation result files. The content of this summary is the same as found in the *FILE.summary* file (see **Post Analysis/Results** section), and is useful for inspecting calculation files that were not submitted using silico.

To analyse a file named 'Benzene.log', the cresult command would be used as follows:

```
> cresult Benzene.log
-----
|      Metadata      |
-----
Name: Benzene/TDA 10 Singlets 10 Triplets Gas Phase PBE1PBE 6-31G(d,p) /
Output/Benzene.log
Computational package: Gaussian (2009+D.01)
Calculations: Excited States"
Methods: DFT
Functional: PBE1PBE
...
```

To save the summary to a file, Linux file redirection can be used with the '>' character. For example, to save the summary to a file named 'Benzene.summary':

```
> cresult Benzene.log > Benzene.summary
```

**NOTE:** The standard redirection character (>) will overwrite the file if it already exists, deleting any existing data. To avoid this behaviour, use the double redirection character instead (>>).

## Analysis of Multiple Results

Although extracting and formatting results from a single calculation output file may be useful, the real strength of the cresult tool is being able to analyse and tabulate multiple calculation result files simultaneously.

This is achieved by specifying multiple files to cresult with the '-a' (table) option. It is recommended that this command be paired with the 'less' command, which allows the considerable quantity of data to be scrolled, as follows:

```
> cresult Benzene.log "Methyl benzene.log" -a | less -S
```



## Silico – Computational Chemistry Management – Quick Reference Guide

**NOTE:** Silico loads calculation result files in parallel, meaning there is little difference in reading one file or two (up to the number of cores in the computer).

The table format is useful for 'quickly' comparing the results from a handful of calculations, but may become cumbersome when comparing/analysing hundreds of results. For extended post analysis, the '-c' (**CSV**) option is recommended, which produces files that can be easily imported into a spreadsheet of choice.

To write tabulated CSV result files, use cresult with the '-c' option and redirect the output using the standard Linux redirection character '>'. For example, to write to a file named 'Results.csv':

```
> cresult Benzene.log "Methyl benzene.log" -c > Results.csv
```

When analysing a great many result files, typing each file name individually is cumbersome. Instead, using the Linux wildcard character (\*) is recommended. For example, to tabulate the results of all calculations in the current directory, the following command would be used:

```
> cresult */*/Output/*.log -c > Results.csv
```

The program may take some time to complete, depending on the number of calculations being analysed. The program will print the name of each file being read to keep track of progress:

```
> cresult */*/Output/*.log -c > Results.csv
silico: INFO: Parsing calculation result file 'Benzene/Opt Freq Gas
Phase B3LYP 6-31G(d,p)/Output/Benzene.log'
silico: INFO: Parsing calculation result file 'Benzene/Opt Freq Gas
Phase PBE1PBE 6-31G(d,p)/Output/Benzene.log'
silico: INFO: Parsing calculation result file 'Benzene/Single Point
Singlet Gas Phase PBE1PBE 6-31G(d,p)/Output/Benzene.log'
silico: INFO: Parsing calculation result file 'Benzene/TDA 10 Singlets
10 Triplets Gas Phase PBE1PBE 6-31G(d,p)/Output/Benzene.log'
silico: INFO: Parsing calculation result file 'Benzene/TD-DFT 10
Singlets 10 Triplets Gas Phase PBE1PBE 6-31G(d,p)/Output/Benzene.log'
```

Once complete, the resulting file can be downloaded from the server and analysed at leisure.

# Generating Reports

## Normal Usage

The PDF reports automatically generated by Silico at the end of a calculation can also be generated manually by using the 'creport' command. Reports require both the main output file (.log) and also the checkpoint file (either .chk or .fchk), as image data is only stored in the latter. However, creport will automatically search for checkpoint files in the same folder as the .log file, so long as they have the same name, so often only the .log file needs to be specified. Creport is used as follows:

```
> creport Benzene.log
```

Creport will search in the same folder as 'Benzene.log' for a file named either 'Benzene.fchk' or 'Benzene.chk'. The checkpoint file can also be given explicitly using:

```
> creport Benzene.log C6H6.fchk
```

The report will be written, by default, to a folder name 'report' **in the same folder as the input (.log) file**. This can be changed using the -o (output) option:

```
> creport Benzene.log C6H6.fchk -o "Benzene Report"
```

Generating the report can take some time, especially for calculations involving excited states (TDA and TD-DFT). The creport program will print status messages periodically to keep track of progress:

```
> creport Benzene.log
silico: INFO: Found 'Benzene.fchk' in input directory; using as
gaussian/formatted-checkpoint file
```

This message indicates that creport found the checkpoint file correctly. Other output will appear similar to the following:

## Silico – Computational Chemistry Management – Quick Reference Guide

```
> creport Benzene.log
silico: INFO: Found 'Benzene.fchk' in input directory; using as
gaussian/formatted-checkpoint file
silico: INFO: Reading result file 'Benzene.log'
silico: INFO: Generating report 'report/Benzene.pdf'
silico: INFO: Converting gaussian/formatted-checkpoint file
'Benzene.fchk' to gaussian/cube file
'report/image/Structure/Benzene.structure.cube'
silico: INFO: Rendering report/image/Structure/Benzene.structure.jpg to
file(s)
silico: INFO: Rendering report/image/Benzene.SCF_graph.png to file(s)
silico: INFO: Rendering report/image/Dipole Moment/Benzene.dipole.jpg
to file(s)
silico: INFO: Rendering report/image/Orbital
Diagram/Benzene.HOMO_LUMO.png to file(s)
silico: INFO: Converting gaussian/formatted-checkpoint file
'Benzene.fchk' to gaussian/cube file
'report/image/HOMO/Benzene.HOMO.cube'
silico: INFO: Rendering report/image/HOMO/Benzene.HOMO.jpg to file(s)
silico: INFO: Converting gaussian/formatted-checkpoint file
'Benzene.fchk' to gaussian/cube file
'report/image/LUMO/Benzene.LUMO.cube'
silico: INFO: Rendering report/image/LUMO/Benzene.LUMO.jpg to file(s)
silico: INFO: Rendering report/image/HOMO LUMO/Benzene.HOMO_LUMO.jpg to
file(s)
silico: INFO: Rendering report/image/Orbital
Diagram/Benzene.orbitals.png to file(s)
silico: INFO: Rendering report/image/Benzene.simulated_frequencies.png
to file(s)
silico: INFO: Done generating report 'report/Benzene.pdf'
```

### Including Emission Energy

Reports do not, by default, include results on emission energy. This is because these results require multiple calculation result files, and creport cannot know where to find these automatically. To include emission data, use the '--emission' option with the path to the emission result file. For example, using Benzene with calculation code 1/1/4 (the standard for organometallic phosphorescent emitters) the full command would be:

```
> creport "Opt Freq Acetonitrile (SCRF-PCM) B3LYP 6-31+G(d,p)_SBKJC-VDZ
(ECP)/Output/Benzene.log" --emission "Opt Freq Unrestricted Triplet
Acetonitrile (SCRF-PCM) uB3LYP 6-31+G(d,p)_SBKJC-VDZ
(ECP)/Output/Benzene.log" -o "Benzene emission"
```

**NOTE:** EZC SOP is to calculate emission data for phosphorescent emitters only (because these calculations are more accessible). The emission calculation is the 'Opt Freq Unrestricted Triplet' calculation. The difference in energy between this calculation and the normal, ground state optimisation gives the **adiabatic** (relaxed) emission energy. SOP is also to calculate the **vertical** (unrelaxed) emission energy, which is the difference between the ground state optimised and a single point calculation at the triplet geometry. These calculations are all performed by code 1/1/4.

## Silico – Computational Chemistry Management – Quick Reference Guide

This will calculate the **adiabatic** emission energy. To also include the **vertical** emission, use the '--vertical\_ground' option:

```
> creport "Opt Freq Acetonitrile (SCRF-PCM) B3LYP 6-31+G(d,p)_SBKJC-VDZ  
(ECP)/Output/Benzene.log" --emission "Opt Freq Unrestricted Triplet  
Acetonitrile (SCRF-PCM) uB3LYP 6-31+G(d,p)_SBKJC-VDZ  
(ECP)/Output/Benzene.log" --vertical_ground "Single Point Singlet  
Acetonitrile (SCRF-PCM) B3LYP 6-31+G(d,p)_SBKJC-VDZ  
(ECP)/Output/Benzene.log" -o "Benzene emission"
```

In this case, the report would be rendered to the “Benzene emission” folder.