

# InSilicoLab Manual

## InSilicoLab for Chemistry

InSilicoLab is a collection of tools intended to assist the researcher in performing complex experiments of computational chemistry. It facilitates preparation of input files and submitting the jobs to the grid infrastructure, controls job execution, collects the output files and performs basic analysis of results. The user-defined experiment can be stored to allow for easy repetition of computational procedure. InSilicoLab integrates with the grid storage providing access to the user data.

Currently, experiments available in InSilicoLab for Chemistry are:

- Trajectory Sculptor
- Quantum-chemical calculations experiment
- Cubegen experiment

The InSilicoLab portal can be accessed at <https://insilicolab.chemia.plgrid.pl>.

**Note:** It is advisable to start reading with the [basic user's manual](#) before advancing to this one.

## Getting access to InSilicoLab

InSilicoLab is accessible to anyone registered as a PL-Grid user or having a valid Grid Certificate and belonging to *vo.plgrid.pl* or *gaussian* Virtual Organisations. PL-Grid users may log in using their PL-Grid credentials by choosing to log in with OpenID or with personal certificate. In case of OpenID login, the user will be redirected to the OpenID page, where they will be prompted for PL-Grid login and password or a Grid Certificate associated with the account. If certificate login is chosen, the InSilicoLab portal will read the certificate installed in the user's browser. **Note:** If the certificate is not present or the user declines to use it for authentication, browser restart would be required to log in with the certificate again.

More information about activating and accessing the service for PL-Grid users can be found [here](#).

InSilicoLab offers also anonymous preview access. It includes manipulation of molecular dynamics trajectories with use of Trajectory Sculptor and viewing sample experiments and their data. Downloading any data is not possible in the anonymous mode.

## Trajectory Sculptor

Sequential MD/QC approach is often used in explicit solvent modeling. In the first step, Molecular Dynamics simulations (classical or *ab initio*) are employed to generate atomistic structures of the system and their evolution in time. In the next step, quantum-chemical calculations are performed for selected part of the system (solute molecule and closest solvent molecules in its solvation shell) to calculate required properties (such as potential energy, excitation energies, chemical shifts, etc.). Such calculations are usually repeated for series of frames chosen from the trajectory.

Trajectory Sculptor is a tool facilitating data preparation in such a procedure. It assists the user in trajectory analysis, selection of the part of studied system and extraction of solute-solvent geometries from selected frames. Data can then be passed to another InSilicoLab experiment to set-up and execute quantum-chemical calculations, monitor their status, collect output files and perform basic post-processing of the results.

Trajectory Sculptor can work with trajectories obtained from many MD packages. With support for multiple types of solvent molecules and flexible construction of the solvation shell (choice of distance metric and selection of solvent molecules) this tool is applicable to a wide range of systems, including solvent mixtures, electrolyte solutions or ionic liquids.

All steps of MD trajectory analysis in Trajectory Sculptor are described in detail in following sections.

### Accessing Trajectory Sculptor

Trajectory Sculptor can be accessed from the main page after entering the portal. After choosing *Trajectory Sculptor* from the list of new experiments to create, the Trajectory Sculptor interface, similar to the one in the picture is shown.

## Trajectory upload

Basic format of the MD trajectory files used by Trajectory Sculptor is the *xyz* format. However, trajectory may be supplied in other format recognized by OpenBabel in which case it will be automatically converted to *xyz* file.

Note: trajectory files generated by some popular MD programs including NAMD, Gromacs, AMBER or Tinker can be saved as *xyz* files using VMD software.

In order to make the analysis of the trajectory file possible in Trajectory Sculptor, several requirements have to be fulfilled:

- correct chemical symbols must be used in the trajectory file (i.e. no nonstandard labels)
- currently only one solute molecule is allowed
- sequence of atoms must be the same in all frames
- identity of molecules must be preserved in the course of MD run (no changes in bonding pattern)
- if periodic boundary conditions are used, all periods must be kept constant; only rectangular boxes are supported

To upload the trajectory file, press the **Browse...** button in the **Load trajectory** panel and select the file. After the file has been successfully loaded, a message appears in the message box on the right of the panel and the number of frames found in the file is displayed:

Uploaded trajectory may be previewed in the Jmol viewer using the Preview frame... button and selecting one of available options

- **Preview first frame**
- **Preview frame...** (on choosing this option a message box appears allowing to enter desired frame number)
- **Preview whole trajectory**

Note: In case the trajectory has only one frame or it is in format other than *xyz* and exceeds the size limit, only the **Preview whole trajectory** option is available. The size limit is introduced due to performance issues of [OpenBabel](#) which is used to extract the chosen frame. The limits are: 10,000 atoms or 100 frames or 200,000 atoms\*frames. In case of *xyz* files these limits do not apply as a more optimal processing is chosen. Still, if the trajectory has only one frame, all the options are equal to **Preview whole trajectory**.

Selected frames may be also downloaded back in the *xyz* format using the **Download frame...** button and providing frame number (the same size limits as for frame preview apply here).

If periodic boundary conditions (PBC) were used in the MD simulation, period lengths must be provided. To do this check the **Use Periodic Boundary Conditions** checkbox, then enter period values (in Å) for three perpendicular directions. If the simulation box was a cube, select **Equal period values** checkbox in which case only one value (for period A) is necessary, the other two are automatically duplicated.

Trajectory solution.xyz loaded. Found 27 frames.

Preview frame...

Download frame...

☒ Use Periodic Boundary Conditions

A =  Å B =  Å C =  Å

☒ Equal period values

PBC information for the system is stored within the Trajectory Sculptor. If the user downloads the same file in the future, PBC settings can be easily recalled using the **Insert saved settings** option.

Trajectory solution.xyz loaded. Found 27 frames.

Preview frame...

Download frame...

☐ Use Periodic Boundary Conditions    **Insert saved settings**

## System specification



The next step of MD trajectory analysis is the specification of molecules (solute and solvent) present in the system. Specification is entered in the **Define the molecules** panel.

Solute molecule is defined by the range of atom numbers within the frame entered in the **Atom Range** field in **Solute** section. Sequence of atom numbers must be continuous and is given as hyphen-separated two integers. Optionally, sequence of atom labels for the solute may be entered in the **Molecular Formula** field. This input is used only for verification.

Solvent molecules are defined by the sequence of chemical symbols entered in the **Atom sequence(s)** field in the **Solvent** section. The sequence of symbols must match exactly sequence of atoms in the solvent molecule in the trajectory, e.g. OHH defines water molecule. All molecules of given type in the trajectory must have the same order of atom labels. If more than one type of solvent is present in the system, another **Atom sequence(s)** field may be created by clicking the **Add...** command. Sequence of symbols for the next solvent may be then entered. Note: solvents should be defined in the order of first appearance in the trajectory file.



On clicking the **Find molecules** button the first frame of the trajectory is analyzed and the information about molecules found in the system is displayed:

## Define the molecules

 **Solute:** Find by *Atom range* 


**Atom Range**

**Molecular Formula**


 **Solvent:** Find by *Atom sequence* 

**Atom Sequence(s)**   [Add...](#)

Found:

1 solute with molecular formula: C<sub>9</sub>N<sub>1</sub>O<sub>2</sub>H<sub>11</sub> [Preview](#) 

200 solvents with molecular formula: O<sub>1</sub>H<sub>2</sub> [Preview sample](#) 

100 solvents with molecular formula: C<sub>2</sub>O<sub>1</sub>H<sub>6</sub> [Preview sample](#) 

[Find molecules](#)

In the above example one solute molecule consisting of 23 atoms with the formula C<sub>9</sub>N<sub>1</sub>O<sub>2</sub>H<sub>11</sub>, 200 water molecules (defined as OHH sequence) and 100 ethanol molecules (defined as OHHCHCHHH) are found in the frame. Solute molecule and example of solvent molecules may be viewed on choosing **Preview** or **Preview sample** option, respectively. Their geometries may be downloaded in the xyz format on pressing the arrow symbol.

## Distance metric definition

To select solvent molecules closest to the solute one must define how the distance between two molecules is defined. Trajectory Sculptor allows several ways of measuring solute-solvent distances.

To define the distance metric for solute-solvent pair two reference points are defined: one anchored at the solute molecule, the other at the solvent. Distance is then defined as the distance measured from one reference point to the other. Reference points are selected from the **Solute: Measure from** and the **Solvent: Measure to** lists.

Four possible reference points may be defined:

- **Chosen atom** - reference point is selected atom of the molecule; its number has to be entered in the atom number field. Atom numbers are counted within the molecule, not in the whole system, i.e. atom no. 2 means the second atom in the molecule, although it may be the atom number 1502 in the frame
- **Nearest atom** - reference point is the atom of the molecule closest to the other reference point (at the other molecule)
- **Geometric center**
- **Center of mass** - reference point is calculated using the atomic masses of elements; isotopes are not taken into account





Any combination of these reference points is allowed.

Examples of distance metric definitions are:

- solute: nearest atom, solvent: nearest atom - distance is defined as the distance between closest atoms of the two molecules
- solute: chosen atom no. 10, solvent: nearest atom - distance is measured between atom no. 10 in the solute molecule and the closest to it atom from the solvent molecule
- solute: center of mass, solvent: chosen atom no. 2 - distance is measured between center of mass of the solute molecule and the second atom of the solvent

If more than one solvent type is present, different distance metric may be defined for each type of solute-solvent pair as in the following example:

## Define distance metric

	<b>Solute:</b>	Measure from	<input type="text" value="Chosen atom"/>	atom number	<input type="text" value="10"/>	<input type="checkbox"/> Use the same metric for each solute-solvent pair
	<b>Solvent OHH:</b>	Measure to	<input type="text" value="Nearest atom"/>			
	<b>Solute:</b>	Measure from	<input type="text" value="Chosen atom"/>	atom number	<input type="text" value="10"/>	
	<b>Solvent OHHCHCHHH:</b>	Measure to	<input type="text" value="Center of mass"/>			

Checking the **Use the same metric for each solute-solvent pair** option facilitates entering the same definition for all solvents. Note, however, that this option is inactive if the **Chosen atom** reference point has been selected for the solute.

## Selection of the solvation shell


After the solute-solvent distance metric has been defined it is possible to specify which solvent molecules should be included in the solvation shell. This is achieved by proper settings of the values in **Choose** option. Solvent molecules can be specified in two ways:


- **Up to specified distance** - all solvent molecules within the specified distance (in Å) are included in the solvation shell
- **Closest molecules** - given number of solvent molecules closest to the solute is selected


In both cases the solute-solvent distances are measured according to the distance metric defined for given solute-solvent pair.


If multiple types of solvent molecules are present, either one or many values may be supplied for both options. If one value is given, it is used for all solvents. Alternatively, different values (separated by semicolons) may be specified for each type of the solvent; the numbers should be given in the same order as used in solvent specification.


#### Define distance metric

 **Solute:** Measure from Chosen atom atom number  ☐ Use the same metric for each solute-solvent pair


 **Solvent OHH:** Measure to Nearest atom

 **Solute:** Measure from Chosen atom atom number

 **Solvent OHHCHCHHH:** Measure to Center of mass

Choose Closest molecules  molecules 


[Reduce frame](#)


The reduced frame contains:  
1 solute with molecular formula: C<sub>9</sub>N<sub>1</sub>O<sub>2</sub>H<sub>11</sub>  
10 solvents (out of total 200) with molecular formula: O<sub>1</sub>H<sub>2</sub>  
5 solvents (out of total 100) with molecular formula: C<sub>2</sub>O<sub>1</sub>H<sub>6</sub>  
[Preview result](#) 


In the above example 10 water molecules and 5 ethanol molecules closest to the solute are selected. Solute-water distance is measured from the atom no. 10 in the solute molecule to the nearest atom of water molecule; solute-ethanol distance is measured from the atom no. 10 of solute to the center of mass of ethanol molecule.


The **Reduce frame** button may be used to analyze a sample frame (the first frame of the trajectory). Number of solvent molecules matching the requested criteria is displayed (in the example 11 water and 12 ethanol molecules were found within the 8 Å distance from the solute):


#### Define distance metric

 **Solute:** Measure from Chosen atom atom number  ☐ Use the same metric for each solute-solvent pair


 **Solvent OHH:** Measure to Nearest atom

 **Solute:** Measure from Chosen atom atom number

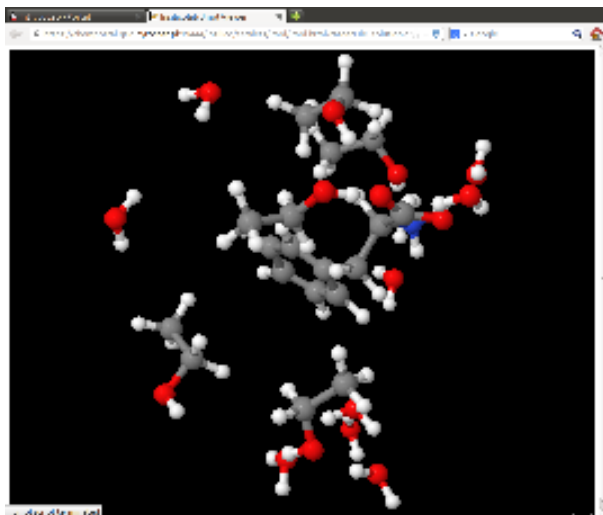
 **Solvent OHHCHCHHH:** Measure to Center of mass

Choose Up to specified distance  Å 

[Reduce frame](#)

The reduced frame contains:  
1 solute with molecular formula: C<sub>9</sub>N<sub>1</sub>O<sub>2</sub>H<sub>11</sub>  
11 solvents (out of total 200) with molecular formula: O<sub>1</sub>H<sub>2</sub>  
12 solvents (out of total 100) with molecular formula: C<sub>2</sub>O<sub>1</sub>H<sub>6</sub>  
[Preview result](#) 

and the sample of reduced frame may be previewed in Jmol using the **Preview result** link



or downloaded as an xyz file on clicking on the arrow symbol.

Distance metric definition and the parameters of the **Choose** option may be changed as many times as necessary, therefore the **Preview result** command helps to find out the best distance metric and the appropriate way of selecting solvent molecules.

### Final analysis of the trajectory

When the molecules in the system have been defined, the distance metric and the way of selecting molecules have been set up, the MD trajectory can be finally analyzed to extract solute with specified solvation shell from selected frames.

The **Choose frames** option allows to specify which frames should be analyzed:

- **Only first frame**
- **List frames** - allows to specify the list of frames
- **Define range** - allows to select numbers of the first and the last frame and the increment
- **All frames**

When the **List frames** has been used the frame numbers separated by commas can be entered

Choose frames

☐ Only first frame
☒ List frames 
☐ Define range  to  step 
☐ All frames
☐ Put resulting frames to one file

If more than one frame has been specified, the **Put resulting frames to one file** option can be checked to combine all reduced frames into one xyz file. If this option has been left unchecked, a separate xyz file is produced for each frame.

Clicking the **Run** button starts the analysis and reduction of the MD trajectory.

If the **Put resulting frames to one file** option has been used, the resulting xyz file may be downloaded by clicking on the arrow.

Choose frames

☐ Only first frame
☐ List frames 
☒ Define range  to  step 
☐ All frames
☒ Put resulting frames to one file

! One result file generated ↓

Otherwise several files are created:

**Choose frames**

☐ Only first frame  
☐ List frames   
☒ Define range  to  step   
☐ All frames  
☐ Put resulting frames to one file

9 files generated [More...](#)

Download

Download all

Use all results

They can be downloaded individually (**Download**), downloaded as one *xyz* file (**Download all**) stored or passed to another InSilicoLab experiment (**Use all results**).

## Quantum-chemical calculations experiment

This experiment facilitates preparation of input files for quantum-chemical calculations and execution of jobs on the PLGrid infrastructure.

### Accessing the QC experiment

QC experiment can be accessed from the main page after entering the portal - by choosing *Quantum Chemistry experiment* from the list of new experiments to create.

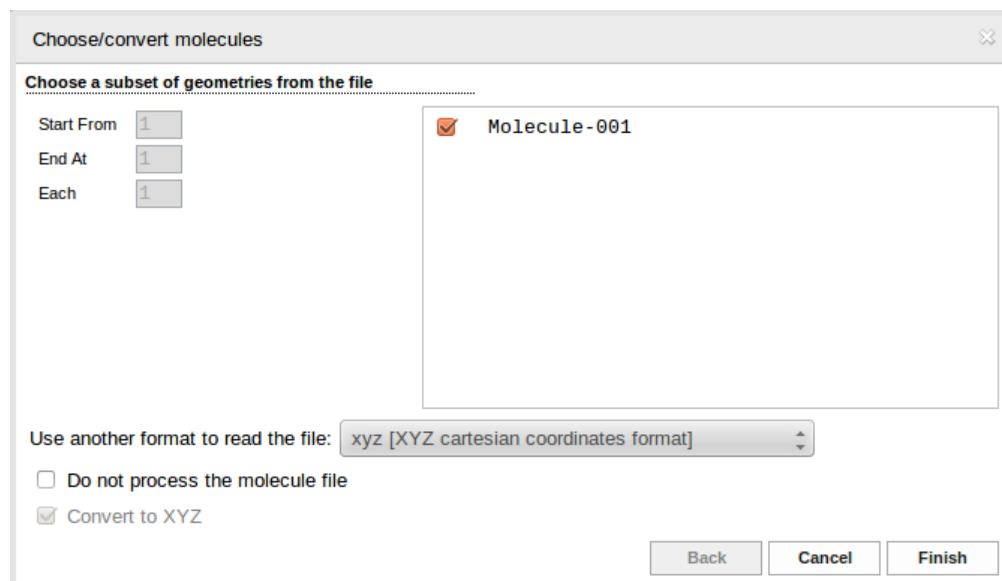
Another possibility is to invoke the QC experiment from another InSilicoLab experiment (e.g. Trajectory Sculptor). In such case geometries of the system may be passed between experiments and they are immediately available as input data in the QC experiment interface.

The screenshot shows the InSilicoLab Portal interface. On the left is a sidebar with 'Your Experiments' (listing 'sculptor-nalewka.xyz') and 'LFC Catalogue'. The main area is titled 'Welcome sculptor-nalewka...' and contains a form for describing the experiment. The form includes fields for 'Short name' and 'Description'. Below these is the 'Input Data Specification' section, which has a 'Gaussian' dropdown, a 'Load prepared input' button, and a 'Browse...' button. It also has fields for 'Output file name' (gaussian.log), 'Checkpoint options' (Use checkpoint files... and Write checkpoints), 'Title Section' (%chk=<output\_file\_name>.chk), 'Link 0 Section', 'Route Section', 'Charge & Multiplicity', and 'Additional Parameters'. There is a 'Cubegen Options' checkbox. The 'Geometry Specification' section has a 'Browse...' button and a list of frames: 1. nalewka.xyz: Frame-025, 2. nalewka.xyz: Frame-010, 3. nalewka.xyz: Frame-005, 4. nalewka.xyz: Frame-020, 5. nalewka.xyz: Frame-015, 6. nalewka.xyz: Frame-030. At the bottom, there are 'Input.inp' and 'Download Input Data' buttons. The footer shows '© 2014 InSilicoLab v1.4' and a 'Report a problem' link.

### Setting-up the QC experiment

The QC experiment is identified by the name entered in the field **Short name**. Optionally, more detailed description may be provided in the **Description** field. If no Short name is supplied, the contents of the **Title** section will be used. If both fields are empty, the "(no title)" will be used as a short name.

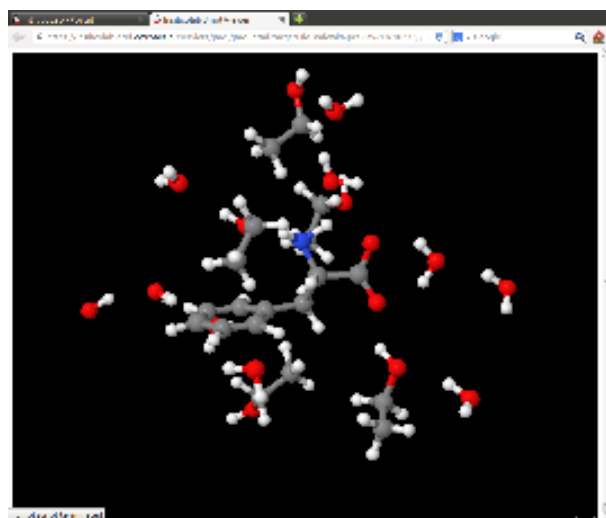
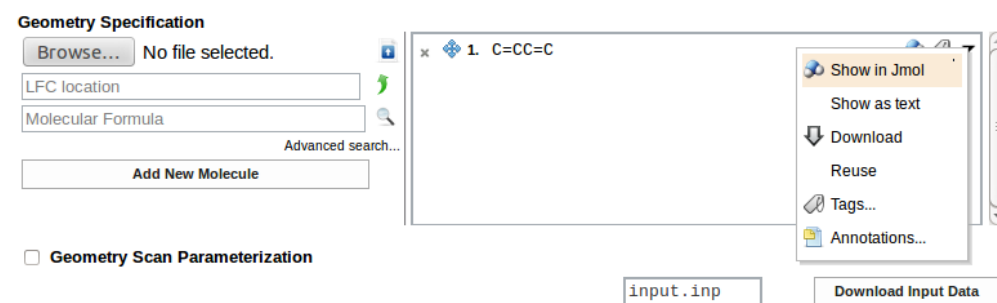
The geometry of the system may be entered from a file loaded in the **Molecule Specification** section. The type of the file may be specified in the opening window.



The system and the requested calculations are specified in the **Input Data Specification** section.

As a default, the file type is determined from the filename extension. Click **Finish** button to finish the specification of the geometry. If the QC calculations were invoked from another InSilicoLab experiment, the list of molecular geometries already contains the data passed from the other experiment.

Structure of the system selected from the list may be visualized as an Jmol applet (**Show in Jmol**):



The software used in calculations can be selected from the list. Currently, three QC programs are supported:

- Gaussian 09
- GAMESS
- Turbomole
- Niedoida (QC package developed at the Faculty of Chemistry of the Jagiellonian University)

The fields needed to properly describe the calculations depend on the software chosen.

In the case of Gaussian 09 the following can be supplied:

- **Title** section - short description of the job (optional)
- **Link 0** section - parameters of program execution such as memory and disk limits and the use of checkpoint files (optional)
- **Route** section - QC method and type of calculations (required)
- **Charge & Multiplicity** of the system (required)
- **Additional parameters** - added to the Gaussian input file after the geometry specification (optional)
- **Cubegen options** - to generate a Gaussian cube file (optional)
- **Geometry Scan Parameterization** - allow performing geometry scan of the structure (optional)

For example, the input for Gaussian ZINDO calculations of 30 lowest electronic excitations for a phenylalanine molecule in a solution using structures prepared in Trajectory Sculptor and requesting 512MB RAM for each job may read:

Describe your experiment

Short name

solution

Description

phenylalanine in water/ethanol solution

Input Data Specification

Gaussian

Load prepared input

Browse...

No file selected.

Output file name

gaussian

.log

Checkpoint options

Use checkpoint files...

☐ Write checkpoints

Title Section

phenylalanine

Link 0 Section

%mem=512MB

Route Section

#zindo(nstates=30)


Charge & Multiplicity

0 & 1

Add

Additional Parameters

☐ Cubegen Options

  
Gaussian  
Gaussian 09 C 01

## Job execution

Clicking the **Run** button creates the input files based on the supplied geometries and parameters entered in the **Input Data Specification**. The jobs are then submitted for execution to the PLGrid.

InSilicoLab - Portal

https://insilicolab.grid.cyfronet.pl/main.html

You are logged in as **plgeilmes**

## in silico LAB

### Your Experiments

Menu

- solution
- sculptor-nalewka.xyz

Filter by name

### LFC Catalogue

experiments

Filter by name

© 2014 InSilicoLab v1.4

Welcome **sculptor-nalewka...** New experiment solution

Experiment **solution**

Started **2014-05-01** at 11:31:47

Status **Running**

Identifier **1398936707810**

LFC directory **/grid/vo.plgrid.pl/insilicolab/c.PL\_o.PL-Grid\_o.Uzytkownik\_o.CYFRONET\_cn.Andrzej\_Eilmes\_cn.plgeilmes/experiments/experiment-1398936707810**

phenylalanine in water/ethanol solution

Reuse experiment input

Experiment Input Data

Output file name **gaussian** .log

Checkpoint options

Use checkpoint files...

Write checkpoints

Title Section

phenylalanine

Link 0 Section

%mem=512MB

Route Section

#zindo(nstates=30)

Charge & Multiplicity

0 & 1

Additional Parameters

Cubegen Options

Geometry Specification

1. nalewka.xyz: Frame-025
2. nalewka.xyz: Frame-010
3. nalewka.xyz: Frame-005
4. nalewka.xyz: Frame-020
5. nalewka.xyz: Frame-015
6. nalewka.xyz: Frame-030

Geometry Scan Parameterization

input.inp

Download Input Data

Report a problem

In the **Job Execution Status** section one may trace the progress of calculations. The job status changes from **EXTERNAL** through **SUBMITTED** and **RUNNING** to **FINISHED**.

### Job Execution Status

Input File

Collapse

Job	Status	Input File
Job 01:	<b>RUNNING</b>	8fc5f4ab34fcb78a31b6c668747ec00d
Job 02:	<b>RUNNING</b>	1bc6390d6afaf6edaa8256aced9904a7
Job 03:	<b>RUNNING</b>	3b5046a9f74b97367d9305da6cfc605a
Job 04:	<b>SUBMITTED</b>	1101488cd701f54964ac5b4d6c573ac7
Job 05:	<b>SUBMITTED</b>	c20a7e6c1e51846b7f3a5bc85e9eff14
Job 06:	<b>SUBMITTED</b>	ab550d4ceb96044e282df667ac0233f4

Collapse

Abort all jobs

Get All Job IDs

Index 1

Job ID https://1b02.grid.cyf-kr.edu.pl:9000/-f10PpyDLzNGLIB0xPZfQg

Status **running**

Additional status **JSAGA:RUNNING\_QUEUED**

LFC location /grid/vo.plgrid.pl/insilicolab-test/c.PL\_o.PL-Grid\_o.Uzytkownik\_o.Uniwersytet\_Jagiellonski\_cn.Andrzej\_Eilmes\_cn.plgeilmes/experiments/experiment-1361823717317/job-01

Download job files

Results of finished jobs are used to update the table and the plot of SCF energy values:



For each job the **Download job files** list may be used to select and download the files associated with the job: input, checkpoint and output files as a single .tar.gz file or the XML file containing the results extracted from the output.

Job Execution Status

**Input File**

Job 01:	8fc5f4ab34fcb78a31b6c668747ec00d	<b>FINISHED</b>
Job 02:	1bc6390d6afaf6edaa8256aced9904a7	<b>FINISHED</b>
Job 03:	3b5046a9f74b97367d9305da6cfc605a	<b>FINISHED</b>
Job 04:	1101488cd701f54964ac5b4d6c573ac7	<b>FINISHED</b>
Job 05:	c20a7e6c1e51846b7f3a5bc85e4eff14	<b>FINISHED</b>
Job 06:	ab550d4ceb96044e282df667ac0233f4	<b>FINISHED</b>

**Index** 3

**Job ID** https://lb02.grid.cyf-kr.edu.pl:9000/8C4fwBZTUqgmLfGeYSQrXQ

**Status** finished

**Additional status** JSAGA:RUNNING\_POST\_STAGING

**LFC location** /grid/vo.plgrid.pl/insilicolab-test/c.PL\_o.PL-Grid\_o.Uzytkownik\_o.Uniwersytet\_Jagiellonski\_cn.Andrzej\_Eilmes\_cn.plgeilmes/experiments/experiment-1361823717317/job-03

**Download job files**

- jobfiles.tar.gz
- parsed\_data.xml

Collapse

Collapse   Abort all jobs   Get All Job IDs

In the above example of ZINDO calculations such file contains the list of calculated transition energies and oscillator strengths.

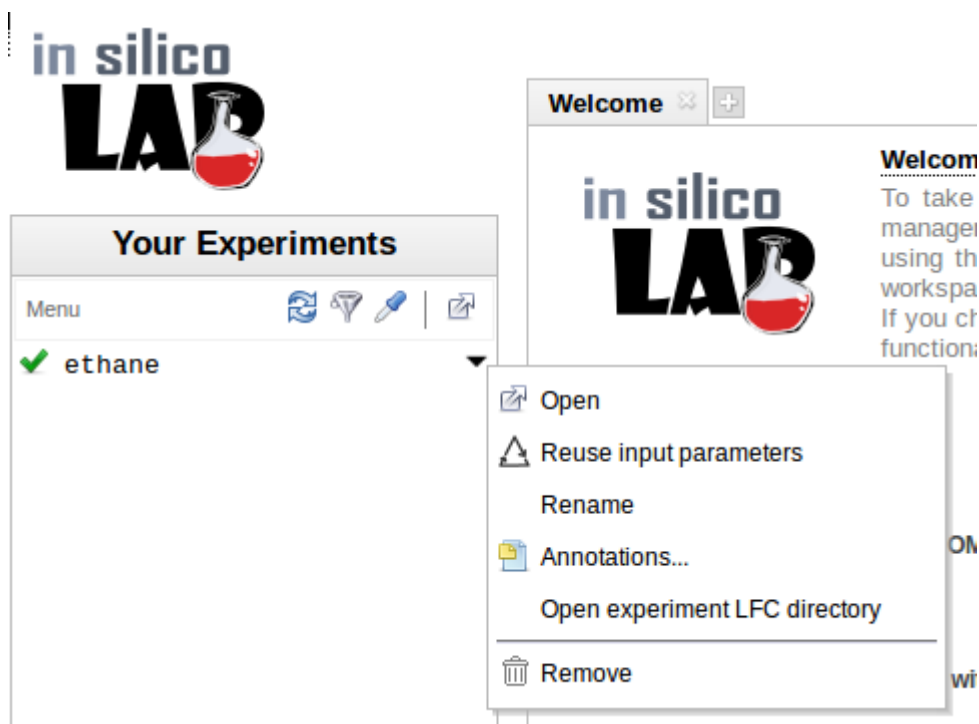
```

InSilicoLab - Portal
file:///tmp/p...ed_data-1.xml
file:///tmp/parsed_data-1.xml

<orbital energy="10.5620988937" number="203" symmetry="A"/>
<orbital energy="10.8491790002" number="204" symmetry="A"/>
<orbital energy="11.3596645734" number="205" symmetry="A"/>
<orbital energy="11.4916397882" number="206" symmetry="A"/>
<orbital energy="12.3191379908" number="207" symmetry="A"/>
<orbital energy="12.4200922273" number="208" symmetry="A"/>
<orbital energy="15.0402764408" number="209" symmetry="A"/>
</orbitals>
-<Electronic_transitions number_of_ET="30" energy_units="cm^-1">
<ET energy="33643.23072" osc_strength="0.0005" number="1" symmetry="Singlet-A"/>
<ET energy="38508.40064" osc_strength="0.0157" number="2" symmetry="Singlet-A"/>
<ET energy="40285.25232" osc_strength="0.0009" number="3" symmetry="Singlet-A"/>
<ET energy="42396.01984" osc_strength="0.1477" number="4" symmetry="Singlet-A"/>
<ET energy="49328.40304" osc_strength="0.5356" number="5" symmetry="Singlet-A"/>
<ET energy="50305.1472" osc_strength="0.4842" number="6" symmetry="Singlet-A"/>
<ET energy="51060.89392" osc_strength="0.1953" number="7" symmetry="Singlet-A"/>
<ET energy="52370.74736" osc_strength="0.1139" number="8" symmetry="Singlet-A"/>
<ET energy="53116.0088" osc_strength="0.2633" number="9" symmetry="Singlet-A"/>
<ET energy="54110.49728" osc_strength="0.0807" number="10" symmetry="Singlet-A"/>
<ET energy="55571.984" osc_strength="0.0626" number="11" symmetry="Singlet-A"/>
<ET energy="55709.0992" osc_strength="0.0019" number="12" symmetry="Singlet-A"/>
<ET energy="56720.52544" osc_strength="0.031" number="13" symmetry="Singlet-A"/>
<ET energy="57310.92736" osc_strength="0.0383" number="14" symmetry="Singlet-A"/>
<ET energy="57593.22336" osc_strength="0.0113" number="15" symmetry="Singlet-A"/>
<ET energy="58277.18624" osc_strength="0.0082" number="16" symmetry="Singlet-A"/>
<ET energy="58778.06" osc_strength="0.0022" number="17" symmetry="Singlet-A"/>
<ET energy="60008.87056" osc_strength="0.0002" number="18" symmetry="Singlet-A"/>
<ET energy="60106.46432" osc_strength="0.0055" number="19" symmetry="Singlet-A"/>
<ET energy="60524.2624" osc_strength="0.0158" number="20" symmetry="Singlet-A"/>
<ET energy="61053.36576" osc_strength="0.0171" number="21" symmetry="Singlet-A"/>
<ET energy="61210.64496" osc_strength="0.0072" number="22" symmetry="Singlet-A"/>
<ET energy="61459.06544" osc_strength="0.0018" number="23" symmetry="Singlet-A"/>
<ET energy="61899.4472" osc_strength="0.0025" number="24" symmetry="Singlet-A"/>
<ET energy="62047.85424" osc_strength="0.0149" number="25" symmetry="Singlet-A"/>
<ET energy="62896.35536" osc_strength="0.0038" number="26" symmetry="Singlet-A"/>
<ET energy="63204.46128" osc_strength="0.0012" number="27" symmetry="Singlet-A"/>
<ET energy="63402.87504" osc_strength="0.005" number="28" symmetry="Singlet-A"/>
<ET energy="63730.3384" osc_strength="0.0039" number="29" symmetry="Singlet-A"/>
<ET energy="64577.2264" osc_strength="0.0072" number="30" symmetry="Singlet-A"/>
</Electronic_transitions>
</computation>
</job>

```

User can store the parameters of the experiment and job files for future retrieval or easy repetition of calculations:



## Using niedoida program in general quantum-chemistry experiment

Niedoida is a quantum-chemical package developed at the Faculty of Chemistry of the Jagiellonian University in collaboration with the Academic Computer Centre ACK Cyfronet AGH.

Program implements standard methods of quantum-chemistry

- Hartree-Fock
- Density Functional Theory and TDDFT
- *post*-HF methods: MP2 and CIS

MP2 calculations can be accelerated by the resolution-of-identity technique (RI-MP2; efficient for large basis sets) or Laplace-Transform MP2 (effective for spatially extended systems).

Program features implementation of "dressed"-TD DFT incorporating corrections from doubly-excited states.

### LT-MP2 calculations example

To use the program its name needs to be selected from the list of programs in the **Input Data Specification** section:

Input Data Specification

Niedoida

Load prepared input

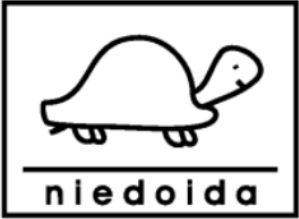
Browse...

No file selected.

Title

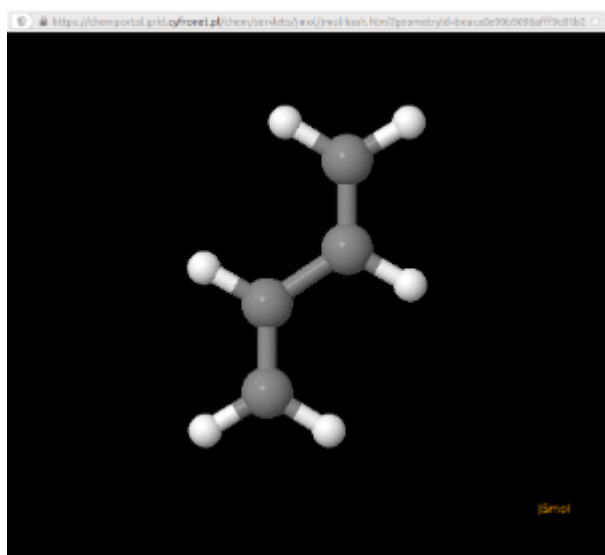
Basis Set

Other Parameters



Niedoida 0.5.3

Coordinates of the system can be entered in the **Geometry Specification** section. Uploaded geometry can be viewed as Jmol applet on clicking the molecule icon next to the file name. Butadiene molecule is used in the example.



Required input data should be entered in the **Input Data Specification** section:

Welcome x New experiment x

Your proxy expires in 10 hours, MyProxy is enabled.

---

Describe your experiment

**Short name** butadiene, LT-MP2

**Description** butadiene molecule, MP2 with Laplace transform

---

Input Data Specification

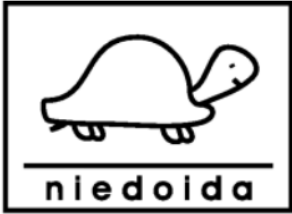
Niedoida Load prepared input Browse... No file selected.

**Title** butadiene, LT-MP2

**Basis Set** 6-31++g\*\*

**Other Parameters**

```
moller_plesset:
order: 2
type: laplace_ao
quadrature: fitted_simple
```



Niedoida 0.5.3

---

**Geometry Specification**

Browse... No file selected.

LFC location

Molecular Formula

Advanced search...

Add New Molecule

1. butadiene.xyz

In the above example 6-31++G\*\* basis is requested for calculations (**Basis Set** field), other data are entered in **Other Parameters** field:

units:

```
length: angstrom
energy: hartree
```

moller\_plesset:

```
order: 2
type: laplace_ao
quadrature: fitted_simple
```

Here we have specified units of distance (default unit is bohr) and energy. MP2 calculations with Laplace-Transform have been requested.

Calculations are initiated on clicking the **Run** button. After successful completion output file (niedoida.log) and parsed XML data can be downloaded.

InSilicoLab - Portal

https://chemportal.grid.cyfronet.pl/chem/main.html

You are logged in as pigeimes

**Your Experiments**

Menu

butadiene, dressed TD DF

butadiene, LT-MP2

Filter by name

**LFC Catalogue**

experiments

Filter by name

Job Execution Status

Input Data

Job 01: **FINISHED** butadiene.xyz

Index 1

Job ID https://lb02.grid.cyf-kr.edu.pl:9000/zMt6GT0sB5HP0U4t60po6Q

Status **finished**

Additional status JSAGA:RUNNING\_POST\_STAGING

Live log https://chemportal.grid.cyfronet.pl/liveLogging/view.php?id=1396043797855-job1&token=9cef1320b35a1761684cda15c1d7f6c7

LFC location /grid/vo.plgrid.pl/insilicob-test/c.PL\_o.PL-Grid\_o.Uzytkownik\_o.CYFRONET\_cn.Andrzej\_Eilmes\_cn.pigeilmes/experiments/experiment-1396043797855/job-01

Download job files

experimentConfig.py

niedoida.inp

niedoida.log

output.txt

parsed\_data.xml

Results

Job No.	SCF Energy [eV]	s <sup>2</sup>	Charge	Multiplicity
1	-4216.02356033	-	0	1

Check for results Export data Change energy unit eV

Experiment Log

```
[DEBUG] pythonConfiguration: {"experimentId":"1396043797855-job1",
"loggerHost":"chemportal.grid.cyf-kr.edu.pl", "loggerToken":"LOGGER_TOKEN",
"gridEngine":"JSAGA", "vo":"vo.plgrid.pl", "lfcHost":"lfc.grid.cyf-kr.edu.pl",
"dpmList":["dpm.cyf-kr.edu.pl","se.reef.man.poznan.pl","se.polgrid.pl"],
"replicaDpmList":[]}, "createReplica":false, "LFCOutputDirectory":"/grid/vo.plgrid.pl
```

Note that in the **Results** window the HF SCF energy is displayed, therefore to read the MP2 energy log file needs to be inspected or (more conveniently) the requested value can be found in the `parsed_data.xml` file:

InSilicoLab - Portal

file:///tmp/p...ed\_data-1.xml

file:///tmp/parsed\_data-1.xml

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<job>
  <computation output_filename="niedoida.log" parse_successful="true">
    <parameters charge="0" multiplicity="1" normal_termination="true" number_of_atoms="10"/>
    <geometry_optimization stable_geometry_found="true">
      <optimization_step number="0">
        <geometry symmetry="C2h" units="Angstroms">
          C 0.601615759066 1.750961305 0.0 C 0.601615759066 0.410956836744 0.0 C -0.601615759066 -0.410956836744 0.0
          C -0.601615759066 -1.750961305 0.0 H 0.325521871565 -2.31855207836 0.0 H -1.52375339202 -2.32251107375 0.0 H
          1.52375339202 2.32251107375 0.0 H -0.325521871565 2.31855207836 0.0 H 1.55104038286 -0.124334951661 0.0 H
          -1.55104038286 0.124334951661 0.0
        </geometry>
        <SCF energy="-4216.02356033" units="eV"/>
        <S2 value="-"/>
        <MP2 energy="-4231.84870698" units="eV"/>
      </optimization_step>
    </geometry_optimization>
  </computation>
  <orbitals computation="restricted" units="eV">
    <orbital energy="-11.24653889" homo="false" number="1"/>
    <orbital energy="-11.24573883" homo="false" number="2"/>
  </orbitals>
</job>
```

## dressed-TD DFT calculations example

Selection of the program and entering the coordinates of butadiene molecule can be done as in the former example.

Welcome | New experiment | butadiene, LT-M... |

Your proxy expires in 10 hours, MyProxy is enabled.

---

Describe your experiment

**Short name** butadiene, dressed TD DFT

**Description** butadiene molecule, dressed TD DFT calculations

---

Input Data Specification

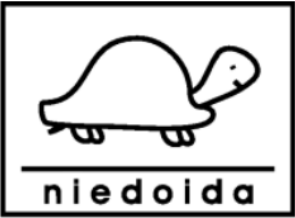
Niedoida Load prepared input Browse... No file selected.

**Title** butadiene, dressed TD DFT

**Basis Set** 6-311++g\*\*

**Other Parameters**

- type: pruned
- accuracy: 1e-5
- max\_no\_radial\_points: 50
- max\_no\_angular\_points: 500



Niedoida 0.5.3

---

Geometry Specification

Browse... No file selected.

LFC location

Molecular Formula

Advanced search...

Add New Molecule

1. butadiene.xyz

Necessary data to be entered as **Input Data Specification** are the **Basis Set** (here 6-311++G\*\*) and **Other Parameters**:

units:

length: angstrom  
energy: hartree

theory: pbe0

td:

type: tda  
multiplicity: 1  
no\_states: 8  
no\_roots: 16  
no\_iterations: 100  
diagonalization\_threshold: 0.00001  
max\_no\_dressing\_iterations: 100  
dress\_delta\_energy: 0.3  
max\_no\_davidson\_dressing\_iterations: 30  
dress\_davidson\_threshold: 0.0001  
dress\_state: 5

integration\_params:

cache\_size: 64

grid:

type: pruned  
accuracy: 1e-5  
max\_no\_radial\_points: 50  
max\_no\_angular\_points: 500

PBE0 functional has been chosen for TD DFT calculations employing Tamm-Dancoff approximation. Eight excited states have been requested. Corrections from double excitations should be calculated for the 5th state using energy window of 0.3 hartree.

The XML file available after job completion contains the uncorrected TDA results:

```

-<Electronic_transitions energy_units="cm^-1" number_of_ET="8">
  <ET energy="48860.1487445" number="1" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="50643.0330996" number="2" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="50936.7691252" number="3" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="51923.4040211" number="4" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="55366.1199073" number="5" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="58371.9825761" number="6" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="58450.3328135" number="7" osc_strength="0.0" symmetry="Singlet"/>
  <ET energy="58605.231402" number="8" osc_strength="0.0" symmetry="Singlet"/>
</Electronic_transitions>
</computation>

```

To access the results of the "dressed" calculations output file (niedoida.log) needs to be checked.

Results of the uncorrected TDA calculations are displayed after the "end restricted tdkc tda" line:

```

niedoida.log ✕
allocated dynamic memory: 418.669 MB

end restricted tdkc tda
wall clock time: 00h07m10.090s
cpu time: 00h07m9.868s
allocated dynamic memory: 255.844 MB

no_states_converged: 8
energies: 8, 5
      1      2      3      4      5
0.22262325 0.23074667 0.23208503 0.23658047 0.25226664
      6      7      8
0.26596236 0.26631935 0.26702512
energies (eV): 8, 5
      1      2      3      4      5
6.05788689 6.27893649 6.31535511 6.43768213 6.86452449
      6      7      8
7.23720405 7.24691834 7.26612312

```

Corrected energies are available after the last "end restricted tdkc dressed tda" statement:

```

niedoida.log ✕
allocated dynamic memory: 436.556 MB

end restricted tdkc dressed tda
wall clock time: 00h00m47.573s
cpu time: 00h00m47.559s
allocated dynamic memory: 273.939 MB

no_states_converged: 8
energies: 8, 5
      1      2      3      4      5
0.22262325 0.23074667 0.23208505 0.23527884 0.23658046
      6      7      8
0.26596235 0.26631935 0.26702512
energies (eV): 8, 5
      1      2      3      4      5
6.05788682 6.27893638 6.31535543 6.40226306 6.43768180
      6      7      8
7.23720379 7.24691833 7.26612330

```

In the above example corrections arising from double excited states lowered the energy of the 5th state from 6.865 to 6.438 eV.

## Cubegen

This experiment is used to generate Gaussian cube files.

Optional description of the experiment consists of its **Short name** and the **Description** used to annotate the experiment.

To generate the cube file it is necessary to provide (in the section **Checkpoint Specification**) a checkpoint file containing the data saved from a previous Gaussian run. The file can be downloaded by the user (Add file location list) or retrieved from a LFC storage (LFC location - open the browser and allows the user to select the file). The file stored in LFC directory can be generated in an earlier general QC experiment (option requesting saving checkpoint files had to be used).

**Input Data Specification** describes the type and parameters of the cube file:

- **Cube filename**

- **Cube type** - on clicking the wand icon a list of possible choices is presented, e.g. electron density, molecular orbital, spin density, potential, gradient, etc.; for detailed description of cube types the User is referred to the Gaussian [cubegen manual](#)
- **Include header** option (yes or no)
- **Cube specification** - controls the size/density and units of the cube grid, as described in the [cubegen manual](#)

The experiment is started on clicking **Run** button.