CBE 60475/40475 Molecular Modeling and Simulation

E. Marin-Rimoldi, J.K. Shah E. J. Maginn

Department of Chemical and Biomolecular Engineering
University of Notre Dame
Notre Dame, IN 46556 USA

Cassandra Tutorial: Monte Carlo Simulation of Liquids

Simulation of Liquids in Cassandra

Goal: Set up a MC simulation of Pentane from scratch.

- United Atom Model
- TraPPE force field (Transferable Potentials for Phase Equilibria)
- Features:
  - fixed bond lengths
  - harmonic angle potentials

$$\nu_\theta = K_\theta(\theta - \theta_0)^2 \tag{1}$$

  - OPLS functional form for dihedral potentials

$$\nu_\phi = \frac{1}{2}K_1[1 + cos(\phi)] + \frac{1}{2}K_2[1 - cos(2\phi)] \tag{2}$$

$$\frac{1}{2}K_3[1 + cos(3\phi)] + \frac{1}{2}K_4[1 - cos(4\phi)] \tag{3}$$

Simulation of Liquids in Cassandra

Motivation: Manual generation of MCF files can be error-prone for large molecules. Automation tools: mcfgen.py. Steps:

- Obtain a PDB (protein data bank) file of the molecule
- Run: mcfgen.py -fffile molecule.pdb OR mcfgen.py -fffile molecule.cml
- Fill out the generated molecule.ff with the parameters found in literature
- Run: mcfgen.py -cassandra molecule.pdb. The script will assume molecule.ff is located in the current directory.
- Check the generated mcf file (molecule.mcf)
- Generate fragment library by running library_setup.py: library_setup.py $path/cassandra.exe inputfile.inp molecule1.pdb molecule2.pdb ...

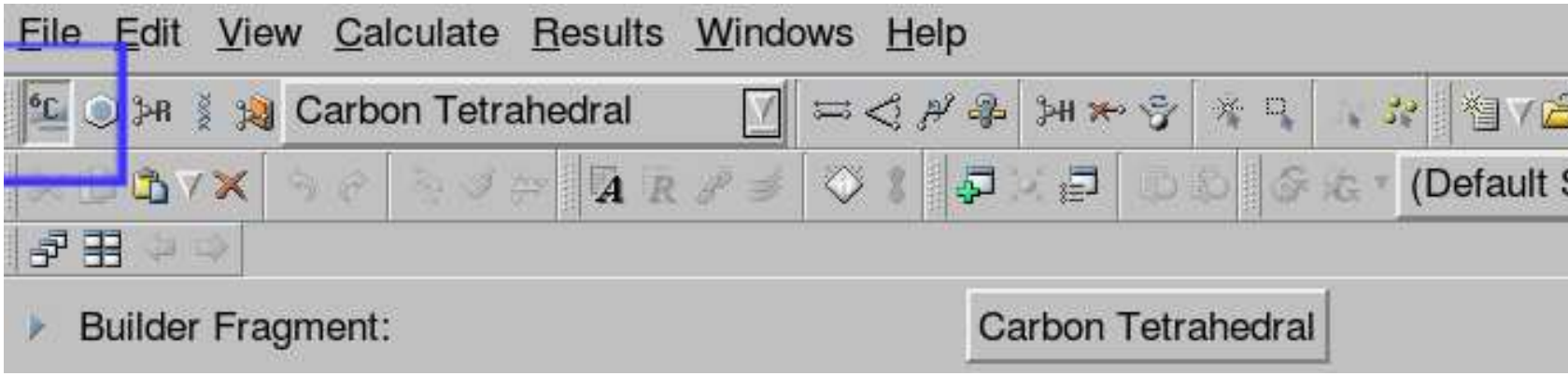
Obtaining a PDB file of the molecule

Options:

- Obtain from databases (i.e. www.rcsb.org).
- Generate using drawing tools (i.e. Gaussview, Avogadro).

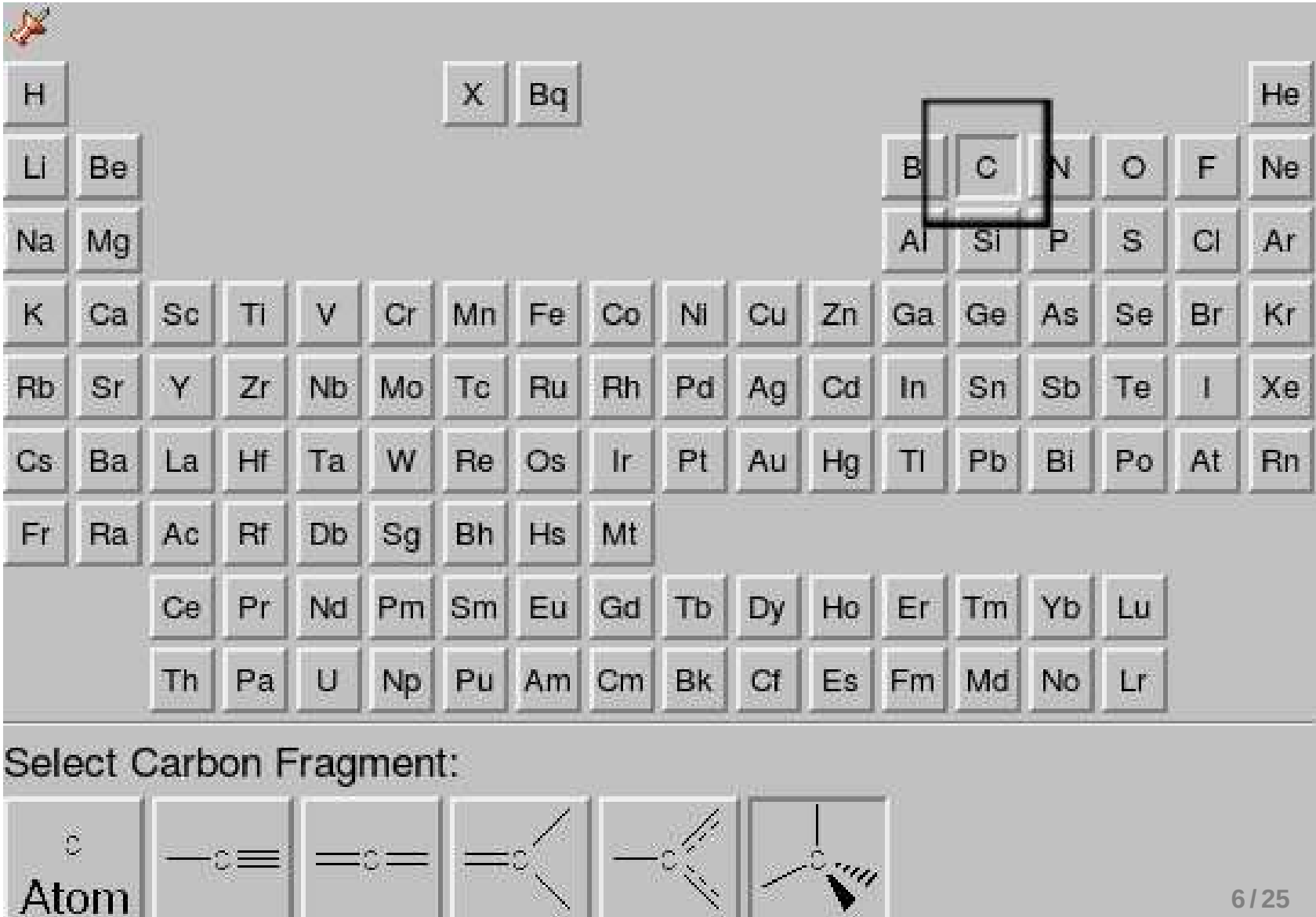
Note that this script has been tested only with PDB files generated by Gaussview 5.08 or Avogadro 1.1.1.

Using Gaussview to generate a PDB file

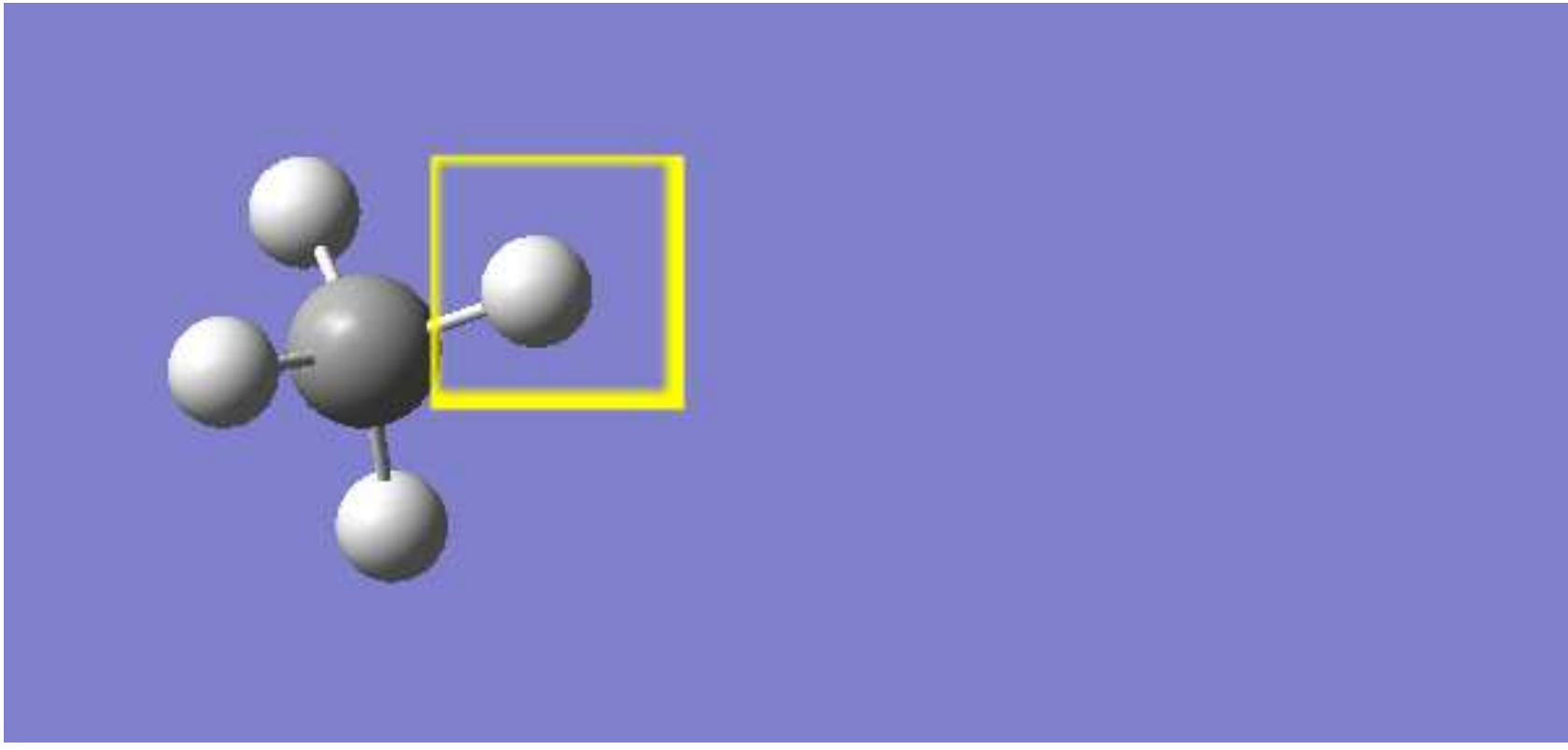


Click on the button ”C” to display the periodic table

Using Gaussview to generate a PDB file

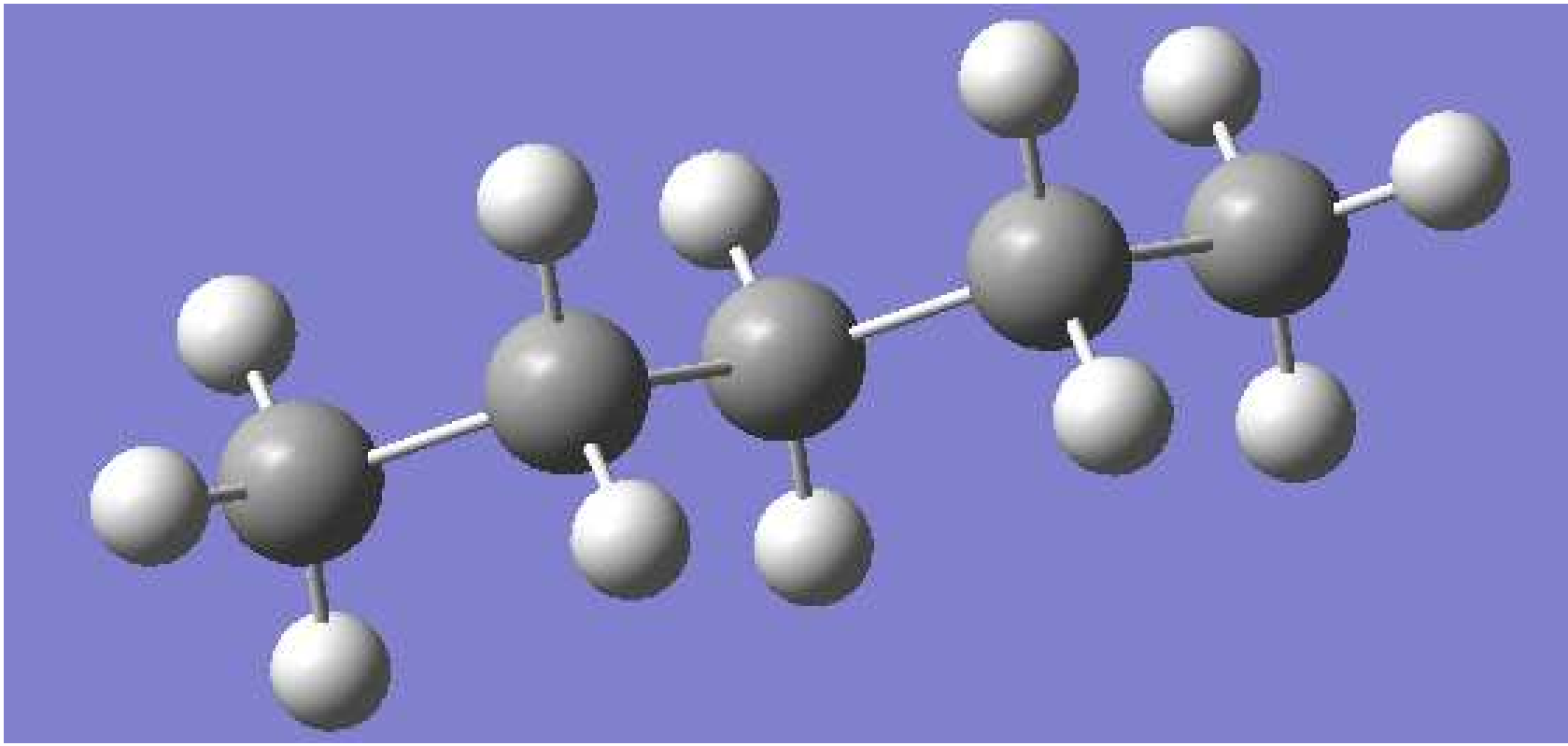


Using Gaussview to generate a PDB file

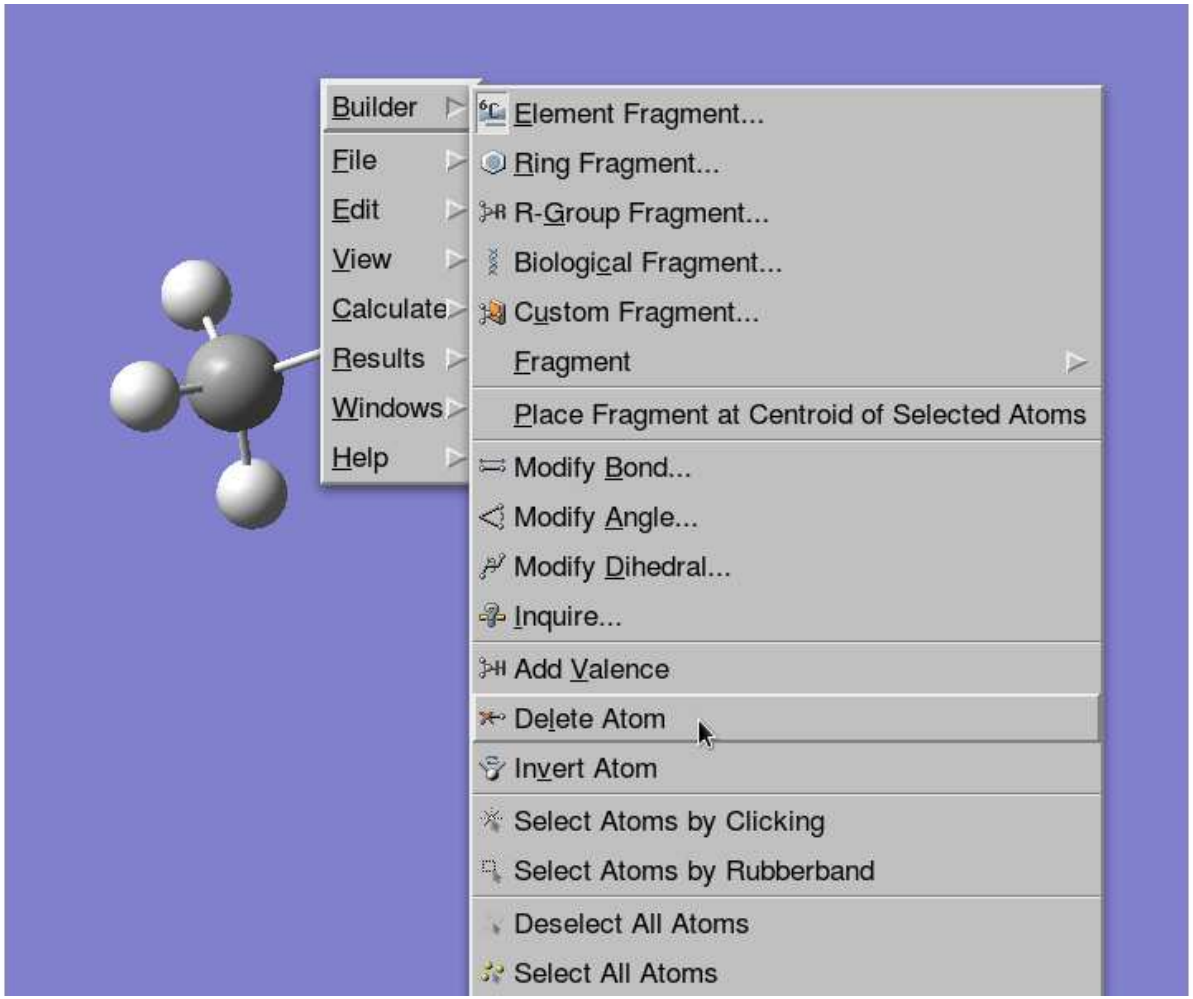


Click on the purple workplace to insert the first CH_4 fragment. To increase the chain length, click on a Hydrogen attached to the Carbon.

Using Gaussview to generate a PDB file

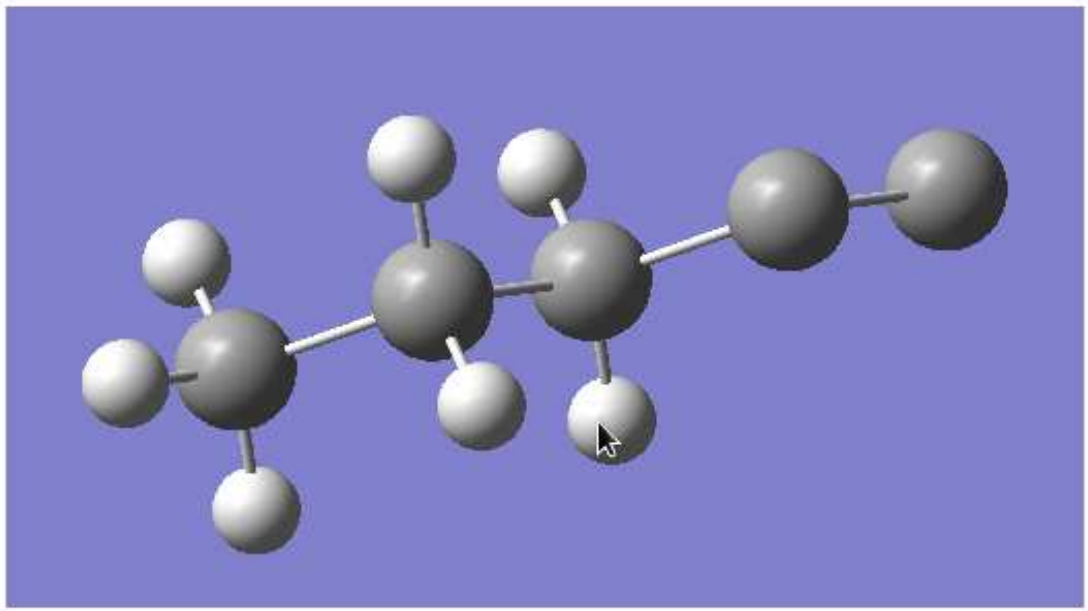


Using Gaussview to generate a PDB file



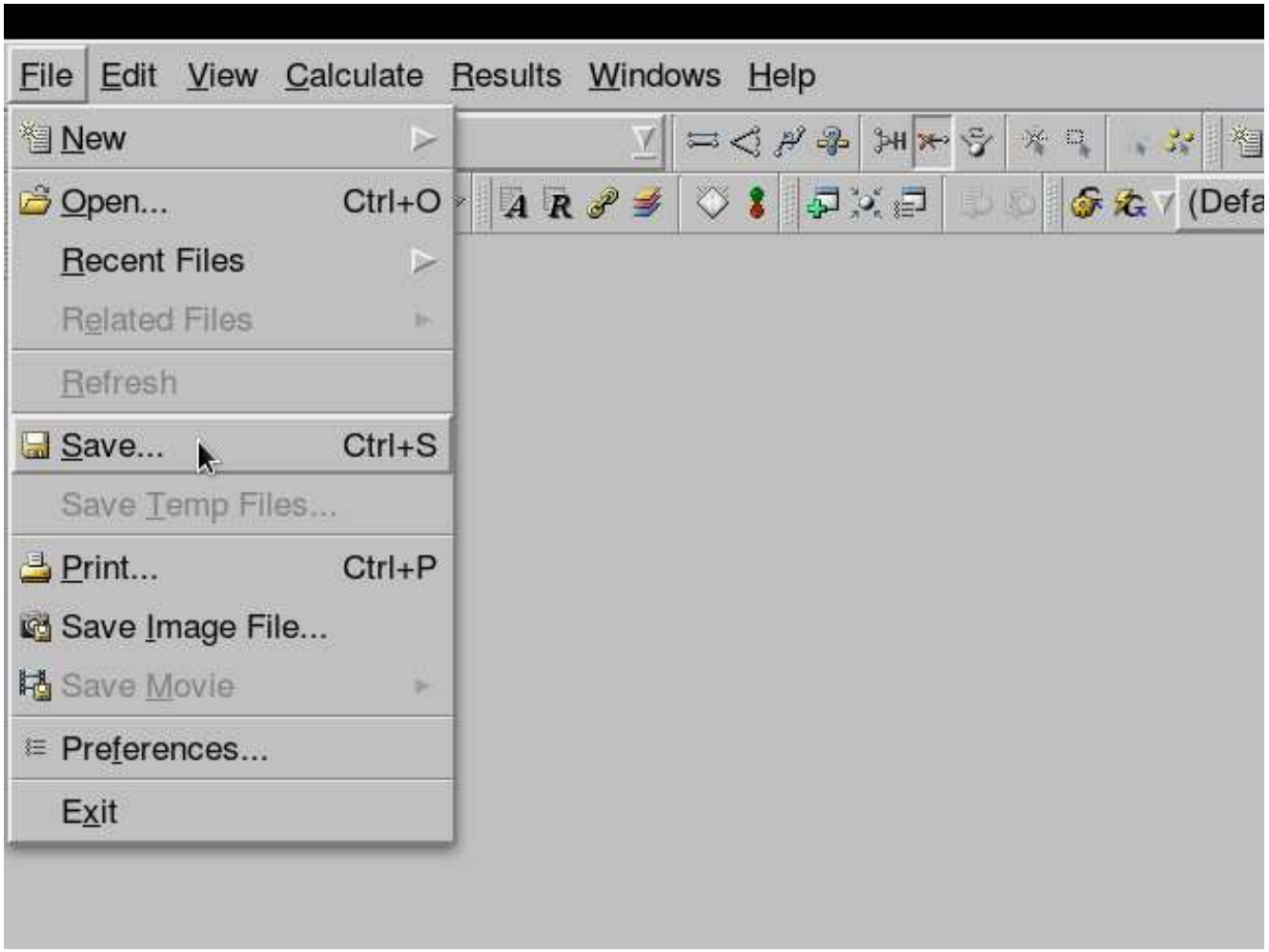
Right Click - Builder - Delete Atoms

Using Gaussview to generate a PDB file



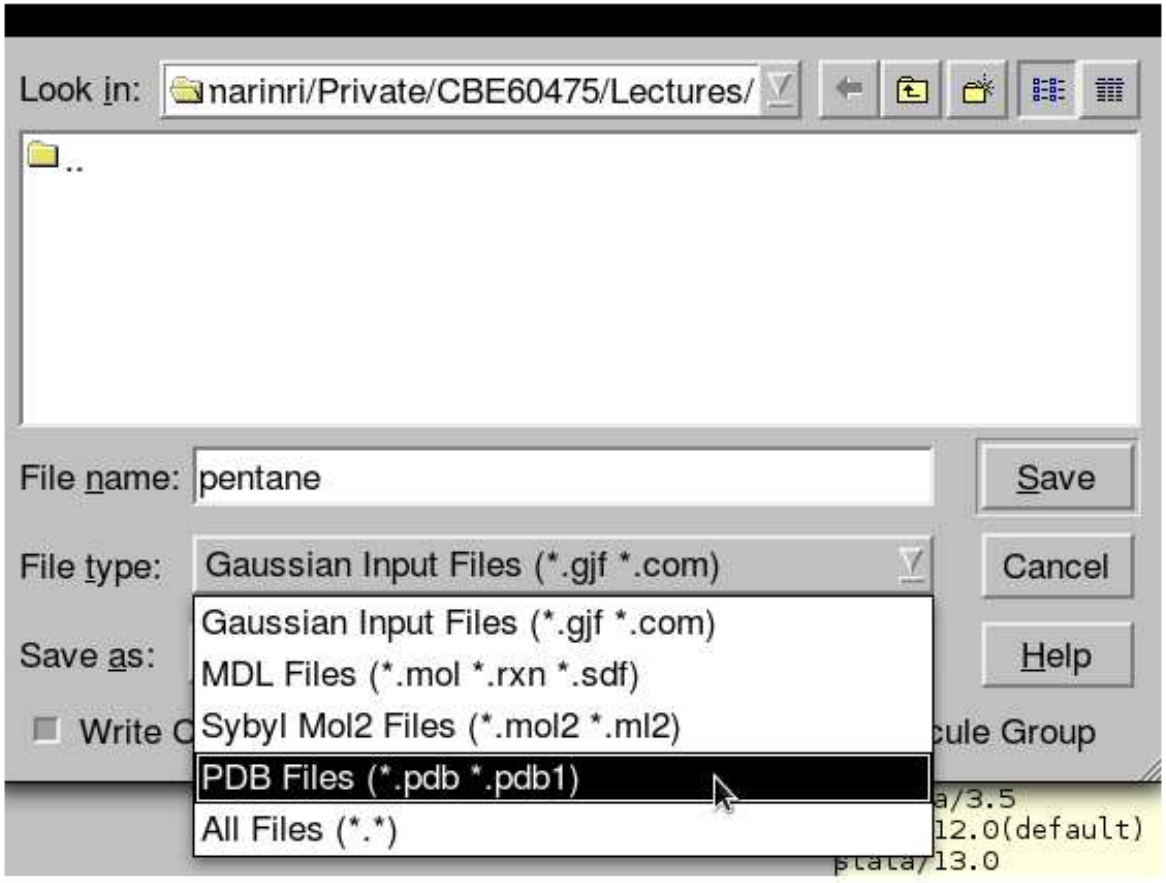
Click on Hydrogen atoms to delete them

Using Gaussview to generate a PDB file



Go back to main menu. File - Save

Using Gaussview to generate a PDB file



Type the name of the file. Select PDB file from the bottom menu.

Using Gaussview to generate a PDB file

```
REMARK    1 File created by GaussView 5.0.8
HETATM    1  C           0       -4.277    0.355    0.000                       C
HETATM    2  C           0       -3.763    1.081    1.257                       C
HETATM    3  C           0       -2.223    1.078    1.259                       C
HETATM    4  C           0       -1.710    1.805    2.516                       C
HETATM    5  C           0       -0.170    1.802    2.517                       C
END
CONECT    1    2
CONECT    2    1    3
CONECT    3    2    4
CONECT    4    3    5
CONECT    5    4
```

Close Gaussview. Back in the terminal, Open the PDB File using your favorite text editor.

Using Gaussview to generate a PDB file

```
REMARK    1 File created by GaussView 5.0.8
HETATM    1  C           0       -4.277    0.355    0.000                       C CH3
HETATM    2  C           0       -3.763    1.081    1.257                       C CH2
HETATM    3  C           0       -2.223    1.078    1.259                       C CH2
HETATM    4  C           0       -1.710    1.805    2.516                       C CH2
HETATM    5  C           0       -0.170    1.802    2.517                       C CH3
END
CONECT    1    2
CONECT    2    1    3
CONECT    3    2    4
CONECT    4    3    5
CONECT    5    4
```

Append a column containing the atom types.

Using Avogadro to generate a CML file

Avogadro can be used to generate CML files. The procedure is analogous to the one previously presented for Gaussview.

Using Avogadro to generate a CML file

```
<molecule>
 <atomArray>
  <atom id="a1" elementType="C" x3="-0.367789" y3="-0.161907" z3="0.206019"/>
  <atom id="a2" elementType="C" x3="-1.354811" y3="-1.178938" z3="-0.372241"/>
  <atom id="a3" elementType="C" x3="-2.735586" y3="-0.597632" z3="-0.678858"/>
  <atom id="a4" elementType="C" x3="-3.435276" y3="0.007943" z3="0.526735"/>
  <atom id="a5" elementType="C" x3="1.027694" y3="-0.340782" z3="-0.372648"/>
 </atomArray>
 <bondArray>
  <bond atomRefs2="a1 a2" order="1"/>
  <bond atomRefs2="a3 a4" order="1"/>
  <bond atomRefs2="a3 a2" order="1"/>
  <bond atomRefs2="a5 a1" order="1"/>
 </bondArray>
</molecule>
```

Pentane united atom CML file generated using Avogadro

Using Avogadro to generate a CML file

```
<molecule>
 <atomArray>
  <atom id="a1" elementType="C" x3="-0.367789" y3="-0.161907" z3="0.206019"/> "CH2"
  <atom id="a2" elementType="C" x3="-1.354811" y3="-1.178938" z3="-0.372241"/> "CH2"
  <atom id="a3" elementType="C" x3="-2.735586" y3="-0.597632" z3="-0.678858"/> "CH2"
  <atom id="a4" elementType="C" x3="-3.435276" y3="0.007943" z3="0.526735"/> "CH3"
  <atom id="a5" elementType="C" x3="1.027694" y3="-0.340782" z3="-0.372648"/> "CH3"
 </atomArray>
 <bondArray>
  <bond atomRefs2="a1 a2" order="1"/>
  <bond atomRefs2="a3 a4" order="1"/>
  <bond atomRefs2="a3 a2" order="1"/>
  <bond atomRefs2="a5 a1" order="1"/>
 </bondArray>
</molecule>
```

Modified pentane united atom CML file. Note the atom type is appended as a last column between quotation marks.

- Run the following command:

```
> module load python
> python mcfgen.py -fffile pentane.pdb
```

- Fill out the newly created .ff file with literature values (see next slide).

```
atomtypes
2

begin atom-atomtype
1 CH3
2 CH2
3 CH2
4 CH2
5 CH3
end atom-atomtype

dihedraltype OPLS
```

The first three sections of the FF file are displayed above. Do not modify these.

```
bonded
CH2 CH2
Length 1.54
Constant fixed

angles
CH3 CH2 CH2
Angle 114.0
Constant 31250.0
```

The force field parameters for non-bonded (not shown), bonded, angle, dihedral (not shown) and coulombic interactions (not shown) must be entered next to the corresponding keyword. For example, the angle type CH3 CH2 CH2 has an angle of 114.0. This value must be placed next to the ‘Angle’ keyword.

For more examples of filled FF files, please refer to the examples contained in the /Scripts/ directory.

- Using the filled .ff file, run:

```
> python mcfgen.py -cassandra pentane.pdb
```

- Check the file pentane.mcf for any possible errors.
- Substitute pentane.pdb by pentane.cml if CML files are being used.
- Note that the FF file must be in the same current directory as the PDB or CML file.

- Prepare an input file, specifying the ensemble to be used, box lengths, probabilities, etc...
- To set up the simulation, run:

```
> python library_setup.py inputfile.inp
species1.pdb species2.pdb
```

Note that the script library_setup.py will assume that the PDB files, MCF files and input file are in the same current directory.

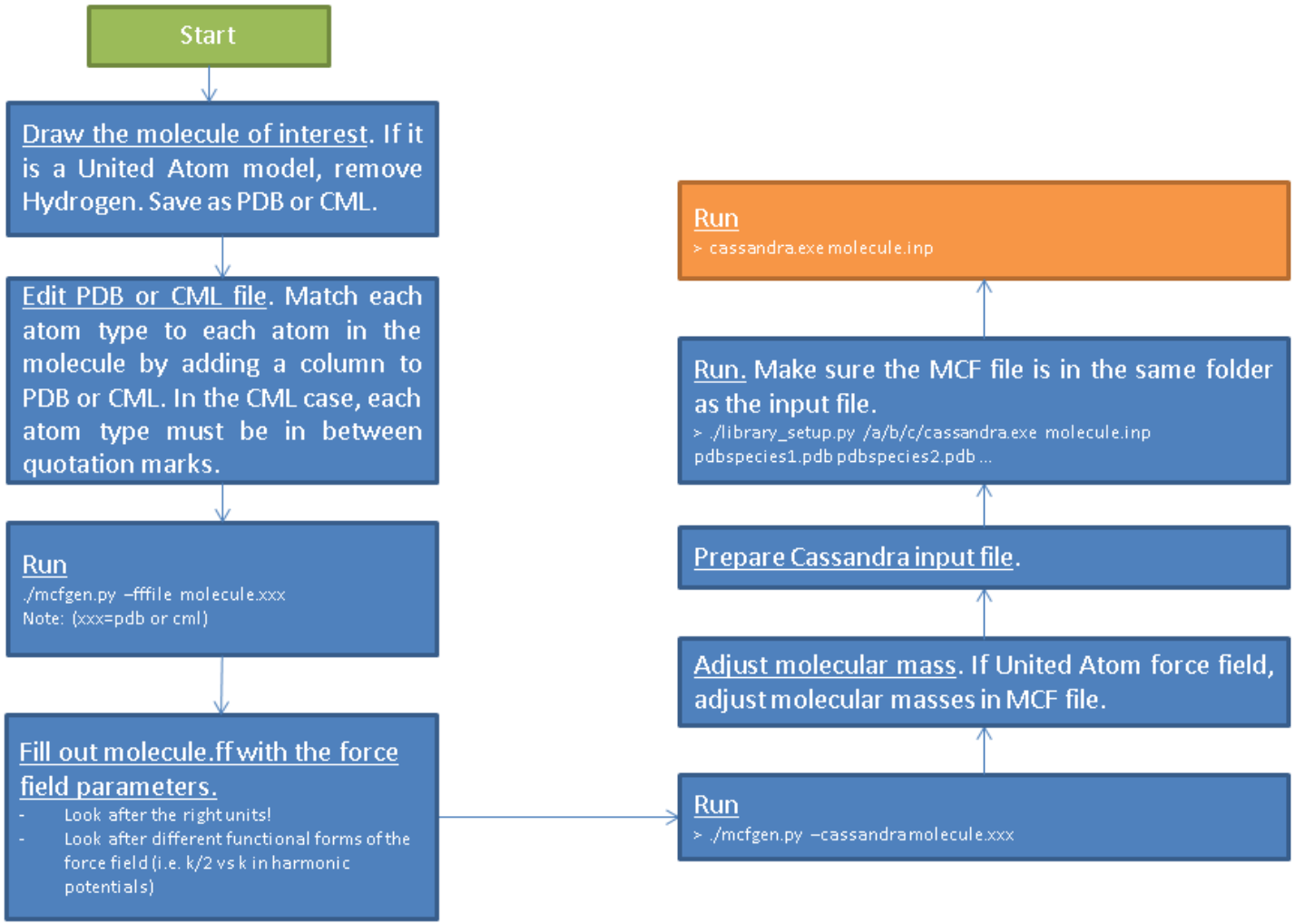
- To run the simulation:

```
> $path/cassandra.exe inputfile.inp
```

Things to take care:

- Units and force field functional forms must be consistent with those required by Cassandra. See documentation for further information.
- Manually change molecular weights for United Atom models in the MCF file (i.e. CH_3 has a weight of 15 g/mol instead of 12 g/mol)
- Some software generate PDB/CML files with a format that this script does not recognize. Please check the provided examples to see how a PDB should look like.
- This script does not recognize byciclic molecules (e.g. decalin).

Flowsheet


Start
Draw the molecule of interest. If it is a United Atom model, remove Hydrogen. Save as PDB or CML.
Edit PDB or CML file. Match each atom type to each atom in the molecule by adding a column to PDB or CML. In the CML case, each atom type must be in between quotation marks.
Run
./mcfgen.py –fffile molecule.xxx
Note: (xxx=pdb or cml)
Fill out molecule.ff with the force field parameters.
- Look after the right units!
- Look after different functional forms of the force field (i.e. k/2 vs k in harmonic potentials)
Run
> ./mcfgen.py –cassandra molecule.xxx
Adjust molecular mass. If United Atom force field, adjust molecular masses in MCF file.
Prepare Cassandra input file.
Run. Make sure the MCF file is in the same folder as the input file.
> ./library_setup.py /a/b/c/cassandra.exe molecule.inp pdbspecies1.pdb pdbspecies2.pdb ...
Run
> cassandra.exe molecule.inp