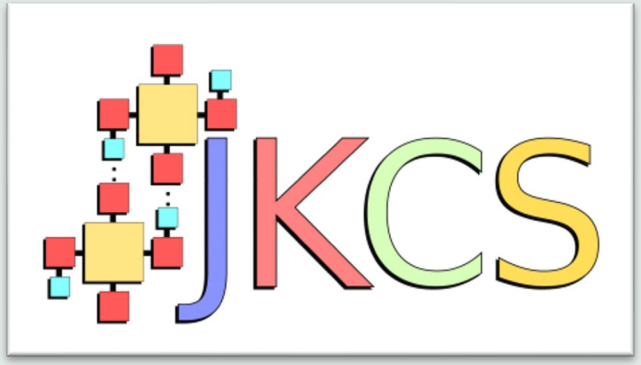
JK

WORKSHOP

Part 3

JKCS



- Configurational Sampling of Molecular Clusters

  -> this is how it all started

- JKCS is basically just collection of my automated scripts for this workflow

- When mastered, it can make your life orders of magnitude easier.

- see the JKCS commands

```
$user: JK + <TAB>
JKCS0_copy          JKQC                JKfitxtb            JKlog2com           JKorcaDLPNO         JKsend8
JKCS1_prepare       JKTS                JKfor               JKlog2xyz           JKout2xyz           JKsubg16
JKCS2_explore       JKacdc              JKformation         JKmov2xyz           JKpython            JKtakebest
JKCS3_run           JKattachstr.py      JKforsend           JKmovetoorigin.py   JKqdelall           JKtar
JKCS4_collect       JKavg               JKg16.sh            JKname              JKqmuch             JKxyz2com
JKCS5_filter        JKcheck             JKgaussstat         JKnms               JKremovefiles       JKxyz2inp
JKCS8_clean         JKchmod             JKgoodvibes         JKoptimizer         JKsacct             JKxyz2minp
JKML                JKex                JKjxyz2xyz          JKorca              JKsend
```

Global minima
ωB97X-D/6-31++G**
Comparison to
[Myllys *et al.* (2018)]

LECTURE_3/TASK_00/submit.sh

Our first submit job

Run submit.sh:

```
sh submit.sh
```

Make submission of submit.sh working.

```
sbatch submit.sh
```

Make JKsend working (use: which JKsend)

```
sbatch JKsend echo Hi
sbatch -p qtest --time=10:00 JKsend echo Hi
```

Modify ~/.JKCSusersetup.txt

```
34 SBATCH_PREFIX="project78566 " #group sbatch req.
```

LECTURE_3/TASK_01/

How to JKCS1_prepare

Examine JKCS0_copy

```
JKCS0_copy -help
```

TODO: Prepare input files for water clusters (use: aq)

TODO: Study input.txt and modify it so $(w)_{1\text{-}5}$ systems can be studied

TODO: When done, run JKCS1_prepare and check whether you got it right.

TODO: In the same folder, prepare $(sa)_{1,2,4}(w)_1$ clusters (yes, overwrite input.txt w JKCS0)

TODO: In the same folder, prepare $(\text{bisulfate}^-)_1(w)_1$ dimer (check the subfolder. Why?)

TODO: From the same folder, check the content of sa.xyz.

TODO: In STR folder, check the org.xyz and run <PATH TO ABCluster>/topgen org.xyz

TODO: In the TASK_01 folder, prepare clusters: $(org)_{1\text{-}6}$.

TODO: In the same folder, prepare cluster containing $(sa)_1(base)_3$, where base=(TMA|AM|DMA)

LECTURE_3/TASK_02/

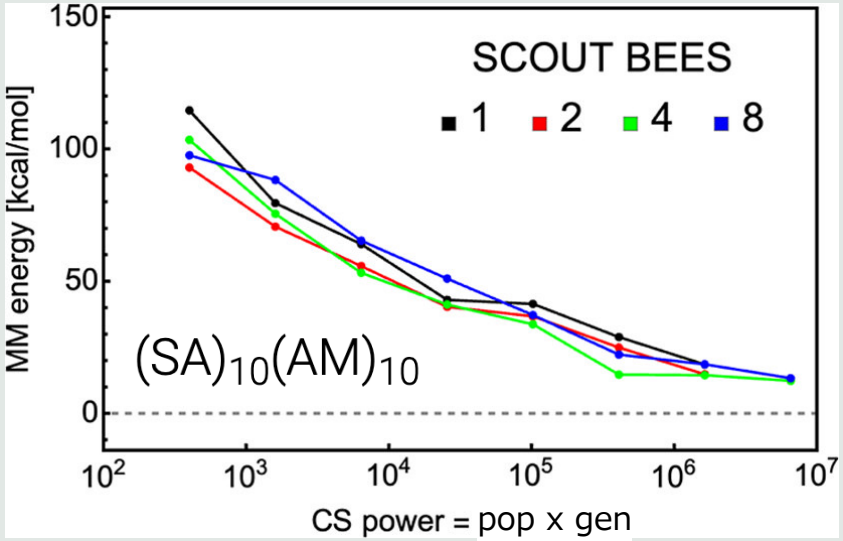
How to JKCS2_explore

Let us explore $(w)_{1\text{-}2}$

```
JKCS0_copy w
JKCS1_prepare
#this will work only for those with linked ABCluster
JKCS2_explore -loc -gen 3 -pop 3 -lm 1
```

In general you want more thorough configurational sampling

(we will see proper paramaterization later)



LECTURE_3/TASK_02/

How to JKCS2_explore

Let us discuss other parameters

```
JKCS2_explore -help
```

-repeat <int> = you can reduce the length of CS but run it more times

-sc <int> = scout bees = max structure lifetime

-pickle = if you use large LM, it is worth saving the structures as pickle

-exploded = if will pickle and remove structures with large radius

-box <float> = defines the simulation box size for scout bees.

-helpxtb = WE WILL EXAMINE THIS LATER

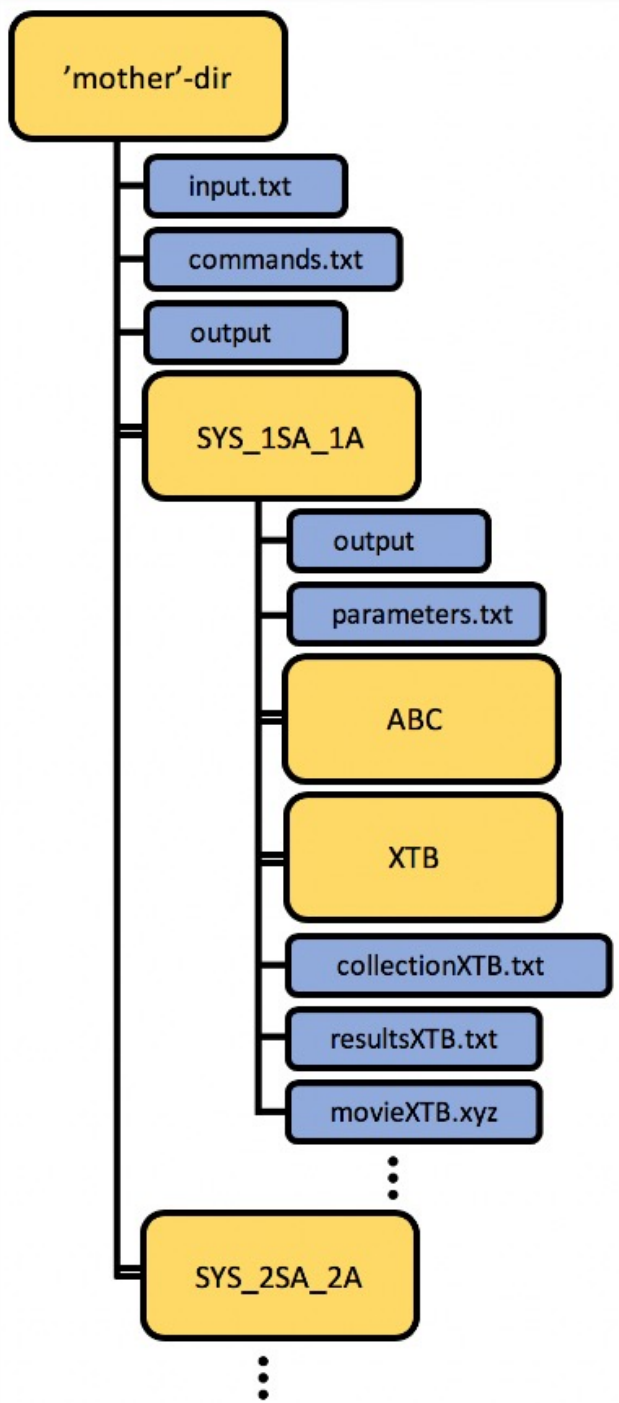
LECTURE_3/TASK_03/

Folder structure

distinguish – mother folder
(due to GE, read: parent folder)

– cluster/system folder

TODO: See the structure of SYS_1SA_1A/ABC folder



LECTURE_3/TASK_03/SYS_1SA_1A

How to JKCS3_run

Any communication with other programs than ABCluster is performer via JKCS3_run

```
JKCS3_run -help
```

All command can run from parent or from system folder.

You can run specific system from parent folder using:

```
JKCS3_run -p XTB -of ABC -nf XTBsp -m "--gfn 1 --sp" -loc SYS_1SA_1A
```

TODO: Run SP XTB calculation only for the 1SA_1A cluster.

TODO: Which command did I use to prepare the ABC folders?

TODO: How many structures are in XTBsp subfolder? Collect them (use: JKQC)

TODO: Modify program_XTB so it returns all XTB files.

LECTURE_3/TASK_03/SYS_1SA_1A

How to JKCS4_collect

JKCS4_collect is just (submittable) envelope for JKQC

```
JKCS4_collect -help
```

TODO: In the SYS_1SA_1A folder, try this:

```
ls
JKCS4_collect XTBsp -loc -oc
ls
JKCS4_collect XTBsp -loc
ls
JKCS4_collect XTBsp -par small -time 10:00 -slurm "-begin=now+60" #sec
squeue -u $USER
JKcheck
```

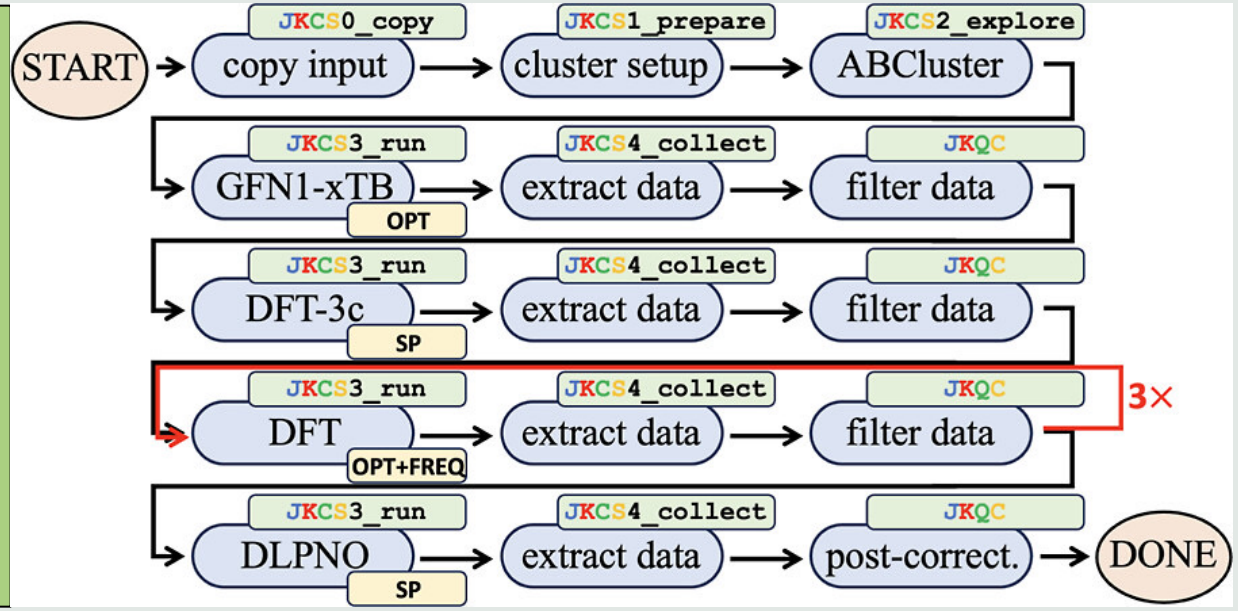
(Note: to collect ORCA calculations you need the –orca argument for now)

LECTURE_3/TASK_03/SYS_1SA_1A

The configuration sampling workflow

Do you see the idea behind workflow:

```
JKCS0_copy …
JKCS1_prepare …
JKCS2_explore …
JKCS3_run …
JKCS4_collect …
JKQC …
JKCS3_run …
JKCS4_collect …
JKQC …
…
```



TODO: To practise, select the lowest 1SA_1A el. energy conformer and calculate its Gibbs free energy using XTB. (use: --ohess)

LECTURE_3/TASK_04/commands.txt

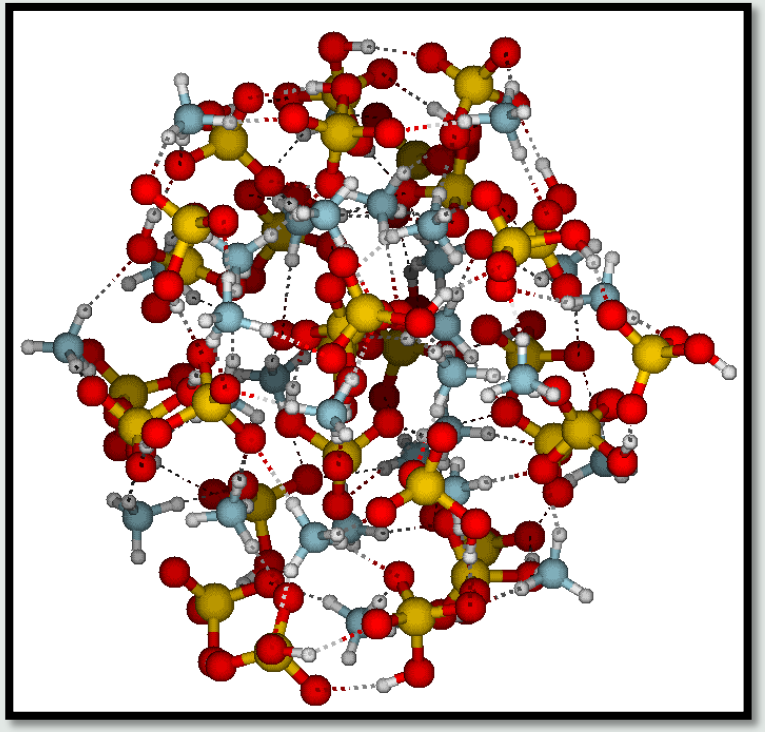
Advanced CS workflow

TODO: Understand all the steps/arguments of this CS workflow.

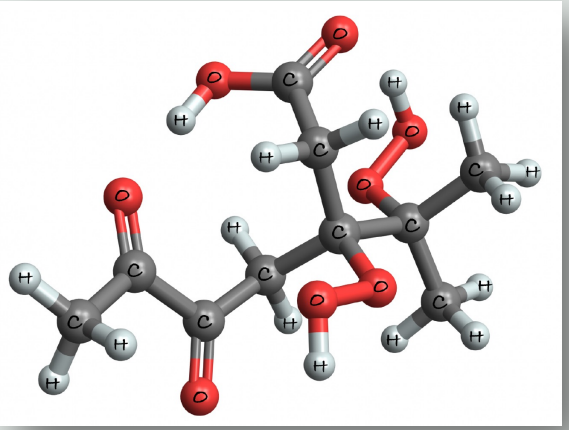
LECTURE_3/TASK_05/commands.txt

CS workflow for large clusters

TODO: Understand all the steps/arguments of this CS workflow.



LECTURE_3/TASK_06/org.xyz



CS of flexible organic molecules

TODO: Visualize org.xyz

TODO: Prepare a command that would submit the org.xyz for XTB optimization using GFN2-xTB on 2 CPUs.

Configurational sampling using CREST:

```
JKCS3_run -p CREST -rf org.xyz -nf CREST_sampling
          -m "--gfn 2 --noreftopo" -cpu 1 -par small
```

see CS or CS8

LECTURE_3/TASK_07/STR/org.xyz

CS of molecular clusters also containing flexible organic molecules

---

TODO: In STR folder, check the org.xyz and run <PATH TO ABCluster>/topgen org.xyz
(If not possible use the STR/PRE-PREPARED folder)

TODO: Examine org-bonding.xyz

TODO: Prepare input.txt file for sampling 3w (using only XYZ)

```
JKCS0_copy -help
JKCS0_copy -helpxyz
JKCS0_copy XYZaq
```

TODO: Modify input.txt for CS of org_1w_3 (use: org-bonding.xyz)

TODO: Run CS using ABCluster coupled with XTB

```
JKCS2_explore -helpxtb
JKCS2_explore -abcxtb -gen 2 -lm 2 -gfn 2 -loc
```

LECTURE_3/TASK_07/STR/org.xyz

Searching for errors

TODO: Examine the ABC-XTB folder. Visualize one of the resultant molecules.

TODO: JKrecenter the molecules and run again:

```
cd LECTURE_3/TASK_07/STR/
JKrecenter org.xyz
<PATH_TO_ABCluster>/topgen org.xyz
cd ../SYS_1org_3w
rm -r ABC_XTB
JKCS2_explore -abcxtb -gen 2 -lm 2 -gfn 2 -loc
```

TODO: Did it finish correctly now? Most likely not long enough simulation for such a large simulation box. Rerun using:

```
JKCS2_explore -abcxtb -gen 2 -lm 2 -gfn 2 -loc -box 5
```

LECTURE_3/TASK_08/HF/

Searching for errors 2

Somebody complained that their calculations did not finish correctly.

```
cat collectionHF.txt
```

COMPETION: Where the user made a mistake?

Write a correct command and call me.

LECTURE_3/TASK_09/

Hydrogen atom

TODO: What is the theoretically correct electronic energy of hydrogen atom?

TODO: Which of the trial methods is closest to the true theoretical value?

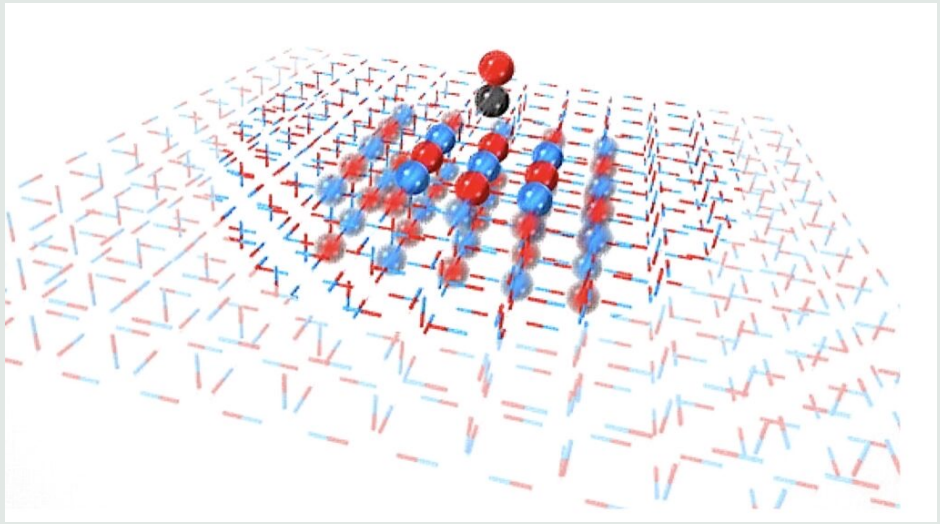
LECTURE_3/TASK_10/

CS of molecular systems on a surface

I think the geom script from ABCluster is perhaps the best for this, but CREST could handle some stuff too.

TODO: Let us examine this folder together.

For surface reaction, go and check e.g. ONIOM or Green's functions.



LECTURE_3/TASK_11/

Build your own workflow inc. ML

---

1) Sample $(w)_{1\text{-}10}$
2) Optimize at XTB
3) For $(w)_{1\text{-}7}$ run a SP DFT calculation with forces
4) Train NN
5) Use NN to pre-optimize all structures
6) Select all structures within X kcal/mol, where X = number of w molecules
7) Optimize structures at DFT + calculate vibrational frequencies
8) Plot IR spectrum for the global free-energy conformers

Discuss some unclear arguments

1) cleaning
2) manager.sh / boss.sh
3) -mf for JKCS3_run
4) -oc -orca for JKCS4_collect
5) JKcheck
6) JKtakebest
7) linking another QC program
8) JKTS
9) something else?