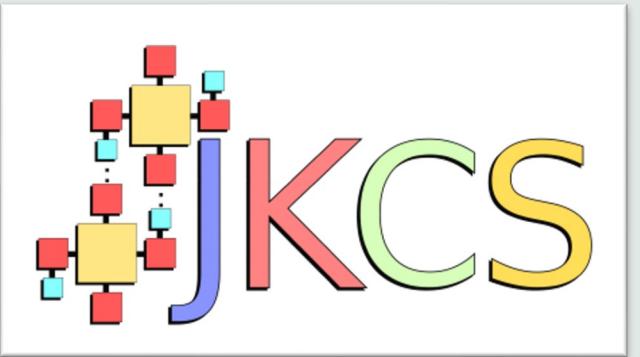




JK  
**WORKSHOP**  
Part 2

# 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	JKcorcaDLPNO	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	JKcorca	JKsend	

Global minima  
 $\omega$ B97X-D/6-31++G\*\*  
Comparison to  
[Myllys et al. (2018)]

BEER CHALLENGE

M

Myllys, Nanna L P

Wed 2/21, 4:35 PM

Kubecka, Jakub



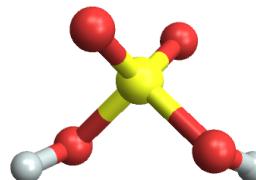
Reply all | v

Inbox

Let's put ABCluster to work! Forget for a while potential energy surface dependence of level of theory, temperature, quasi-harmonic/anharmonic corrections etc. and only focus **Gibbs free energies of guanidine-sulfuric acid clusters at 298.15 K using wB97X-D/6-31++G\*\*.**

In wb97xd folder I sent previously, you have log files for SA(1-4)-GUA(1-4) clusters and desired energies can grep by "Sum of electronic and thermal Free Energies". Your task is simple: **find lower Gibbs free energy structures for those clusters** I offer a beer per every cluster for which you have found at least one better structure than me by end of March!

Good luck :D



0

1

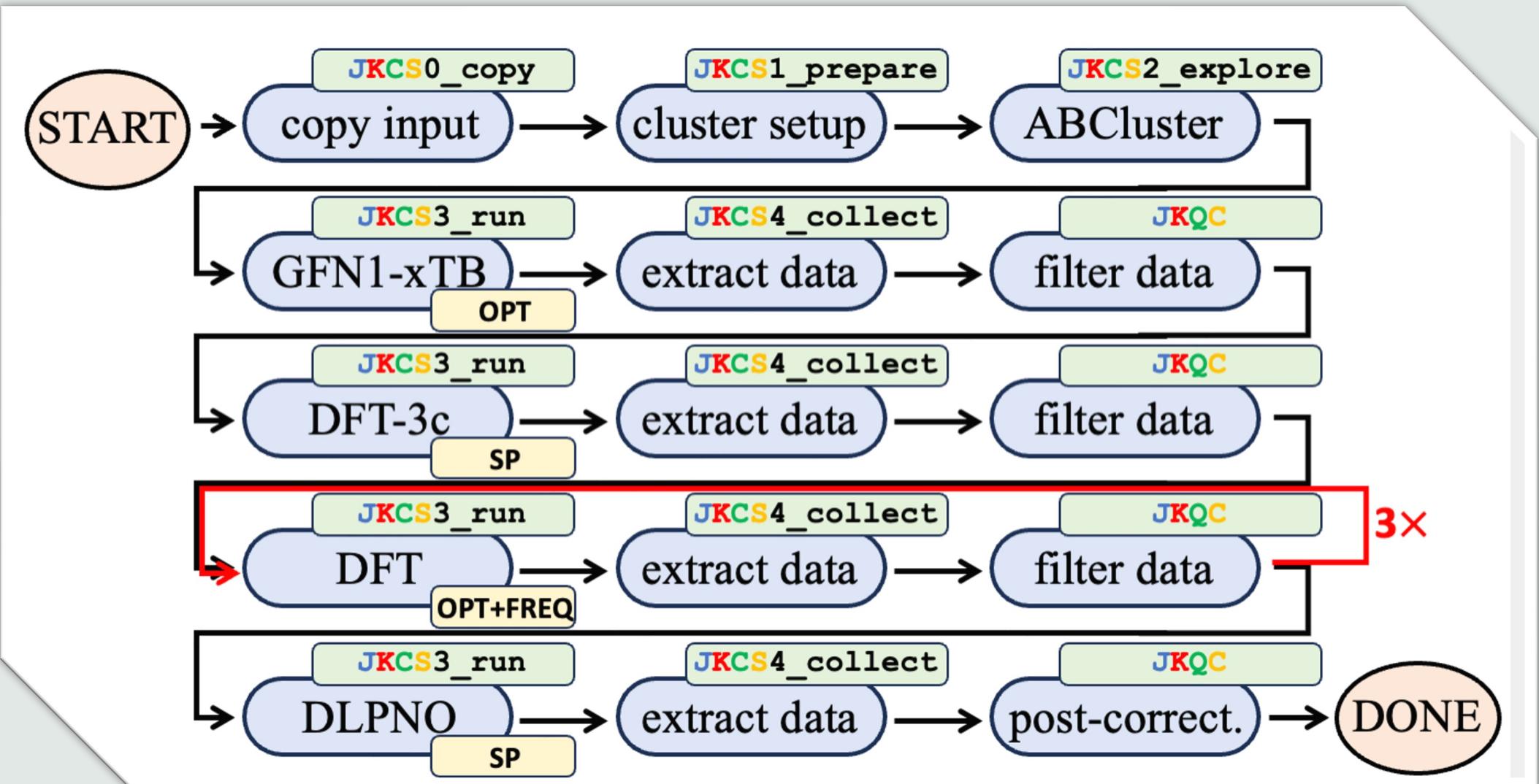
2

3

4

Sulphuric acid's molecules

# THE WORKFLOW





**Let's understand  
how to perform  
this on computer  
cluster**

# 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="--account=project78566 " #group SLURM id.
```

# Useful SLURM ‘words’

---

- sbatch, scancel, squeue, sinfo, showpartitions, showuserlimits, bj, sacct, seff, scontrol

```
squeue --me
```

- -c,-n,-cpus-per-tasks,-time,-e,-o,--mem,--mem-per-cpu,-N,-n  
--tasks-per-node,-t,-array,-dependency,-J,--mail-user,--begin

```
sbatch --test-only JKsend echo hi
```

- Go through this page once:

<https://docs.rc.fas.harvard.edu/kb/convenient-slurm-commands/>

or check this cheat sheet:

<https://www.carc.usc.edu/user-information/user-guides/hpc-basics/slurm-cheatsheet>

# 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  $(w)_{1-5}$  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:  $(\text{org})_{1-6}$ .

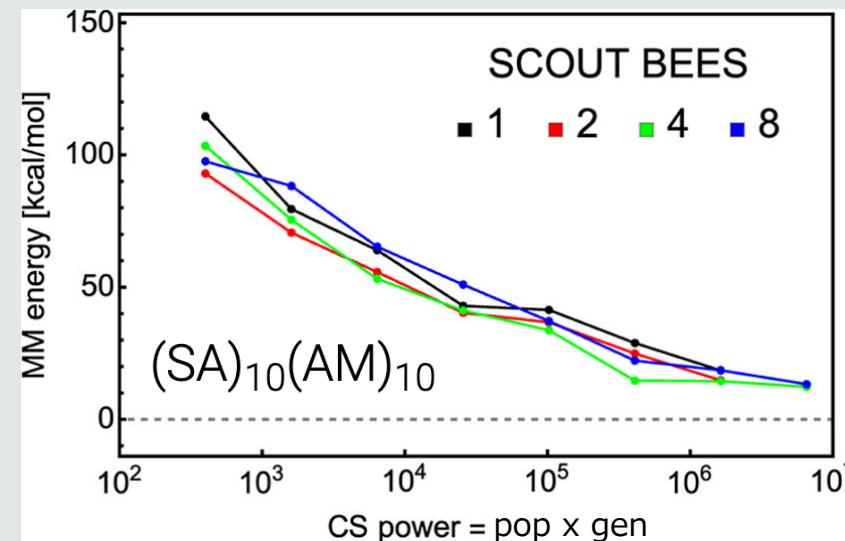
**GAME:** In the same folder, use **JKCS0\_copy sa tma am dma**, and prepare cluster containing  $(sa)_1(\text{base})_3$ , where base=(TMA|AM|DMA)

# How to JKCS2\_explore

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

```
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  
parameterization later)



# How to JKCS2\_explore

---

Let us discuss other parameters

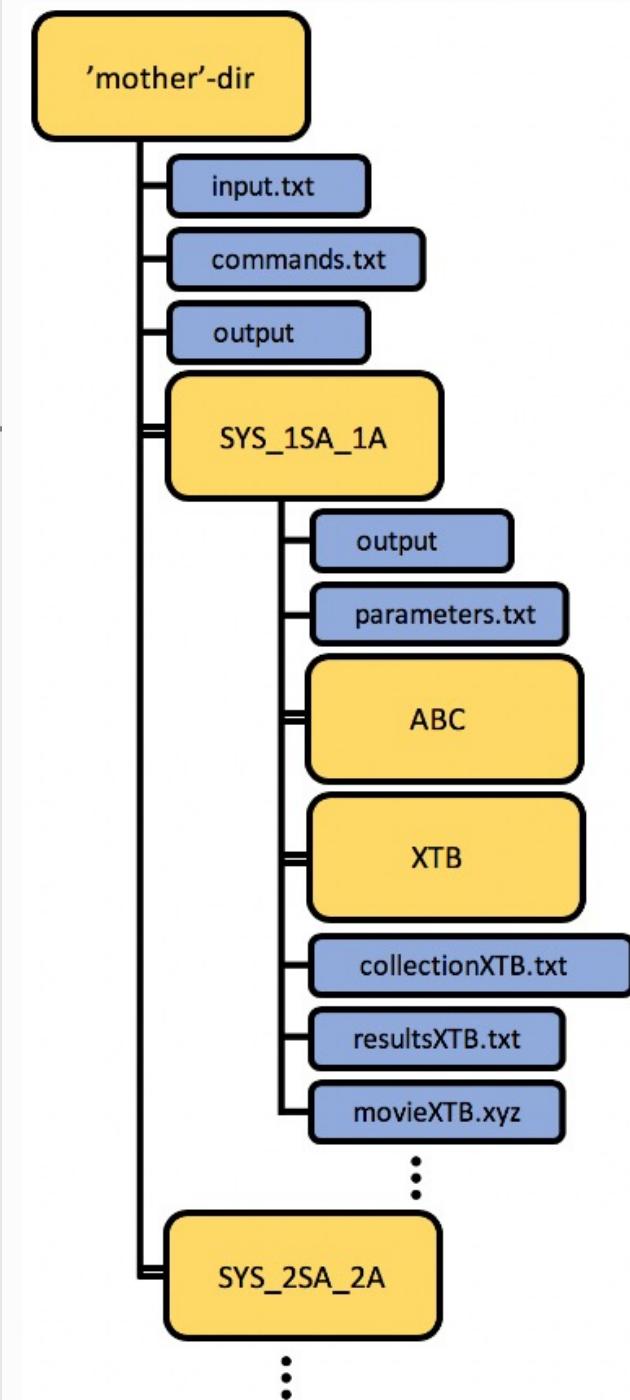
```
JKCS2_explore -help
```

- 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

# Folder structure

distinguish - **GAME:** what is wrong with the right figure  
(2 mistakes)  
– cluster/system folder

**TODO:** See the structure of SYS\_1SA\_1A/ABC folder



# How to JKCS3\_run

---

Any communication with other programs than ABCluster is performed 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.

**GAME:** How long simulation was used for ABCluster, i.e. what was the command?

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

**TODO:** Modify program\_XTB so it returns all XTB files.

# How to JKCS4\_collect

---

JKCS4\_collect is just (submittable) envelope for JKQC

```
JKCS4_collect -help
```

**TODO:** In the SYS\_1SA\_1A folder, try this:

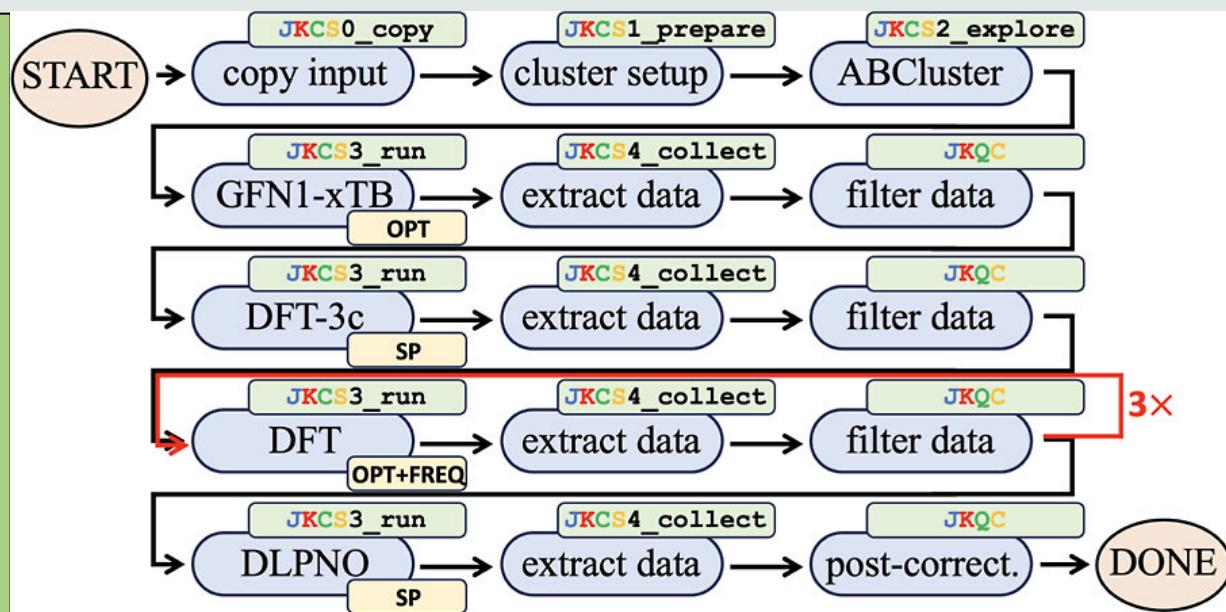
```
ls
JKCS4_collect XTBsp -loc -oc
ls
JKCS4_collect XTBsp -loc
ls
JKCS4_collect XTBsp -par small -time 10:00 -slurm "-begin=now+60" #sec
squeue -u $USER
JKcheck
```

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

# 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)

# Advanced CS workflow

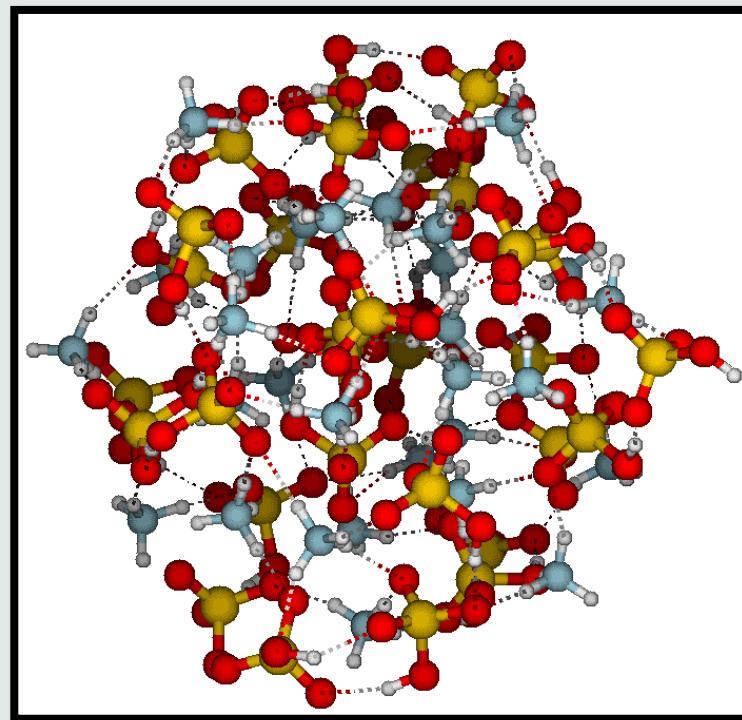
---

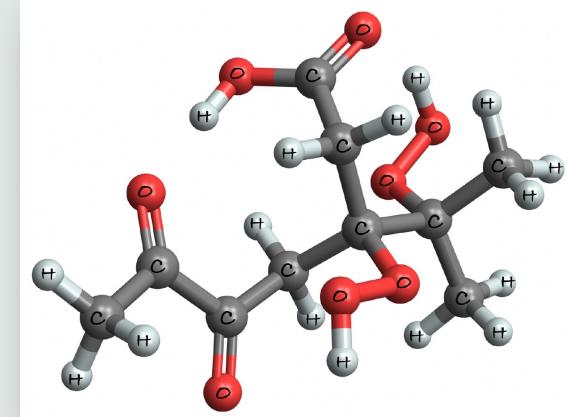
**TODO:** Understand all the steps/arguments of this CS workflow.

# CS workflow for large clusters

---

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





# 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

# 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<sub>1</sub>w<sub>3</sub> (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
```

# 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
```

**GAME:** Did it finish correctly now? Why not?

Most likely not long enough simulation for such a

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

# Searching for errors 2

---

Somebody complained that their calculations did not finish correctly.

```
cat collectionHF.txt
```

**GAME:** Where the user made a mistake?

Write a correct command and call me.

# Hydrogen atom

---

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

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

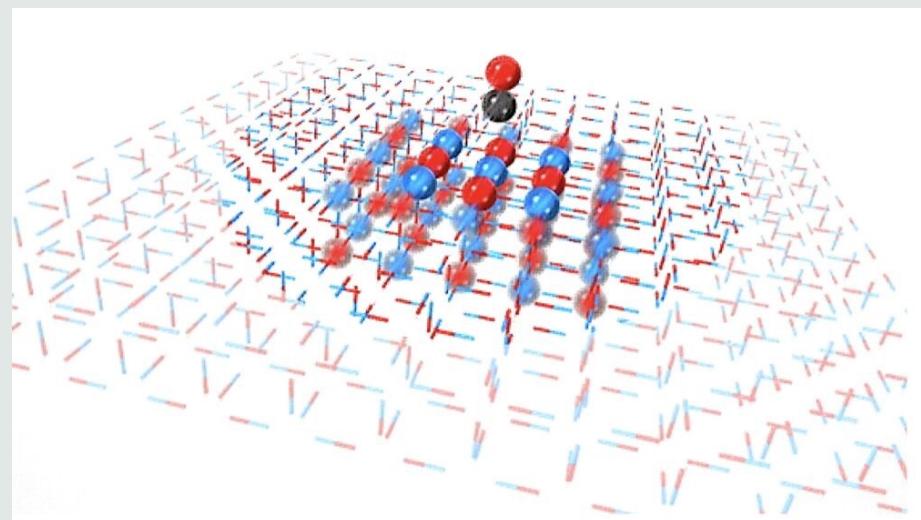
# CS of molecular systems on a surface

---

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 reactions, go and check e.g. ONIOM or Green's functions.



# Discuss some unclear arguments

---

- 1) Cleaning
- 2) ACDB
- 3) manager.sh / boss.sh
- 4) -mf for JKCS3\_run
- 5) -oc -orca -forces for JKCS4\_collect
- 6) JKcheck
- 7) JKtakebest
- 8) linking another QC program
- 9) JKTS
- 10)something else?

## **Build your own workflow inc. ML**

---

- 1) Sample  $(w)_{1-10}$
- 2) Optimize at XTB
- 3) For  $(w)_{1-7}$  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  $X = \text{number of } w \text{ molecules}$
- 7) Optimize structures at DFT + calculate vibrational frequencies
- 8) Plot IR spectrum for the global free-energy conformers