

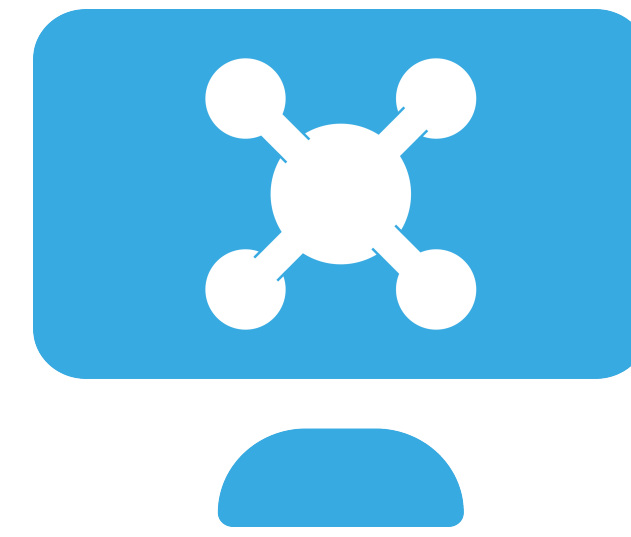
COMPUTATIONAL  
CHEMISTRY  
DAY 2023

# Use of Scripting in Computational Chemistry

Making everything faster and consistent

doc. dr. sc. Davor Šakić  
Sveučilište u Zagrebu Farmaceutsko-biokemijski fakultet





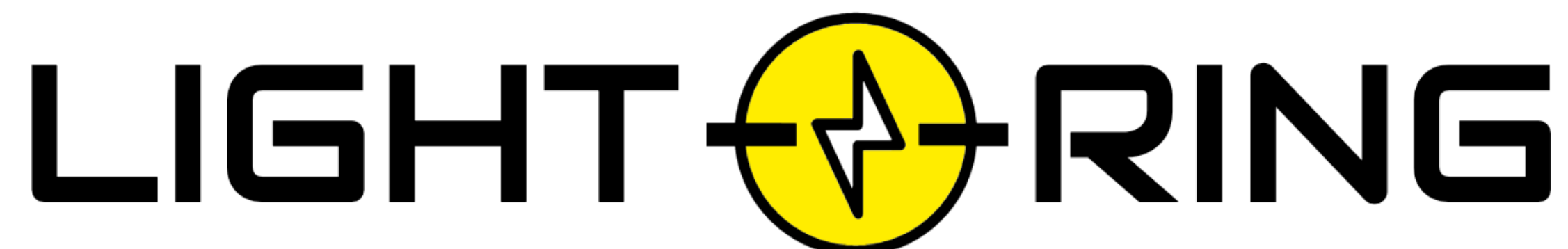
COMPUTATIONAL  
CHEMISTRY  
DAY 2023

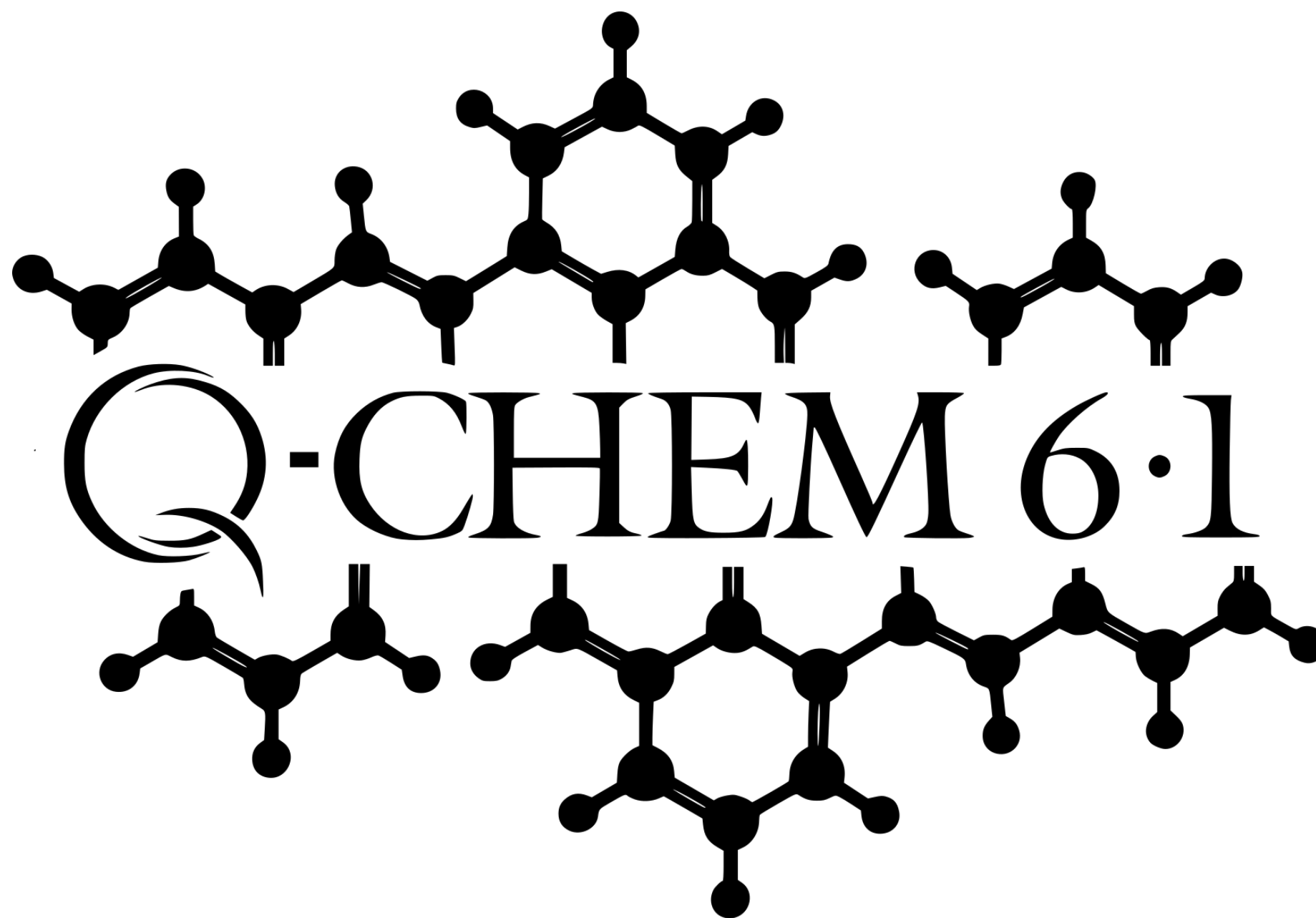
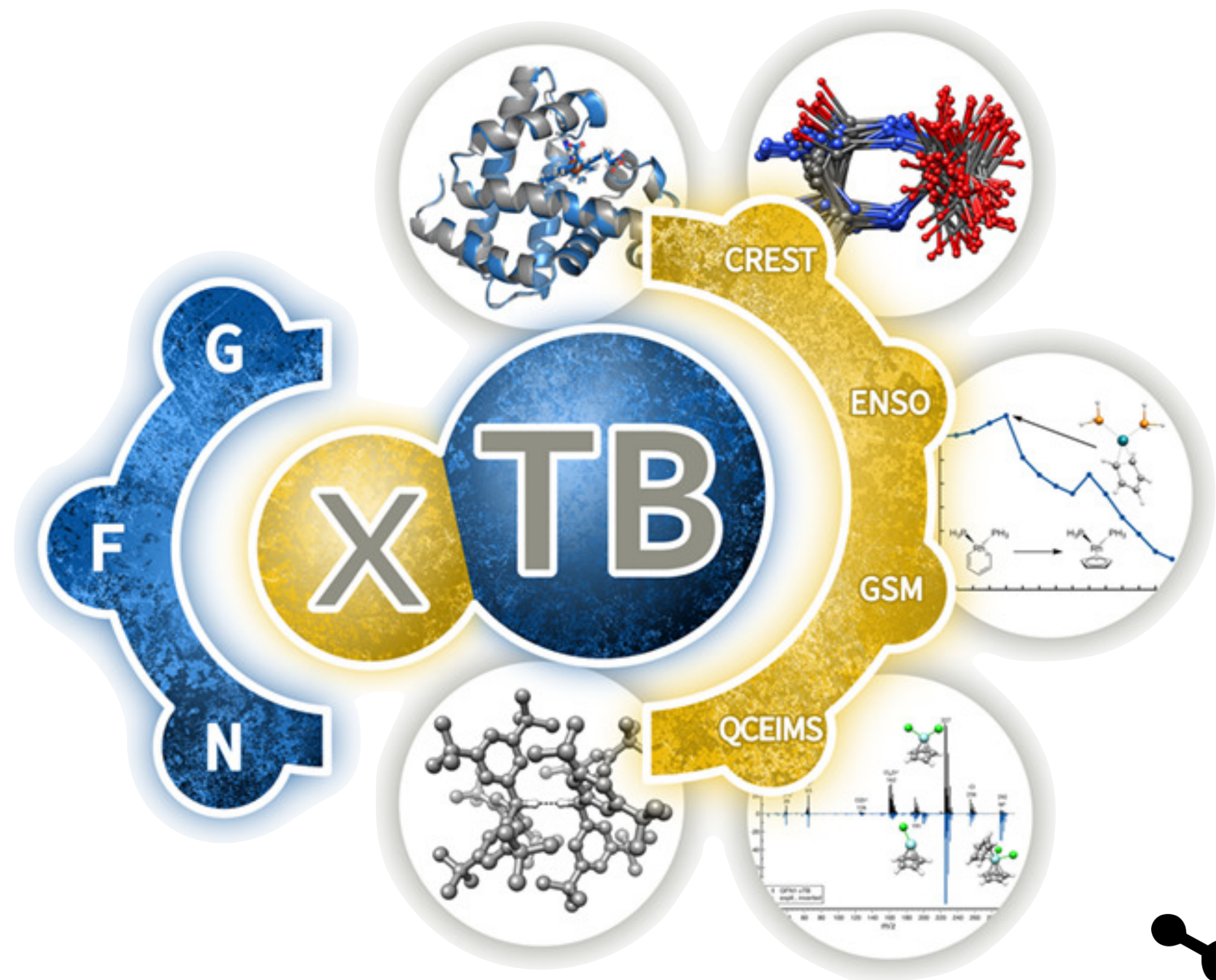
# Use of Scripting in Computational Chemistry

Making everything faster and consistent

doc. dr. sc. Davor Šakić

Sveučilište u Zagrebu Farmaceutsko-biokemijski fakultet





eta  chem



# CRTANJE

## Avogadro



- besplatan
- interface s drugim programima
- uključeni FF optimizator
- vizualizacija orbitala/frekvencija
- konformacijska analiza



**GAUSSVIEW**



## IQmol

besplatan

uključen FF optimizator

direktno pokretanje računa  
QChem

.xyz format

# Formati

**.xyz**

Broj atoma

Naslov/prazni red

Element x y z

Položaji u angs

```
24
N      2.47853    0.21632    0.04817
C      1.80729   -1.02331    0.41525
C      0.27138   -0.90081    0.40757
C     -0.40452   -0.19974    1.60585
C     -0.13598    1.30206    1.75733
C     -1.00658    1.90240    2.85924
C      3.24995    0.34913   -1.06774
O      3.41030   -0.58626   -1.83902
C      3.91446    1.66084   -1.36366
H      2.39153    1.03499    0.65793
H      2.16405   -1.38063    1.40468
H      2.05733   -1.81560   -0.32309
H     -0.11720   -1.94312    0.43902
H     -0.06444   -0.45524   -0.55404
H     -0.12845   -0.72263    2.54777
H     -1.49951   -0.32829    1.45423
H     -0.36186    1.82318    0.80187
H      0.91291    1.48012    2.05031
H     -2.08304    1.77534    2.61589
H     -0.79182    2.98780    2.95432
H     -0.79246    1.41114    3.83255
H      3.14369    2.45228   -1.47318
H      4.50268    1.60209   -2.30450
H      4.59809    1.92887   -0.53112
```

**.pdb**

**.sdf**

**.com**

**.mol**

**.inp**

```
module load scientific/xtb/6.6.0-gnu  
module load scientific/crest/2.12-gnu
```

```
module load scientific/gaussian/16-01
```

```
module load scientific/orca/5.0.4
```



# Brza optimizacija i konformacijska analiza

- semi-empirijske metode
- AM1, PM3, PM6
- GFN-1, GFN-2
- CREST, MD, MTD
- GFN-FF
- ENSO, CENSO, QCxMS

ime.xyz

```
xtb ime.xyz -opt --chrg 0 --uhf 0 --namespace ime
```

ime.xtbopt.xyz

```
xtb ime.xtbopt.xyz -hess --chrg 0 --uhf 0 --namespace ime.xtbopt
```

ime.xtbopt.g98.out

```
crest ime.xtbopt.xyz -T 4 --v4 --gfn2 --chrg 0 --uhf 0
```

crest\_conformers.xyz

# Optimizacija

```
%nproc=4
%mem=8gb
%chk=ime.chk
# opt freq b3lyp/SVP

naslov

0 1
A x.x y.y z.z
```

ime.com

```
#!/bin/sh
#PBS -q cpu
#PBS -l ncpus=4
#PBS -N ime

xtb ime.xyz -T 4 -opt
--chrg 0 --uhf 0
--namespace ime
```

ime.pbs

```
#PBS -q cpu
#PBS -l ncpus=8
#PBS -l mem=2gb
#PBS -N ime

cd ${PBS_0_WORKDIR}

module load scientific/
gaussian/16-C01

dog16-C01 ime.com
```

ime.pbs

```
! B3LYP SV(P) OPT FREQ
%pal nprocs 4 end
%maxcore 8000
* XYZ 0 1
A x.x y.y z.z
*
```

ime.inp

```
#!/bin/sh
#PBS -q cpu
#PBS -l select=16:mem=2gb
#PBS -l place=pack
#PBS -N ime

cd ${PBS_0_WORKDIR}
module load scientific/orca/5.0.4

${ORCA_ROOT}/orca ime.inp > ime.out
```

ime.pbs

qsub ime.pbs



# Odabir metoda i baznih skupova

Basis Set	Applies to	Polarization Functions	Diffuse Functions	Local and gradient					
					Hybrid functionals				
3-21G	H-Xe		+	HFS					
6-21G	H-Cl	* or **		LDA or LSD	B1LYP				
4-31G	H-Ne	* or **		VWN or VWN5	TPSS				
6-31G	H-Kr	through (3df,3pd)	+,++	VWN3					
6-311G	H-Kr	through (3df,3pd)	+,++	PWLDA	B3LYP and B3LYP/G				
D95	H-Cl except Na and Mg	through (3df,3pd)	+,++	BP86 or BP	TPSSh				
D95V	H-Ne	(d) or (d,p)	+,++	BLYP	TPSS0				
SHC	H-Cl	*		OLYP					
CEP-4G	H-Rn	* (Li-Ar only)		GLYP	O3LYP				
CEP-31G	H-Rn	* (Li-Ar only)		XLYP	X3LYP				
CEP-121G	H-Rn	* (Li-Ar only)		PW91	B1P				
LanL2MB	H-La, Hf-Bi			mPWPW	B3P				
LanL2DZ	H, Li-La, Hf-Bi			mPWLYP	B3PW				
SDD, SDDAll	all but Fr and Ra			PBE	PW1PW				
cc-pVDZ	H-Ar, Ca-Kr	included in definition		RPBE	mPW1PW				
cc-pVTZ	H-Ar, Ca-Kr	included in definition		REVPBE	mPW1LYP				
cc-pVQZ	H-Ar, Ca-Kr	included in definition		RPW86PBE	PBE0				
cc-pV5Z	H-Ar, Ca-Kr	included in definition		PWP	REVPBE0				
cc-pV6Z	H, B-Ne	included in definition			REVPBE38				
SV	H-Kr				BHANDHLYP				
SVP	H-Kr	included in definition							
TZV and TZVP	H-Kr	included in definition							
QZVP and Def2	H-La, Hf-Rn	included in definition							
MidIX	H, C-F, S-Cl, I, Br	included in definition							
EPR-II, EPR-III	H, B, C, N, O, F	included in definition							
UGBS	H-Lr	UGBS(1,2,3)P							
MTSmall	H-Ar								
DGDZVP	H-Xe								
						wB97			
							wB97X		
								wB97X-D3	
								wB97X-D4	
									B2PLYP
								wB97X-V	mPW2PLYP
								wB97X-D3BJ	
									B2GP-PLYP
								wB97M-V	B2K-PLYP
								wB97M-D3BJ	B2T-PLYP
									PWPB95
								wB97M-D4	
									B97M-D4
								CAM-B3LYP	
								LC-BLYP	PBE-QIDH
								LC-PBE	
									PBE0-DH
								SCANfunc	

# Traženje prijelaznog stanja - iz nacrtane strukture

## G16 i ORCA input

```
%nproc=4  
%mem=8gb  
%chk=ime.chk  
# opt=(calcfc,ts,noeigentest)  
freq b3lyp/SVP
```

naslov

```
0 1  
A x.x y.y z.z
```

ime.com

ime.script

```
! B3LYP SV(P) OPTTS FREQ  
%pal nprocs 4 end  
%maxcore 8000  
* XYZ 0 1  
A x.x y.y z.z  
*
```

ime.inp

qsub ime.script

# Traženje prijelaznog stanja - iz reaktanata/produkata

## G16 i ORCA input

```
%nproc=4  
%mem=8gb  
%chk=ime.chk  
# opt=(qst2) freq b3lyp/SVP
```

ime.com

reaktant

```
0 1  
A x.x y.y z.z
```

produkt

```
0 1  
A x.x y.y z.z
```

ime.script

```
%nproc=4  
%mem=8gb  
%chk=ime.chk  
# opt=(qst3) freq b3lyp/SVP
```

reaktant

```
0 1  
A x.x y.y z.z
```

produkt

```
0 1  
A x.x y.y z.z
```

TS

```
0 1  
A x.x y.y z.z
```

```
! B3LYP SV(P) NEB-TS FREQ  
%pal nprocs 4 end  
%maxcore 8000  
%NEB NEB_END_XYZFILE "produkt.xyz"  
END  
* XYZfile 0 1 reaktant.xyz
```

ime.inp

```
! B3LYP SV(P) NEB-TS FREQ  
%pal nprocs 4 end  
%maxcore 8000  
%NEB NEB_END_XYZFILE "produkt.xyz"  
NEB_TS_XYZFILE "mislitTS.xyz" END  
* XYZfile 0 1 reaktant.xyz
```

ime.script

U skripti obavezno navesti koji podaci i gdje se prebacuju

# Traženje prijelaznog stanja - pretraživanje PES-a

## G16 i ORCA input

```
%nproc=4
%mem=8gb
%chk=ime.chk
# opt=(modredundant) b3lyp/SVP

naslov

0 1
A x.x y.y z.z

B 1 2 S 15 -0.01
```

ime.com

ime.script

Brojanje atoma u ORCA-i ide od 0!

```
! B3LYP SV(P)
! ScanTS
%geom
scan B 1 0 = 2.0, 1.0, 10 end
end
%pal nprocs 4 end
%maxcore 8000
* XYZ 0 1
A x.x y.y z.z
*
```

ime.inp

Više o skriptiranju  
<https://tldp.org/LDP/abs/html/>

# A gdje su skripte?

xyz\_s

```
file=`echo $1 | sed 's,.log,,g' `
echo $file
gile=`echo $file".log" | sed 's,.log,.xyz,g' `
echo $gile
natom=`grep "NAtoms" $file.log | awk '{print $2}' | head -n 1`
echo natom $natom
nlin=`echo "$natom+3" | bc`
echo nlin $nlin
echo $natom > $gile
echo >> $gile
less $file.log | grep --after-context=$nlin
"Center      Atomic      Atomic
Coordinates (Angstroms)" | sed 's, 29 , Cu ,g'
| sed 's, 79 , Au ,g' | sed 's, 53 , I ,g' |
sed 's, 3 , Li ,g' | sed 's, 14 , Si ,g' | sed
's, 35 , Br ,g' | sed 's, 15 , P ,g' | sed 's,
5 , B ,g' | sed 's, 26 , Fe ,g' | sed 's, 9 , F
,g' | sed 's, 25 , Mn ,g' | sed 's, 16 , S ,g'
| sed 's, 17 , Cl ,g' | sed 's, 6 , C ,g' |
sed 's, 1 , H ,g' | sed 's, 7 , N ,g' | sed 's,
8 , O ,g' | sed
's,-----
-----,,g' | sed
's,--, "$natom $natom",g' | grep -v "Center" |
grep -v "Number" | awk '{print $2,$4,$5,$6}' |
sed 's, , ,g' | tail -n `echo "$natom+1" | bc`
>> $gile
echo >> $gile
cat $gile
```

chmod 777 xyz\_s  
./xyz\_s ime.log

Zadnja geometrija  
-> .xyz

sdf2com

```
ls *sdf
echo -n "name: "
read name
ls *sdf | grep $name
echo -n "notname: "
read notname
ls *sdf | grep $name | grep -v $notname
echo -n "new_name: "
read new_name
echo -n "chg&mult: "
read chg
echo -n "nproc: "
read nproc
mem=`echo "$nproc" | awk '{print $1*2"GB"}'`
ls *sdf | grep $name | grep -v $notname | sed
's,.sdf,,g' | while read file; do
num=`grep V200 $file.sdf | head -n 1 | awk
'{print $1}'`
cat $file.sdf | grep --after-context=$num V200 >
temp
csplit -s temp /V20/ {*}
ls xx* | while read gile; do
cat $gile | grep -v V2 | awk '{print
$4,$1,$2,$3}' | grep -v "\-\-" >
$new_name$gile.xyz
echo "%nproc=$nproc
%mem=$mem
# opt freq b3lyp/6-31g(d)

$file

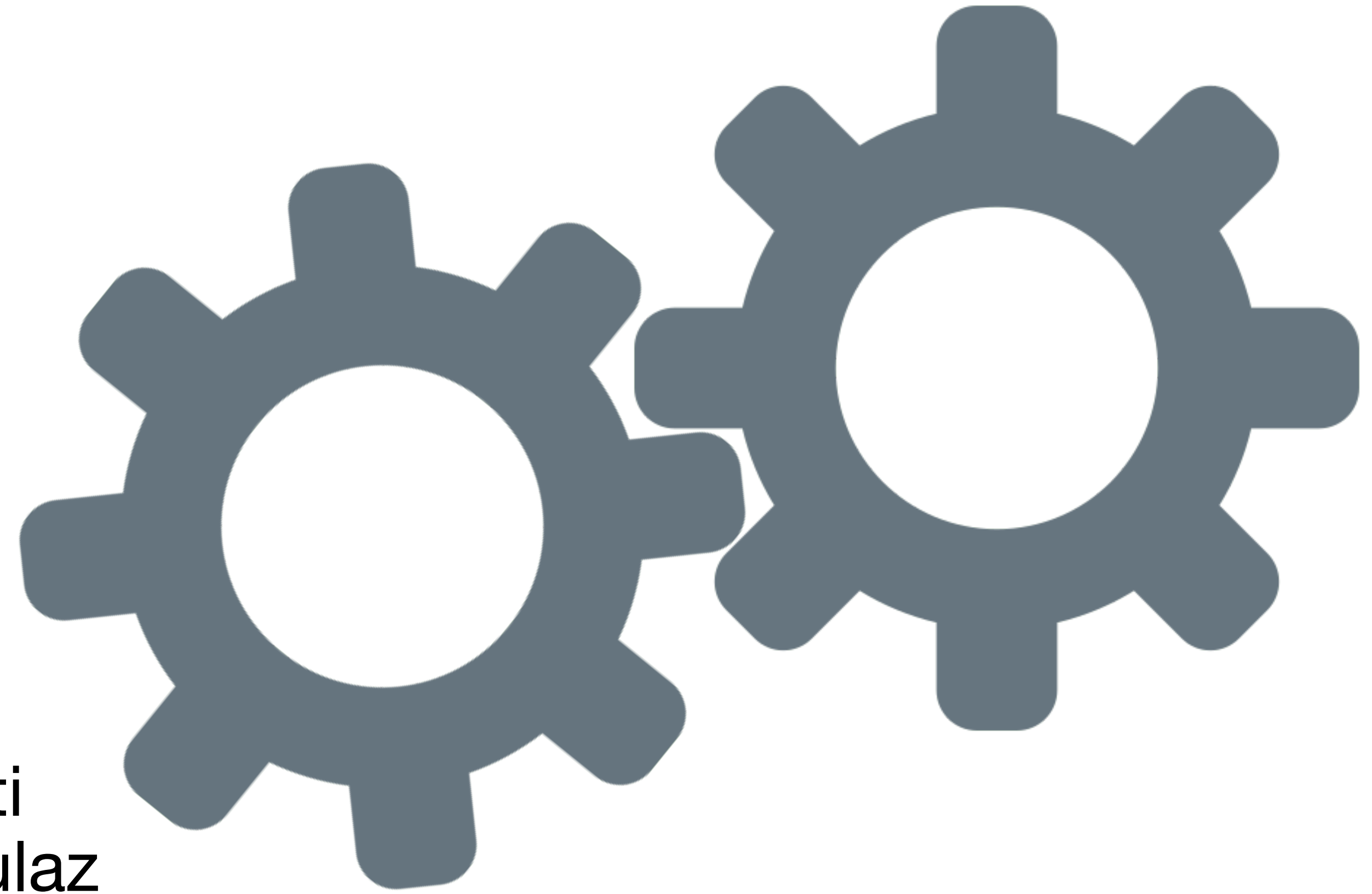
$chg
`cat $new_name$gile.xyz `

" > $new_name$gile.com; only_run.sh
$new_name$gile.com $nproc;
rm $gile
done

done
```

# Što su skripte?

- niz sistemskih komandi
- pokreću se nizom kojima su zadane
- svaka komanda bi trebala imati ulaz (input) i nekakav izlaz (output)
- komande se međusobno mogu spajati tako da izlaz prve komande postane ulaz druge komande
- **NAJVAŽNIJI ZNAK U SKRIPTAMA**



```
ls *log  
ls *log | grep "ime"  
ls *log | grep "ime" | grep -v "nije"
```

# Prednost skripti

- sve se odvija po ISTOJ proceduri
- izbacuje se ljudski faktor
- PETLJE!!!
- while; do; done
- if; then; fi
- if; then; else; then; fi
- jedna u drugu u treću - ne se izgubiti!
- varijable s \$var, \$1, \$2...

```
ls *log | while read file
do
    if [ $file == "ime" ]
    then
        echo $file
    else
        echo "nije "$file
    fi
done
```

# Korištenje skripti

- priprema datoteka za pokretanje
- pokretanje datoteka
- obrada podataka
  - izvlačenje važnih parametara
    - energije, geometrije, distribucija naboja...
- pripremanje podataka u oblik za publikaciju/  
daljnju obradu
- spremanje podataka i arhiviranje

```
ls *log | while read file
do
    gibbs=`grep "Free Energies"
    $file | awk '{print $8}'`
    echo $file $gibbs;
done | sort --key=2 -nr > out

best=`tail -n 1 out |
awk '{print $2}'`

cat out |
awk '{print $1", "$2, "($2-
"$best")*627.509*4.184}'
> out.csv

scp out.csv
sstudent@133.78.119.180:~/
```



# Synthesis of Chlorinated Oligopeptides via $\gamma$ - and $\delta$ -Selective Hydrogen Atom Transfer Enabled by the *N*-Chloropeptide Strategy

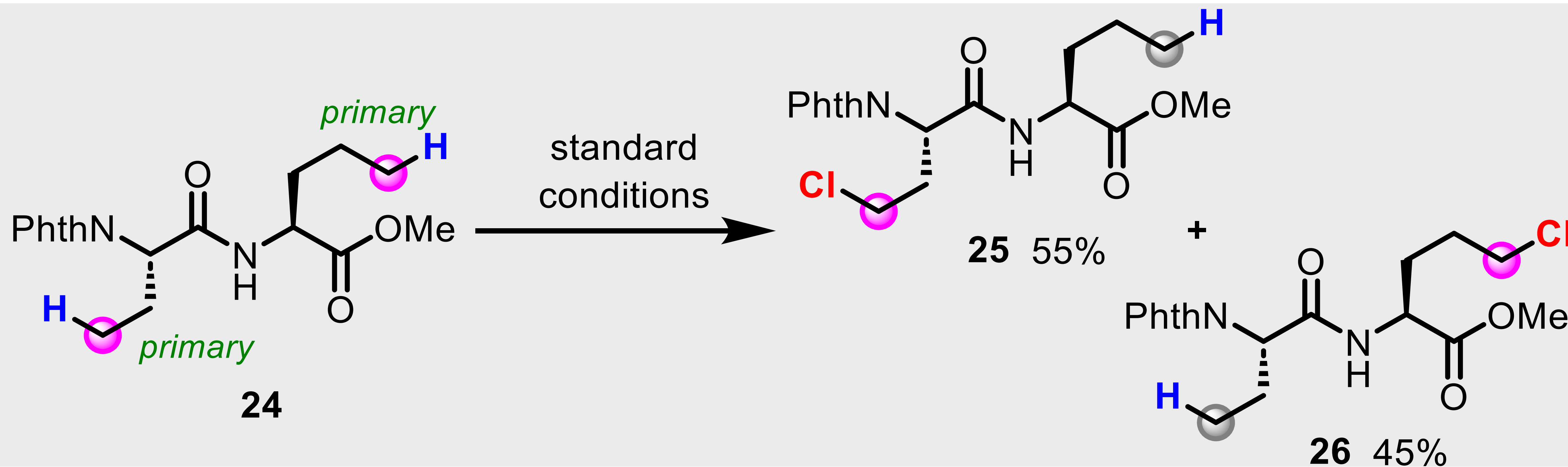
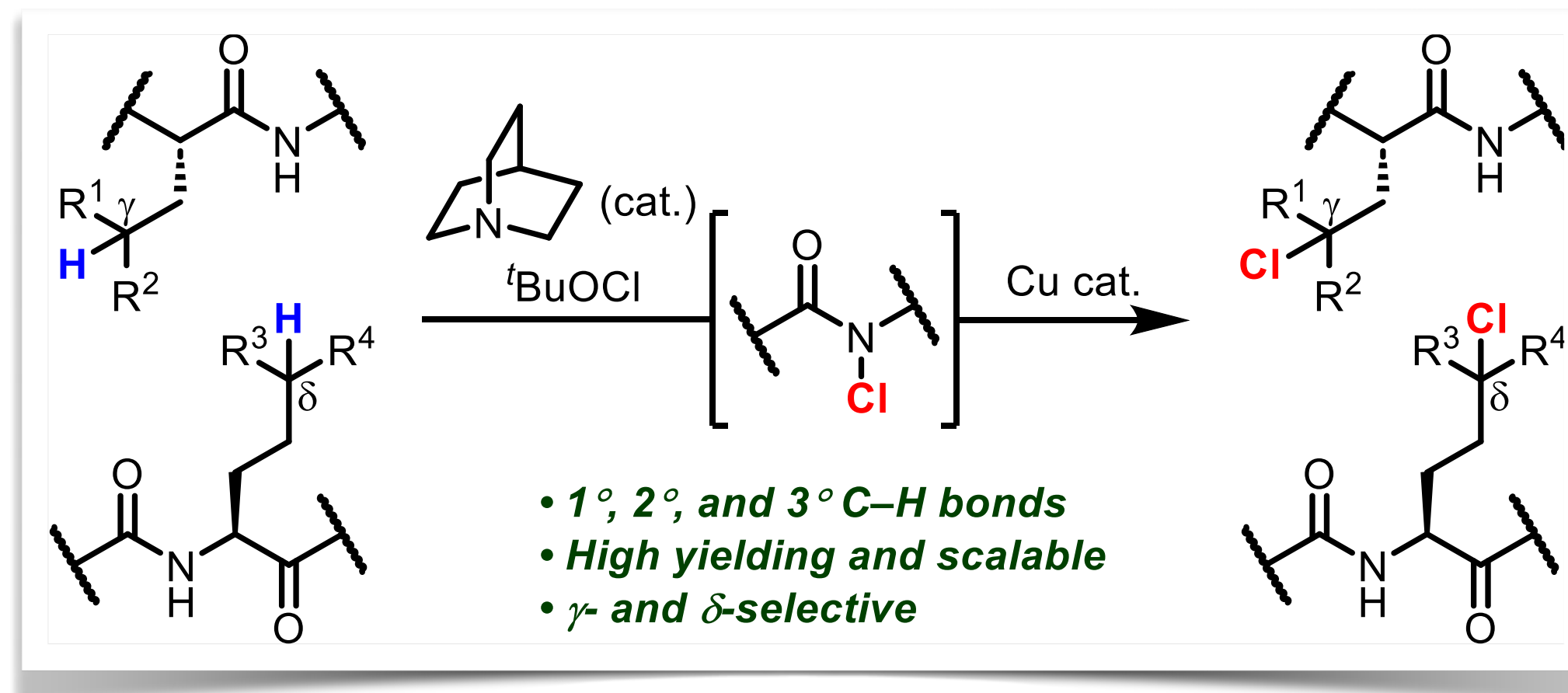
Takeshi Nanjo,\* Ayaka Matsumoto, Takuma Oshita, and Yoshiji Takemoto\*

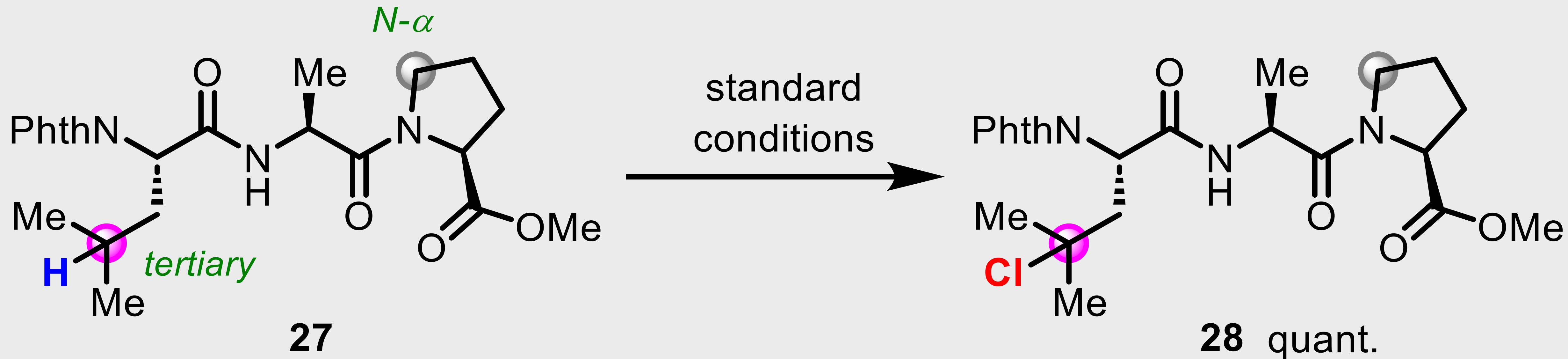


Cite This: *J. Am. Chem. Soc.* 2023, 145, 19067–19075



Read Online





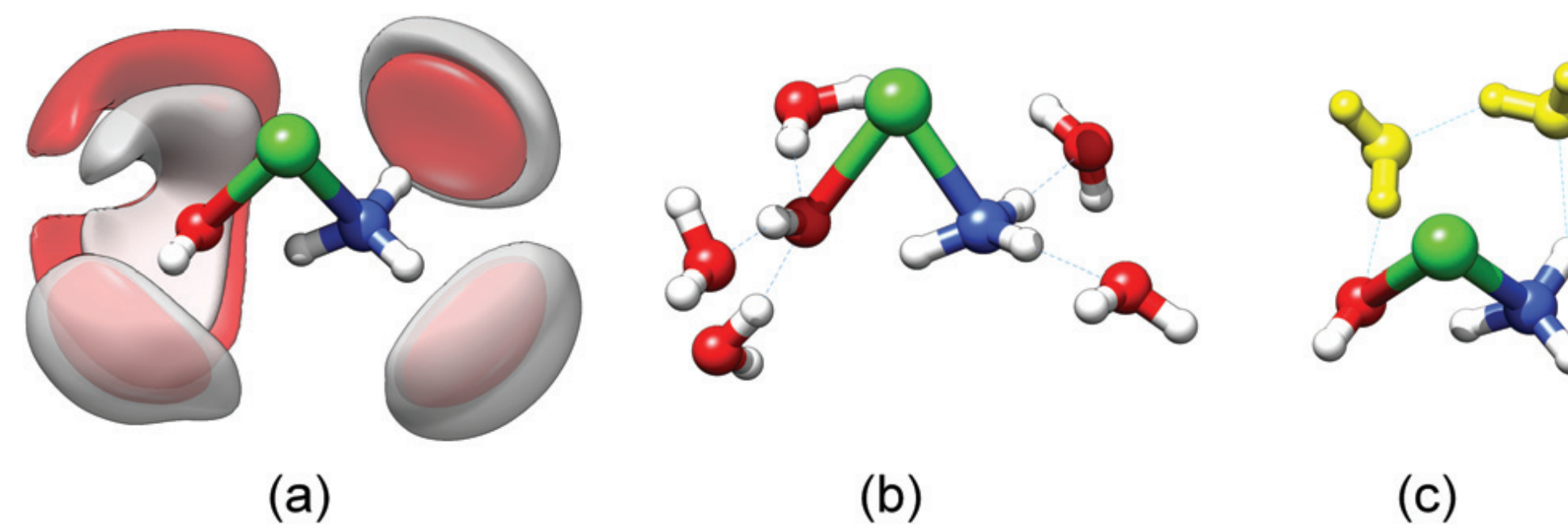
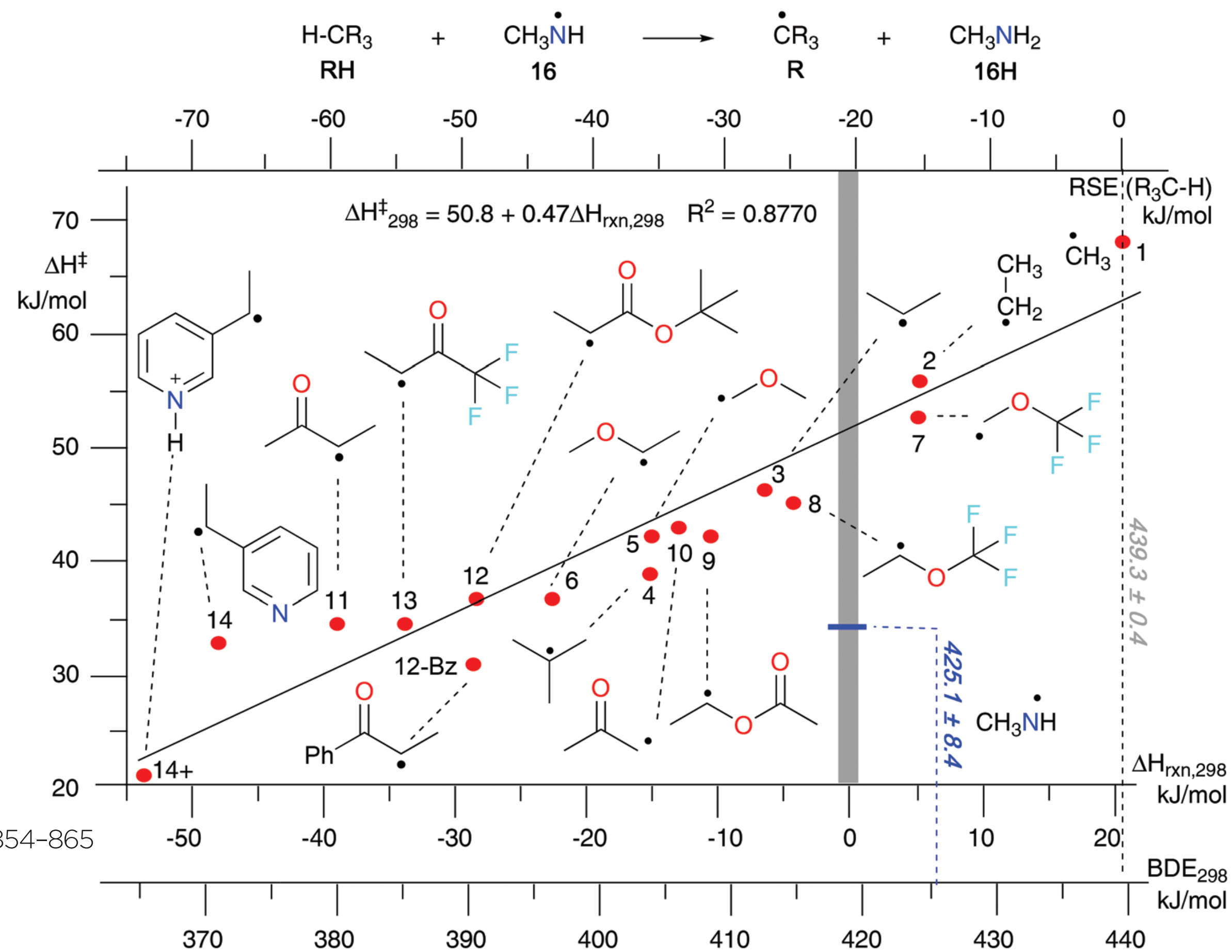
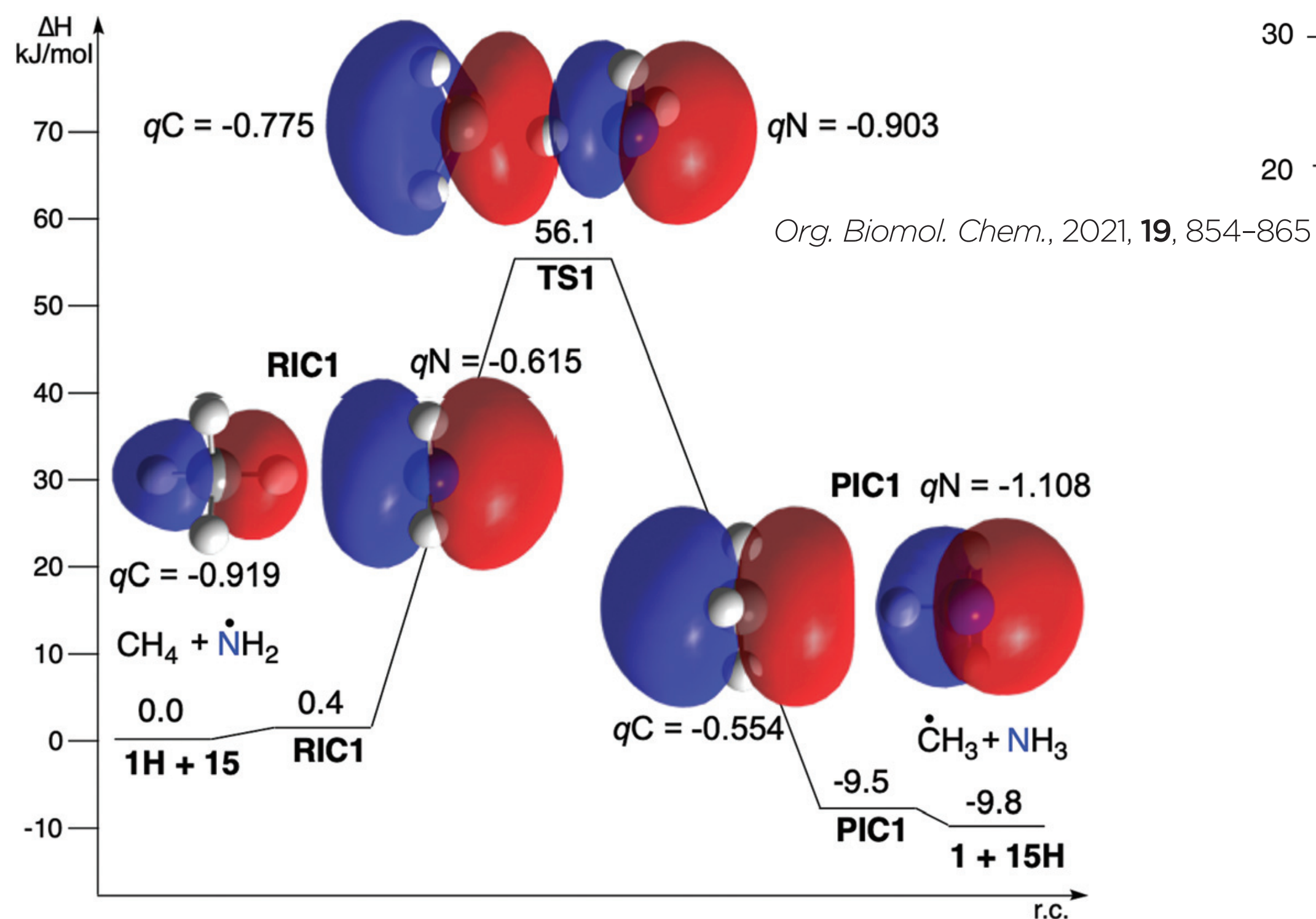
**Zašto?**

Kako riješiti to pitanje?

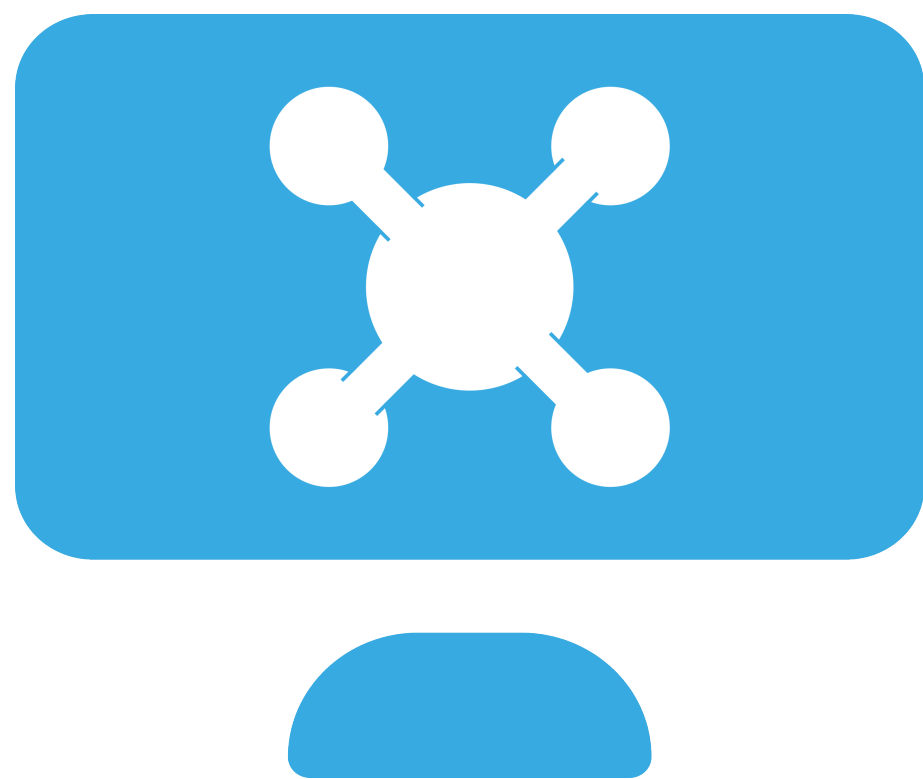
# Prikaz rezultata

**Table 3** RSE,  $\Delta H_{\text{rxn},298}$ , and  $\Delta H_{298}^\ddagger$  values for all systems shown in Chart 3

Species	RSE (N-rad)	RSE (C-rad)	BDE (N-H) <sub>calc.</sub>	BDE (C-H) <sub>calc.</sub>	$\Delta H_{\text{rxn},298}$	$\Delta H_{298}^\ddagger$
L1	-33.8	-71.0	416.3	368.3	-47.9	45.7
L1+	-40.5	-89.4	409.6	349.9	-59.6	46.0
L2	-58.4	-71.1	391.7	368.3	-23.4	54.2
L2+	-71.7	-89.0	378.4	350.3	-28.0	54.0
L3	-26.1	-66.7	424.0	372.7	-43.0	30.1
L4	-2.0	-63.5	448.1	375.8	-82.9	18.5
L5	2.0	-75.2	452.0	364.1	-87.2	16.2
L6	15.4	-73.7	465.5	365.6	-99.3	4.8



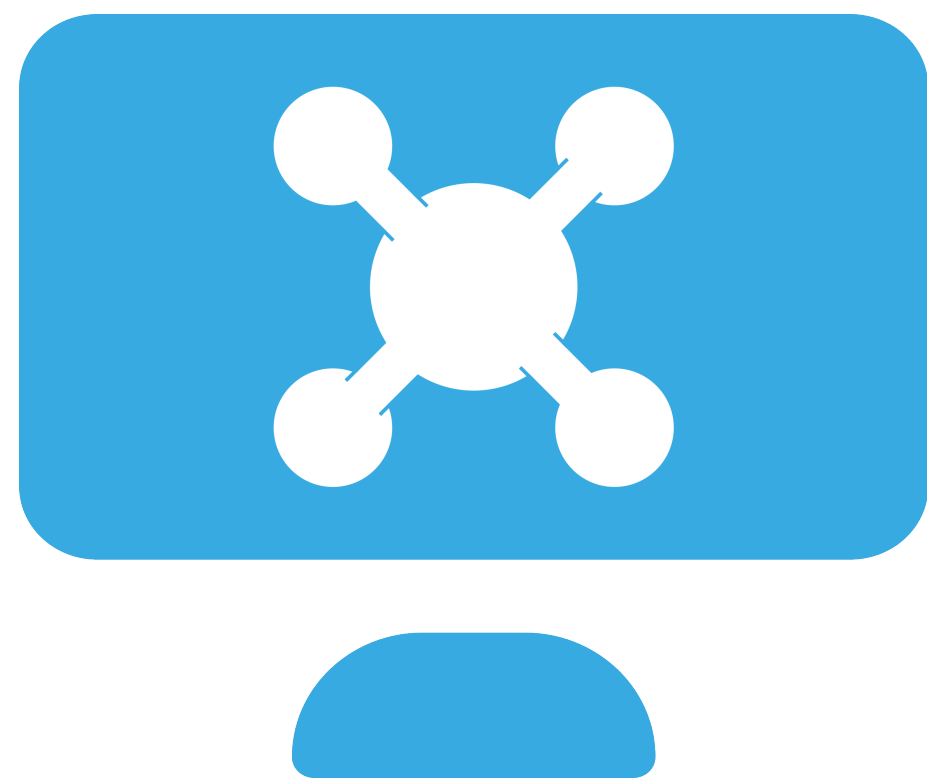
*Org. Biomol. Chem.*, 2015, **13**, 11740–11752



# COMPUTATIONAL CHEMISTRY DAY 2023

# Hvala





COMPUTATIONAL  
CHEMISTRY  
DAY 2023

Hvala

