

Introduction to CP2K

Vaibhav Modi

Department of Chemistry, University of Jyväskylä

CP2K

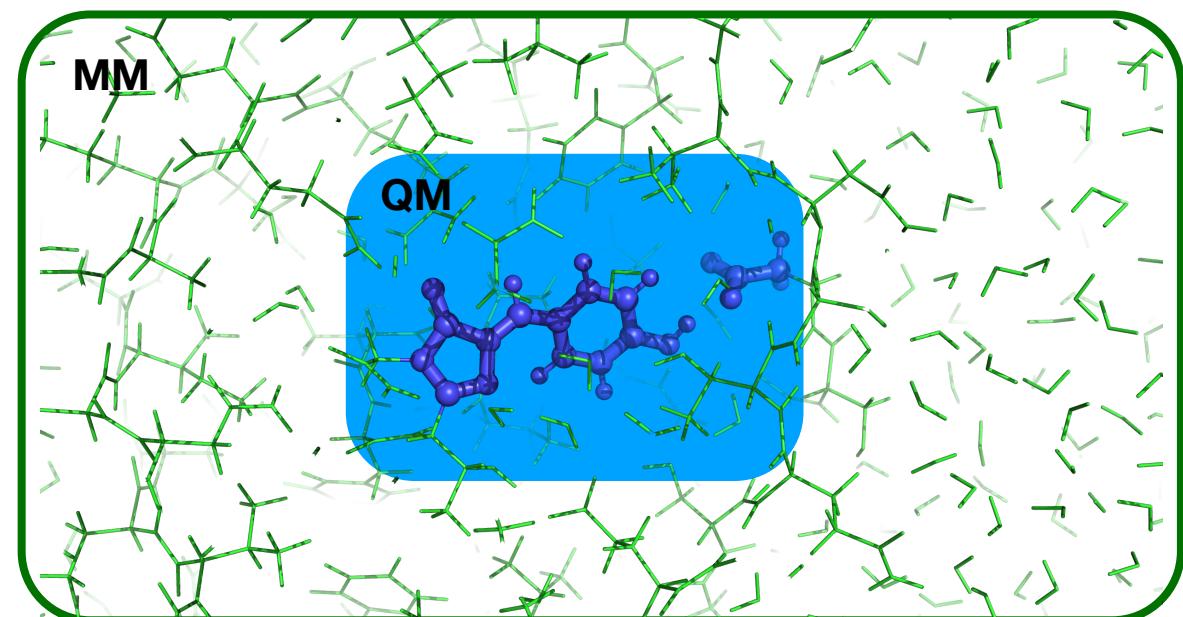
Quickstep: Mixed Gaussian and Plane wave
basis implementation of
Density Functional Theory (DFT)



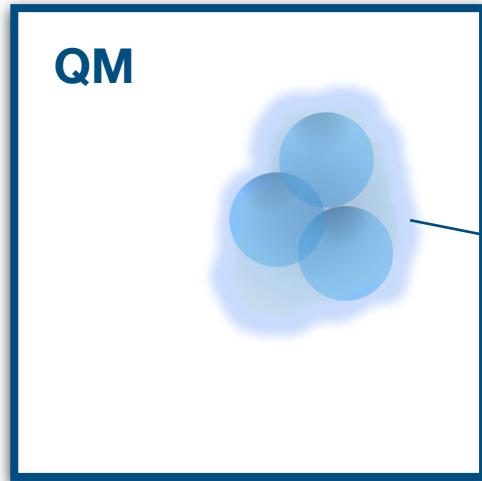
Gaussian Expansion of Electrostatic
Potential (GEEP) to compute the QM/MM
coupling

$$H = H_{MM} + \underline{H_{QM}} + H_{QM/MM}$$

Quickstep & GEEP



Density Functional Theory



Electron density

$$\rho(r) = \sum_{a,b}^n P_{ab} \phi_a(r) \phi_b^*(r)$$

Variational principle

$$E[\rho(r)] \geq E_0$$

Kohn-Sham DFT energy

$$E[\rho(r)] = \sum_{ab} P_{ab} \langle \phi_a | -\frac{1}{2} \nabla^2 | \phi_b \rangle - \sum_A Z_A \int \frac{\rho(r)}{|r - r_A|} dr + \frac{1}{2} \int \int \frac{\rho(r)\rho(r')}{|r - r'|} dr dr' + \frac{1}{2} \sum_{A \neq B} \frac{Z_A Z_B}{|r_A - r_B|} + \int \epsilon^{xc} \rho(r) dr$$

Kinetic

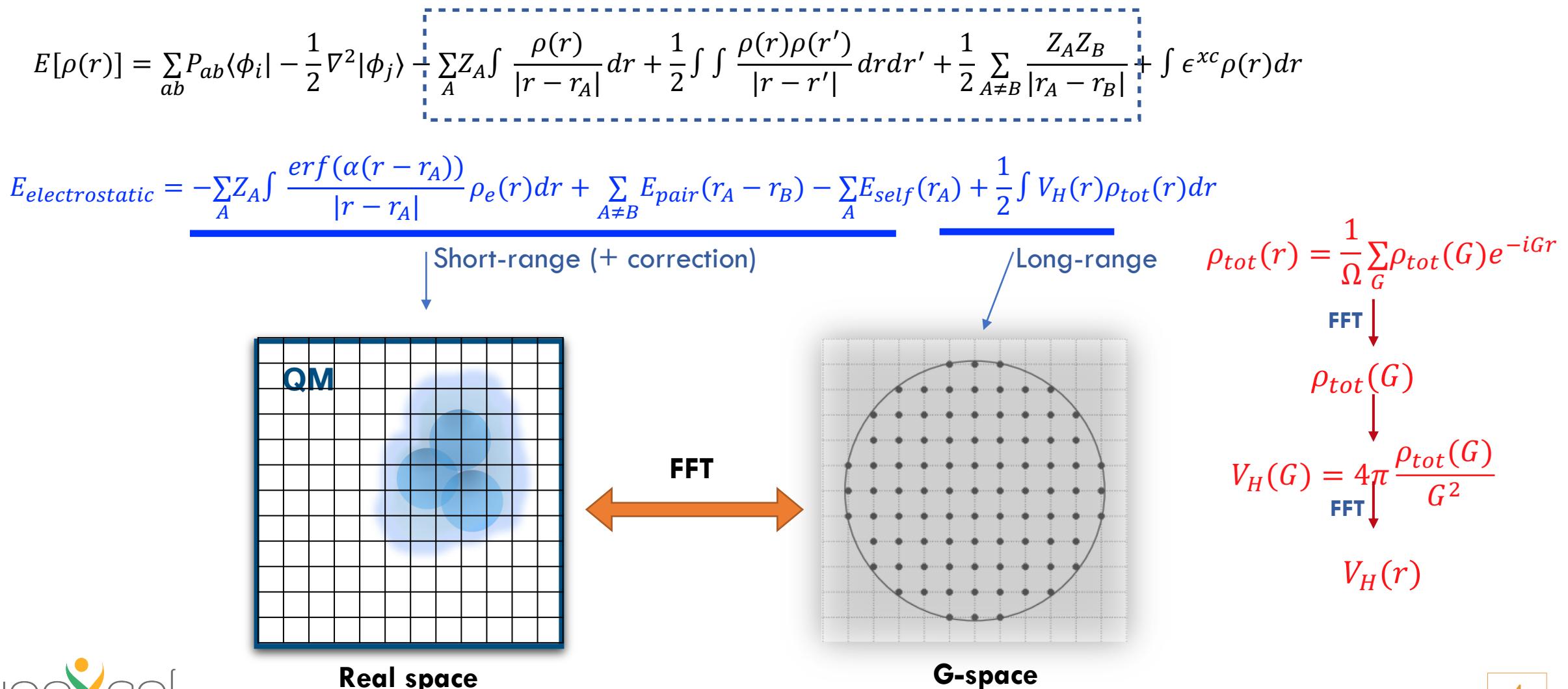
electron-nuclei

Classical coulomb

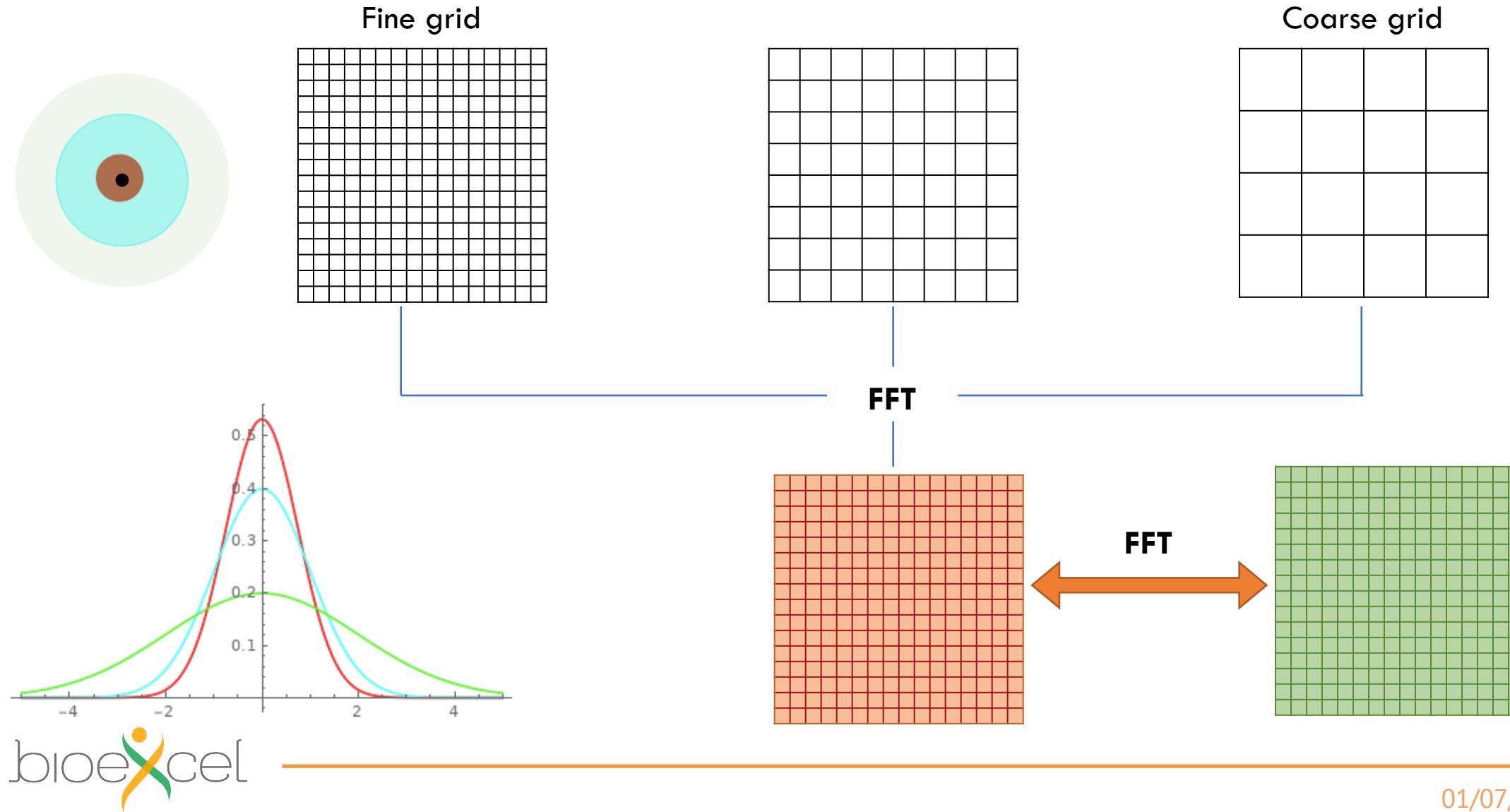
Ion-ion

Exchange & correlation

Gaussian and Plane Waves method (GPW)



Multigrid Framework



CP2K: Quickstep (QM)

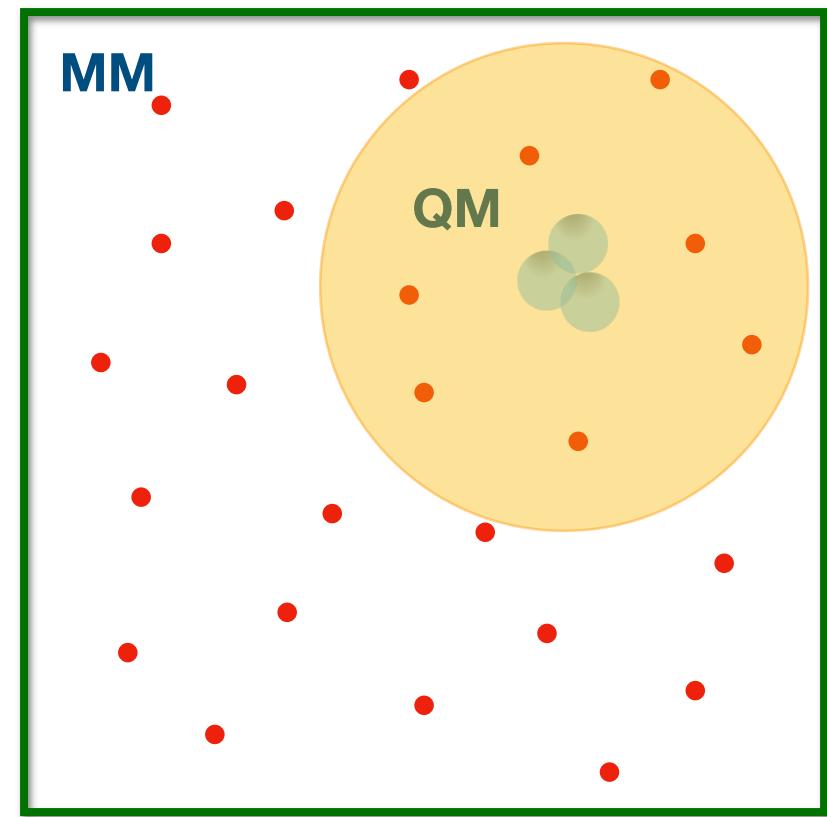
- Construction of the Kohn-Sham energy functional (Guess density from Gaussian basis)
 - Mapping of operators represented using density in Gaussian basis onto the RS multi-grids (collocation & interpolation)
 - FFT to map functions defined on each level of the RS multi-grid into the corresponding plane-wave basis; and its reverse operation to RS
 - Minimisation of the Kohn-Sham energy with respect to the electronic
 - Self-consistent cycle for the electronic charge density

QM/MM Electrostatic scheme

$$E_{tot}(r_a, r_A) = E_{MM}(r_A) + \underbrace{E_{QM}(r_a)}_{\text{Quickstep}} + \underbrace{E_{QM/MM}(r_a, r_A)}_{??}$$

- Spherical cutoff

$$E_{QM/MM} = \int V_c^{MM} \cdot \rho_{tot}^{QM} \cdot d\vec{r}$$

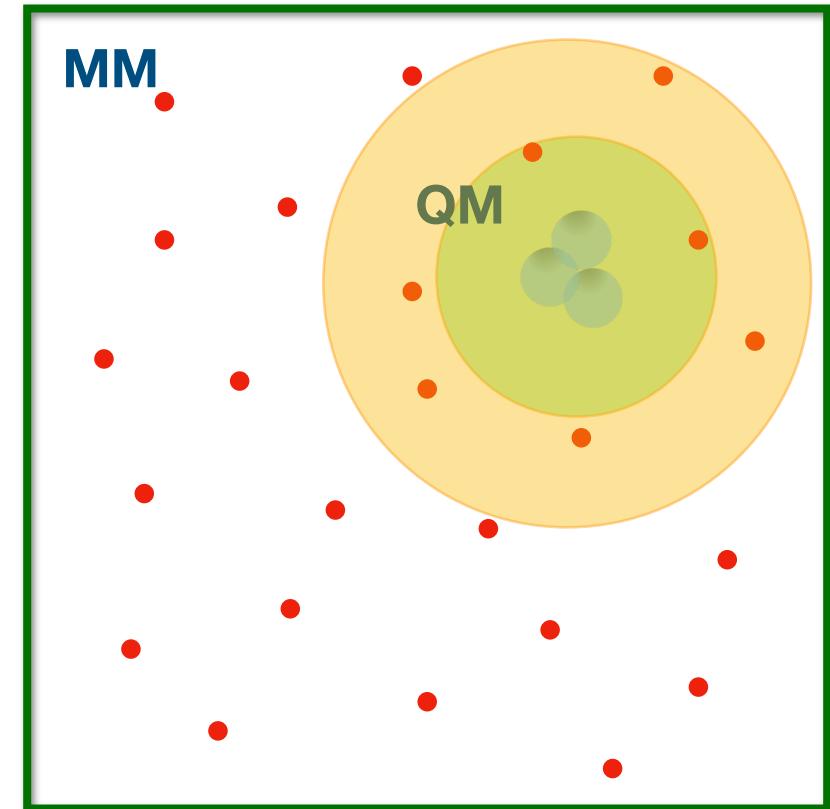


QM/MM Electrostatic schemes

- Multipole expansion for long range interactions

$$E_{QM/MM} = \sum_{A \in MM} q_A \int \frac{\rho(r)}{r - r_A} \cdot dr + \sum_{A \in MM, a \in QM} v_{vdw}(r_{aA})$$

Short-range term
(modified coulomb potential) + Mid-range term
(Potential based D-ESP charges) + Long-range term
(Multipole expansion of charge density
Of QM subsystem)



QM/MM Electrostatic schemes - GEEP

$$E_{tot}(r_a, r_A) = E_{MM}(r_A) + \underbrace{E_{QM}(r_a)}_{\text{Quickstep}} + \underbrace{E_{QM/MM}(r_a, r_A)}_{\text{GEEP}}$$

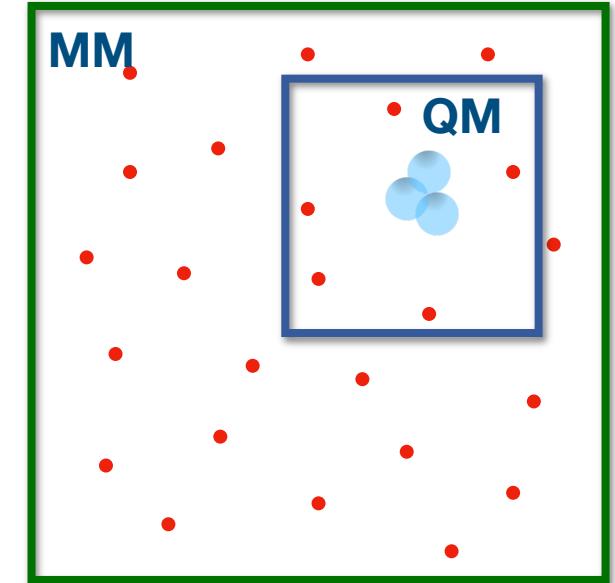
(Gaussian expansion of electrostatic potential)

- MM atoms as gaussian charge distribution (avoid electron-spilling)
- Electrostatic potential generated from MM atoms is fitted using gaussian expansion with different cutoffs

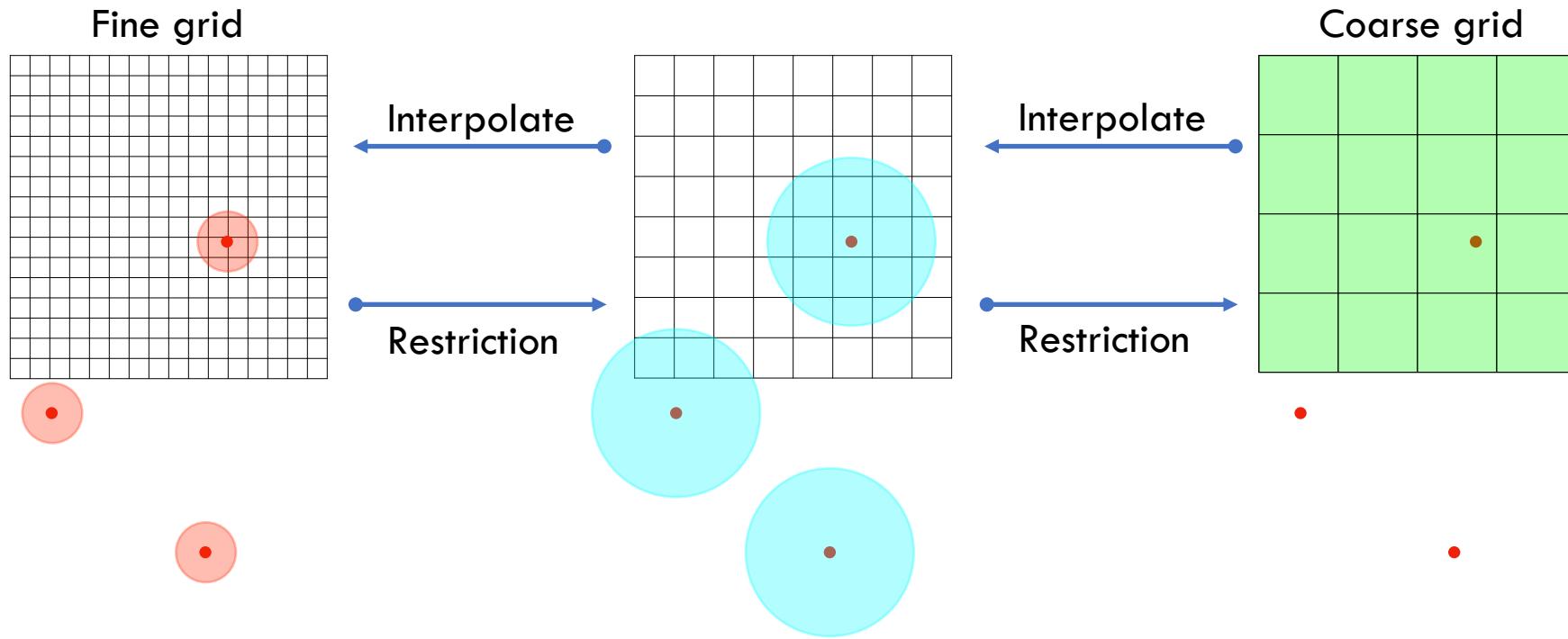
$$v_A(r, r_A) = \frac{Erf\left(\frac{|r - r_A|}{r_{c,A}}\right)}{|r - r_A|} = \sum_{N_g} A_g \exp\left(-\frac{|r - r_A|^2}{G_g^2}\right) + R_{low}(|r - r_A|)$$

$$E^{QM/MM}(r_a, r_A) = \int V_{QM/MM}(r, r_A) \rho(r, r_a) dr$$

$$V^{QM/MM}(r, r_A) = \sum_A q_A v_A(r, r_A)$$



Collocation & Interpolation



Real space interpolation
Coarse → Fine
20-40%

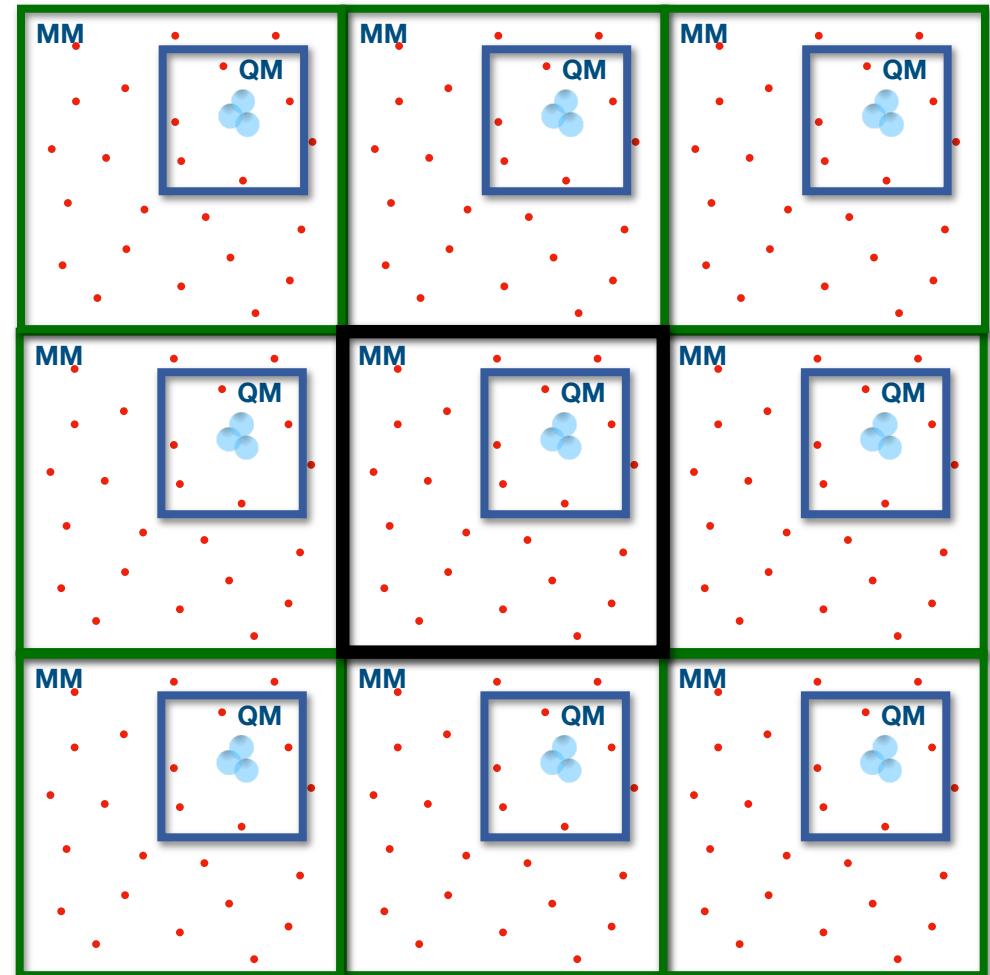
$$V^{QM/MM}(r, r_A) = \sum_{i=coarse}^{fine} \prod_{k=1}^{fine-1} I_{k-1}^k V_i^{QM/MM}(r, r_A)$$

Periodic QM/MM with GEEP

How to handle electrostatic potential over PBC?

EWALD Summation

$$V^{QM/MM}(r, r_A) = V_{real}^{QM/MM}(r, r_A) + V_{recip}^{QM/MM}(r, r_A)$$



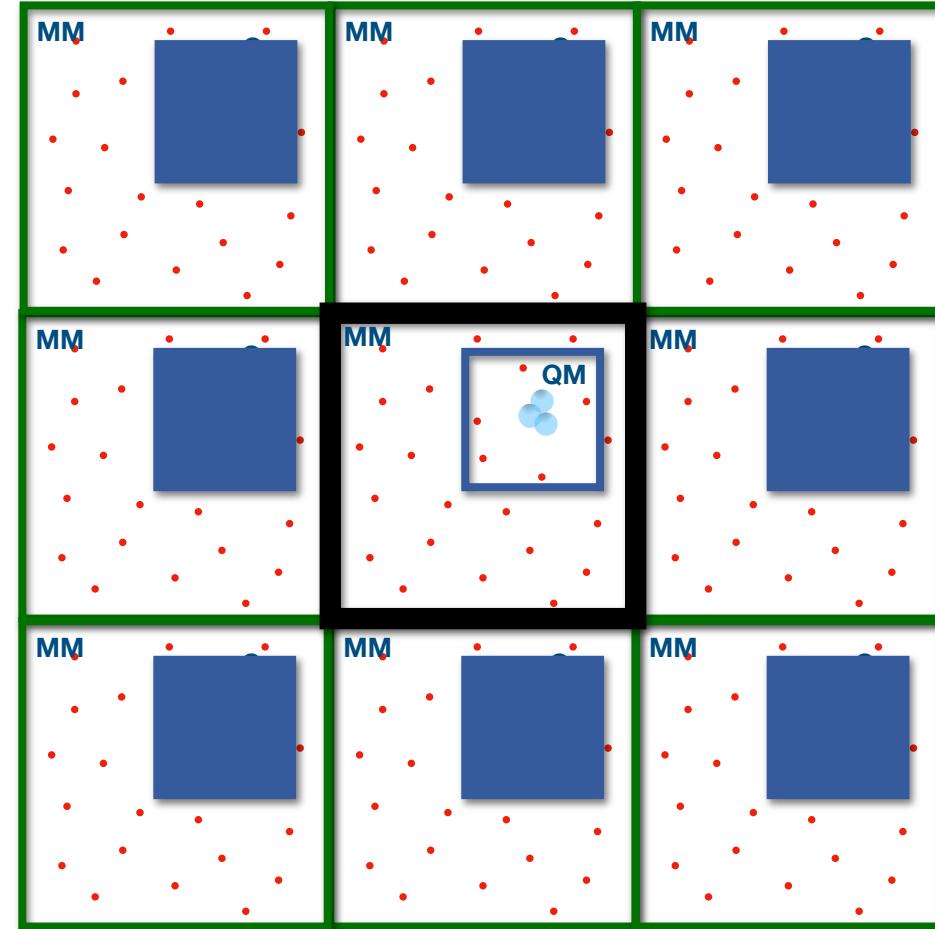
Periodic QM/MM with GEEP

How to handle electrostatic potential over PBC?

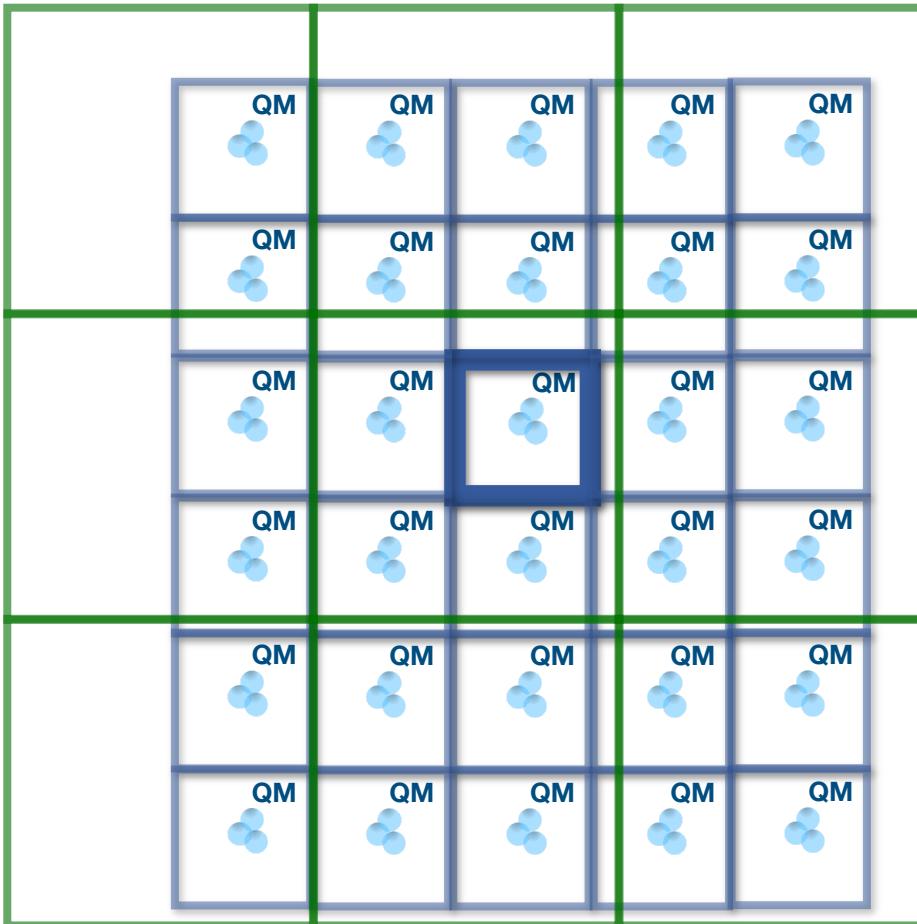
EWALD Summation

$$V^{QM/MM}(r, r_A) = V_{real}^{QM/MM}(r, r_A) + V_{recip}^{QM/MM}(r, r_A)$$

$$v_A(r, r_A) = \underbrace{\sum_{|L| \leq L_{cut}} \sum_{N_g} A_g \exp\left(-\frac{|r - r_A + L|^2}{G_g}\right)}_{\text{Real-space/short-range}} + \underbrace{\sum_L^\infty R_{low}(|r - r_A + L|)}_{\text{Reciprocal space/long-range}}$$

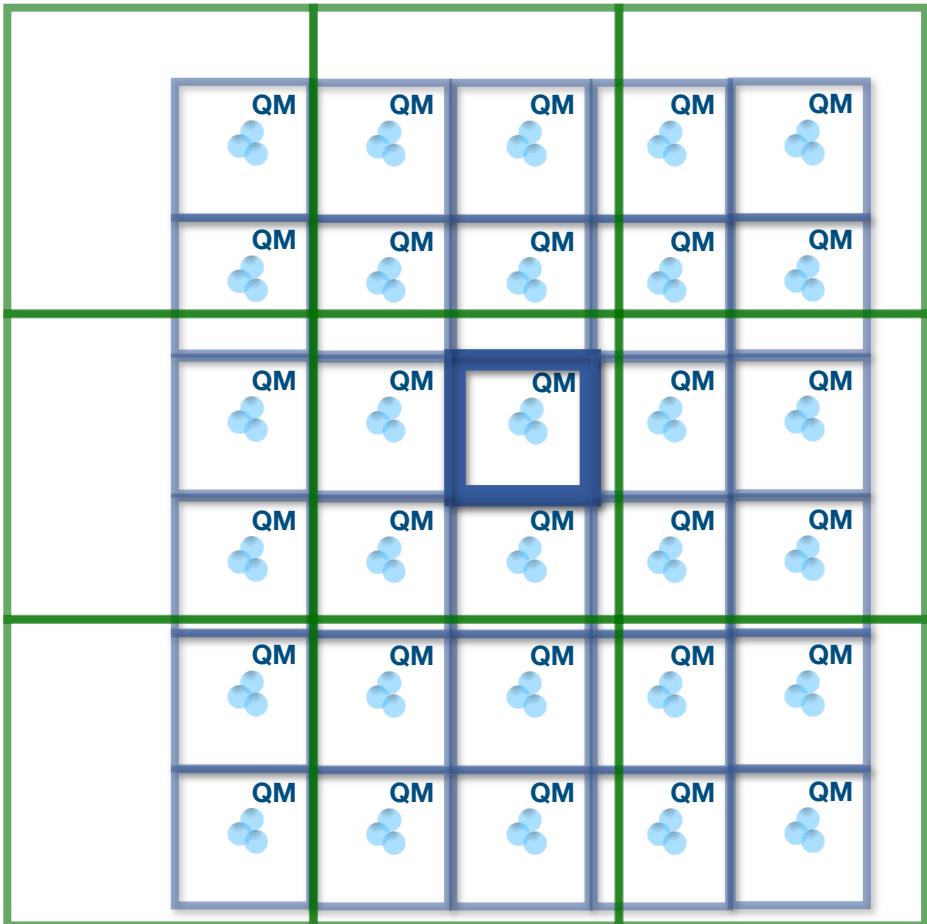


Periodic Coupling QM-QM



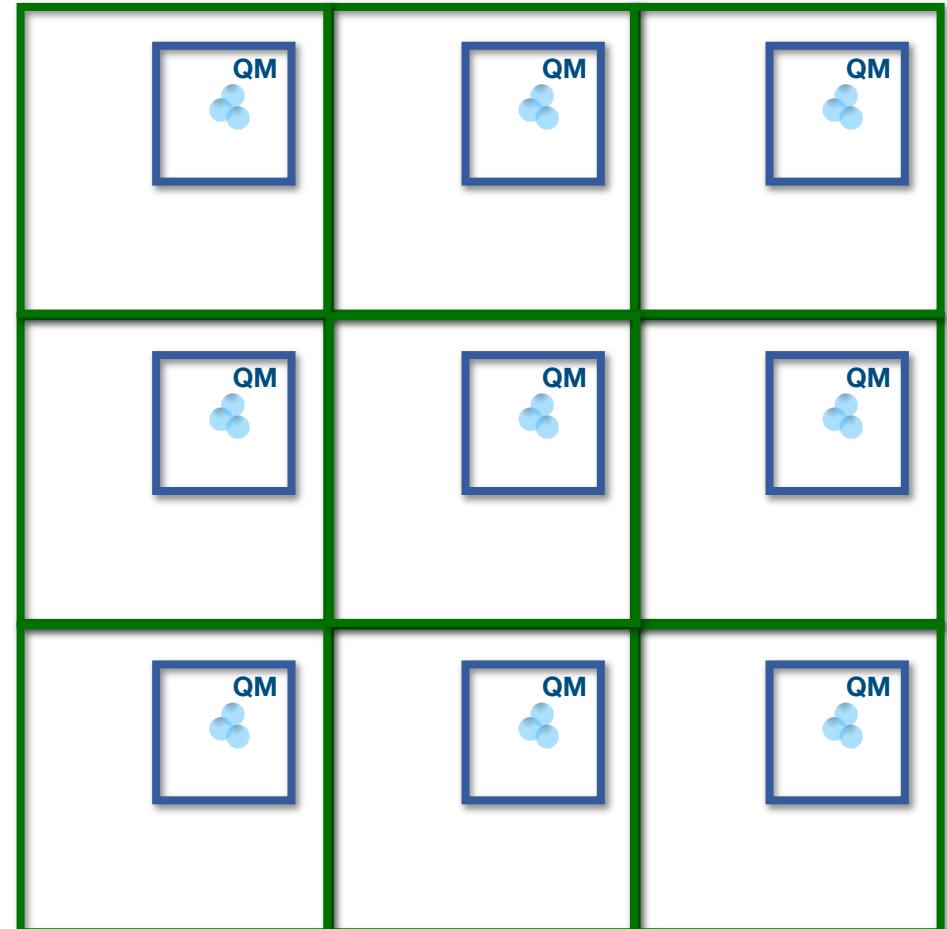
- Unless the QM and MM box have same dimensions the QM images over PBC will have wrong periodicity

Periodic Coupling QM-QM



Blöchl
Scheme

De-coupling &
re-coupling

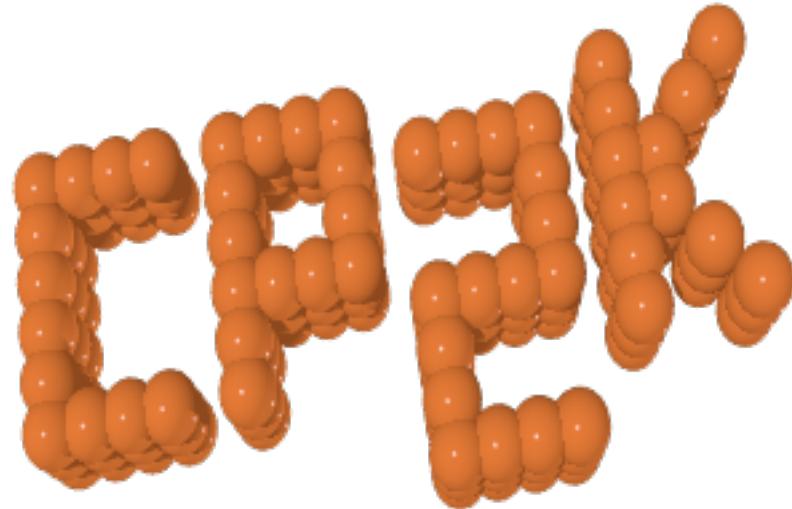


Summary

- GEEP coupling for QM/MM electrostatics
- Fast linear scaling of electrostatic coupling
- Fully periodic QM/MM electrostatics
- Fully periodic QM

CP2K Features

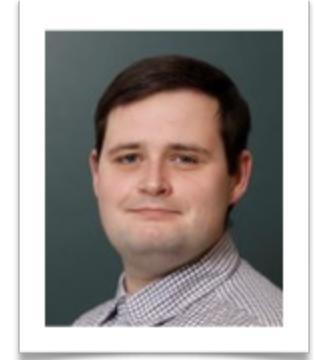
- **Energy and Forces**
- **Optimisation**
 - Geometry optimisation
 - Nudged elastic band
- **Molecular Dynamics**
 - Born-Oppenheimer MD
- **Properties**
 - Atomic charges (RESP, Mulliken..)
 - Spectra
 - Frequency calculations



<https://www.cp2k.org/features>

Hybrid QM/MM Interface

- Hybrid interface where **GROMACS** will drive **CP2K** modules (**Quickstep** and **GEEP**)

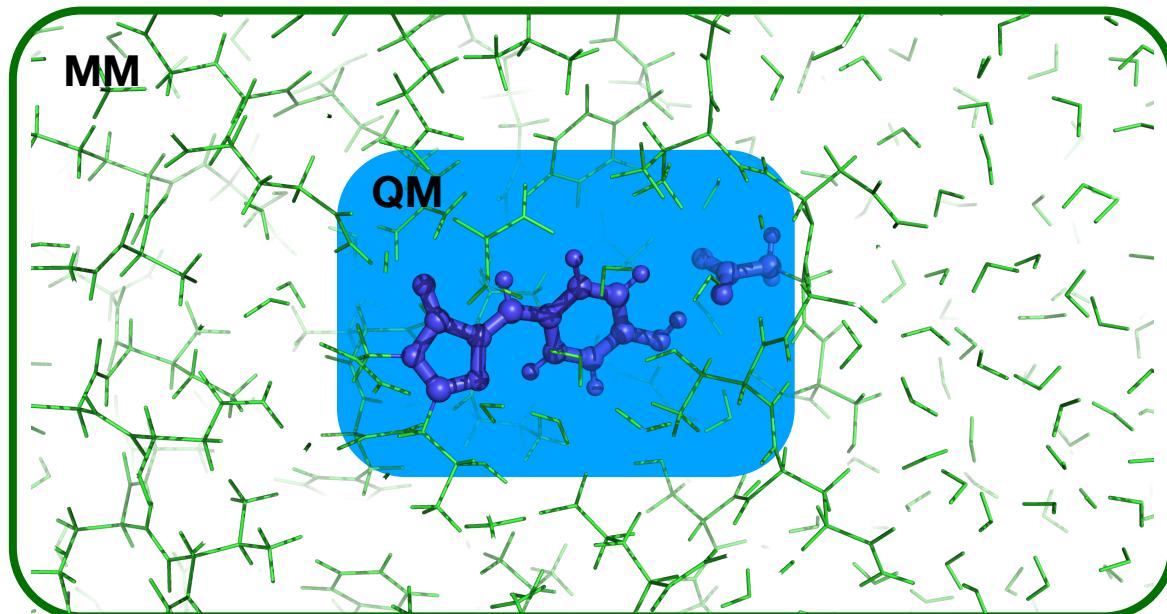


Gerrit Groenhof

Dmitry Morozov

$$E_{Tot} = \underbrace{E_{MM}}_{\text{mdrun}} + \underbrace{E_{QM}}_{\text{Quickstep \& GEEP}} + E_{QM/MM}$$

- Currently under development



Literature

- Lippert, G; Hutter, J; Parrinello, M. [Molecular Physics, 92, 477-487 \(1997\)](#). A hybrid Gaussian and plane wave density functional scheme.
- VandeVondele, J; Krack, M; Mohamed, F; Parrinello, M; Chassaing, T; Hutter, J. [COMPUTER PHYSICS COMMUNICATIONS, 167 \(2\), 103-128 \(2005\)](#). QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach.
- Laino, T; Mohamed, F; Laio, A; Parrinello, M. [JOURNAL OF CHEMICAL THEORY AND COMPUTATION, 1 \(6\), 1176-1184 \(2005\)](#). An efficient real space multigrid QM/MM electrostatic coupling.
- Laino, T; Mohamed, F; Laio, A; Parrinello, M. [JOURNAL OF CHEMICAL THEORY AND COMPUTATION, 2 \(5\), 1370-1378 \(2006\)](#). An efficient linear-scaling electrostatic coupling for treating periodic boundary conditions in QM/MM simulations.

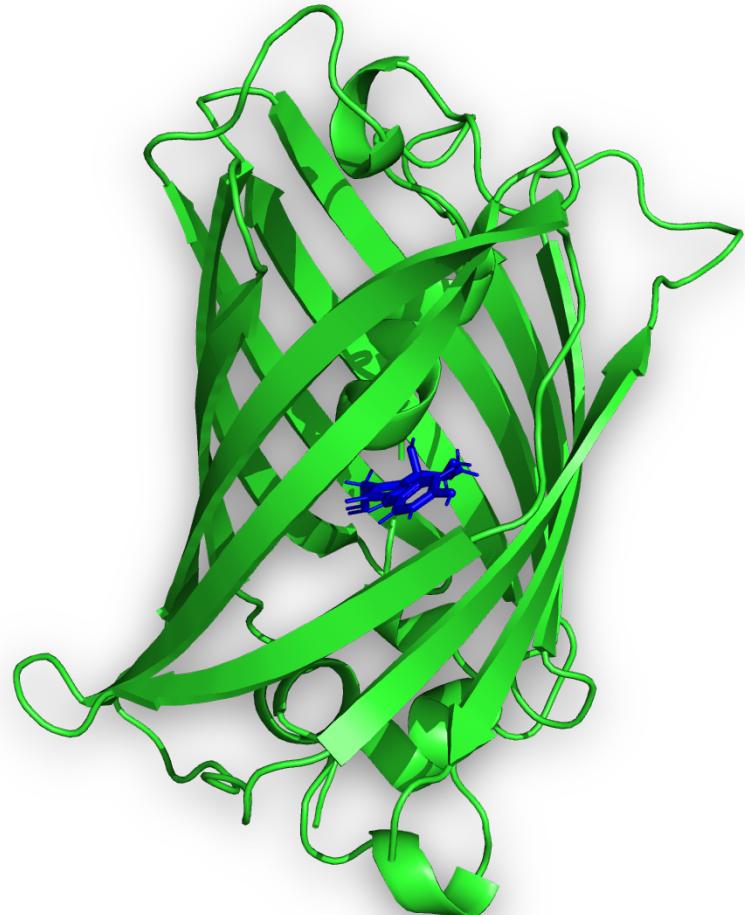
Running CP2K

`~/software/cp2k/exe/local/cp2k.ssmp input.inp > input.out &`

- **sdbg** - Serial with single core (testing)
- **sopt** - Serial with single core
- **ssmp** - multicore with single node
- **pdbg** - multi-node testing
- **popt.** - Parallel with MPI
- **psmp** - Parallel with MPI+OpenMP

-
- ```
graph TD; A[~/software/cp2k/exe/local/cp2k.ssmp input.inp > input.out &] --> B["”.inp” input parameter file
• Input structure file (PDB, XYZ, …)
• Topology Files (Amber, Charmm,..)"]; A --> C["Output structure file (PDB, XYZ, …)
• Wavefunction
• Restart files
• Energy log"]
```

# Exercises



Green Fluorescent Protein (GFP)  
Mutant – S65T/H148D

<https://www.cp2k.org/howto>

- Single pointe energy
- Cutoff convergence
- Geometry optimization
- QM/MM
- Nudged Elastic Band (NEB)

# Exercise: Single point energy-QM

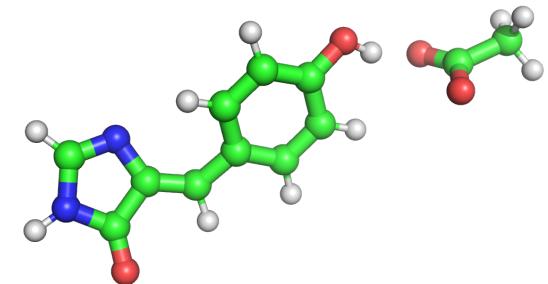
~ /Tutorials /CP2K/Summer-school-2019/single\_point/input.inp

## INPUT

```
&GLOBAL
PROJECT SP
RUN_TYPE ENERGY
PRINT_LEVEL MEDIUM
&END GLOBAL

&FORCE_EVAL
METHOD Quickstep ! Electronic structure method (DFT)
&DFT
CHARGE -1
...
&XC ! Parameters needed to compute the exchange potential
 &XC_FUNCTIONAL PBE
 &END XC_FUNCTIONAL
&END XC
&POISSON
PERIODIC NONE
PSOLVER WAVELET ! system in non-periodic conditions (vacuum)
&END POISSON
&END DFT
```

```
&SUBSYS
&CELL
ABC 20 20 20
PERIODIC NONE
&END CELL
&TOPOLOGY
 &CENTER_COORDINATES
 &END
COORD_FILE_FORMAT xyz
COORD_FILE_NAME init_vacuum.xyz
&END
....
&KIND C
ELEMENT C
BASIS_SET DZVP-MOLOPT-GTH
POTENTIAL GTH-PBE-q4
&END KIND
....
&END SUBSYS
&END FORCE_EVAL
```



29 atoms

# Exercise: Single point energy-QM

~/Tutorials/CP2K/Summer-school-2019/single\_point/input.out

|                                        |                |               |
|----------------------------------------|----------------|---------------|
| Electronic density on regular grids:   | -93.9999992966 | 0.0000007034  |
| Core density on regular grids:         | 92.9999999999  | -0.0000000001 |
| Total charge density on r-space grids: | -0.9999992967  |               |
| Total charge density g-space grids:    | -0.9999992967  |               |

## OUTPUT

|                                                 |                            |
|-------------------------------------------------|----------------------------|
| Overlap energy of the core charge distribution: | 0.00000625979298           |
| Self energy of the core charge distribution:    | -391.34137466423209        |
| Core Hamiltonian energy:                        | 118.27936011171798         |
| Hartree energy:                                 | 158.57689365278532         |
| Exchange-correlation energy:                    | -44.21542888876060         |
| <br>                                            |                            |
| <b>Total energy:</b>                            | <b>-158.70054352869639</b> |

# Exercise: CUTOFF Convergence

- Too low CUTOFF → Only coarse grids → inaccurate calculations
- Too low REL\_CUTOFF → All gaussians will be mapped to the coarsest grid

```
&FORCE_EVAL
 METHOD Quickstep
 &DFT
 ...
 &MGRID
 NGRIDS 4
 CUTOFF 350
 REL_CUTOFF 60
 &END MGRID
```

! or QMMM

E<sub>cut</sub>: Plane wave cutoff (Rydberg) for finest grid

$$E_{cut}^i = \frac{E_{cut}^1}{\alpha^{(i-1)}}, \quad i = 1 \dots N$$

Cutoff for mapping density in Gaussian basis onto the grids

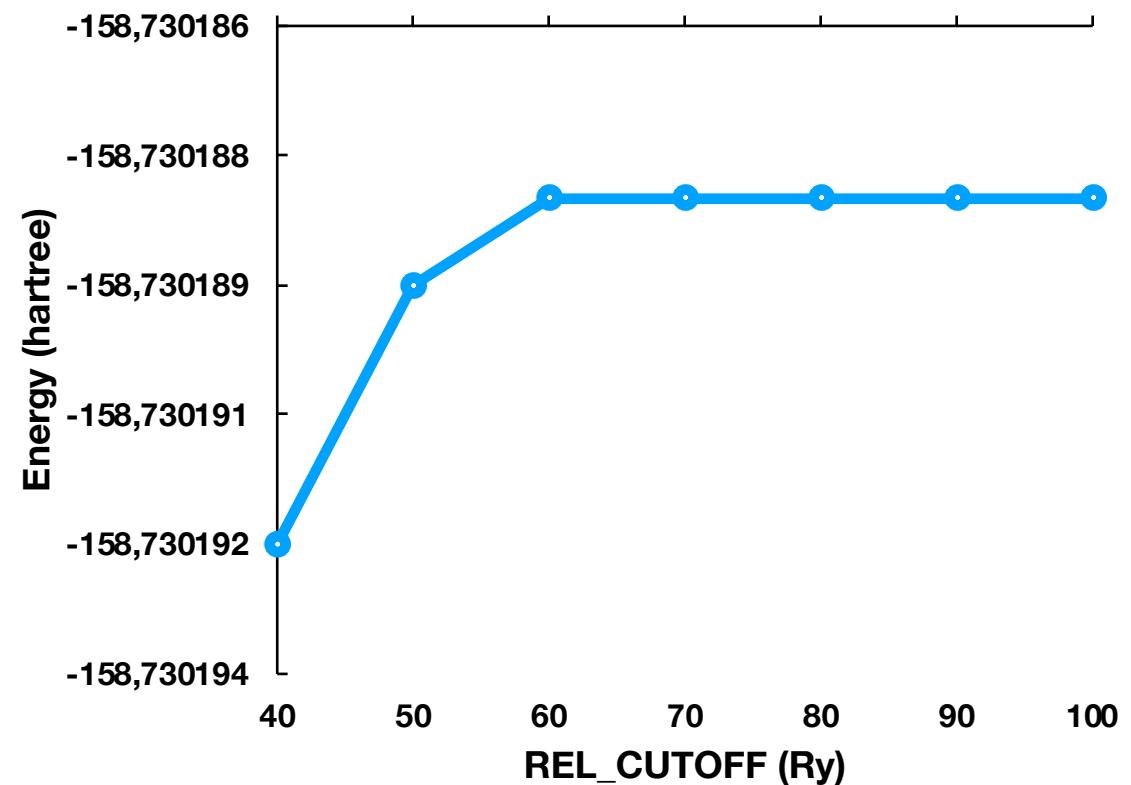
[https://www.cp2k.org/howto:converging\\_cutoff](https://www.cp2k.org/howto:converging_cutoff)

# Exercise: CUTOFF Convergence

~/Tutorials/CP2K/Summer-school-2019/cutoff\_conv/rel\_cutoff/relcutoff\_data.dat

```
Grid rel_cutoff vs total energy
CUTOFF = 300
Rel_Cutoff (Ry) | Total Energy (Ha)
```

|        |                 |        |        |        |       |
|--------|-----------------|--------|--------|--------|-------|
| 40.00  | -158.7301924360 | 404322 | 175605 | 115772 | 90338 |
| 50.00  | -158.7301885768 | 443764 | 163386 | 110037 | 68850 |
| 60.00  | -158.7301883345 | 464321 | 166975 | 98795  | 55946 |
| 70.00  | -158.7301883346 | 502823 | 147207 | 90268  | 45739 |
| 80.00  | -158.7301883345 | 515555 | 145052 | 84243  | 41187 |
| 90.00  | -158.7301883345 | 532292 | 140983 | 76887  | 35875 |
| 100.00 | -158.7301883345 | 551766 | 125435 | 81741  | 27095 |



# Exercise: CUTOFF Convergence

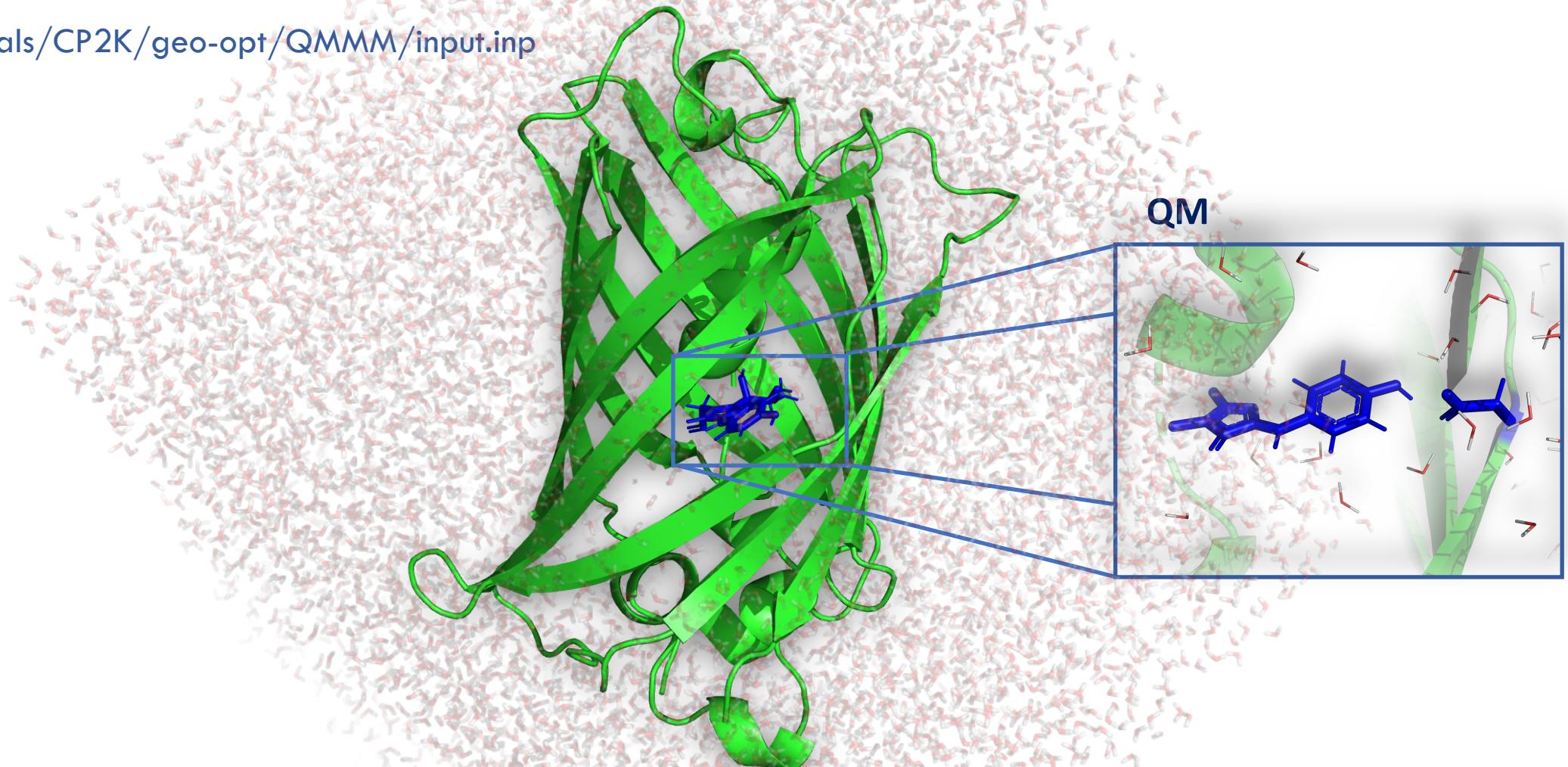
~/Tutorials/CP2K/Summer-school-2019/cutoff\_conv/rel\_cutoff/relcutoff\_data.dat

```
Grid cutoff vs total energy
REL_CUTOFF = 60
Cutoff (Ry) | Total Energy (Ha) | NG on grid 1 | NG on grid 2 | NG on
grid 3 | NG on grid 4
 150.00 -158.7268552410 579927 115772 73077 17261
 200.00 -158.7303067388 532292 140983 76887 35875
 250.00 -158.7299138948 507673 143819 88806 45739
 300.00 -158.7301883345 464321 166975 98795 55946
 350.00 -158.7291445839 446184 167408 108467 63978
 400.00 -158.7289938973 419922 176070 113065 76980
 450.00 -158.7287833145 404322 175605 115772 90338
 500.00 -158.7288760444 377350 182963 123366 102358
 550.00 -158.7286769177 362839 186479 127883 108836
 600.00 -158.7286914646 355451 176841 140983 112762
 650.00 -158.7287572297 347625 170486 144920 123006
 700.00 -158.7287572273 332138 179433 143532 130934
 800.00 -158.7287336794 576610 412859 295889 268775
 900.00 -158.7287291347 548614 371083 329455 304981
```

**0.0001 hartree = 0.06275 kcal/mol**

# Exercise: Geometry Optimization

~/Tutorials/CP2K/geo-opt/QMMM/input.inp



# Exercise: Geometry Optimization

~/Tutorials/CP2K/Summer-school-2019/geo-opt/QMMM/input.inp

## INPUT

```
&GLOBAL
 PROJECT GEO_OPT-INIT
 RUN_TYPE GEO_OPT
 PRINT_LEVEL LOW
&END GLOBAL

&FORCE_EVAL
 METHOD QMMM
 ...
 &QM_KIND O
 MM_INDEX 970 979 2306 2307
 &END QM_KIND
 &QM_KIND N
 MM_INDEX 962 980
 &END QM_KIND

 &CELL
 ABC 20 20 20
 PERIODIC XYZ
 &END CELL
 &COUPL GAUSS
 USE_GEEP_LIB 12
 &PERIODIC
 &MULTIPOLE ON
 &END
 &END PERIODIC
 &QM_KIND O
 MM_INDEX 970 979 2306 2307
 &END QM_KIND
 &QM_KIND N
 MM_INDEX 962 980
 &END QM_KIND

 &MOTION
 &GEO_OPT
 OPTIMIZER LBFGS
 MAX_ITER 5000
 MAX_DR 5.0E-03
 MAX_FORCE 5.0E-03
 RMS_DR 5.0E-03
 RMS_FORCE 5.0E-03
 &END
 &PRINT
 ...
 &END MOTION
```

# Exercise: Geometry Optimization

~/Tutorials/CP2K/Summer-school-2019/geo-opt/QMMM/input.inp

Optimized geometry → gfp-s65t\_h148d-init-opt.pdb

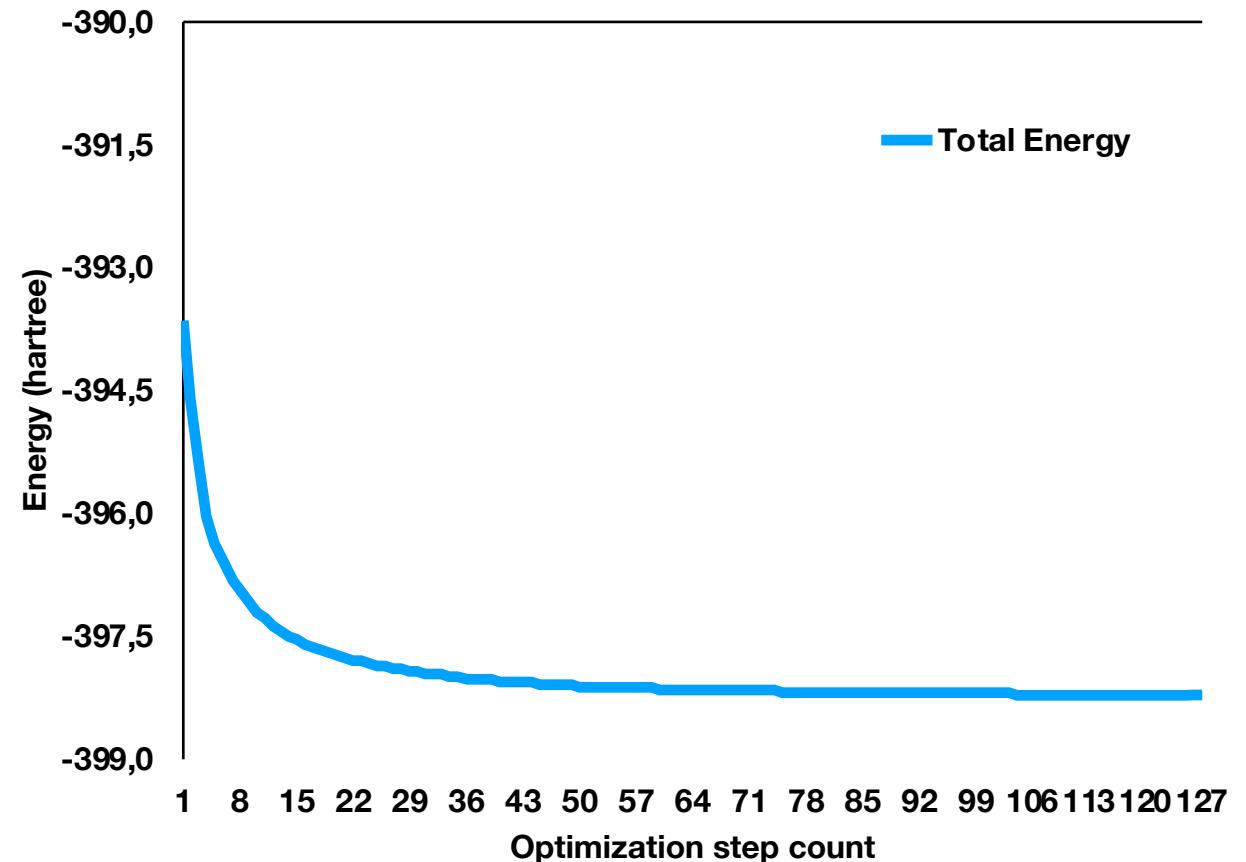
----- Informations at step = 123 -----

Optimization Method = LBFGS  
Total Energy = -398.2149383490  
Real energy change = -0.0004052117  
Decrease in energy = YES  
Used time = 11.339

Convergence check :

Max. step size = 0.0035795584  
Conv. limit for step size = 0.0050000000  
Convergence in step size = YES  
RMS step size = 0.0001035692  
Conv. limit for RMS step = 0.0050000000  
Convergence in RMS step = YES  
Max. gradient = 0.0023741218  
Conv. limit for gradients = 0.0050000000  
Conv. in gradients = YES  
RMS gradient = 0.0000365315  
Conv. limit for RMS grad. = 0.0050000000  
Conv. in RMS gradients = YES

## OUTPUT



# Exercise: QMMM

~/Tutorials/CP2K/Summer-school-2019/qmmm/qmmm.inp

```
&GLOBAL
PROJECT input
RUN_TYPE MD
PRINT_LEVEL LOW
&END GLOBAL
```

```
&FORCE_EVAL
METHOD QMMM
...
&QMMM
&MM
&SUBSYS
&END FORCE_EVAL
```

```
&MOTION
&MD
ENSEMBLE NVT
TIMESTEP [fs] 1
STEPS 1000 ! 1ps
TEMPERATURE 298
&THERMOSTAT
...
&END MD
```

# Exercise: QMMM

~/Tutorials/CP2K/Summer-school-2019/qmmm/input-pos-1.pdb

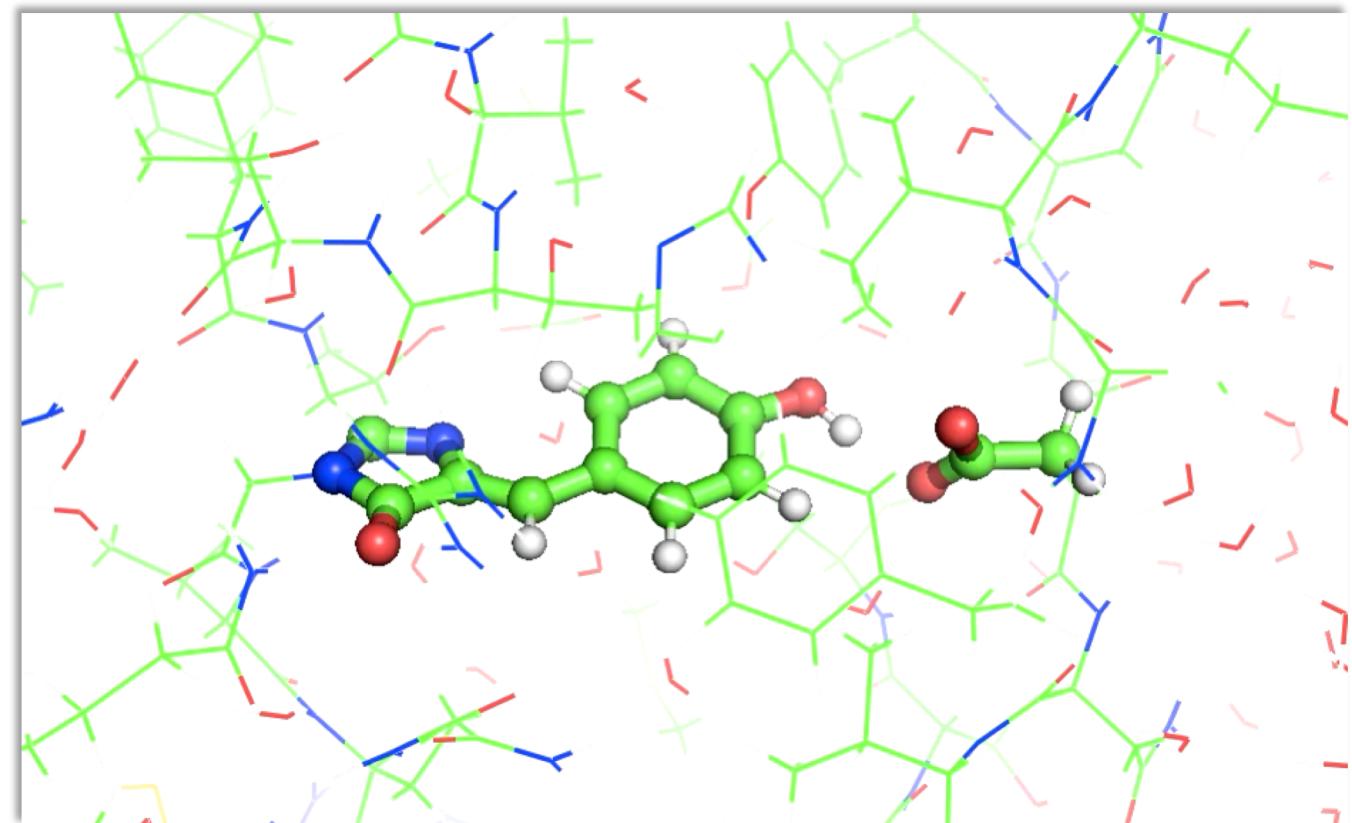
**Objective:** Proton delocalization

**QM subsystem :** Chromophore +

Asp148

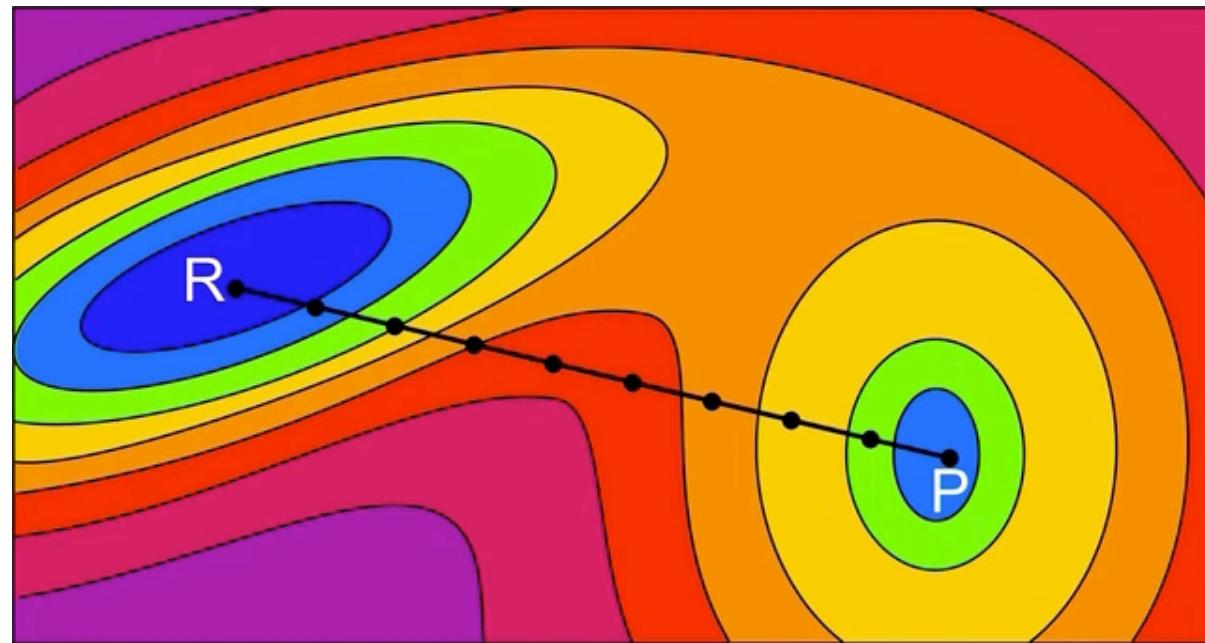
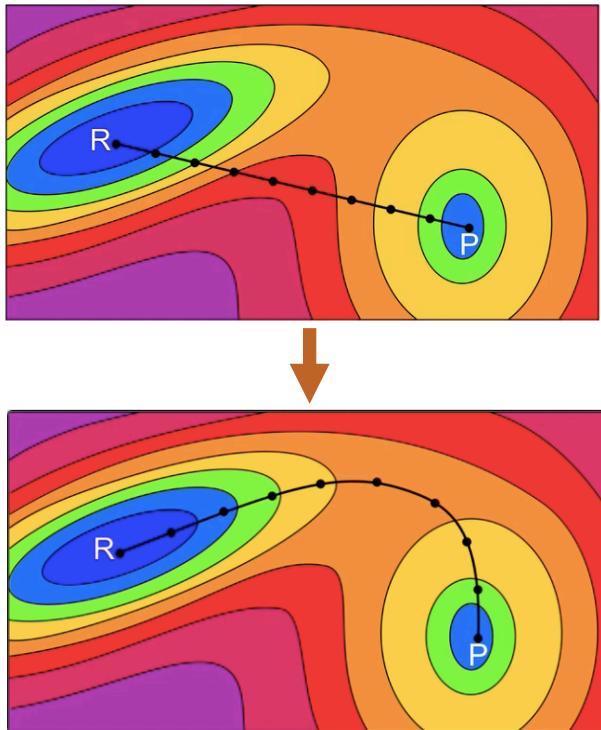
**Functional:** PBE

**Basis:** DZVP-MOLOPT-GTH



# Nudged Elastic Band Method (NEB)

- Find a minimum energy pathway connecting two local minima
- String constrained points w.r.t adjacent configuration points
- Gradients perpendicular to the spring nudge the band



# Exercise: Nudged Elastic Band calculation

~/Tutorials/CP2K/Summer-school-2019/neb/ethane/input.inp

```
&GLOBAL
 RUN_TYPE BAND
 ...
&END GLOBAL

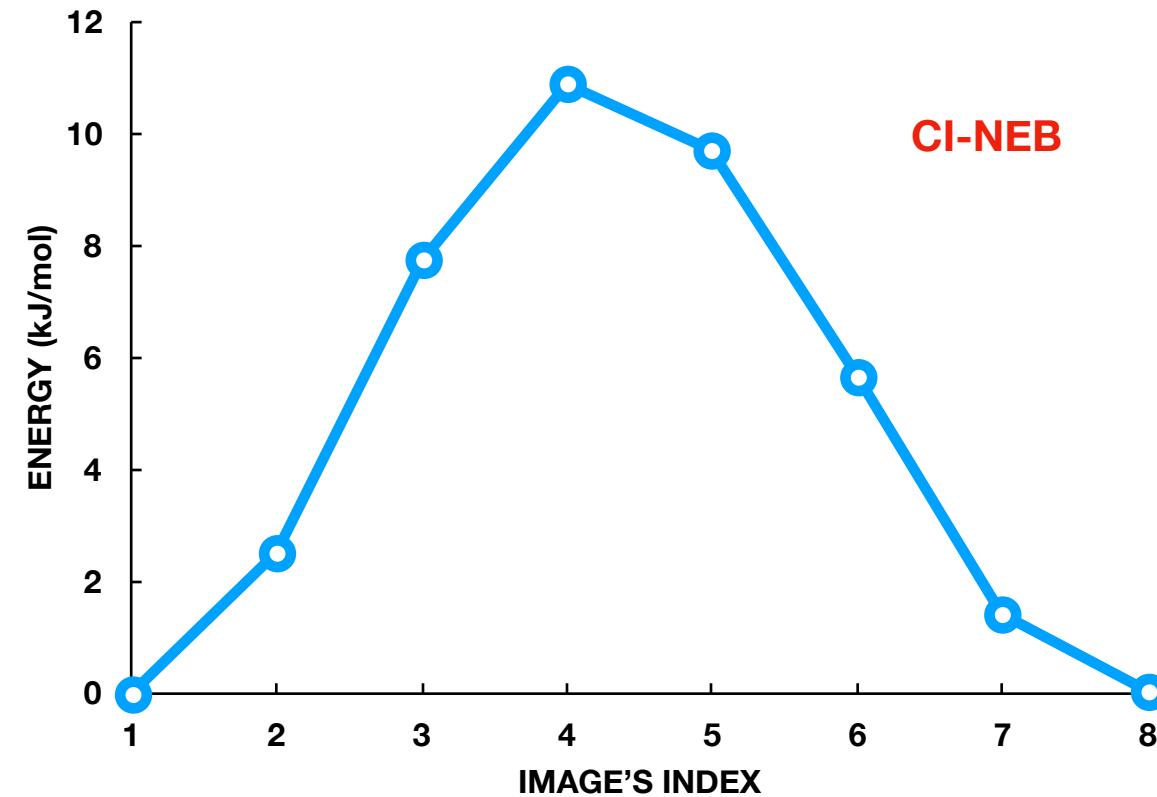
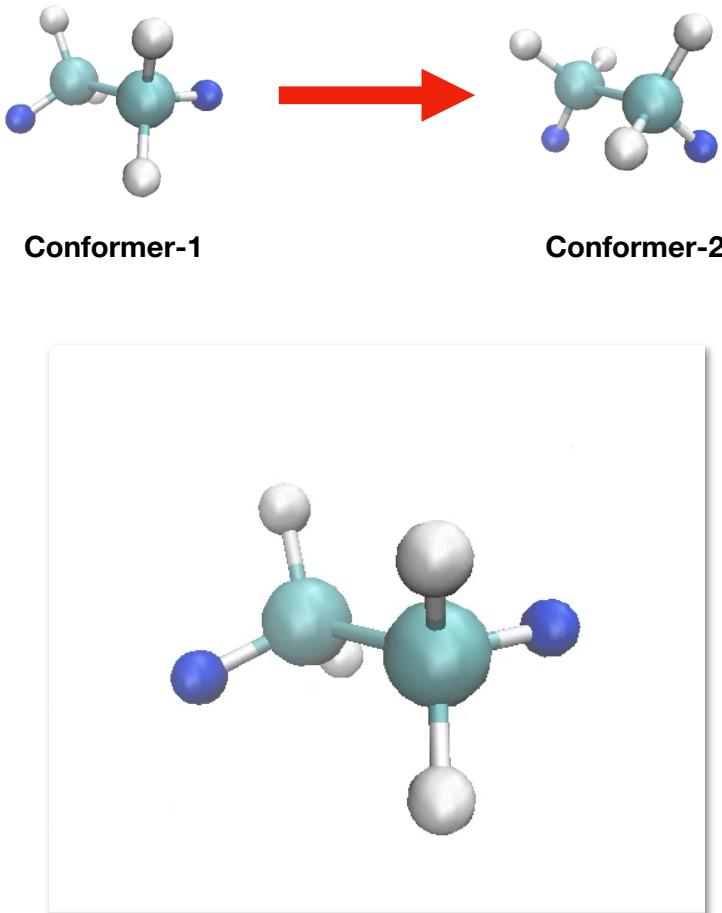
&MOTION
 &BAND
 BAND_TYPE CI-NEB
 NUMBER_OF_REPLICA 8
 K_SPRING 0.05
 ROTATE_FRAMES TRUE
 ALIGN_FRAMES TRUE

 &CONVERGENCE_CONTROL
 MAX_FORCE 0.0010
 RMS_FORCE 0.0050
 &END
```

```
&OPTIMIZE_BAND
 OPT_TYPE DIIS
 OPTIMIZE_END_POINTS FALSE
&END OPTIMIZE_BAND

&REPLICA
 COORD_FILE_NAME initial-mol.xyz
&END
&REPLICA
 COORD_FILE_NAME ts-mol.xyz
&END
&REPLICA
 COORD_FILE_NAME final-mol.xyz
&END
&REPLICA
 ...
&END BAND
&END MOTION
```

# Exercise: NEB - Ethane



# Questions?



## BioExcel Partners 2019



NORMAN  
CONSULTING

Utrecht University



EMBL-EBI



Horizon 2020  
European Union Funding  
for Research & Innovation

BioExcel is funded by the European Union  
Horizon 2020 program under grant  
agreements 675728 and 823830.