ChemShell#
www.chemshell.org

L6

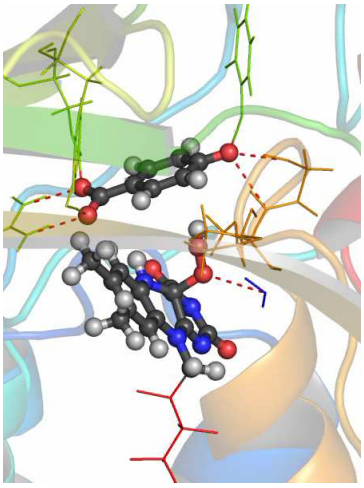
Geometry Optimization and Dynamics with ChemShell

Dr. Johannes Kästner
Computational Science and Engineering Department
STFC Daresbury Laboratory
www.cse.clrc.ac.uk/ccg/j.kaestner

ChemShell#
www.chemshell.org

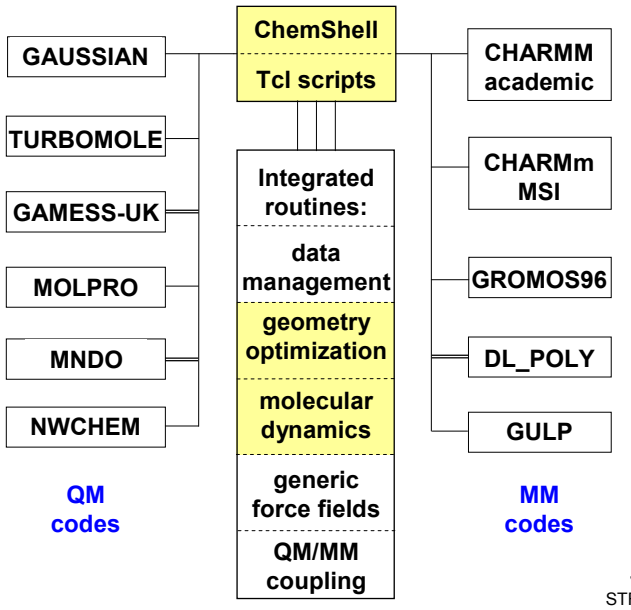
Outline

- Introduction
- Optimisation strategies
  - Coordinate systems
  - Optimisation algorithms
- Molecular Dynamics: Free-energy calculations
  - Umbrella sampling
  - Thermodynamic integration
- System preparation (biological systems)
- Conclusion



ChemShell#
www.chemshell.org

ChemShell

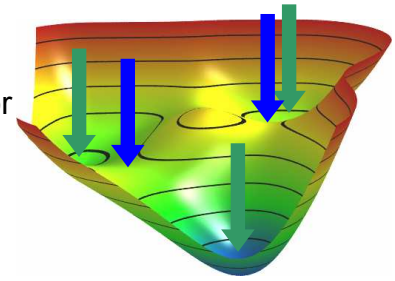


Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Why Optimisation?

- Find minima on the potential energy surface
- Structure: equilibrium structure at stationary points (reactants or products)
- Energy differences: $\Delta E \approx \Delta U$
- Saddle points: transition states
  - Activation energies
  - Reaction rates
- All optimisation results: T=0



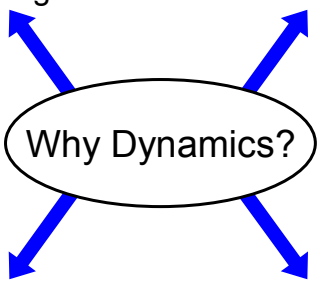
Optimisation – Potential Energy

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Free-energy differences:
- Umbrella sampling
- Thermodynamic integration

Finite-temperature effects



To establish a minimum as being “relatively global”

Global averages

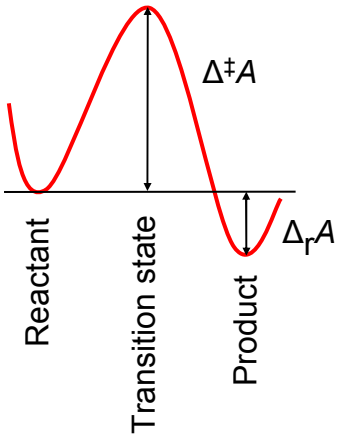
Molecular Dynamics – Free Energy

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Free Energy

- Driving force of chemical reaction
- Helmholtz energy $A = U - TS$ at constant volume
- Gibbs energy $G = H - TS$ at constant pressure
- $\Delta_r A$ defines the equilibrium constant: $\Delta_r A = -RT \ln K$
- $\Delta^{\ddagger} A$ determines the reaction rate



Johannes Kästner
STFC Daresbury Lab

Optimisation in ChemShell

ChemShell#
www.chemshell.org

Established Standards

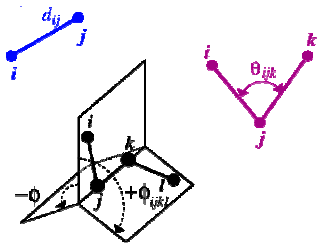
- Z-matrix optimisation with a quasi-Newton optimiser (standard in quantum chemistry)
- Cartesian coordinates with a conjugate gradient optimiser (standard in MD)

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Coordinate Systems

- Cartesian coordinates:
  - O(N^0)
  - Highly coupled (torsions)
- Z-Matrix:
  - O(N)
  - Less coupled
  - Biased (different ways to set up a Z-matrix)
- Redundant internal coordinates
  - O(N^3)
  - Less coupled
  - Nowadays standard for small systems

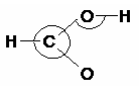


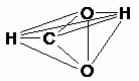
Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

HDLC coordinates

- Hybrid DeLocalised internal Coordinates
- Divide and conquer: large molecule is split into fragments
- Internal coordinates within each fragment
- Coupling of the fragments via Cartesian coordinates
- O(N)
- Less coupled





- Available in ChemShell in **`HDLCopt`**

Billeter, Turner, Thiel *Phys. Chem. Chem. Phys.* **2**, 2177 (2000)

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Optimisation Algorithms

- Determination of the search direction
- First-order methods: gradient
  - Steepest descent
  - Conjugate gradient
  - Damped dynamics
- Second order methods: gradient and Hessian
  - Quasi-Newton methods
  - RFO
  - BFGS (accumulated Hessian)
  - L-BFGS (accumulated Hessian, O(*N*))

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Transition-State Search Algorithms

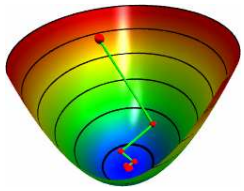
- First order saddle points:
  - Zero gradient
  - One negative eigenvalue of the Hessian, all others positive
- P-RFO (requires accurate Hessian)
- Dimer method (no Hessian required)
- Newton-Raphson (finds any extremum)

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Optimisers in ChemShell

Four optimisers available:



- **`Newopt`**: flexible optimizer.

  Methods: BFGS, (P)RFO, DIIS,
  conjugate gradient, steepest descent, …

  Coordinate systems: Cartesian or z-matrix

- **`HDLCopt`**: fastest for large systems, O(*N*)

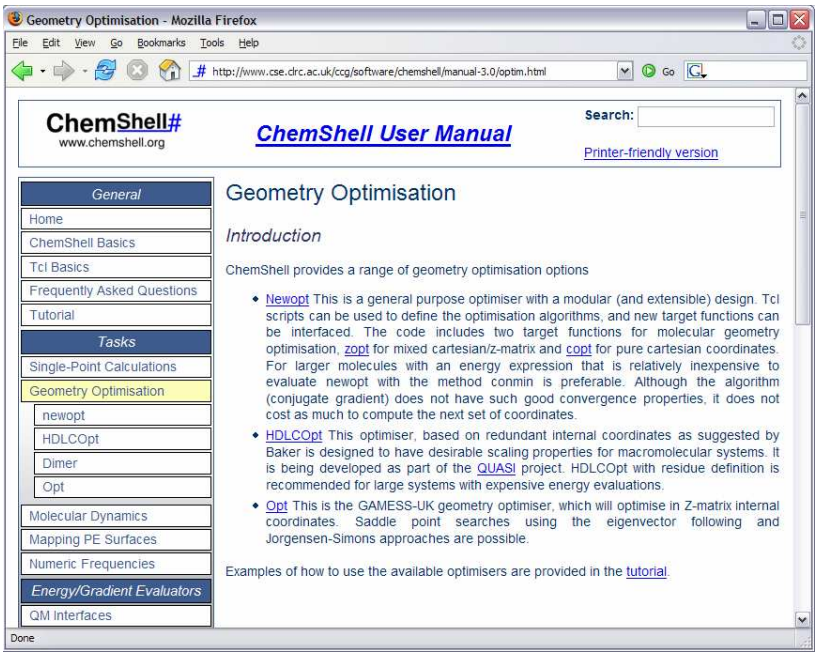
  Methods: L-BFGS, P-RFO (for transition state search)

  Coordinate systems: hybrid delocalized internal coordinates

- **`dimer`**: TS search with the dimer method

- **`opt`**

Johannes Kästner
STFC Daresbury Lab



ChemShell#
www.chemshell.org

Small Molecules: newopt – zopt

Optimisation of a molecule given as z-matrix using the BFGS method:

```
newopt \
    function= zopt : { \
       zmatrix= water.zz4 \
       theory= mndo } \
    method= bfgs
```

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Large Molecules: HDLCopt

- Small molecule, redundant internals:

```
hdlcopt coords=c \
residues= [ res_selectall coords=c ] \
theory= mndo
```

- Large molecule, provided as pdb file, HDLCs:

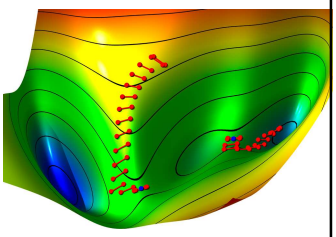
```
set residues [ pdb_to_res 1PBE.pdb ]
hdlcopt coords=c \
   residues= $residues \
   constraints= {{bond 3 5} {angle 3 5 6}} \
   theory= mndo
```

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Dimer Method for TS search

- Dimer: two images of the system, constant distance
- Rotation: difference of the forces
- Movement:
  - In dimer direction: against the force
  - Perpendicular to dimer direction: along the force
- Converges to first-order saddle points
- No Hessian required: for large systems



Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Dimer method in ChemShell

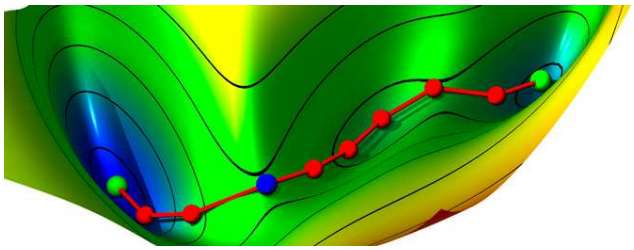
```
dimer coords=c mode=3 \
      maxcyc=100 \
      theory= gamess : $gamess_args
```

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Outlook: Nudged Elastic Band

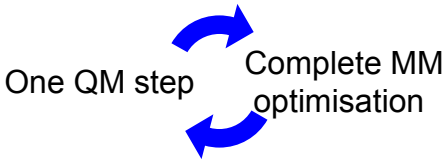
- Multiple images, connected by “springs”
- Converges to the minimum-energy path
- Climbing image: transition state
- Evolved, but can cover difficult reactions
- Implementation in ChemShell is in “beta stage”



Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Exploiting QM/MM capabilities: Micro-iterative QM/MM optimisation



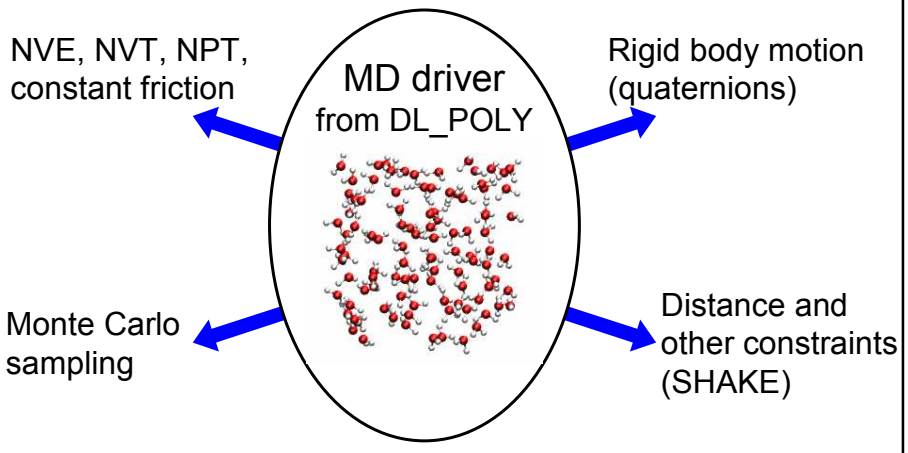
- Electrostatic embedding: ESP charges calculated on the fly
- Optimisation effort becomes independent of the system size
- Saves a factor of 2–10 in CPU time
- **`hdlcopt`**
  `inner_atoms= { 1 2 14 15 } mcore=true`

Kästner, S. Thiel, Senn, Sherwood, W. Thiel, *J. Chem. Theory Comput.* online (2007)

Johannes Kästner
STFC Daresbury Lab

Molecular Dynamics in ChemShell

ChemShell#
www.chemshell.org



Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Free-Energy Methods I

- Fixed constraint
  - Thermodynamic integration: sampling the mean force
- Continuously changing constraint
  - Slow growth: sampling the mean force
  - Fast growth: fast changing constraint, exponential average of the energy change
- Free-energy perturbation: instantaneous changes, exponential average of the energy change

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Free-Energy Methods II

- Fixed restraint (bias)
  - Umbrella sampling: sampling the distribution of the reaction coordinate
- Continuously changing restraint
  - Steered molecular dynamics
- Hessian: quadratic approximation, anharmonicities neglected
- …

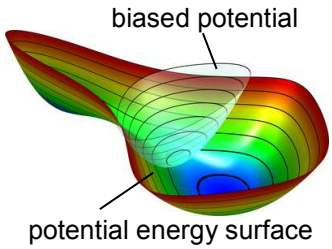
Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Umbrella Sampling

- A number of windows, each with a restraint (bias) centred at different points along the reaction coordinate
- Distribution along the reaction coordinate is sampled
- Analysis by
  - WHAM or
  - Umbrella Integration

  provides the free-energy difference
- Restraints available in ChemShell: bond, angle, torsion, difference between 2, 3, and 4 bond lengths
- Use the NHC thermostat: **`nosehoover=4`**

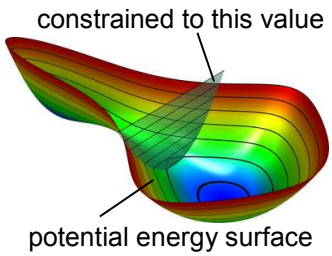


Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Thermodynamic Integration

- A number of windows, each constrained to a particular value of the reaction coordinate
- Shake-like constraint (Lagrange multiplier)
- Force on the constraint is averaged
- Integration of these mean forces provides an approximate free-energy change
- Metric tensor corrections
- Constraints available in ChemShell: bond, torsion, difference between 2 bond lengths



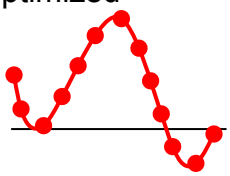
Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Exploiting QM/MM functionality: QM/MM-FEP

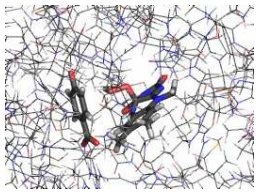
Reaction profile:

- Full QM/MM calculations
- QM and MM atoms optimized



Sampling:

- Frozen QM part
- MM-dynamics



$$\Delta A = \Delta E_{qm} + \Delta A_{qm/mm}$$

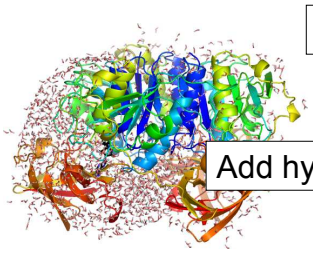
Zhang, Liu, Yang *J. Chem. Phys.* **112**, 3483 (2000)

Johannes Kästner
STFC Daresbury Lab

System Preparation

ChemShell#
www.chemshell.org

Setup of QM/MM calculations of proteins



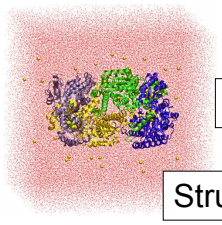
Starting point: **pdb** file

Check errors: *What If*

Add hydrogen

Pronation state (*PropKa*)

Solvate the protein in water



Equilibrate solvate

Equilibrate the whole (MM) system

Structure optimisation at the QM/MM level

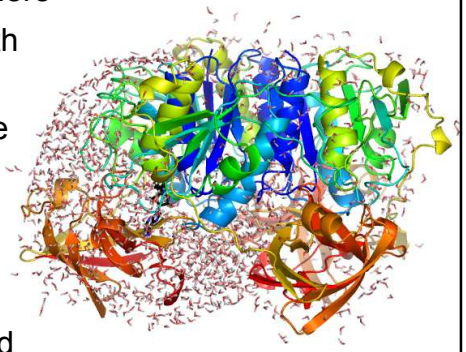
Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Fluorinase: Setup

QM/MM setup:

- Solvated in a 25 Å water sphere
- QM method: DFT (BP86) with Turbomole
- MM method: CHARMM force field with DL_POLY in ChemShell
- Electrostatic embedding (charge shift)
- HDLCopt, shell of 8 Å around the active site optimized



Senn, O'Hagan, Thiel *J. Am. Chem. Soc.* **127**, 13643 (2005)

Johannes Kästner
STFC Daresbury Lab

ChemShell#
www.chemshell.org

Conclusions

Optimisation:

- Small molecules: newopt – zopt
- Large systems: HDLCopt (possibly micro-iterative)

Free-energy sampling:

- Umbrella sampling (NHC thermostat)

Further information:

www.chemshell.org/**manual**

www.chemshell.org/**tutorial**

Johannes Kästner
STFC Daresbury Lab