QM/MM

Dr. Marek Freindorf
Southern Methodist University
QM/MM Workshop
December 2014

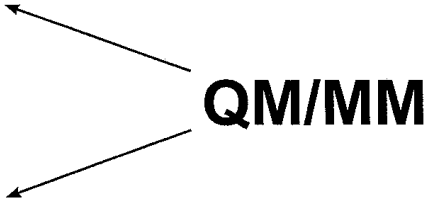
2013 Nobel Prize in Chemistry
M. Karplus, M. Levitt, A. Warshel

Warshel, A., Karplus, M. "Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization" *J. Am. Chem. Soc.* **92** (1972) 5612–5625.

Warshel, A., Levitt, M. "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme". *J. Mol. Biol.* **103** (1976) 227–249.

QM - Calculations (electrons involved)

Advantage: Very accurate, based on first principles (ab initio, DFT - there are not empirical parameters involved)
Disadvantage: Time consuming, limited to small molecular systems (~ 100 atoms, Born-Oppenheimer approximation)



MM - Calculations (electrons not involved)

Advantage: Very fast, capable to calculate entire proteins, DNA or solutions (~ $4*10^{12}$ atoms, molecular dynamics)
Disadvantage: Less accurate, based on empirical parameters, not capable to calculate chemical reactions (bond breaking or formation)

QM part (an active site of a protein) is calculated at the quantum-mechanical level of theory

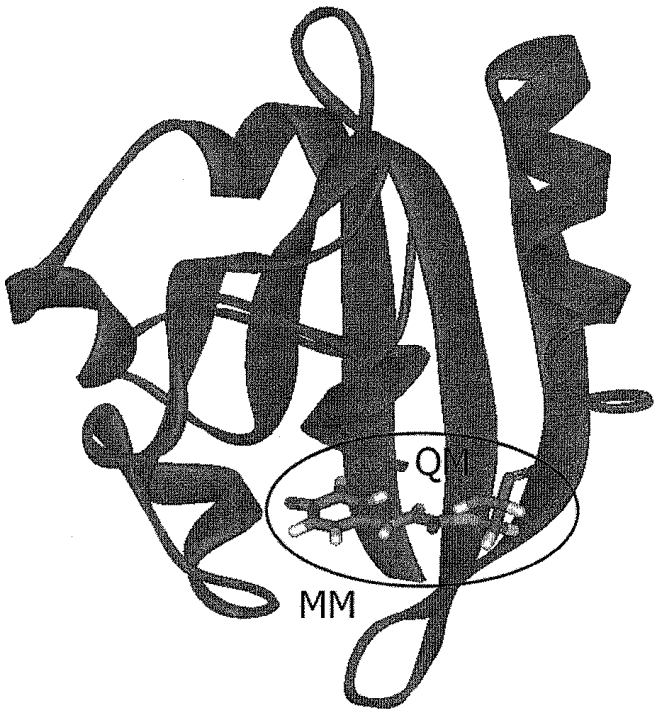
MM part (the rest of the protein) is calculated at the molecular-mechanical level of theory

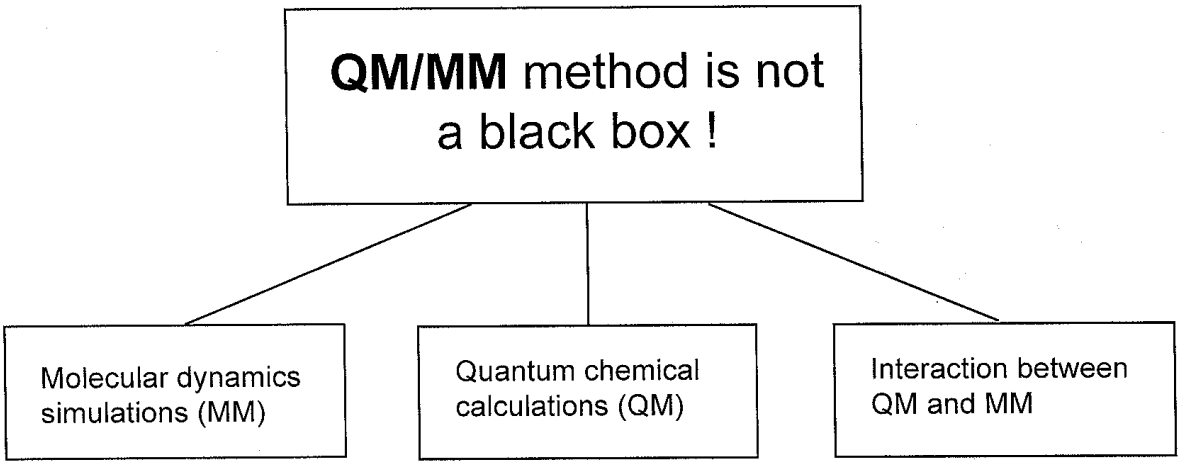
The idea: the QM calculation is done in the MM enviromental perturbation (point charges with vdW spheres)

$$E_{QM/MM} = E_{ele} + E_{vdW}$$

$$E_{ele} = -\sum \frac{q_s}{r_{si}} + \sum \frac{q_s Z_m}{R_{sm}}$$

$$E_{vdW} = \sum 4\varepsilon_{sm} \left[\left(\frac{\sigma_{sm}}{R_{sm}} \right)^{12} - \left(\frac{\sigma_{sm}}{R_{sm}} \right)^{6} \right]$$




QM/MM method is not a black box !
Molecular dynamics simulations (MM)
Quantum chemical calculations (QM)
Interaction between QM and MM

AMBER (Assisted Model Building with Energy Refinement)
(http://ambermd.org/)

What is AMBER

- A collective name for a suite of programs that allow users to carry out molecular dynamic simulations
- A set of molecular mechanical force fields for the simulation of biomolecules
- Along with CHARMM, AMBER is the most used molecular mechanical software for protein dynamics

What can AMBER do

- Classical molecular dynamics simulations (NVT, NPT, etc)
- Explicit Solvent Models with particle-mesh Ewald sum (PME)
- Implicit Solvent models with Poisson-Boltzmann
- Generalized Born approach
- Enhanced sampling (replica exchange MD, Locally Enhanced Sampling)
- Free energy calculation (MM/PBSA, etc.)
- Structural and trajectory analysis
- .....
- Well parallelized up to 64 processors

Molecular dynamics simulations

- Calculates the motion of atoms in a molecular system using Newtonian dynamics, to determine the net force and acceleration experienced by each atom
- Each atom i at position r_i, is treated as a point with a mass m_i, a fixed charge q_i, and and a fixed van der Walls potential
- Interactions between atoms (bonded and non-bonded) are calculated according to a particular force field, which is specific for each molecular mechanical program

$$E_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]$$

$$+ \sum \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right]$$

AMBER parameters

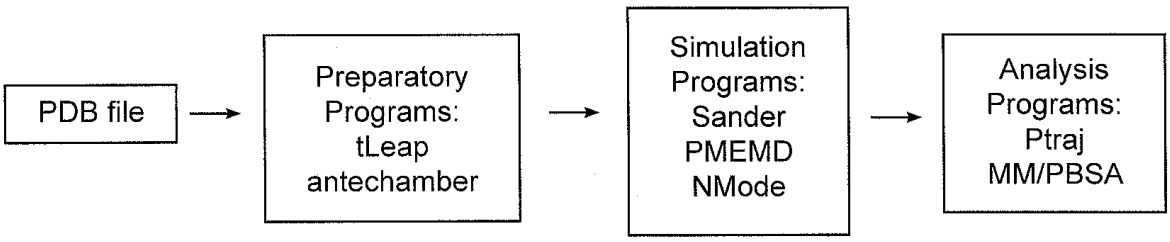
/grid/software/amber/amber14/dat/leap/parm/parm99EP.dat

```
PARM99 for DNA,RNA,AA, organic molecules, TIP3P wat. Polariz.& LP incl.23/06/99
C  12.01         0.616  !            sp2 C carbonyl group
CA 12.01         0.360               sp2 C pure aromatic (benzene)
CB 12.01         0.360               sp2 aromatic C, 5&6 membered ring junction
CC 12.01         0.360               sp2 aromatic C, 5 memb. ring HIS
CD 12.01         0.360               sp2 C atom in the middle of: C=CD-CD=C

C   H   HO  N   NA  NB  NC  N2  NT  N2  N3  N*  O   OH  OS  P   O2
OW-HW  553.0    0.9572    ! TIP3P water
HW-HW  553.0    1.5136      TIP3P water
C -C   310.0    1.525       Junmei et al, 1999
C -CA  469.0    1.409       JCC,7,(1986),230; (not used any more in TYR)
C -CB  447.0    1.419       JCC,7,(1986),230; GUA
C -CM  410.0    1.444       JCC,7,(1986),230; THY,URA
C -CT  317.0    1.522       JCC,7,(1986),230; AA
C -N   490.0    1.335       JCC,7,(1986),230; AA

HW-OW-HW   100.0      104.52    TIP3P water
HW-HW-OW     0.0      127.74    (found in crystallographic water with 3 bonds)
C -C -O     80.0      120.00    Junmei et al, 1999 acrolein
C -C -OH    80.0      120.00    Junmei et al, 1999
CA-C -CA    63.0      120.00    changed from 85.0  bsd on C6H6 nmodes; AA
CA-C -OH    70.0      120.00    AA (not used in tyr)
CB-C -NA    70.0      111.30    NA
```

0.1367

AMBER MD protocol

- Preparation of AMBER residue parameter and topology file (hydrogen addition, water solvation, neutralization by counter-ions)
- Initial energy minimization
- Heating dynamics to a temperature 300K
- Equilibration dynamics at a constant temperature
- Production dynamics (collecting data)
- Analyzing equilibrium trajectory
- Annealing dynamics (cooling) to a temperature 0K



ONIOM (Own N-layered Integrated Orbital and Molecular Mechanics)
(http://www.gaussian.com/g_whitepap/oniom_technote.htm)

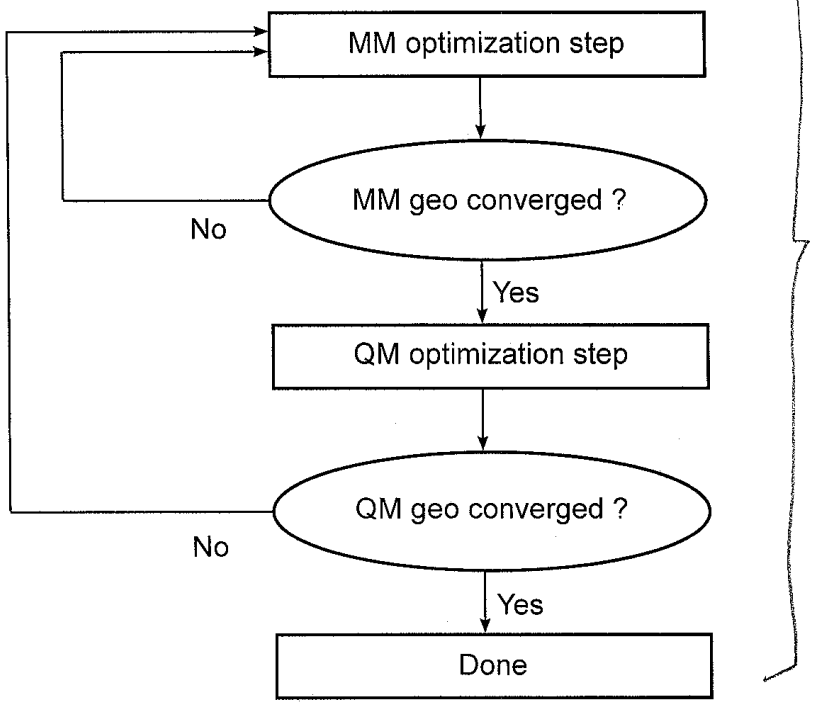
What is ONIOM

- A computational technique models large molecules by defining two or three layers within the structure that are treated at different levels of accuracy
- Usually the High level is a quantum-mechanical method (for example DFT) and the Low level is a molecular-mechanical method (for example AMBER), which leads to the combined QM/MM approach.
- Distributted as a part of Gaussian programing suite

What can ONIOM do

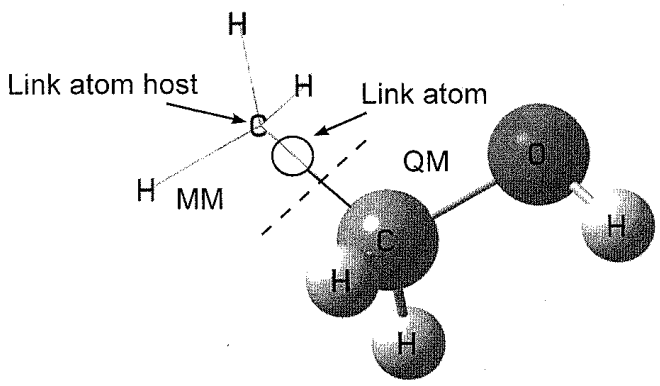
- Quantum chemistry style implementation of the QM/MM method
- Analytical 1st and 2d energy derivatives
- Internal force fields: Amber, UFF, Dreiding
- MM force field parameters can be specified via input
- Library of potential functions
- .....
- Well parallelized up to 16 processors

ONIOM QM/MM geometry optimization with microiterations



There is a new implementation of geometry optimization where QM is optimized in the full QM/MM space (keyword: **quadmacro**)

Partition in QM and MM parts in ONIOM



- The link atom substitutes the link atom host automatically
- The bond length for the link atom is scaled
- Double bonds should not be broken

Potential energy surface in ONIOM

ONIOM energy

$$E_{ONIOM} = E^{MM}_{Real} - E^{MM}_{Model} + E^{QM}_{Model}$$

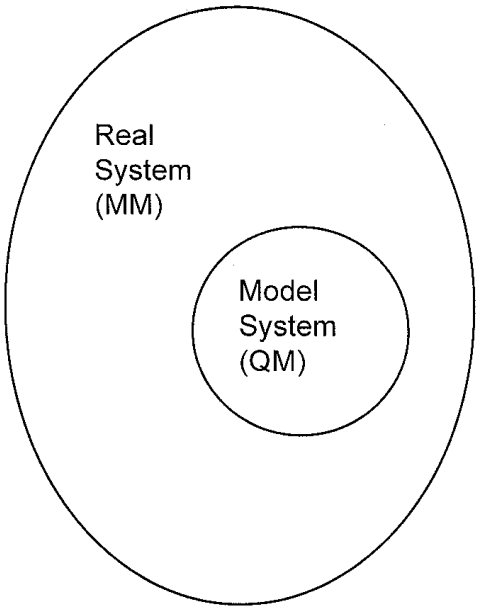
ONIOM gradient

$$G_{ONIOM} = G^{MM}_{Real} - G^{MM}_{Model} \times J + G^{QM}_{Model} \times J$$

ONIOM hessian

$$H_{ONIOM} = H^{MM}_{Real} - J^{Tr} \times H^{MM}_{Model} \times J + J^{Tr} \times H^{QM}_{Model} \times J$$

Jacobian ***J*** projects the forces on the link atoms onto the link atoms hosts. ***J*** is the function of the atomic coordinates of the model system and link atoms hosts



QM/MM Examples

GAFF
Source leaprc.gaff

A) AMBER minimization of HCN in water solution

```
# .bash_profile

# Get the aliases and functions
if [ -f ~/.bashrc ]; then
        . ~/.bashrc
fi


source /grid/software/amber/amber14/amber.sh
```

.bashrc_profile modification

1. Preparation file for HCN (*hcn.in*)

$ *antechamber -i hcn-ini.pdb -fi pdb -o hcn.in -fo prepi -c bcc* Not using GAFF

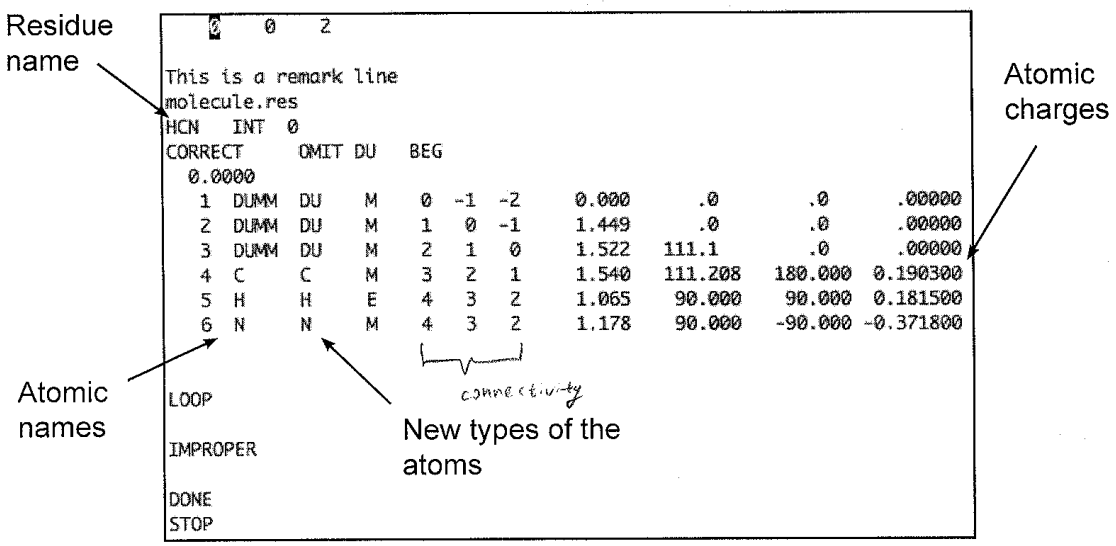
Geometry of HCN in a PDB format

Preparation file for HCN

Method for charge calculations

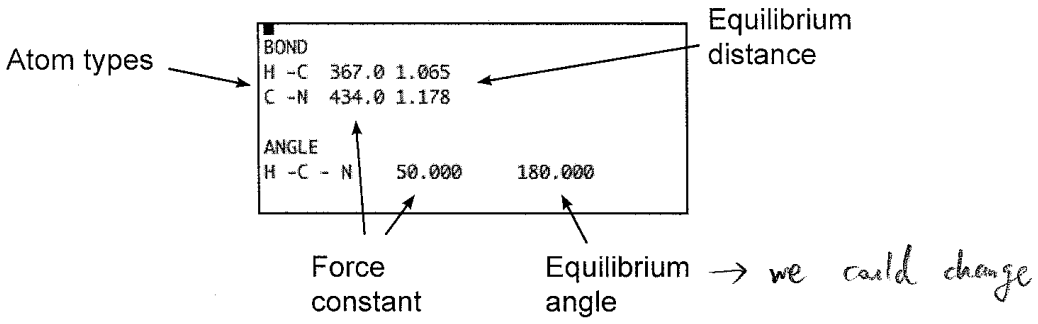
```
ATOM      1  C   HCN     1       0.000   0.000  -0.513
ATOM      2  N   HCN     1       0.000   0.000   0.665
ATOM      3  H   HCN     1       0.000   0.000  -1.578
END
```

2. Modification of the preparation file for HCN (*hcn.in*)




```
   0    0    2

This is a remark line
molecule.res
HCN   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522    111.1       .0      .00000
   4  C     C     M    3   2   1     1.540    111.208   180.000  0.190300
   5  H     H     E    4   3   2     1.065     90.000    90.000  0.181500
   6  N     N     M    4   3   2     1.178     90.000   -90.000 -0.371800

LOOP

IMPROPER

DONE
STOP
```

3. Parameter file for HCN (*hcn.par*)



Force constant is express *kcal/mol/A*^{2}, equilibrium angle in *Degree*, and equilibrium distance in *A*

The parameters can be taken from experimental data, from ab-initio calculations, or from a similar molecular system of the AMBER database

4. Preparation input files for AMBER calculations (*hcn.top, hcn.xyz*)

$ *tleap*

> *loadAmberParams hcn.par*

> *loadAmberPrep hcn.in*

> *P = loadPdb hcn-ini.pdb*

> *solvateCap P TIP3PBOX { 0.0, 0.0, 0.0 } 12.0*

> *savePdb P hcn.pdb*

> *saveAmberParm P hcn.top hcn.xyz*

> *quit*

Topology file

Coordinates file

```
Checking Unit.
Building topology.
Building atom parameters.
Building bond parameters.
Building angle parameters.
Building proper torsion parameters.
Building improper torsion parameters.
old PREP-specified impropers:
 total 0 improper torsions applied
 0 improper torsions in old prep form
Building H-Bond parameters.
Incorporating Non-Bonded adjustments.
Not Marking per-residue atom chain types.
Marking per-residue atom chain types.
  (Residues lacking connect0/connect1 -
   these don't have chain types marked:

        res     total affected

        WAT     212
  )
 (no restraints)
> quit
        Quit
```

4. PDB file of the starting structure of the molecular system (*hcn.pdb*)

```
ATOM      1  C   HCN     1       0.000   0.000  -0.513  1.00  0.00
ATOM      2  H   HCN     1       0.000   0.000  -1.578  1.00  0.00
ATOM      3  N   HCN     1       0.000   0.000   0.665  1.00  0.00
TER
ATOM      4  O   WAT     2       2.674   6.039   3.843  1.00  0.00
ATOM      5  H1  WAT     2       1.785   6.381   3.741  1.00  0.00
ATOM      6  H2  WAT     2       2.958   5.839   2.951  1.00  0.00
TER
ATOM      7  O   WAT     3       4.850   4.623   4.877  1.00  0.00
ATOM      8  H1  WAT     3       4.956   5.523   5.184  1.00  0.00
ATOM      9  H2  WAT     3       3.939   4.574   4.586  1.00  0.00
TER
ATOM     10  O   WAT     4       8.338   1.430   7.807  1.00  0.00
ATOM     11  H1  WAT     4       8.646   0.875   7.091  1.00  0.00
ATOM     12  H2  WAT     4       7.408   1.565   7.625  1.00  0.00
TER
ATOM     13  O   WAT     5       3.033   8.916   4.826  1.00  0.00
ATOM     14  H1  WAT     5       2.912   9.642   4.214  1.00  0.00
ATOM     15  H2  WAT     5       2.723   8.146   4.349  1.00  0.00
```

Atomic names

Residue name

Avogadro can be used for visualization of this molecular system

5. Input file for energy minimization (*min.in*)

```
2000 steps of minimization of HCN
 &cntrl
    imin=1, ntmin=2, drms=0.03,
    ntb=0, cut=12,
    ntc=1, ntf=1,
    ntpr=100,
    maxcyc=2000,
 /
```

6. Running energy minimization with SANDER

$ *condor_submit hcn.min.submit*

```
Universe        = Parallel
Executable      = hcn.min.cmd
Output          = hcn.min.dat
Error           = hcn.min.error
Log             = hcn.min.log
Machine_Count   = 1
Queue
```

hcn.min.submit file

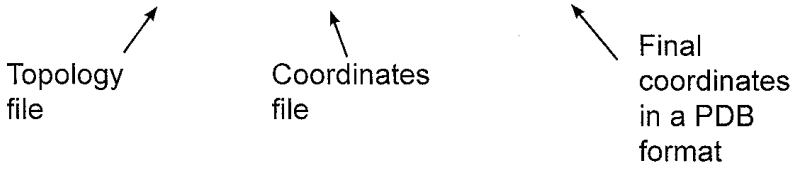
hcn.min.cmd file

```
#!/bin/bash

# Amber
export PATH=/grid/software/amber/amber14/bin:$PATH

# Number of Threads
export MKL_NUM_THREADS=1

# Program Command
/grid/software/amber/amber14/bin/sander -O -i min.in -o hcn.min.out -c hcn.xyz
   -ref hcn.xyz -p hcn.top -r hcn.min.xyz
```

Coordinates file | Topology file | Final coordinates file | Input file | Output file

7. Converting the AMBER coordination file into a PDB file

$ *ambpdb -p hcn.top < hcn.min.xyz > hcn.min.pdb*



Avogadro can be used for visualization of the final PDB file

B) ONIOM calculations of HCN in water solution

Installation of toolkit to assist ONIOM (TAO)
(http://www.chem.wayne.edu/schlegel/Software.html)

- Move this toolkit package to a location you usually install application softwares.
- Edit *install.sh* in the home folder of TAO (taopackage). Change path ~/bin in line 14 *USERPATH=~/bin* to the path that you want the symbolic links to TAO scripts to be installed. e.g. *USERPATH=/home/myhome/bin/oniomtool.* Please make sure this path is in your search *PATH* of the *.bash_profile* file.
- Run *./install.sh* in the home folder of TAO (taopackage) to install this package.

1. Geometry optimization in mechanical embedding

$ *pdb2oniom −i hcn.min.pdb −o hcn-opt.g09*

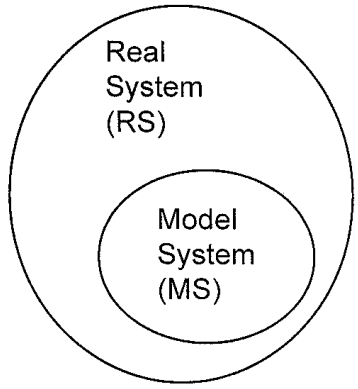
Initial geometry in a PDF format

Initial ONIOM input

There are unknown atom types for ONIOM/ AMBER force filed

```
%chk=hcn-opt.g09.chk
%mem=3700MB
%nprocshared=4

#P ONIOM(B3LYP/6-31G(d):AMBER=HardFirst) geom=connectivity nosymm iop(2/15=3) test

ONIOM inputfile generated by pdb2oniom from PDB file hcn.min.pdb. No connectivity generated.
Please use GaussView read hcn-opt.g09, and generate connectivity information.

0 1 0 1 0 1
 C-UDF-0.000000   0       0.2610000      -0.6550000      0.5010000   L
 H-UDF-0.000000   0       0.6110000       0.2910000      0.1550000   L
 N-UDF-0.000000   0      -0.1260000      -1.7050000      0.8810000   L
 O-OW--0.834000   0       2.7090000       5.7990000      3.8930000   L
 H-HW-0.417000    0       1.7480000       5.9510000      3.6990000   L
 H-HW-0.417000    0       3.0300000       5.7190000      2.9650000   L
 O-OW--0.834000   0       4.9420000       4.6180000      5.0030000   L
 H-HW-0.417000    0       5.1180000       5.3910000      5.5860000   L
 H-HW-0.417000    0       4.0440000       4.8640000      4.6780000   L
 O-OW--0.834000   0       7.9150000       1.7270000      7.2140000   L
```

The ONIOM input file needs to be modified by hand



High level of theory

Low level of theory

Providing additional MM parameters for AMBER

Charge and multiplicity of RS (low-level theory)

MS (high-level theory)

MS (low-level theory)

```
%Chk=hcn-opt
%NProcShared=4
%mem=4GB

# ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst) int(ultrafine) opt(maxcycles=100

Title

0 1 0 1 0 1
 C-C-0.130300     0       0.2610000      -0.6550000      0.5010000   H
 H-H-0.181500     0       0.6110000       0.2910000      0.1550000   H
 N-N--0.371800    0      -0.1260000      -1.7050000      0.8810000   H
 O-OW--0.834000   0       2.7090000       5.7990000      3.8930000   L
 H-HW-0.417000    0       1.7480000       5.9510000      3.6990000   L
 H-HW-0.417000    0       3.0300000       5.7190000      2.9650000   L
 O-OW--0.834000   0       4.9420000       4.6180000      5.0030000   L
 H-HW-0.417000    0       5.1180000       5.3910000      5.5860000   L
 H-HW-0.417000    0       4.0440000       4.8640000      4.6780000   L
```

high-level theory

low-level theory

Atomic symbol

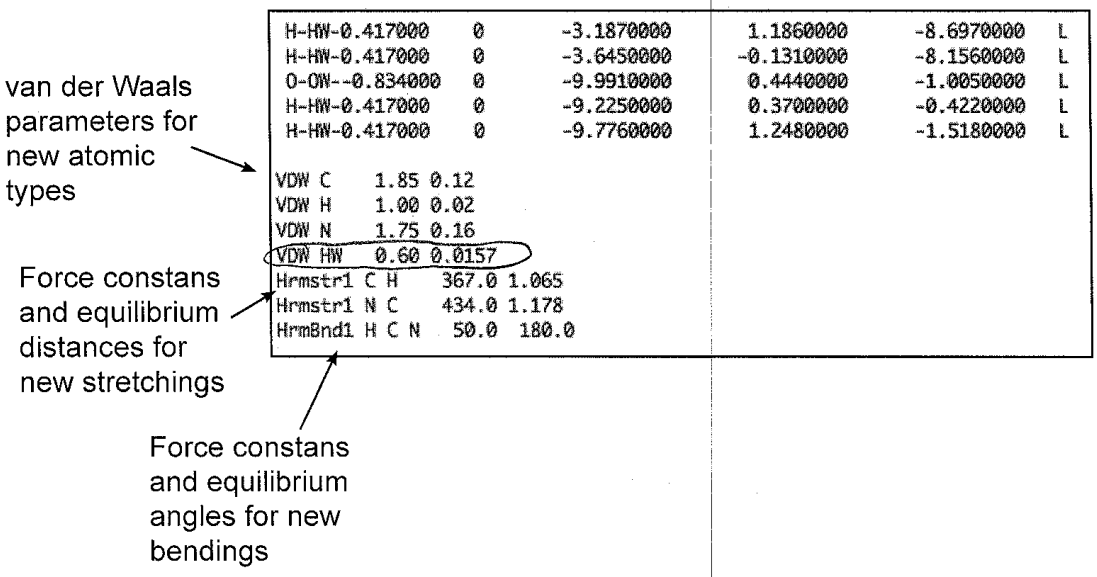
Atomic type

Atomic charge

Atomic constraint (-1)

= dash

2. Molecular mechanical parameters for HCN in ONIOM input



3. Geometry optimization in mechanical embedding

 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst) int(ultrafine) opt(maxcycles=100)

↓ 自改参数优先

4. Frequency calculation in mechanical embedding

 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst) geom(allcheck) guess(read) int(ultrafine) freq

5. Geometry optimization in mechanical and electronic embedding

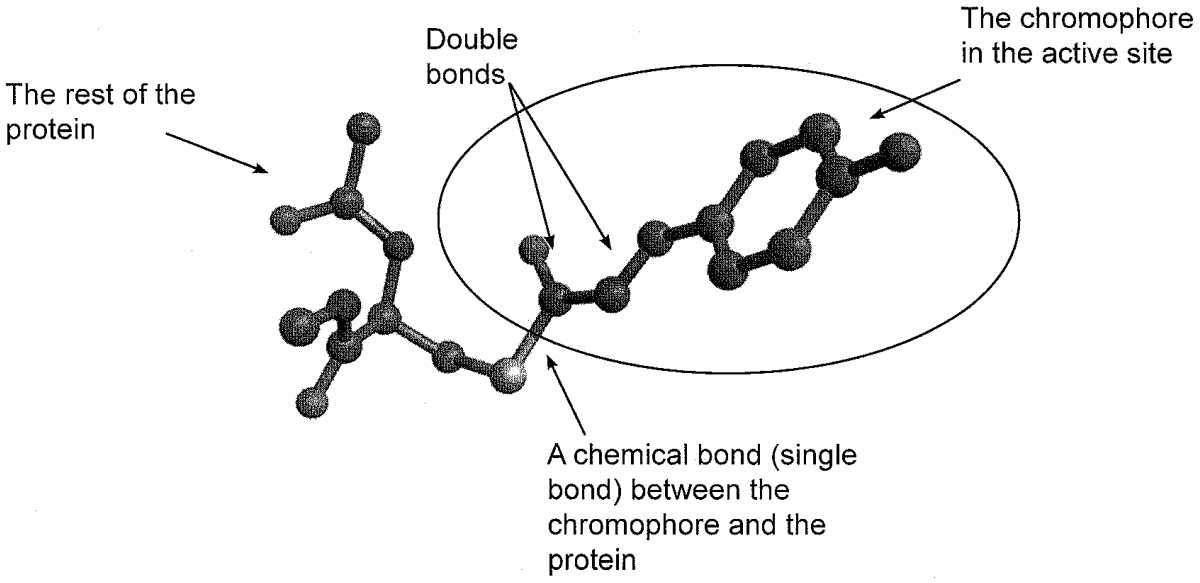
 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst)=EmbedCharge geom(allcheck) guess(read) int(ultrafine) opt(maxcycles=100)

6. Frequency calculation in mechanical and electronic embedding)

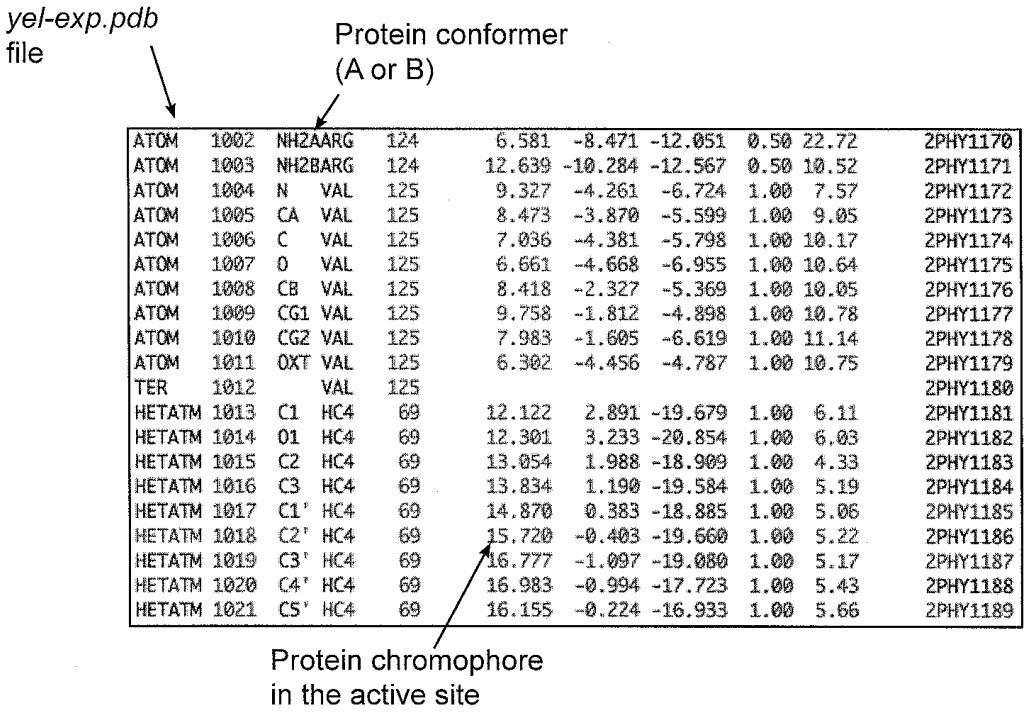
 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst)=EmbedCharge geom(allcheck) guess(read) int(ultrafine) freq

C) AMBER minimization of yellow protein (2PHY)

- There are two protein conformers A and B
- Initial geometry of the protein does not have hydrogen atoms (will be added automatically by *tleap*)
- Atoms, atomic types and parameters of the chromophore in the active site are unknown by AMBER force field (must by created by hand)
- There is a chemical bond between the chromophore and the protein



Initial geometry of yellow protein based on the x-ray PDB file (2PHY entry, *yel-exp.pdb*)



ATOM	1002	NH2A	ARG	124	6.581	-8.471	-12.051	0.50	22.72	2PHY1170
ATOM	1003	NH2B	ARG	124	12.639	-10.284	-12.567	0.50	10.52	2PHY1171
ATOM	1004	N	VAL	125	9.327	-4.261	-6.724	1.00	7.57	2PHY1172
ATOM	1005	CA	VAL	125	8.473	-3.870	-5.599	1.00	9.05	2PHY1173
ATOM	1006	C	VAL	125	7.036	-4.381	-5.798	1.00	10.17	2PHY1174
ATOM	1007	O	VAL	125	6.661	-4.668	-6.955	1.00	10.64	2PHY1175
ATOM	1008	CB	VAL	125	8.418	-2.327	-5.369	1.00	10.05	2PHY1176
ATOM	1009	CG1	VAL	125	9.758	-1.812	-4.898	1.00	10.78	2PHY1177
ATOM	1010	CG2	VAL	125	7.983	-1.605	-6.619	1.00	11.14	2PHY1178
ATOM	1011	OXT	VAL	125	6.302	-4.456	-4.787	1.00	10.75	2PHY1179
TER	1012		VAL	125						2PHY1180
HETATM	1013	C1	HC4	69	12.122	2.891	-19.679	1.00	6.11	2PHY1181
HETATM	1014	O1	HC4	69	12.301	3.233	-20.854	1.00	6.03	2PHY1182
HETATM	1015	C2	HC4	69	13.054	1.988	-18.909	1.00	4.33	2PHY1183
HETATM	1016	C3	HC4	69	13.834	1.190	-19.584	1.00	5.19	2PHY1184
HETATM	1017	C1'	HC4	69	14.870	0.383	-18.885	1.00	5.06	2PHY1185
HETATM	1018	C2'	HC4	69	15.720	-0.403	-19.660	1.00	5.22	2PHY1186
HETATM	1019	C3'	HC4	69	16.777	-1.097	-19.080	1.00	5.17	2PHY1187
HETATM	1020	C4'	HC4	69	16.983	-0.994	-17.723	1.00	5.43	2PHY1188
HETATM	1021	C5'	HC4	69	16.155	-0.224	-16.933	1.00	5.66	2PHY1189

Atoms of the B protein conformer should be removed by hand from the PDB file

1. Making a complete structure of the chromophore (with Avogadro)

```
HETATM 1013  C1  HC4     69      12.122   2.891 -19.679
HETATM 1014  O1  HC4     69      12.301   3.233 -20.854
HETATM 1015  C2  HC4     69      13.054   1.988 -18.909
HETATM 1016  C3  HC4     69      13.834   1.190 -19.584
HETATM 1017  C1' HC4     69      14.870   0.383 -18.885
HETATM 1018  C2' HC4     69      15.720  -0.403 -19.660
HETATM 1019  C3' HC4     69      16.777  -1.097 -19.080
HETATM 1020  C4' HC4     69      16.983  -0.994 -17.723
HETATM 1021  C5' HC4     69      16.155  -0.224 -16.933
HETATM 1022  C6' HC4     69      15.106   0.456 -17.515
HETATM 1023  O4' HC4     69      17.939  -1.579 -17.197
```

Experimental coordinates of the chromophore

```
ATOM      1  C1  HC4     69      12.122   2.891 -19.679
ATOM      2  O1  HC4     69      12.301   3.233 -20.854
ATOM      3  C2  HC4     69      13.054   1.988 -18.909
ATOM      4  C3  HC4     69      13.834   1.190 -19.584
ATOM      5  C1' HC4     69      14.870   0.383 -18.885
ATOM      6  C2' HC4     69      15.720  -0.403 -19.660
ATOM      7  C3' HC4     69      16.777  -1.097 -19.080
ATOM      8  C4' HC4     69      16.983  -0.994 -17.723
ATOM      9  C5' HC4     69      16.155  -0.224 -16.933
ATOM     10  C6' HC4     69      15.106   0.456 -17.515
ATOM     11  O4' HC4     69      17.939  -1.579 -17.197
HETATM   12  H   LIG      1      11.282   3.245 -19.195
HETATM   13  H   LIG      1      13.074   2.002 -17.877
HETATM   14  H   LIG      1      13.732   1.118 -20.608
HETATM   15  H   LIG      1      15.563  -0.471 -20.678
HETATM   16  H   LIG      1      17.399  -1.682 -19.660
HETATM   17  H   LIG      1      16.319  -0.157 -15.916
HETATM   18  H   LIG      1      14.484   1.030 -16.924
```

Complete chromophore structure

1. Making (by hand) an initial geometry of the chromophore (*pct.pdb*)

ATOM	1	C1	HC4	69	12.122	2.891	-19.679
ATOM	2	O1	HC4	69	12.301	3.233	-20.854
ATOM	3	C2	HC4	69	13.054	1.988	-18.909
ATOM	4	C3	HC4	69	13.834	1.190	-19.584
ATOM	5	C1'	HC4	69	14.870	0.383	-18.885
ATOM	6	C2'	HC4	69	15.720	-0.403	-19.660
ATOM	7	C3'	HC4	69	16.777	-1.097	-19.080
ATOM	8	C4'	HC4	69	16.983	-0.994	-17.723
ATOM	9	C5'	HC4	69	16.155	-0.224	-16.933
ATOM	10	C6'	HC4	69	15.106	0.456	-17.515
ATOM	11	O4'	HC4	69	17.939	-1.579	-17.197
HETATM	12	H	LIG	1	11.282	3.245	-19.195
HETATM	13	H	LIG	1	13.074	2.002	-17.877
HETATM	14	H	LIG	1	13.732	1.118	-20.608
HETATM	15	H	LIG	1	15.563	-0.471	-20.678
HETATM	16	H	LIG	1	17.399	-1.682	-19.660
HETATM	17	H	LIG	1	16.319	-0.157	-15.916
HETATM	18	H	LIG	1	14.484	1.030	-16.924

pct.pdb - Initial geometry of the chromophore

ATOM	1	O1	PCT	1	12.301	3.233	-20.854
ATOM	2	C1	PCT	1	12.122	2.891	-19.679
ATOM	3	H23	PCT	1	11.282	3.245	-19.195
ATOM	4	C2	PCT	1	13.054	1.988	-18.909
ATOM	5	H25	PCT	1	13.074	2.002	-17.877
ATOM	6	C3	PCT	1	13.834	1.190	-19.584
ATOM	7	H27	PCT	1	13.732	1.118	-20.608
ATOM	8	C1'	PCT	1	14.870	0.383	-18.885
ATOM	9	C2'	PCT	1	15.720	-0.403	-19.660
ATOM	10	H29	PCT	1	15.563	-0.471	-20.678
ATOM	11	C3'	PCT	1	16.777	-1.097	-19.080
ATOM	12	H31	PCT	1	17.399	-1.682	-19.660
ATOM	13	C4'	PCT	1	16.983	-0.994	-17.723
ATOM	14	O4'	PCT	1	17.939	-1.579	-17.197
ATOM	15	C5'	PCT	1	16.155	-0.224	-16.933
ATOM	16	H33	PCT	1	16.319	-0.157	-15.916
ATOM	17	C6'	PCT	1	15.106	0.456	-17.515
ATOM	18	H35	PCT	1	14.484	1.030	-16.924

New atomic names (H) and a new residue name

2. Combining (by hand) the experimental PDB file of the protein (*yel-exp.pdb*) with the initial geometry of the chromophore (*pct.pdb*) and making the initial PDB file of the entire system (*yel-ini.pdb*)

yel-ini.pdb file, the initial structure of the entire system →

```
ATOM   1008  CB  VAL   125       8.418  -2.327  -5.369
ATOM   1009  CG1 VAL   125       9.758  -1.812  -4.898
ATOM   1010  CG2 VAL   125       7.983  -1.605  -6.619
ATOM   1011  OXT VAL   125       6.302  -4.456  -4.787
TER    1012      VAL   125
ATOM      1  O1  PCT     1      12.301   3.233 -20.854
ATOM      2  C1  PCT     1      12.122   2.891 -19.679
ATOM      3  H23 PCT     1      11.282   3.245 -19.195
ATOM      4  C2  PCT     1      13.054   1.988 -18.909
ATOM      5  H25 PCT     1      13.074   2.002 -17.877
ATOM      6  C3  PCT     1      13.834   1.190 -19.584
ATOM      7  H27 PCT     1      13.732   1.118 -20.608
ATOM      8  C1' PCT     1      14.870   0.383 -18.885
ATOM      9  C2' PCT     1      15.720  -0.403 -19.660
ATOM     10  H29 PCT     1      15.563  -0.471 -20.678
ATOM     11  C3' PCT     1      16.777  -1.097 -19.080
ATOM     12  H31 PCT     1      17.399  -1.682 -19.660
ATOM     13  C4' PCT     1      16.983  -0.994 -17.723
ATOM     14  O4' PCT     1      17.939  -1.579 -17.197
ATOM     15  C5' PCT     1      16.155  -0.224 -16.933
ATOM     16  H33 PCT     1      16.319  -0.157 -15.916
ATOM     17  C6' PCT     1      15.106   0.456 -17.515
ATOM     18  H35 PCT     1      14.484   1.030 -16.924
TER
HETATM 1024  O   HOH   200      21.132   0.254  -3.329
HETATM 1025  O   HOH   201      23.447   4.398   0.293
HETATM 1026  O   HOH   202       2.697 -10.636 -17.802
HETATM 1027  O   HOH   203       6.966   5.068 -16.994
```

3. Preparation file for the chromophore (*pct.in*)

$ *antechamber -i pct.pdb -fi pdb -o pct.in -fo prepi -c bcc -nc -1*

Total charge

pct.in file for the chromophore

```
   2    0    2

This is a remark line
molecule.res
PCT   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0       .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0       .0      .00000
   3  DUMM  DU    M    2   1   0     1.523   111.21      .0      .00000
   4  O1    O     M    3   2   1     1.540   111.208  -180.000 -0.665300
   5  C1    C     M    4   3   2     1.237    81.678   -73.772  0.586100
   6  H23   H1    E    5   4   3     1.032   117.936   -31.407 -0.055800
   7  C2    CA    M    5   4   3     1.509   124.117   148.548 -0.430000
   8  H25   HA    E    7   5   4     1.032   120.930   157.255  0.114000
   9  C3    CA    M    7   5   4     1.304   118.130   -22.707  0.086000
  10  H27   HA    E    9   7   5     1.032   119.824    -6.397  0.108000
  11  C1'   CA    M    9   7   5     1.488   120.347   173.606 -0.295000
  12  C2'   CA    M   11   9   7     1.393   118.005  -176.887 -0.014000
  13  H29   HA    E   12  11   9     1.032   119.535    -5.194  0.092500
  14  C3'   CA    M   12  11   9     1.391   120.870   174.806 -0.346700
  15  H31   HA    E   14  12  11     1.032   120.418   179.621  0.110000
  16  C4'   CA    M   14  12  11     1.376   119.172    -0.371  0.569500
  17  O4'   O     E   16  14  12     1.238   119.938  -178.218 -0.700100
  18  C5'   CA    M   16  14  12     1.379   121.104     0.457 -0.346700
  19  H33   HA    E   18  16  14     1.032   120.338   179.848  0.110000
  20  C6'   CA    M   18  16  14     1.379   119.354    -0.174 -0.014000
  21  H35   HA    E   20  18  16     1.032   119.402   179.800  0.092500


LOOP
  C6'  C1'
```

New atomic types (modified by hand)

4. Parameter file for the chromophore (*pct.par*)

The parameters for the chromophore are taken from the AMBER database, based on the same (similar) atomic types

```
BOND
C -H1      370.0        1.01         Junmei et al, 1999
CA- O      570.0        1.26         JCC,7,(1986),230; AA,CYT,GUA,THY,URA
S - C      230.0        1.79         changed from 222.0 based on dimethylS nmodes

ANGLE
O -CA-CA    70.0      123.00         replacement in tyr
H1- C-CA    70.0      120.00         AA (not used in tyr)
O - C-H1    50.0      120.00         Junmei et al, 1999
O - C-CA    80.0      128.00
S - C- O    50.0      118.00         AA cyx (SCHERAGA  JPC 79,1428)
S - C-CA    50.0      114.00         AA cyx (SCHERAGA  JPC 79,1428)
CT- S -C    70.0      110.00         AA cyx (SCHERAGA  JPC 79,1428)
2C- S -C    70.0      110.00         AA cyx (SCHERAGA  JPC 79,1428)

DIHEDRAL
X -C -S -X    2       3.00      180.000       2.000
```

pct.par file for the chromophore

5. Preparation input files for AMBER calculations

```
$ tleap
> loadAmberParams frcmod.ionsjc_spce
> loadAmberParams pct.par
> loadAmberPrep pct.in
> P = loadPdb yel-ini.pdb
> desc P.126
> remove P P.126.3
> desc P.69
> bond P.69.8 P.126.2
> addIons P Na+ 7
> savePdb P yel.pdb
> saveAmberParm P yel.top yel.xyz
> quit
```

Loadin extra parameters for counter ions (Na^+, Cl^-)

Information about chromophore (resid 126)

Removing extra hydrogen atom from chromophore

Information about cysteine (resid 69)

Making a bond between chromophore and cysteine

Neutralizing the entire protein by Na^+

6. The input file for AMBER minimization of the yellow protein (*min.in*)

```
2000 steps of hydrogen minimization
 &cntrl
    imin=1, ntmin=2, drms=0.03,
    ntb=0, cut=12,
    ntc=1, ntf=1,
    ntpr=100,
    maxcyc=2000,
    ntr=1,
    restraint_wt=500.0,
    restraintmask='(!@H=)',
 /
```

Minimization with a constrained non-hydrogen atoms

7. Running energy minimization with SANDER

$ *condor_submit yel.min.submit*

```
Universe        = Parallel
Executable      = yel.min.cmd
Output          = yel.min.dat
Error           = yel.min.error
Log             = yel.min.log
Machine_Count   = 1
Queue
```

yel.min.submit file

yel.min.cmd file

```
#!/bin/bash

# Amber
export PATH=/grid/software/amber/amber14/bin:$PATH

# Number of Threads
export MKL_NUM_THREADS=1

# Program Command
/grid/software/amber/amber14/bin/sander -O -i min.in -o yel.min.out -c yel.xyz
  -ref yel.xyz -p yel.top -r yel.min.xyz
```

8. Converting the AMBER coordination file into a PDB file

$ *ambpdb -p yel.top < yel.min.xyz > yel.min.pdb*

Topology file

Coordinates file

Final coordinates in a PDB format

Avogadro can be used for visualization of the final PDB file

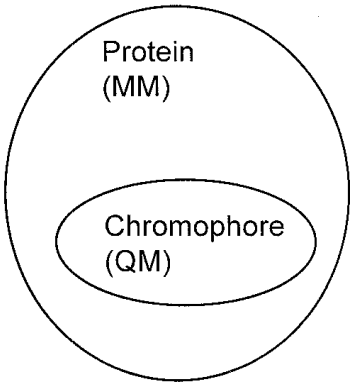
D) ONIOM calculations of yellow protein

1. Geometry optimization in mechanical embedding

$ *pdb2oniom −i yel.min.pdb −o yel-opt.g09*

Initial geometry in a PDF format

Initial ONIOM input



ONIOM input file generated by the *pdb2oniom* script

```
%chk=yel-opt.g09.chk
%mem=3700MB
%nprocshared=4

#P ONIOM(B3LYP/6-31G(d):AMBER=HardFirst) geom=connectivity nosymm iop(2/15=3) test

ONIOM inputfile generated by pdb2oniom from PDB file yel.min.pdb. No connectivity
Please use GaussView read yel-opt.g09, and generate connectivity information.

0 1 0 1 0 1
 N-N--0.394918    0       26.3500000     -16.3130000      -3.5000000   L
 UDF-UDF-0.000000         0       26.5140000     -16.4700000      -2.5140000   L
 UDF-UDF-0.000000         0       27.2430000     -16.3690000      -3.9810000   L
 UDF-UDF-0.000000         0       25.7160000     -17.0030000      -3.8720000   L
 C-CT--0.087681   0       25.8230000     -14.9520000      -3.6780000   L
 H-H1-0.123080    0       25.8960000     -14.6880000      -4.7340000   L
```

There are atoms which are unknown by ONIOM

The original input file needs to be modified by hand

Charge and multiplicity of the entire protein

Chromophore QM level

Chromophore MM level

Connectivity between atoms

```
%Chk=yel-opt
%NProcShared=4
%mem=8GB

# ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst) geom=connectivity int(ultrafine) opt(maxcycles=100)

Title

0 1 -1 1 -1 1
 N-N--0.394918    0       26.3500000     -16.3130000     -3.5000000   L
 H-H-0.2278053    0       26.5140000     -16.4700000     -2.5140000   L
 H-H-0.2278053    0       27.2430000     -16.3690000     -3.9810000   L
 H-H-0.2278053    0       25.7160000     -17.0030000     -3.8720000   L
 C-CT--0.087681   0       25.8230000     -14.9520000     -3.6780000   L
 H-H1-0.123080    0       25.8960000     -14.6880000     -4.7340000   L
 C-CT-0.019227    0       24.3310000     -14.9010000     -3.2890000   L
```

2. Connectivity between atoms in ONIOM

The connectivity between atoms can be provided by GaussView, by reading the original ONIOM input file *yel-opt.g09,* and saving as a new input file

yel-connect.txt file with connectivity between atoms

```
O-OW--0.834000   0       2.9920000      -3.1830000      -7.9670000   L
H-HW-0.417000    0       2.6140000      -3.9860000      -7.5630000   L
H-HW-0.417000    0       3.9240000      -3.2670000      -7.6840000   L

1 2 1.0 3 1.0 4 1.0 5 1.0
2
3
4
5 6 1.0 7 1.0 18 1.0
6
7 8 1.0 9 1.0 10 1.0
8
```

3. Molecular mechanical parameters for chromophore in ONIOM input

vdW parameters for some hydrogen types

MM parameters for atomic types of the chromophore

```
 2202
 2203 2204 1.0 2205 1.0
 2204
 2205

VDW     HO  0.6000  0.015700
VDW     HW  0.6000  0.015700
HrmStr1     S   C   227.0    1.810
HrmStr1     CA  O   450.0    1.364
HrmStr1     C   H1  367.0    1.080
HrmBnd1     N   CT  HC     49.78   109.50
HrmBnd1     CT  S   C      62.0     98.90
HrmBnd1     S   C   O      50.0    114.70
HrmBnd1     S   C   C      50.0    114.70
HrmBnd1     S   C   CA     50.0    114.70
HrmBnd1     O   C   CA     80.0    120.00
HrmBnd1     CA  CA  O      80.0    120.00
HrmBnd1     H1  C   O      50.0    120.00
HrmBnd1     H1  C   CA     50.0    120.00
```

MM atomic types of the chromophore should correspond to the *pct.in* and *pct.par* from AMBER calculation

4. A link atom for chromophore in ONIOM input

```
H-H1-0.078951    0        9.9670000        4.3300000      -20.7460000   L
S-S--0.132272    0       11.0400000        3.8150000      -18.6640000   L H-H1 1913
C-C-0.624788     0       11.5580000        6.8850000      -19.0320000   L
```

There is a link atom *H* of the *H1* type (can have a charge) which links the *S* atom of the protein with the *C* atom of the chromophore (atom number 1913)

5. Geometry optimization in mechanical embedding

 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst) int(ultrafine) opt(maxcycles=100)

6. Frequency calculation in mechanical embedding

 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst) geom(allcheck) guess(read) int(ultrafine) freq

7. TD-DFT calculation (4-roots) in mechanical embedding

 # ONIOM(B3LYP/6-31G(d,p)/~~Auto~~ TD=(NStates=4):AMBER=SoftFirst) int(ultrafine) geom(allcheck) guess(read) sp

8. Geometry optimization in mechanical and electronic embedding

 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst)=EmbedCharge geom(allcheck) guess(read) int(ultrafine) opt(maxcycles=100)

9. Frequency calculation in mechanical and electronic embedding)

 # ONIOM(B3LYP/6-31G(D,P):AMBER=SoftFirst)=EmbedCharge geom(allcheck) guess(read) int(ultrafine) freq

10. TD-DFT calculation (4-roots) in mechanical and electronic embedding

 # ONIOM(B3LYP/6-31G(d,p)/Auto TD=(NStates=4):AMBER=SoftFirst)=EmbedCharge int(ultrafine) geom(allcheck) guess(read) sp

QM/MM Projects

- All calculations will be done on *schem1a1.systems.smu.edu* cluster
- Modify your *.bashrc_profile* file for AMBER calculations, logout and login again to update your shell environment
- Install TAO package (from the internet) on your cluster account, and modify your *.bashrc_profile* file for TAO scripts, logout and login again to update your shell environment
- Download initial files for the HCN and yellow protein projects from CATCO wiki website (*QM/MM Workshop*)
- Ask Rob to install on your local computer GaussView (if you consider you will be doing QM/MM calculations in the future)

Project A: QM/MM calculations of HCN in TIP3P water solution

1. Starting from an initial geometry of HCN (*hcn-ini.pdb*) make the preparation file (*hcn.in*) using *antechamber*
2. Based on the AMBER database, make a parameter file for HCN (*hcn.par*)
3. Using *tleap,* solvate HCN with TIP3P water molecules and save the topology (*hcn.top*) and coordinate (*hcn.xyz*) files
4. Run energy minimization of HCN in a water sphere using AMBER input file (*min.in*)
5. Using TAO script (*pdb2oniom*) generate an initial ONIOM input file for HCN in water (*hcn-opt.g09*)
6. Modify the ONIOM input file for HCN in water (*hcn-opt.g09*)
7. Run ONIOM geometry optimization of HCN in water in mechanical embedding
8. Based on the optimal geometry, run ONIOM frequency calculation of HCN in water in mechanical embedding
9. Based on the optimal geometry from mechanical embedding, run ONIOM geometry optimization of HCN in water in electronic embedding
10. Based on the optimal geometry from electronic embedding, run ONIOM frequency of HCN in water in electronic embedding
11. Report in a table interatomic distances and frequencies of HCN in the gas phase and in water with mechanical and electronic embedding
12. Compare the results

Project B: QM/MM calculations of the chromophore in yellow protein

1. Starting from geometry of the protein chromophore (*pct.pdb*) make the preparation file (*pct.in*) using *antechamber*
2. Based on the AMBER database, make a parameter file for the chromophore (*pct.par*)
3. Using *tleap,* make the topology (*yel.top*) and coordinate (*yel.xyz*) files for the entire proten
4. Run energy minimization of the protein using AMBER input file (*min.in*)
5. Using TAO script (*pdb2oniom*) generate an initial ONIOM input file for the protein (*yel-opt.g09*)
6. Modify the ONIOM input file (*yel-opt.g09*) including atomic connectivity (*yel-connect.txt*)
7. Run ONIOM geometry optimization of the protein in mechanical embedding
8. Based on the optimal geometry, run ONIOM frequency calculation of the protein in mechanical embedding
9. Based on the optimal geometry, run ONIOM TD-DFT calculation of the protein in mechanical embedding
10. Based on the optimal geometry from mechanical embedding, run ONIOM geometry optimization of the protein in electronic embedding
11. Based on the optimal geometry from electronic embedding, run ONIOM frequency of the protein in electronic embedding
12. Based on the optimal geometry from electronic embedding, run ONIOM TD-DFT calculation of the protein in electronic embedding
13. Report in a table interatomic distance and frequency of C_{Tail}=C_{Tail}, and electronic excitation ($^1X \rightarrow {}^1A$) of the chromophore the gas phase and in the protein with mechanical and electronic embedding
14. Compare the results