



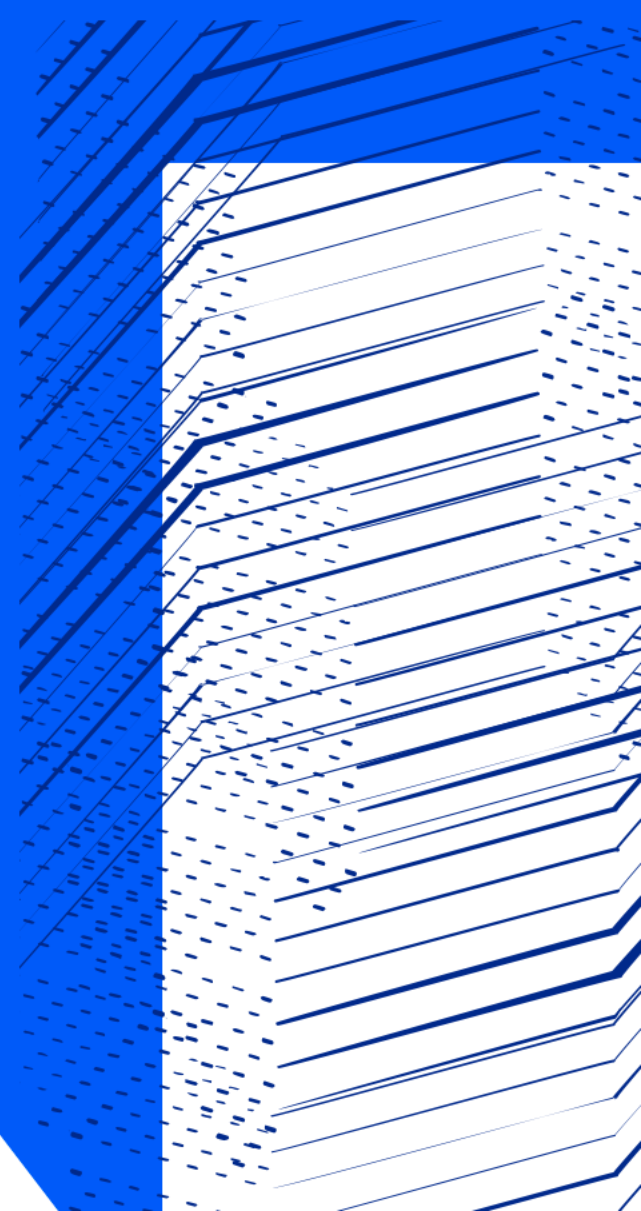
Science and  
Technology  
Facilities Council



# Biomolecular simulation with QM/MM

Kakali Sen  
UKRI-STFC Daresbury Laboratory

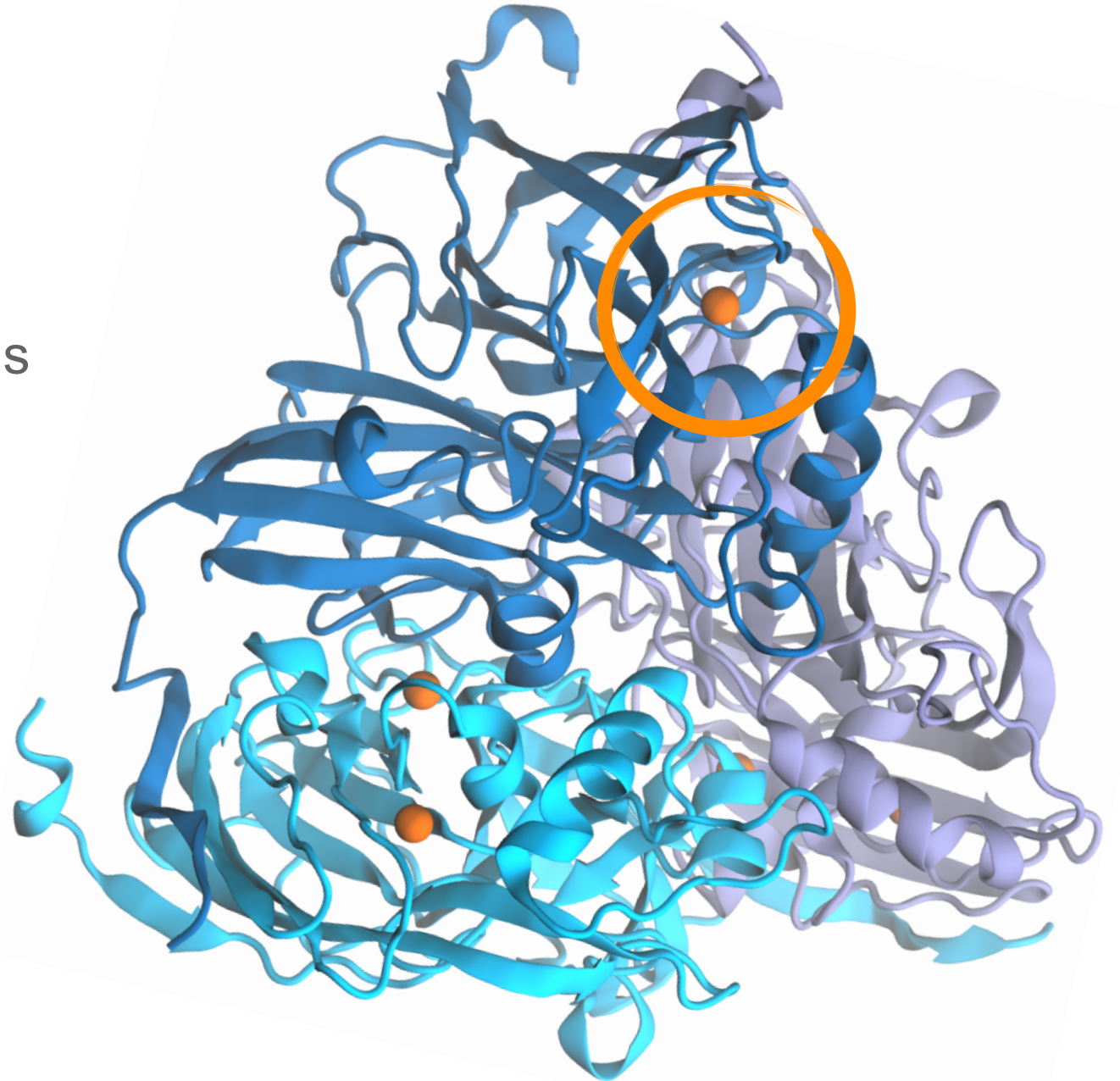
*CCPBioSim Training Week 2021  
1 October 2021*



# QM/MM for enzymes

Reaction occurs in small core: Electrons involved - Electronic structure calculations

Scaffold: Explicit electronic interactions can be ignored - Lower level of theory



# QM/MM for enzymes

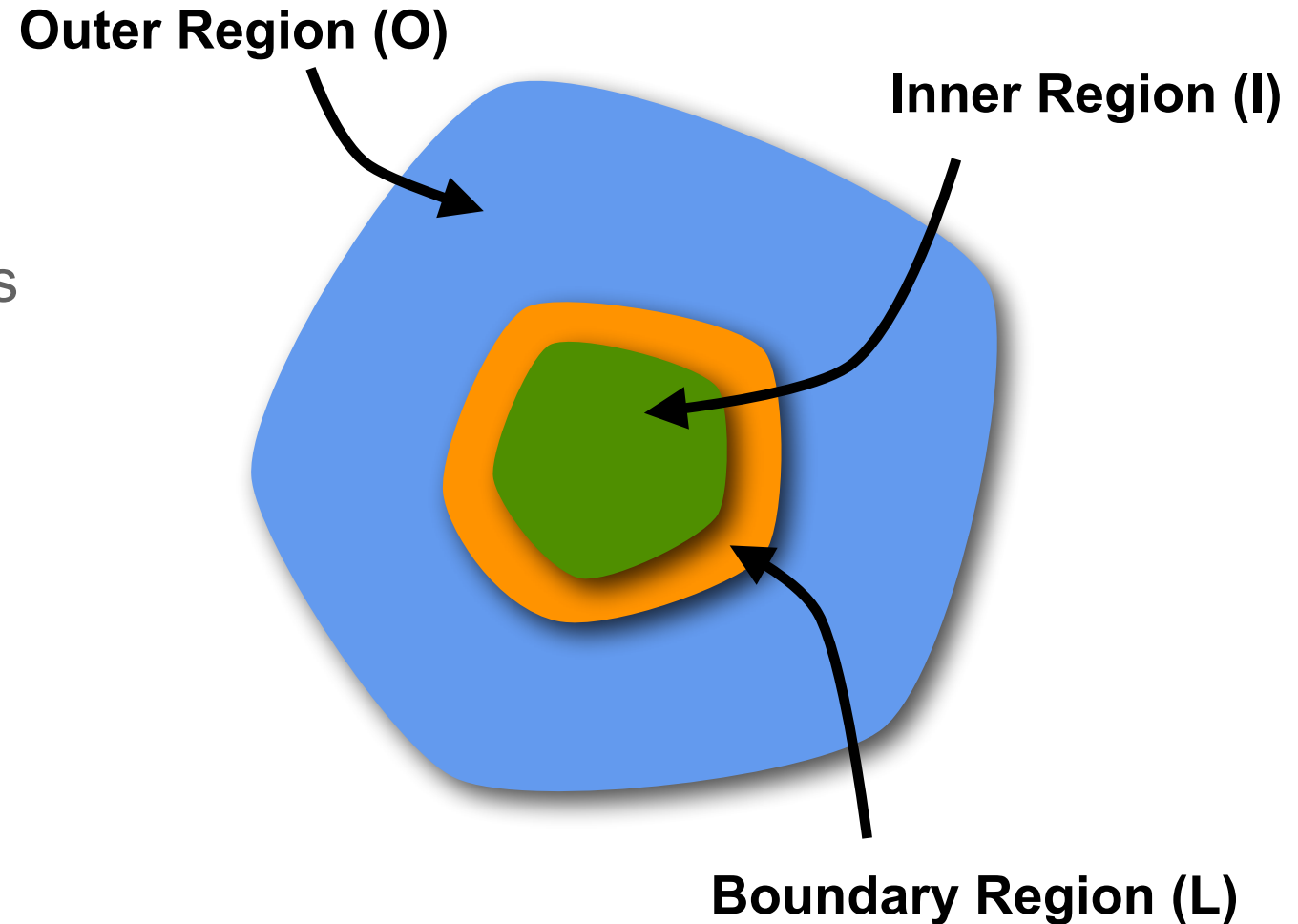
Reaction occurs in small core: Electrons involved - Electronic structure calculations

Scaffold: Explicit electronic interactions can be ignored - Lower level of theory

# QM/MM for enzymes

Reaction occurs in small core: Electrons involved - Electronic structure calculations

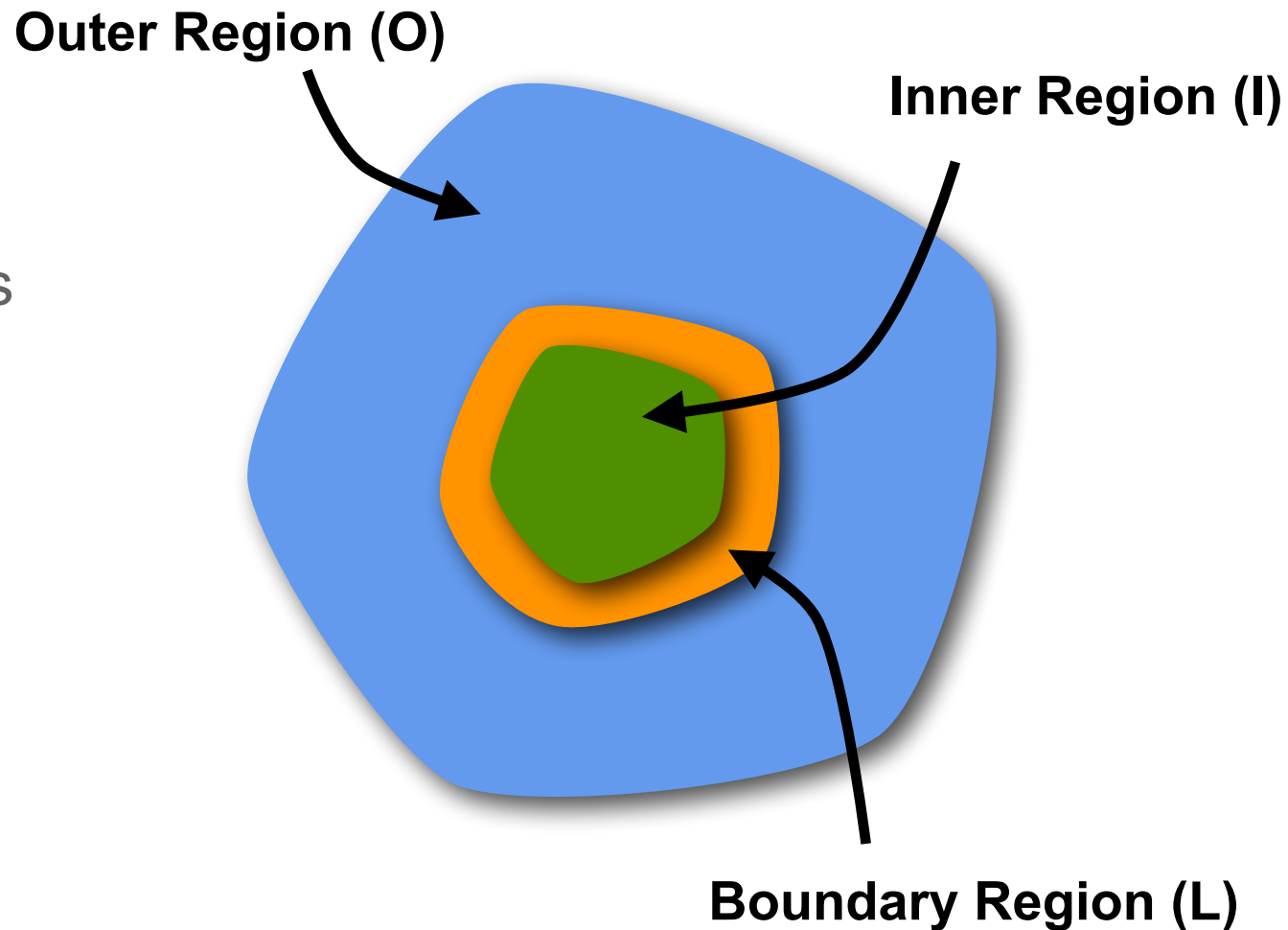
Scaffold: Explicit electronic interactions can be ignored - Lower level of theory



# QM/MM for enzymes

Reaction occurs in small core: Electrons involved - Electronic structure calculations

Scaffold: Explicit electronic interactions can be ignored - Lower level of theory

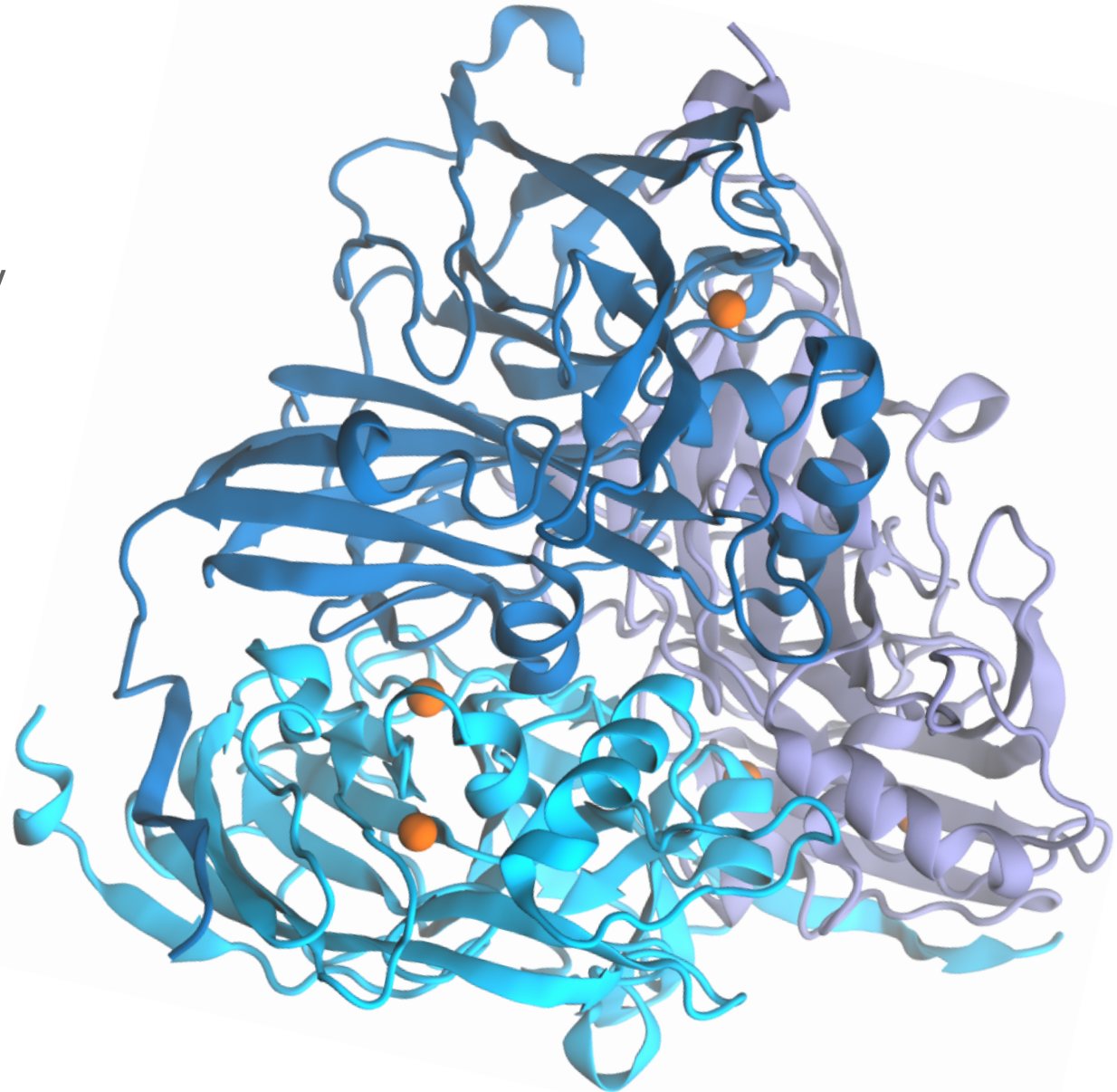


$$E_{\text{qm/mm}}(\text{S}) = E_{\text{qm}}(\text{I} + \text{L}) + E_{\text{mm}}(\text{O}) + E_{\text{qm-mm}}(\text{I}, \text{O})$$

Electrostatic embedding

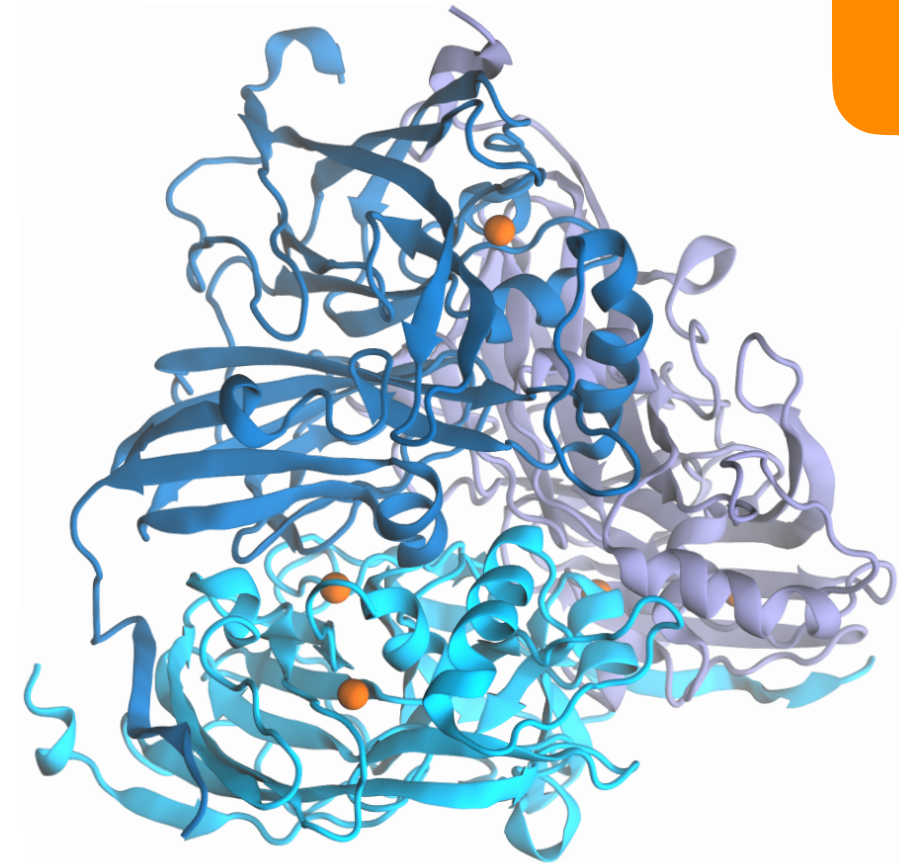
# QM/MM for enzymes

- Investigate mechanisms
- Characterise intermediates that are experimentally elusive; e.g., transition states
- Make predictions of selectivity
- Understand the stabilising interactions around the active site/binding pockets
- Emission properties of chromophoric substrates/cofactor
- Look at effect of mutants
- Free energy of binding

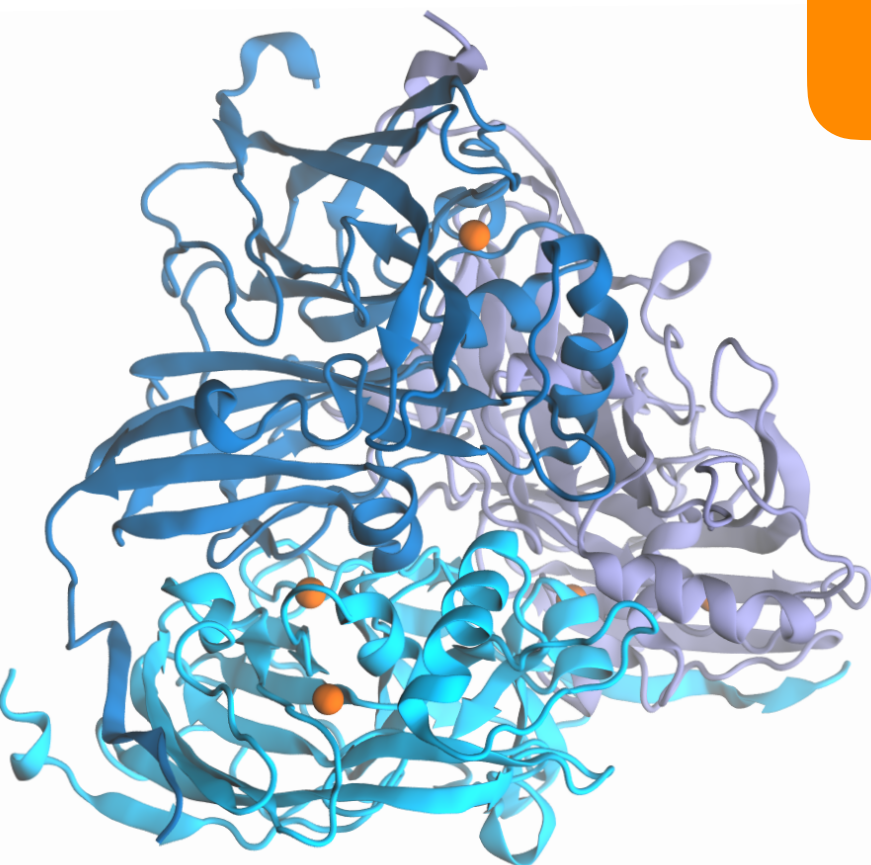
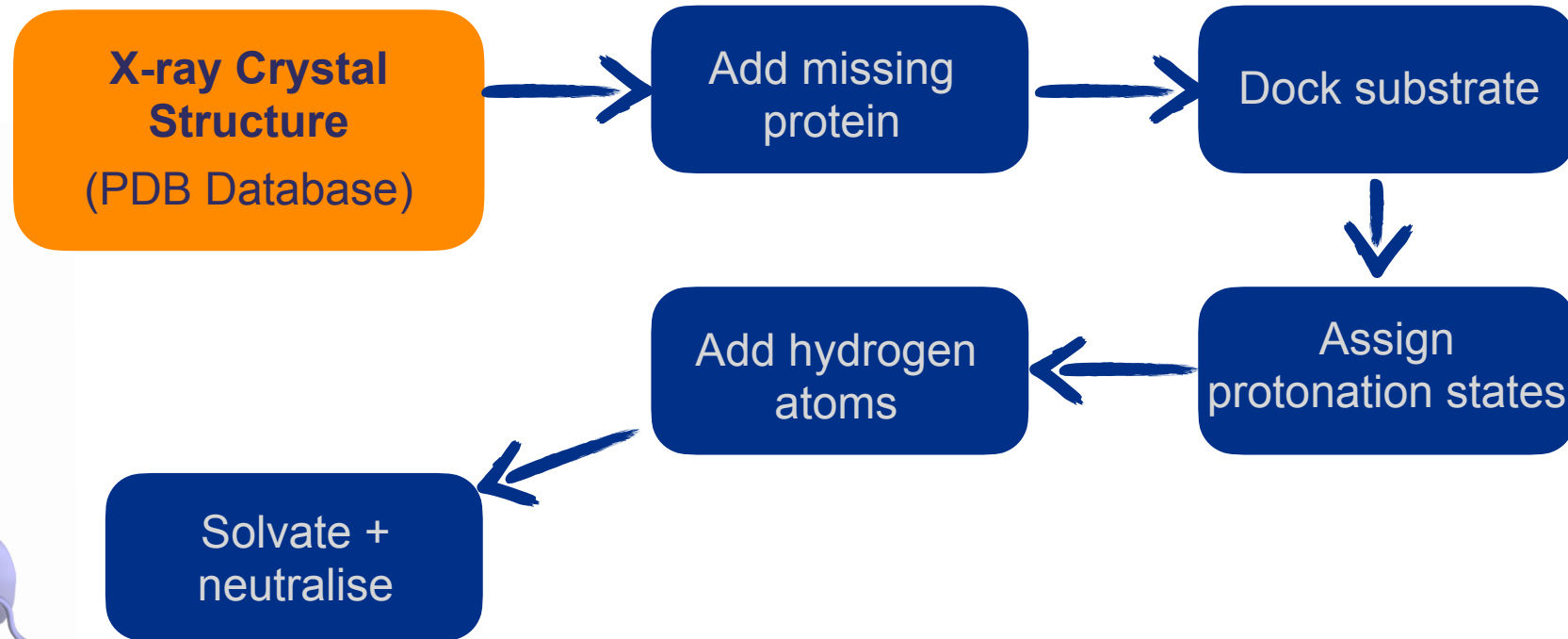


# WorkFlow

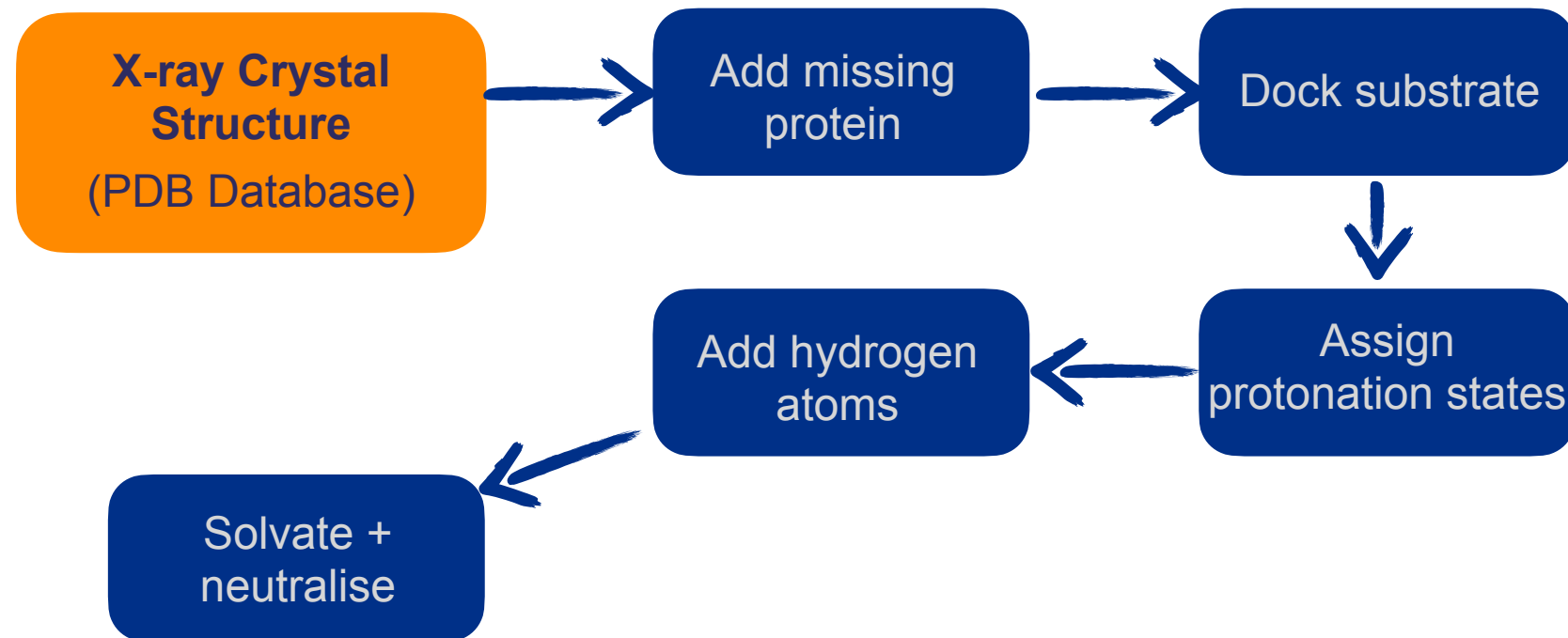
**X-ray Crystal  
Structure**  
(PDB Database)



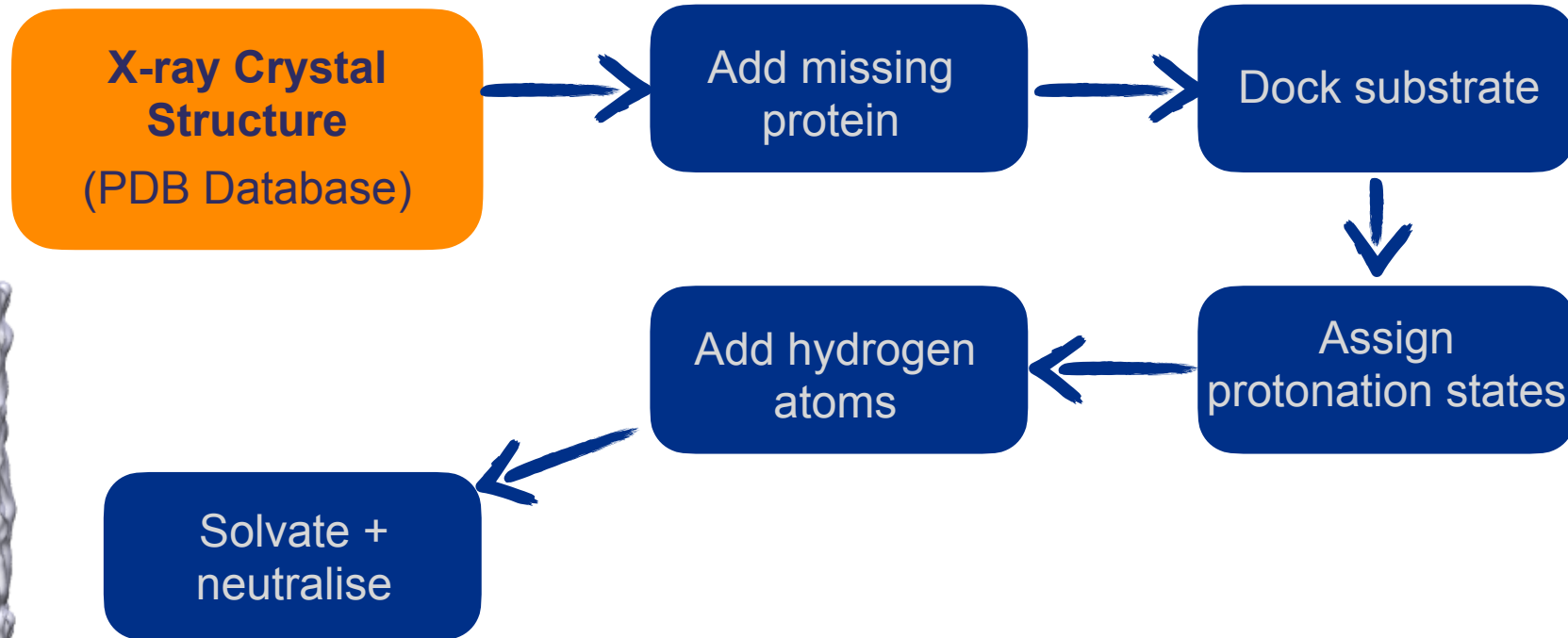
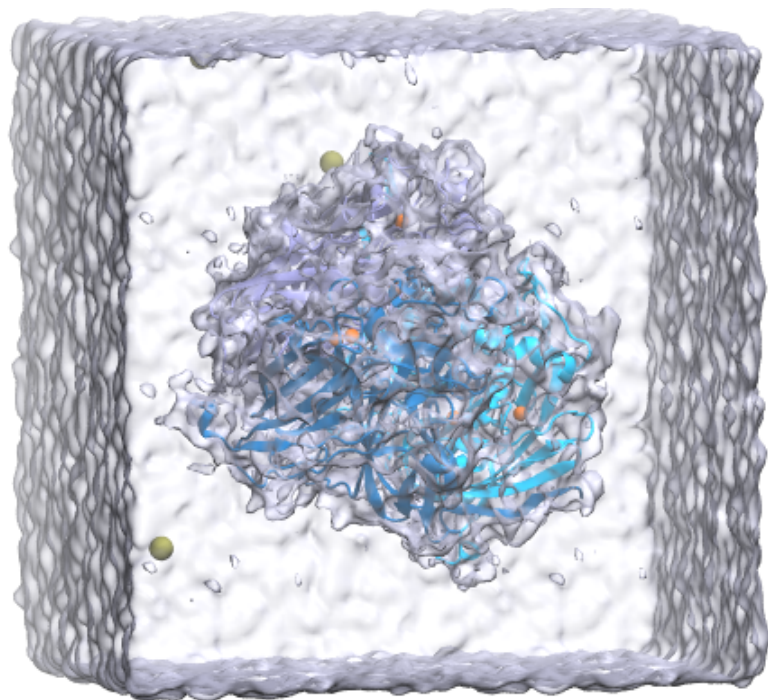
# WorkFlow



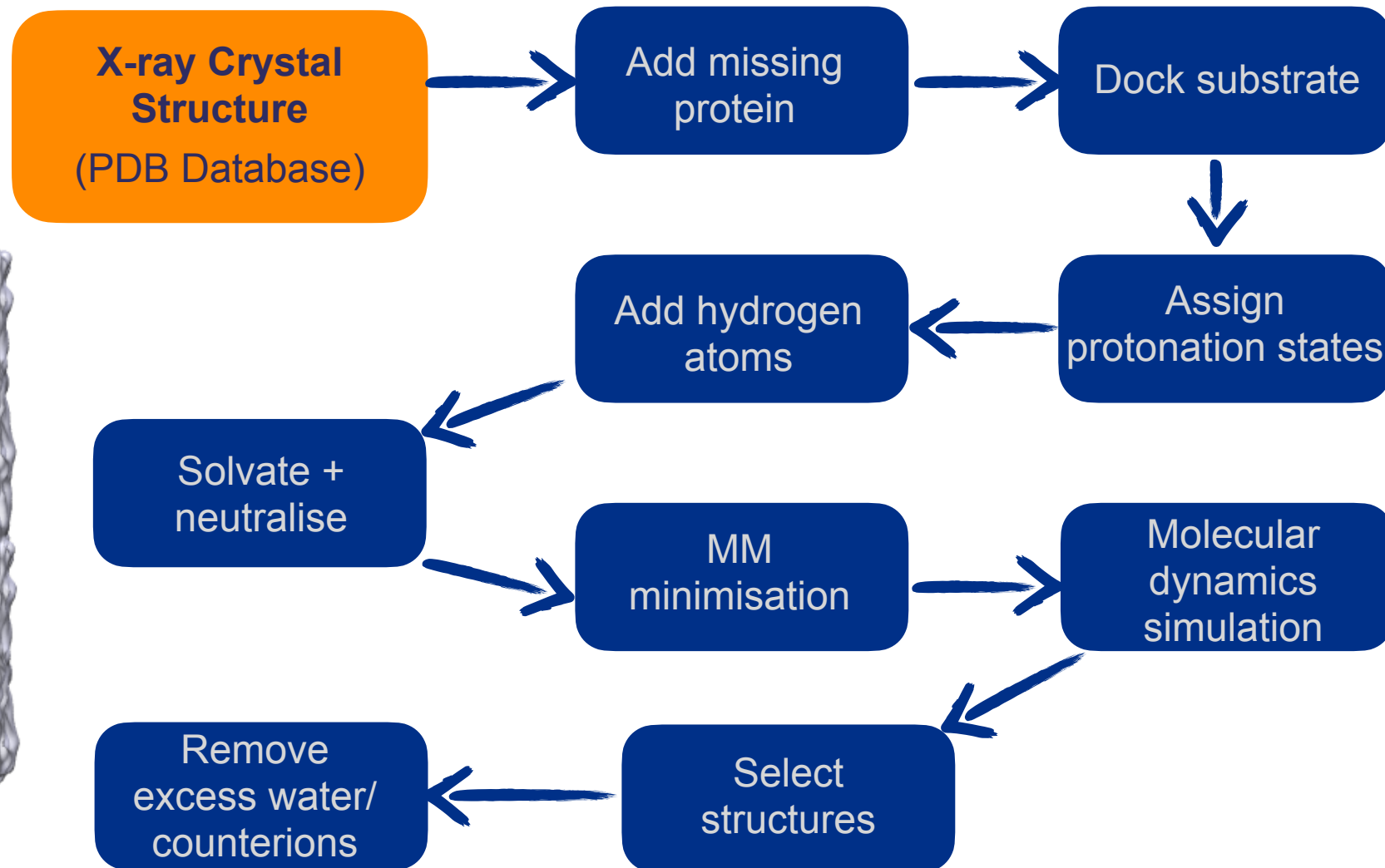
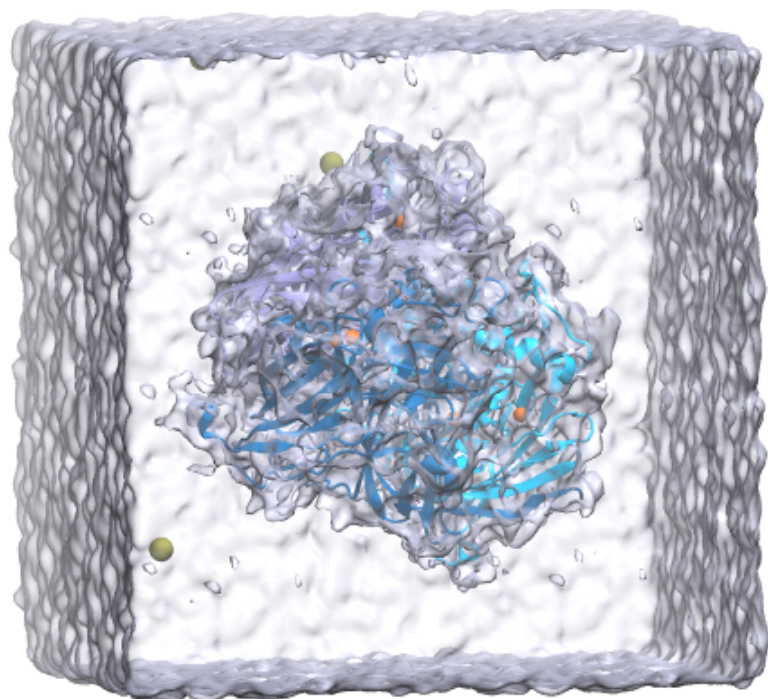
# WorkFlow



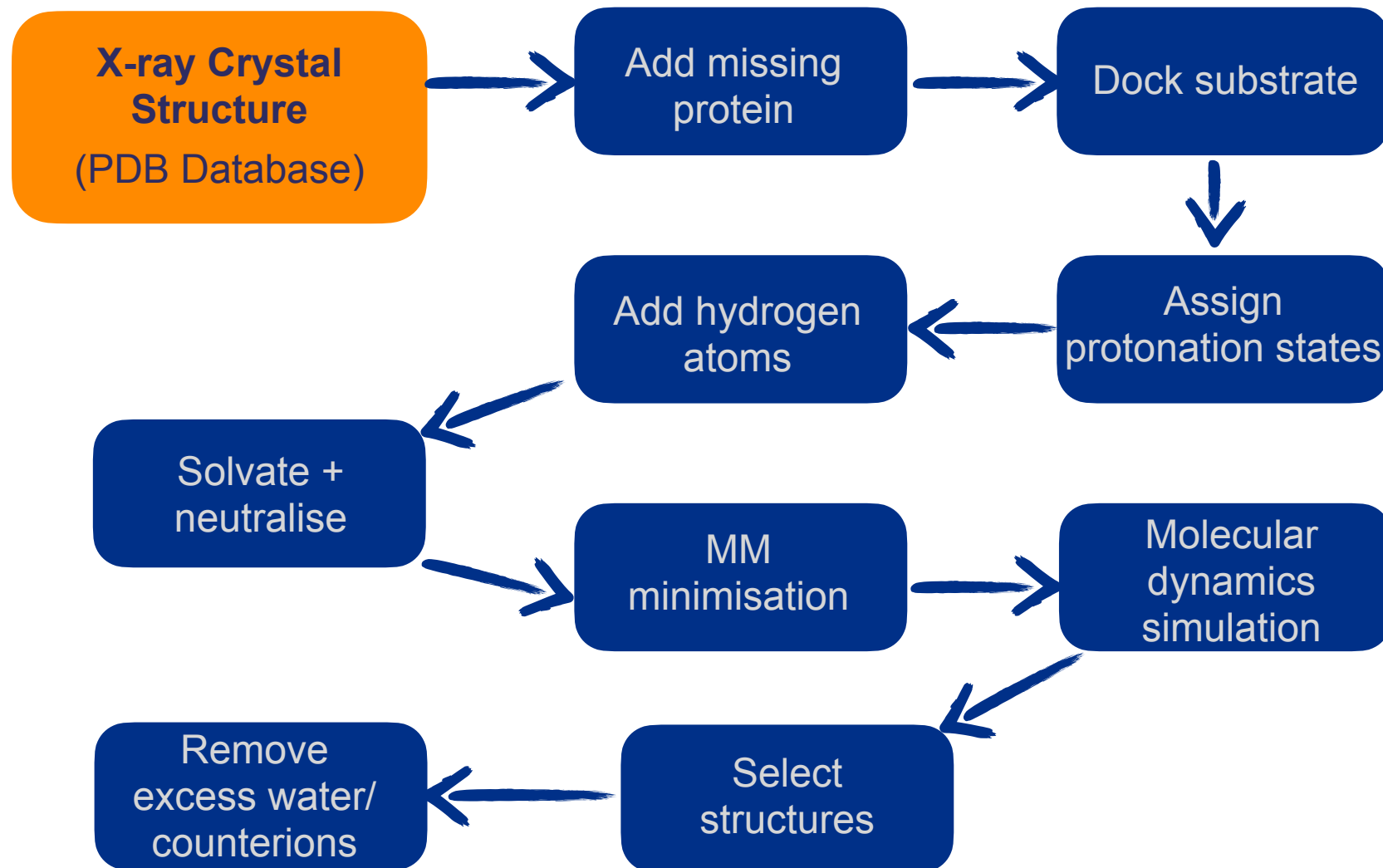
# WorkFlow



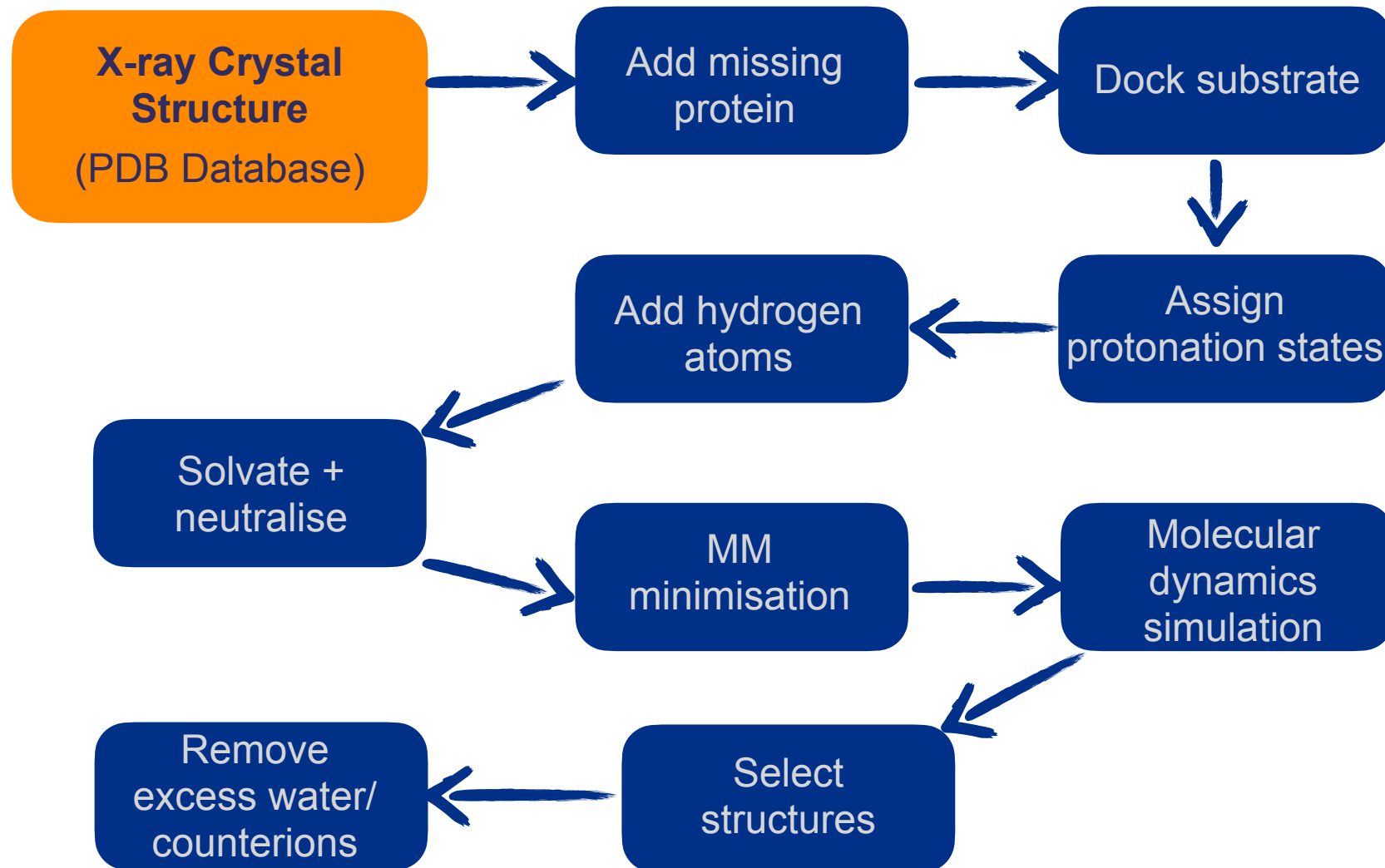
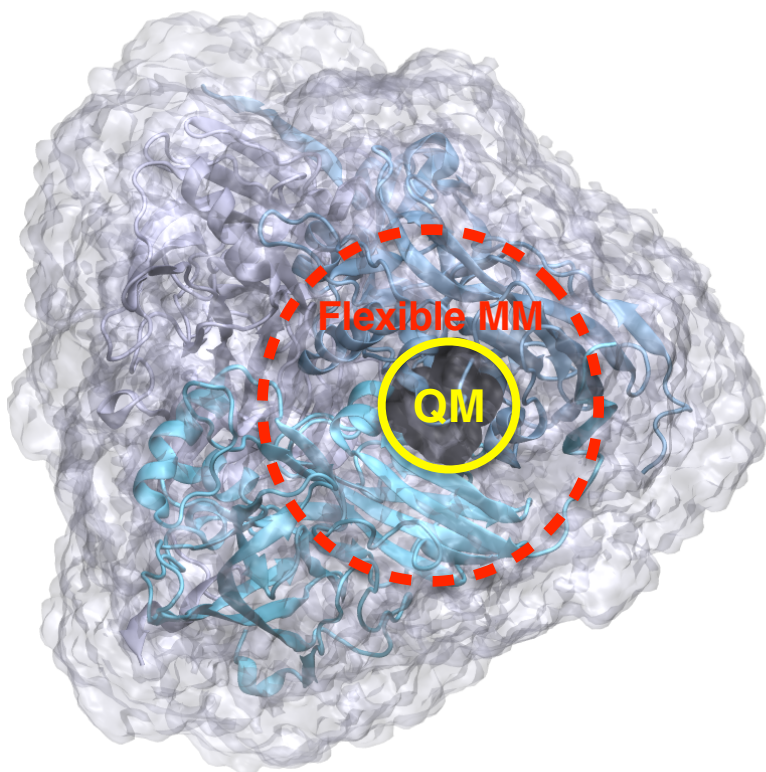
# WorkFlow



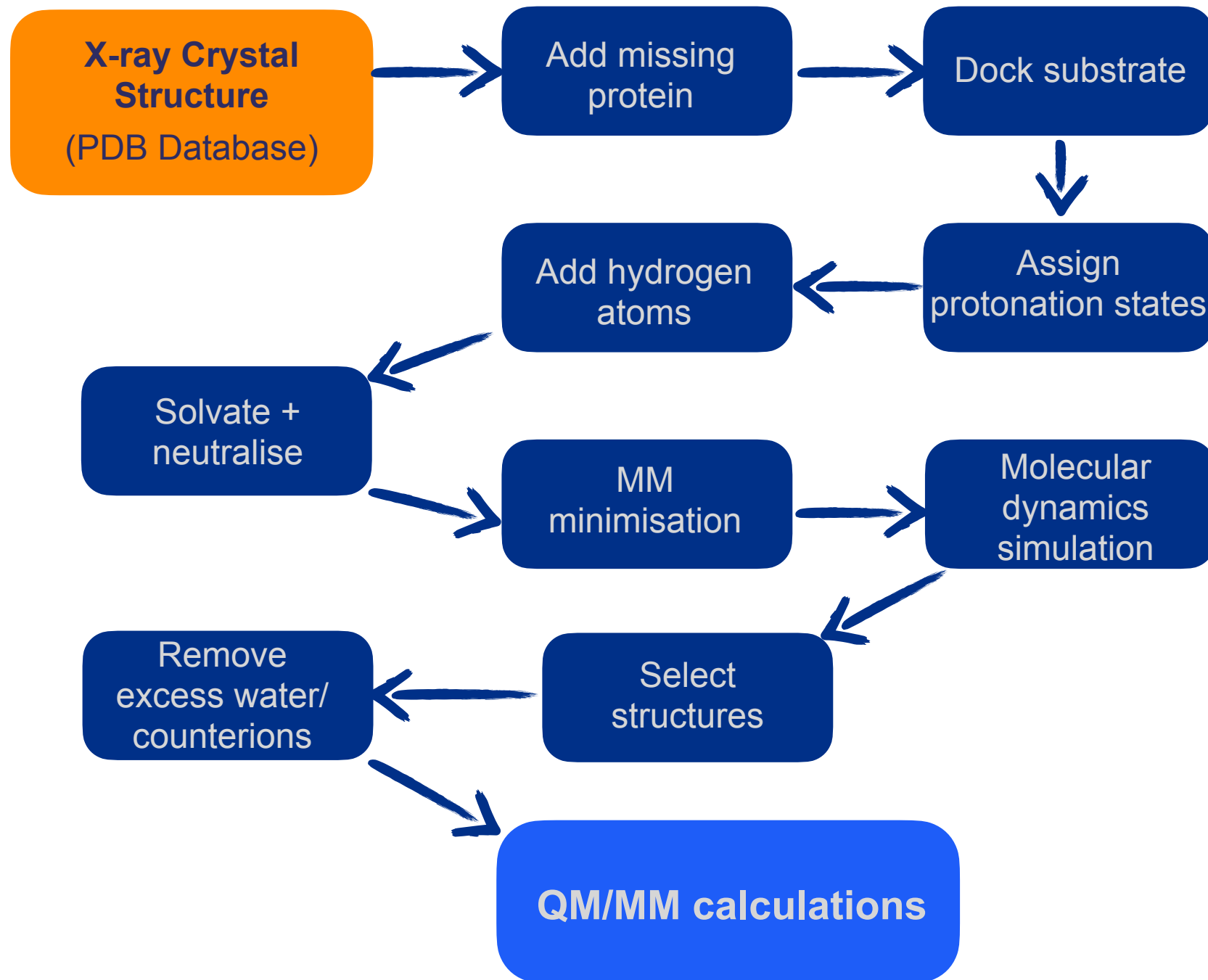
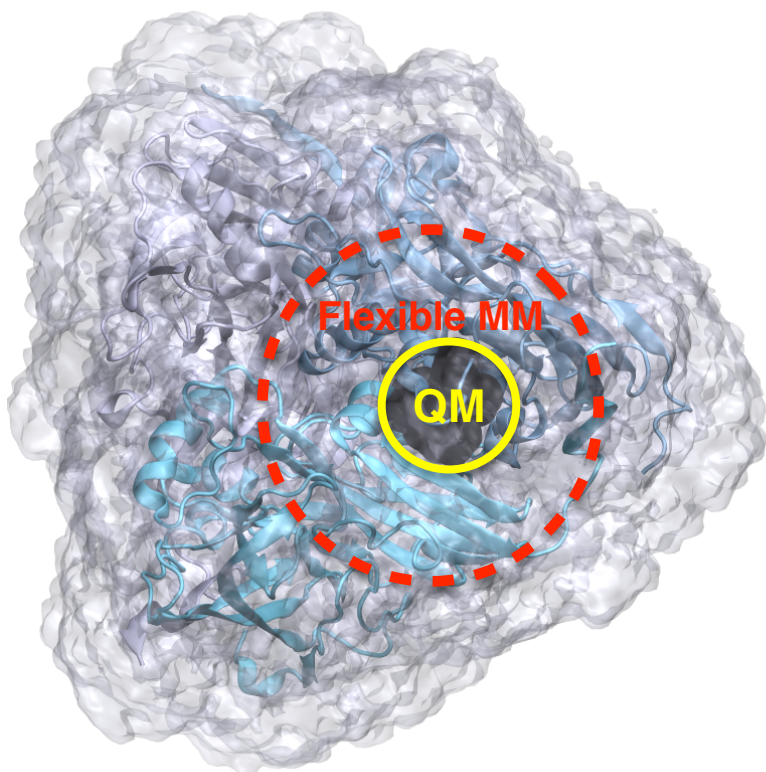
# WorkFlow



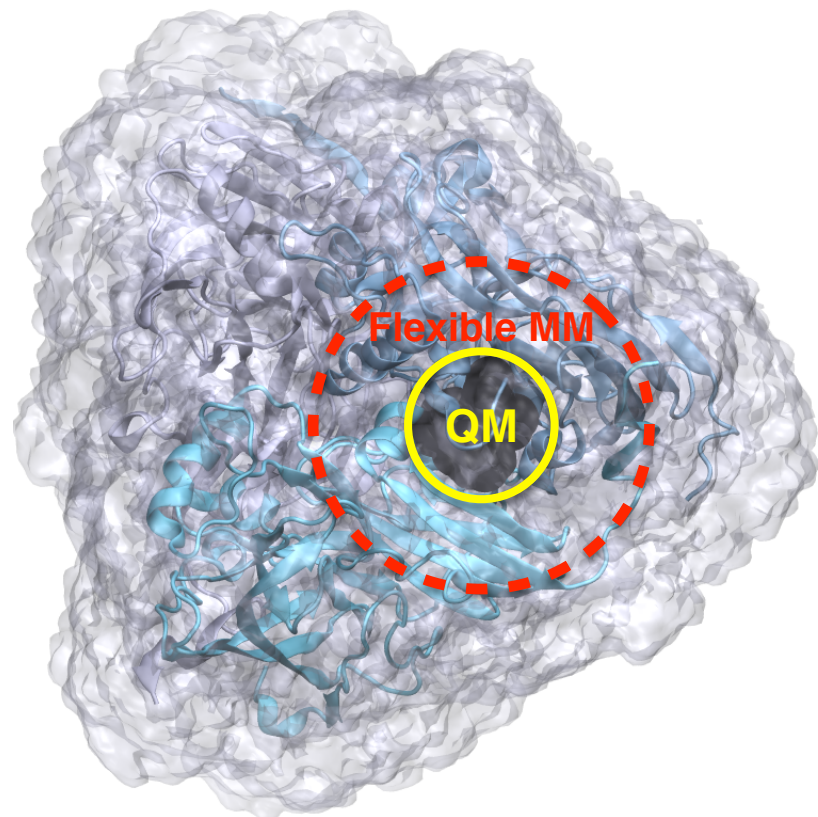
# WorkFlow



# WorkFlow



# QM/MM: General Considerations



QM/MM  
partitioning

Choice of QM  
method

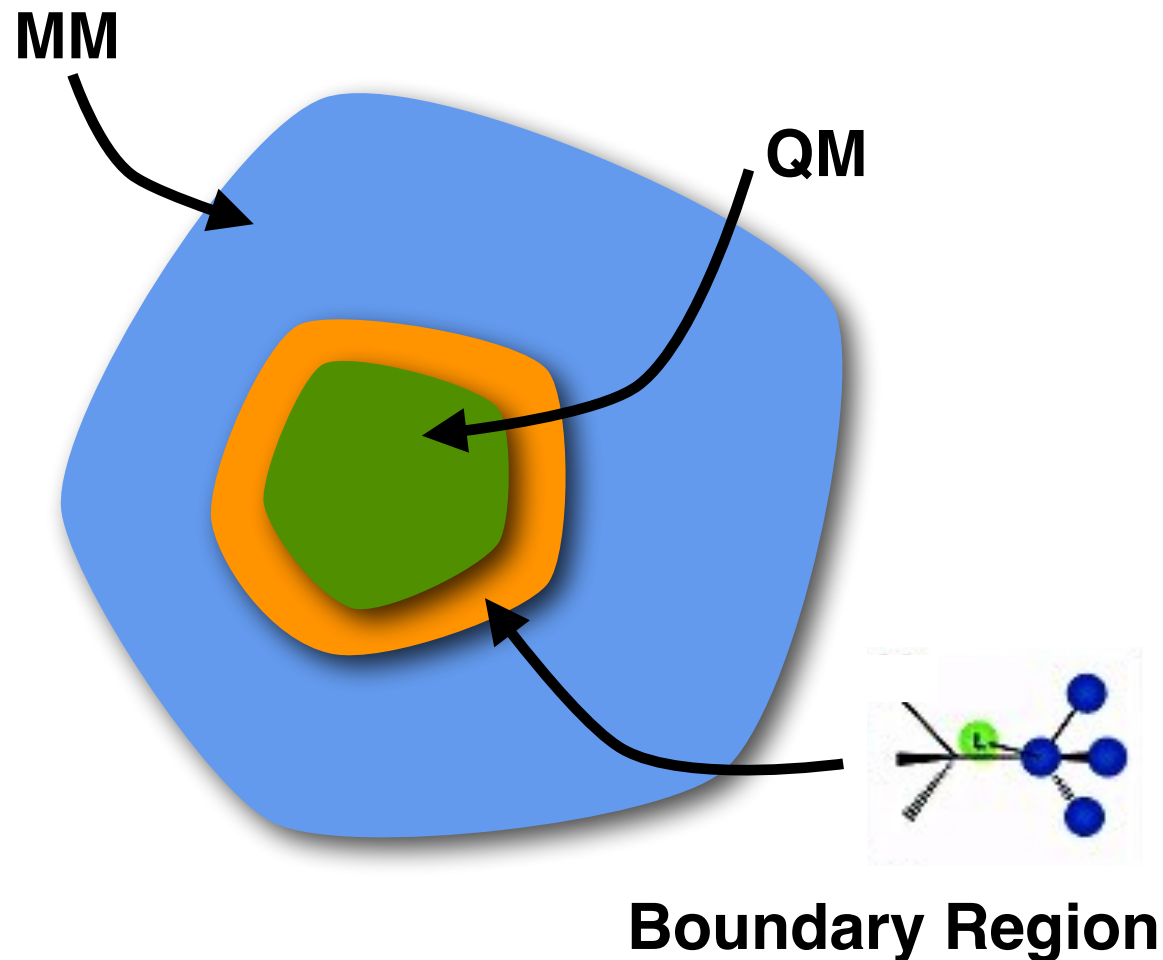
Active region  
during  
minimisation

Amount of  
conformational  
sampling

Choice of  
reaction  
coordinate

# QM/MM partitioning

- QM region must be large enough to capture reactivity
- Experiment with different QM regions
- Try to avoid placing QM-MM boundary across polar bond
  - ▶ C-C bond often best choice



# Choice of QM method

Dictated by:

- ▶ Size of QM region
- ▶ Elements present in active site (e.g. transition metals)
- ▶ Type of modelling (e.g. umbrella sampling)
- ▶ Level of accuracy required

## Semi-empirical methods

- Fast (for QM/MM MD – umbrella sampling profiles)
- Less accurate than other types of method
- Many not parameterised for d-metals

## Density functional theory

- Reasonable accuracy (better when dispersion included)
- Accurate geometries
- Good for transition metals
- Faster than Post Hartree-Fock methods

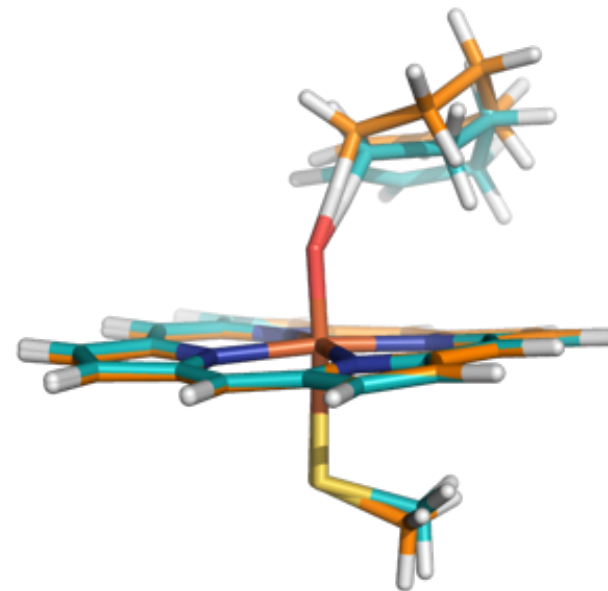
## Post Hartree-Fock ab initio methods

- Very accurate
- Computationally demanding
- Too expensive for d-metals
- Applicable to small QM regions

# Choice of QM method

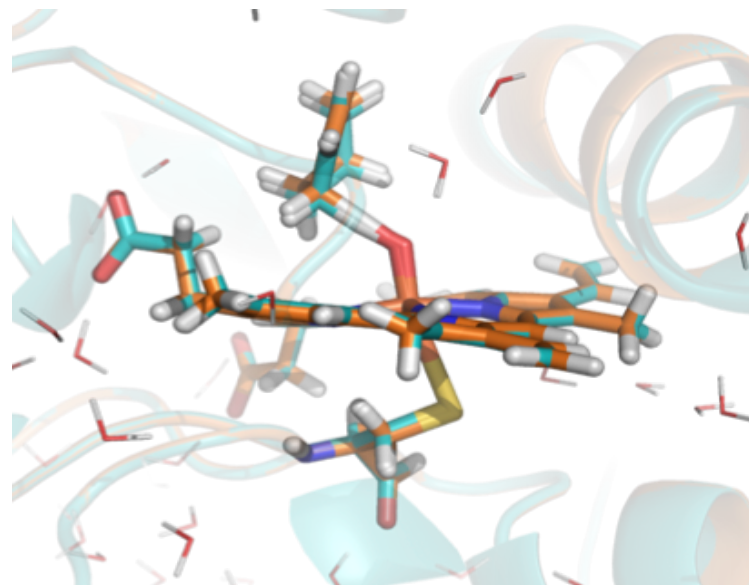
## Effect of Dispersion

- ▶ Missing from most functionals
- ▶ Attractive interaction
- ▶ Large effect on QM-only optimised geometries
- ▶ Smaller effect on QM/MM geometries



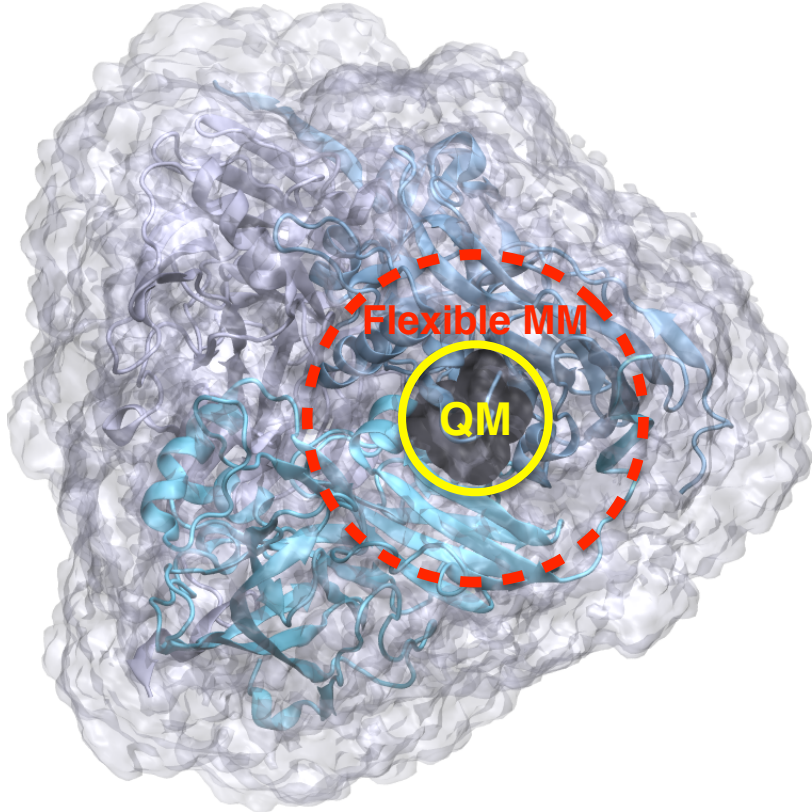
QM/MM

B3LYP  
B3LYP-D



QM/MM

# Active atoms

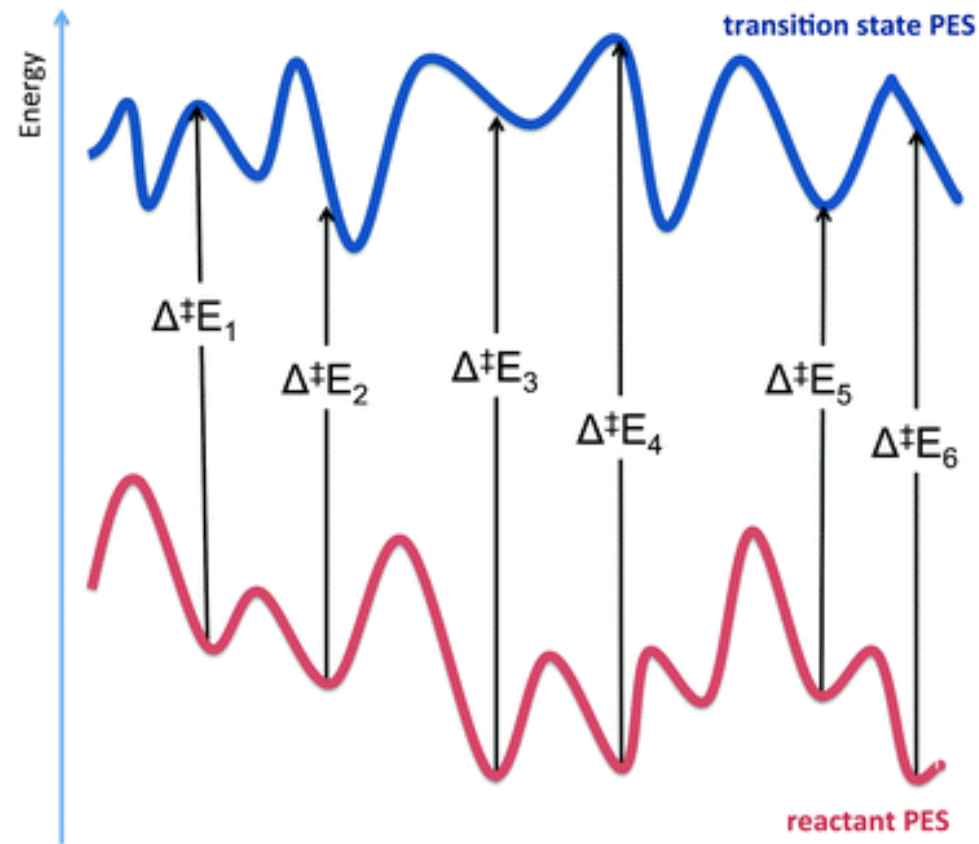
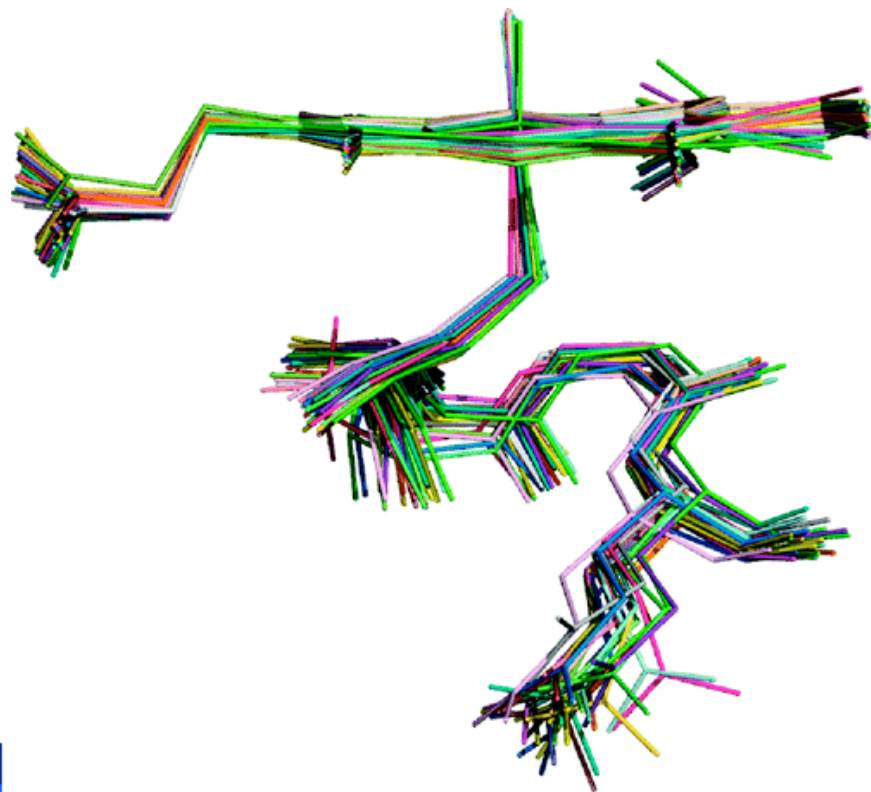


Selection of MM atoms included in optimisation can have a large effect on the results (and time required for calculation) — need benchmarking

- ▶ Shell around QM region
- ▶ Sphere of a defined radius from a QM atom
- ▶ QM atoms always active

# Conformational Sampling

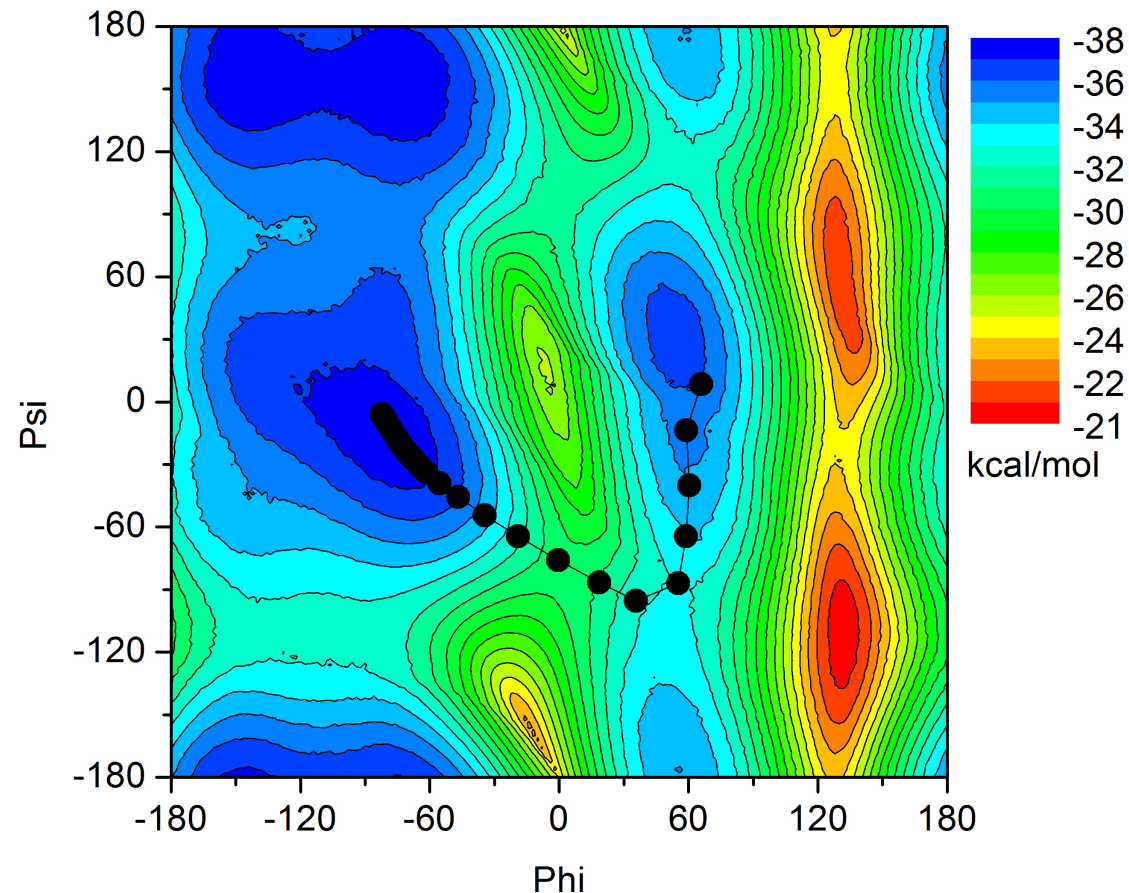
QM/MM optimised structures for an intermediate (from MD simulation)



Calculated barriers vary depending on starting structure from MD

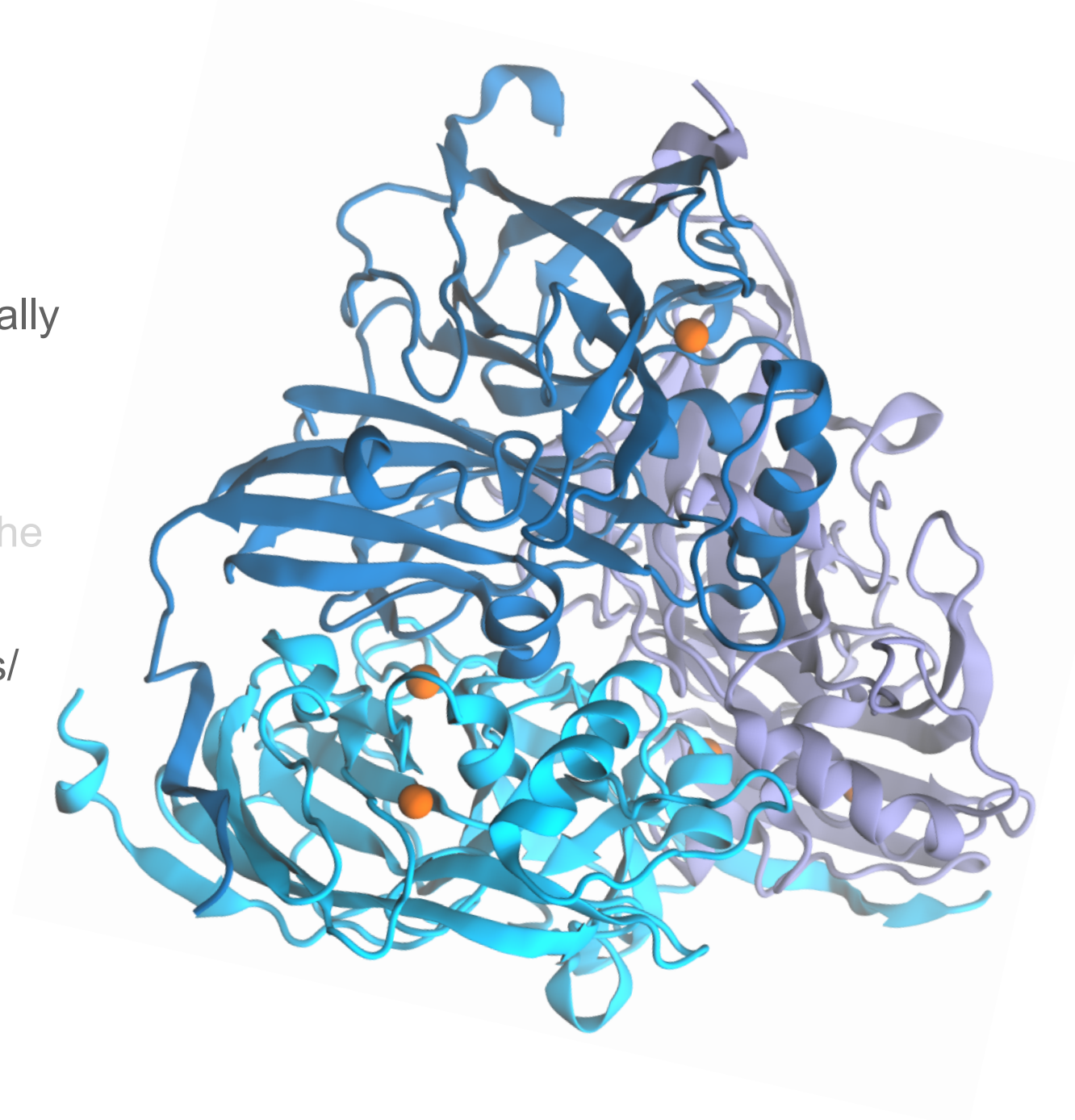
# Reaction coordinate

- Simple: distance, angle, dihedrals etc
- Combination of simple ones
- Concerted mechanism
  - ▶ 2-dimensional energy surface
- Not obvious
  - ▶ Nudged elastic band



# QM/MM for enzymes

- Investigate mechanisms
- Characterise intermediates that are experimentally elusive; e.g., transition states
- Make predictions of selectivity
- Understand the stabilising interactions around the active site/binding pockets
- Emission properties of chromophoric substrates/cofactor
- Look at effect of mutants
- Free energy of binding



# QM/MM for enzymes

🌐 Investigate mechanisms

# QM/MM for enzymes

🌐 Investigate mechanisms



# QM/MM for enzymes

🌐 Investigate mechanisms



Reaction coordinate: Simple distance

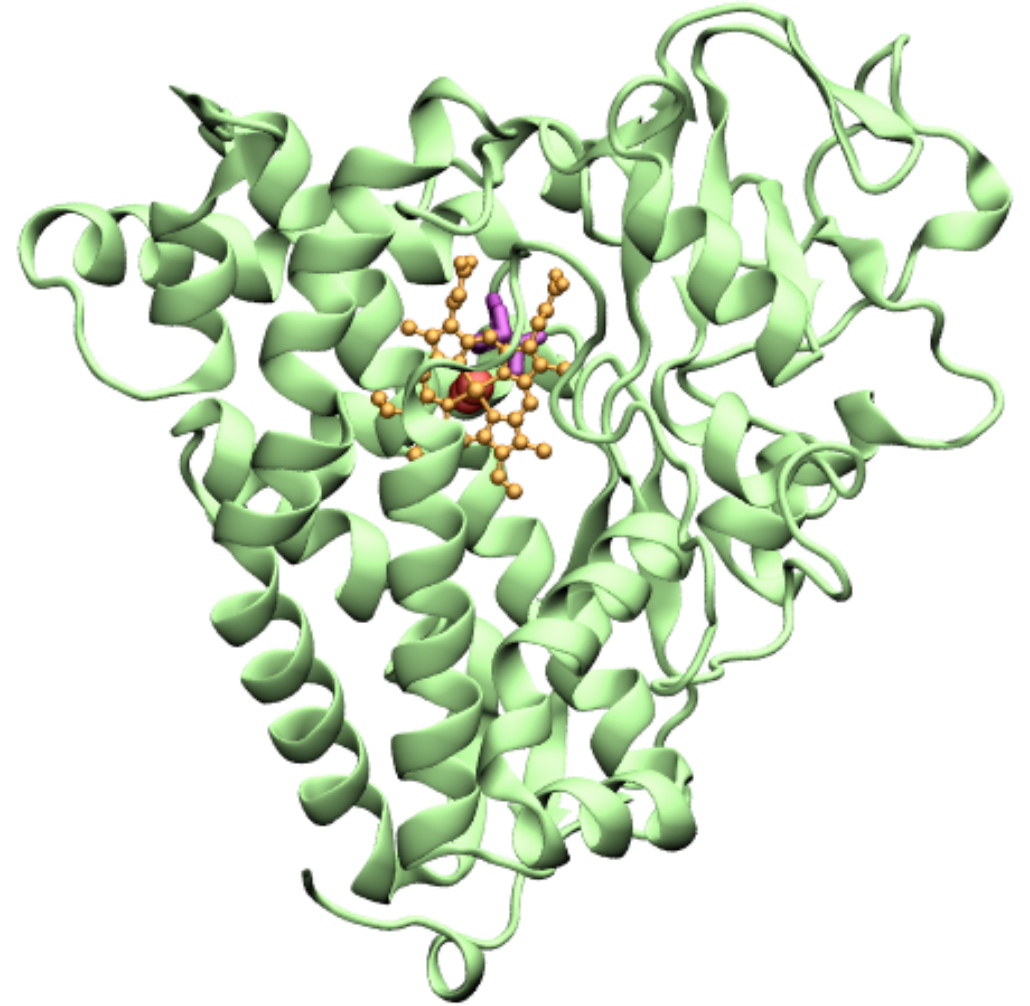
# QM/MM for enzymes

🌐 Investigate mechanisms



Reaction coordinate: Simple distance

## P450 Enzymes



# QM/MM for enzymes

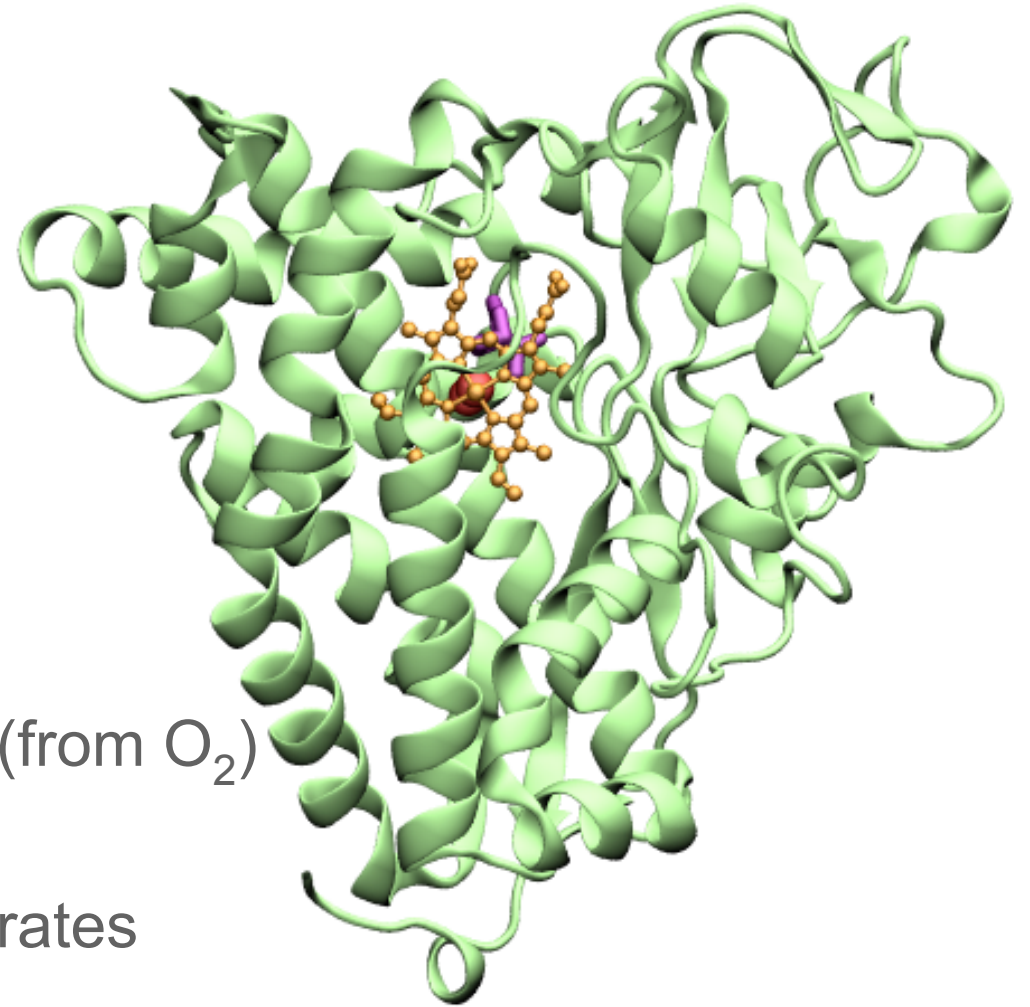
🌐 Investigate mechanisms



Reaction coordinate: Simple distance

- Heme containing enzyme
- Important in drug metabolism
- Insert a single atom of oxygen (from  $O_2$ ) into substrate
- React with wide range of substrates

## P450 Enzymes



# QM/MM for enzymes

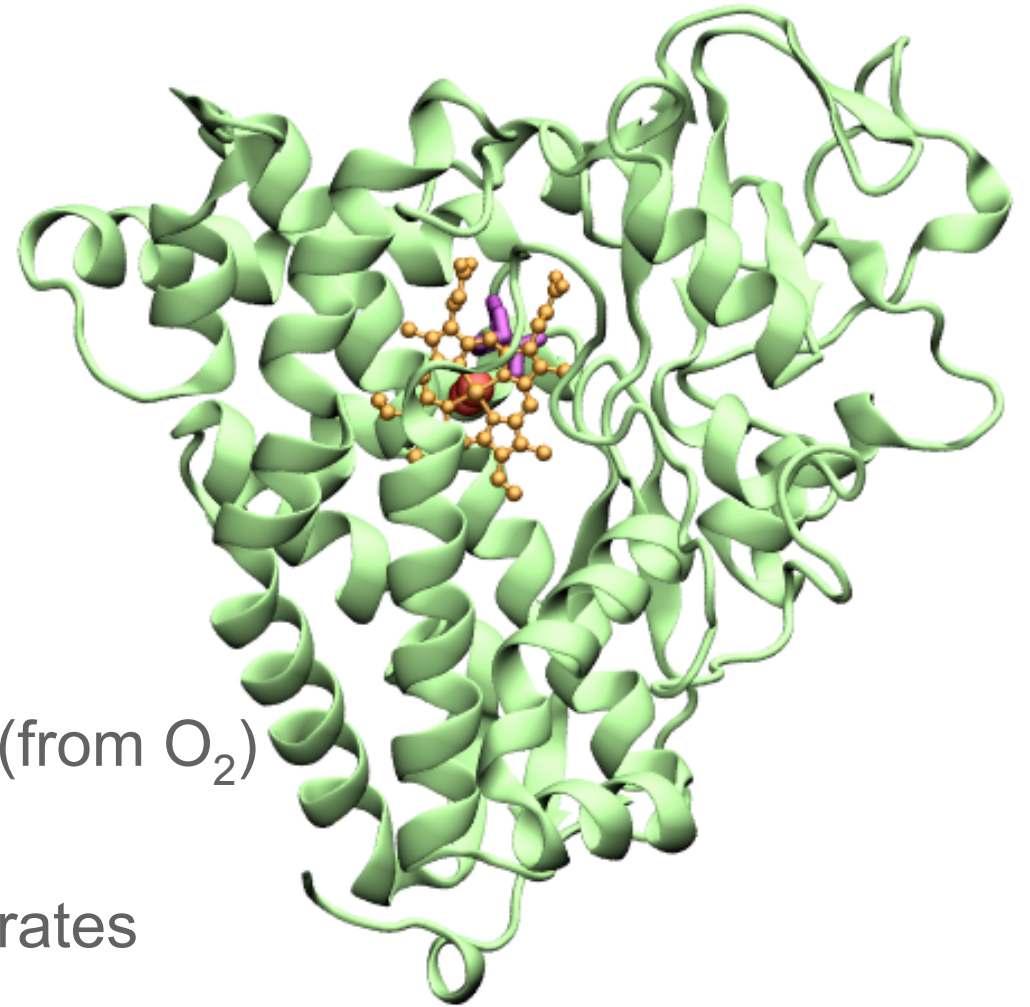
Investigate mechanisms



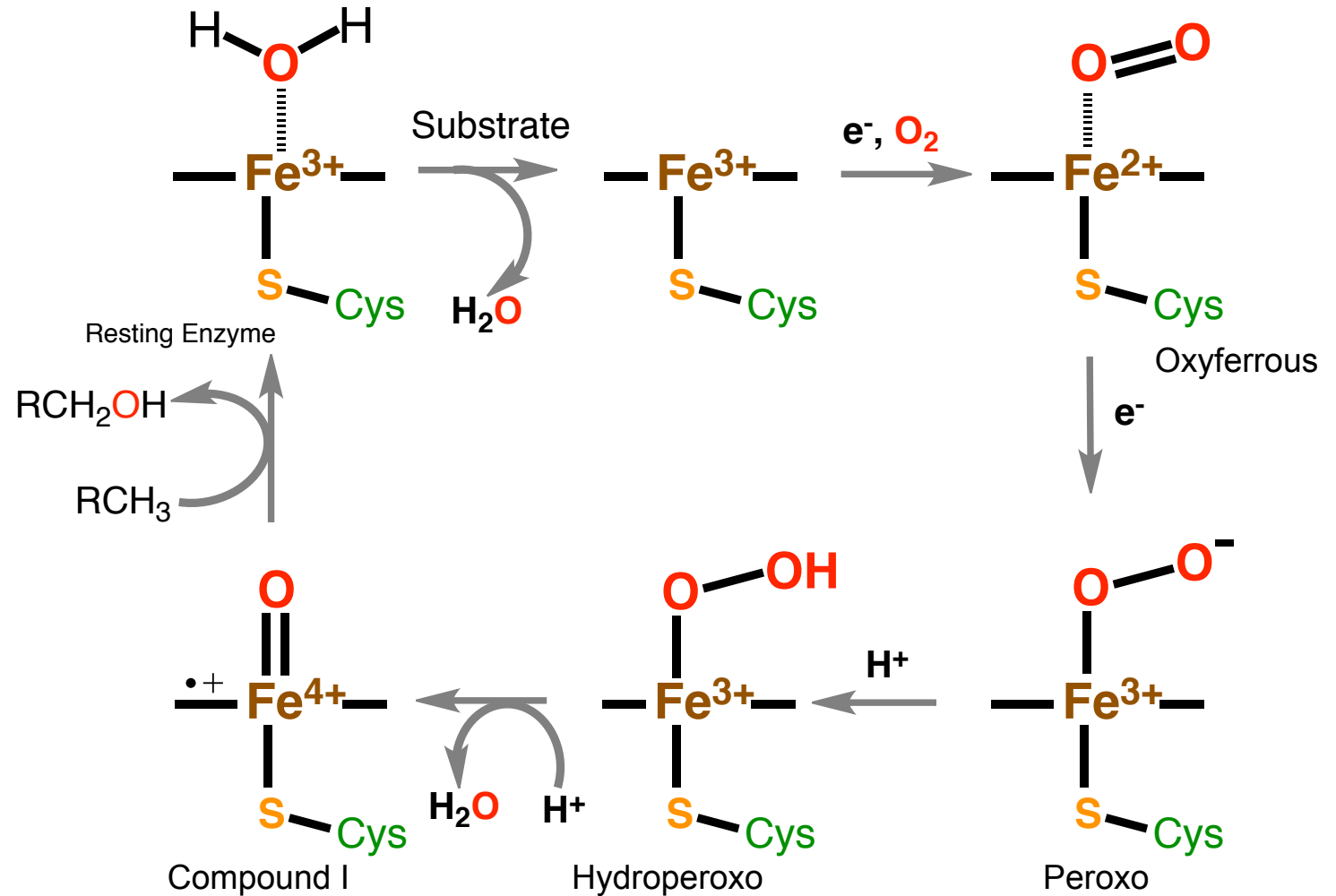
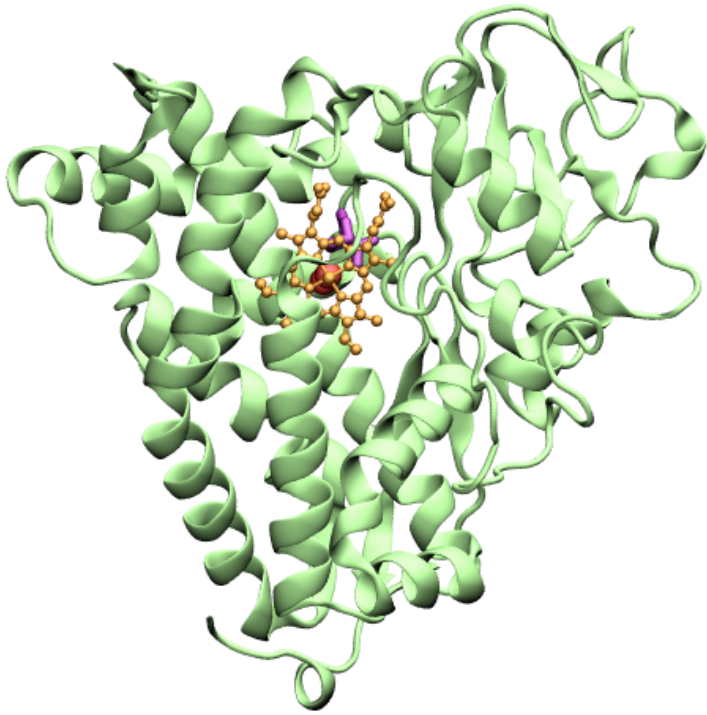
Reaction coordinate: Simple distance

- Heme containing enzyme
- Important in drug metabolism
- Insert a single atom of oxygen (from O<sub>2</sub>) into substrate
- React with wide range of substrates

## P450 Enzymes

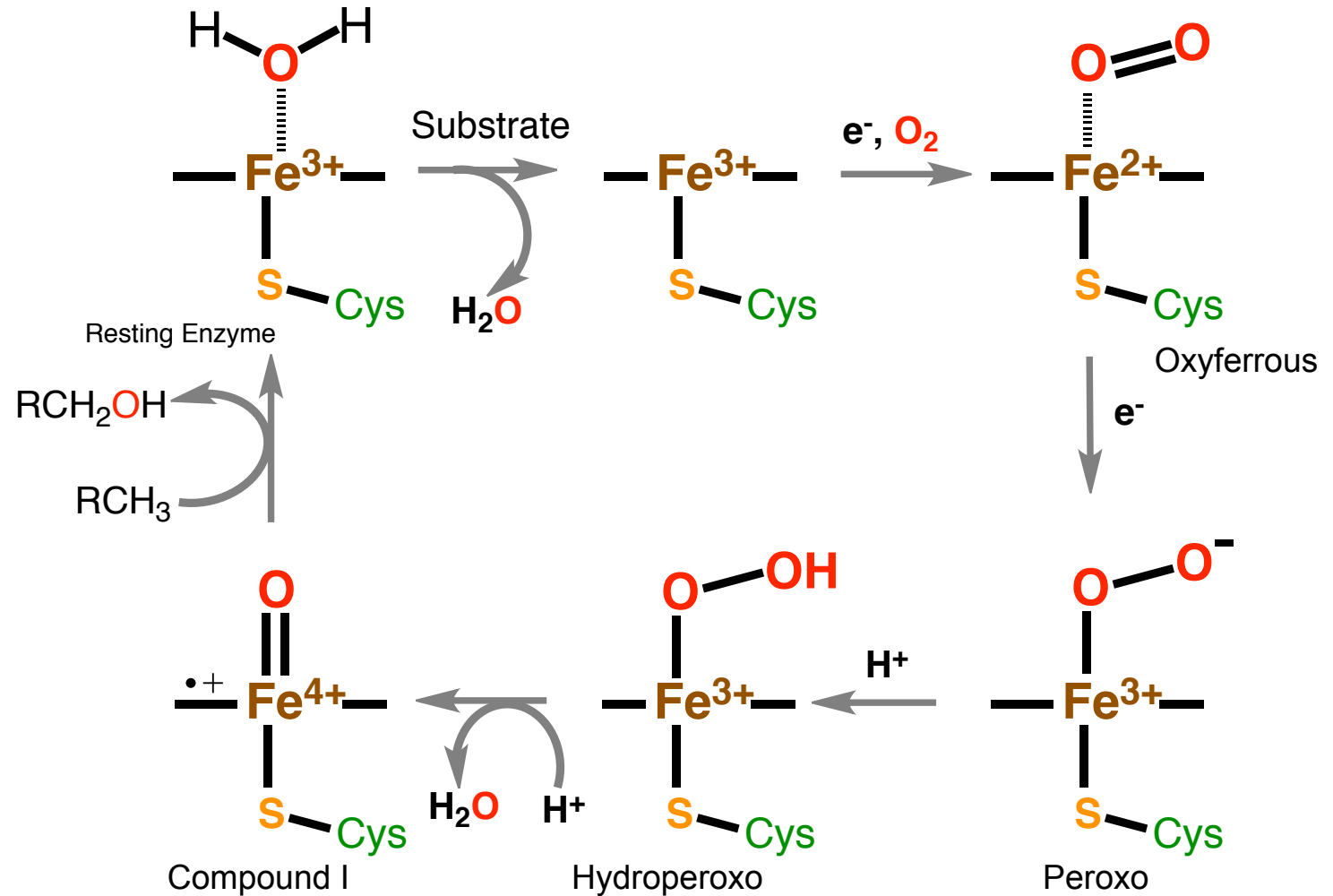
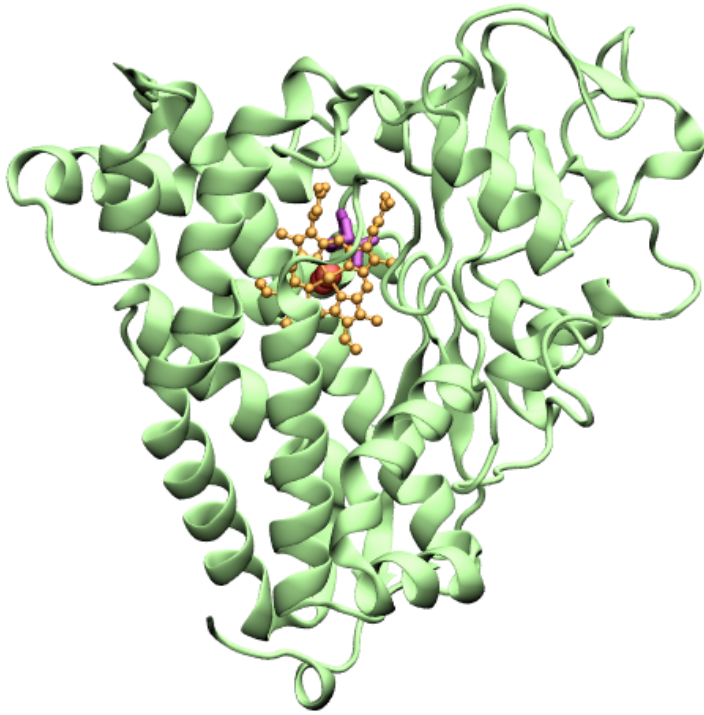


# P450 Catalysis



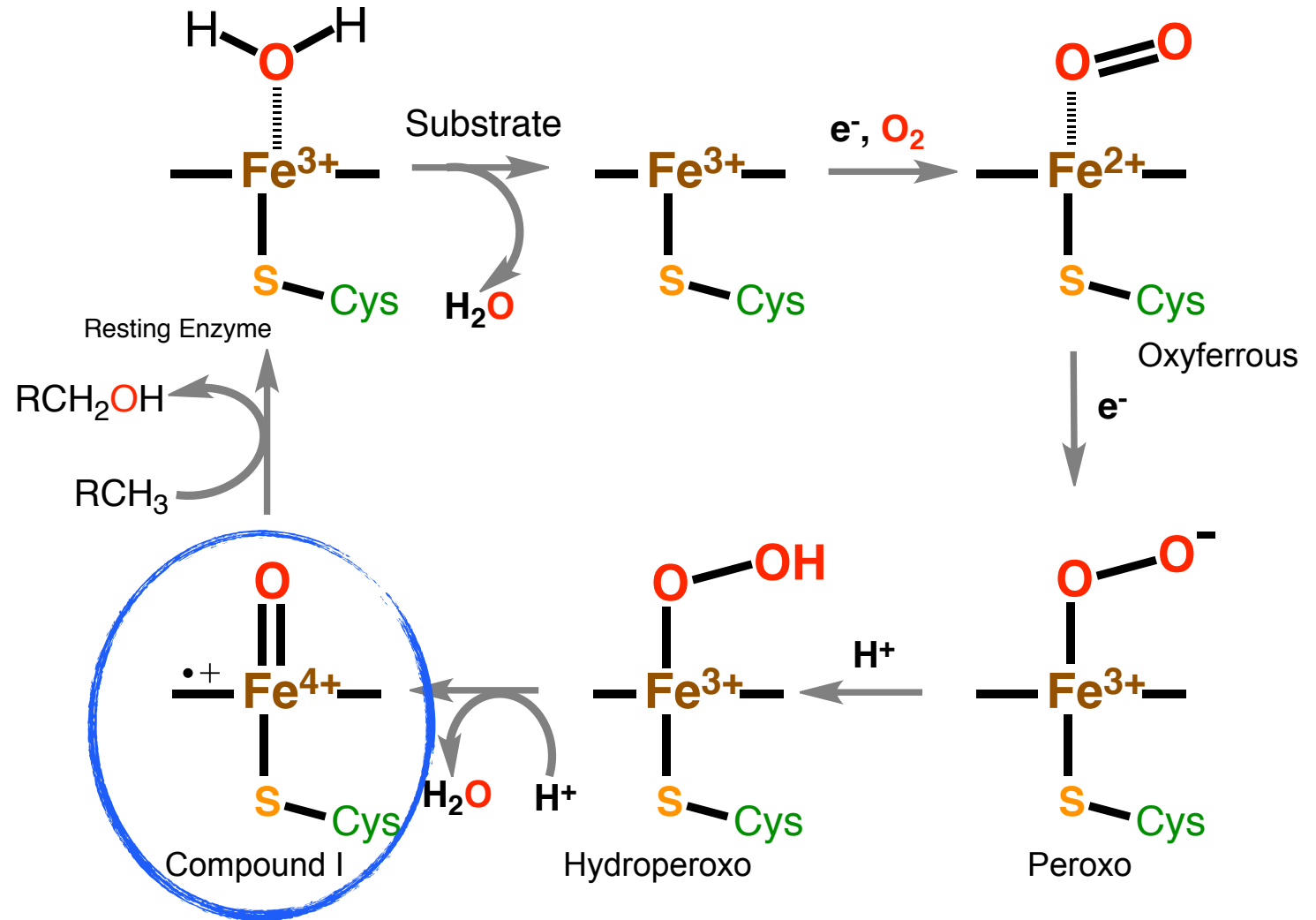
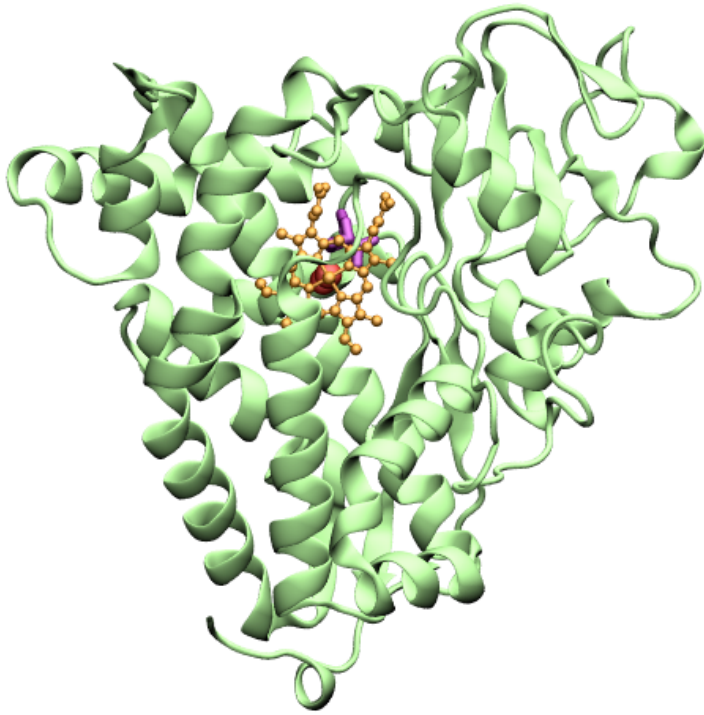
# P450 Catalysis

## Activation of the catalytic cycle



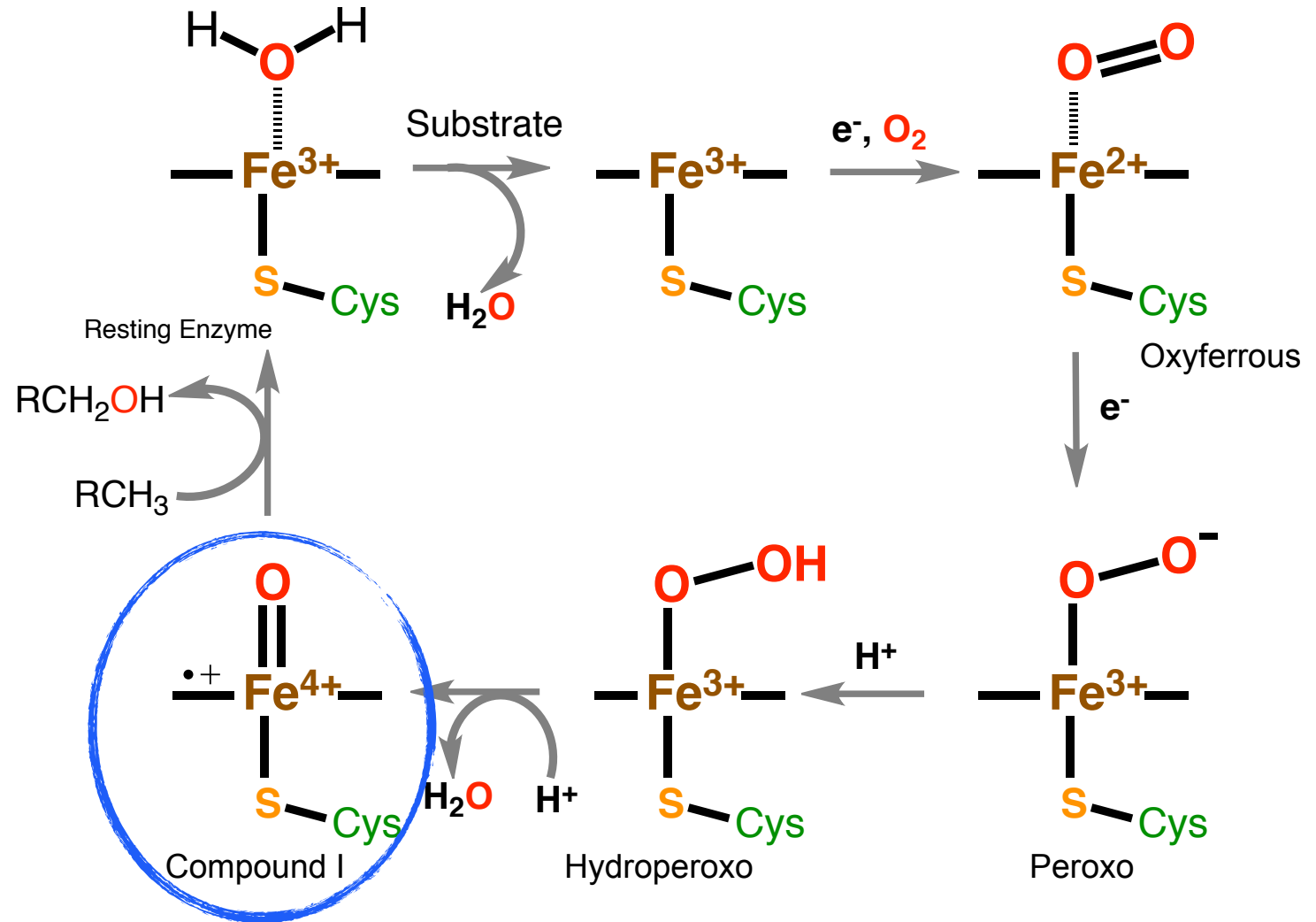
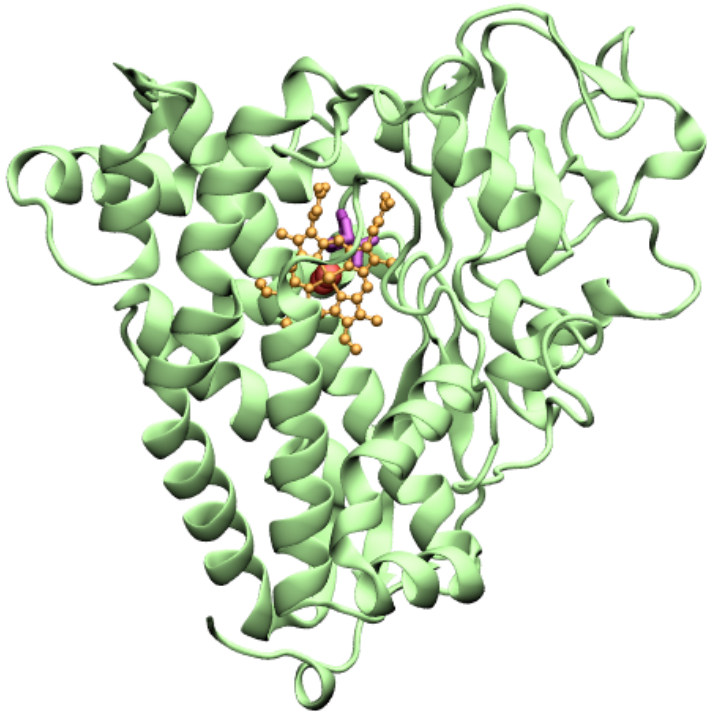
# P450 Catalysis

## Activation of the catalytic cycle



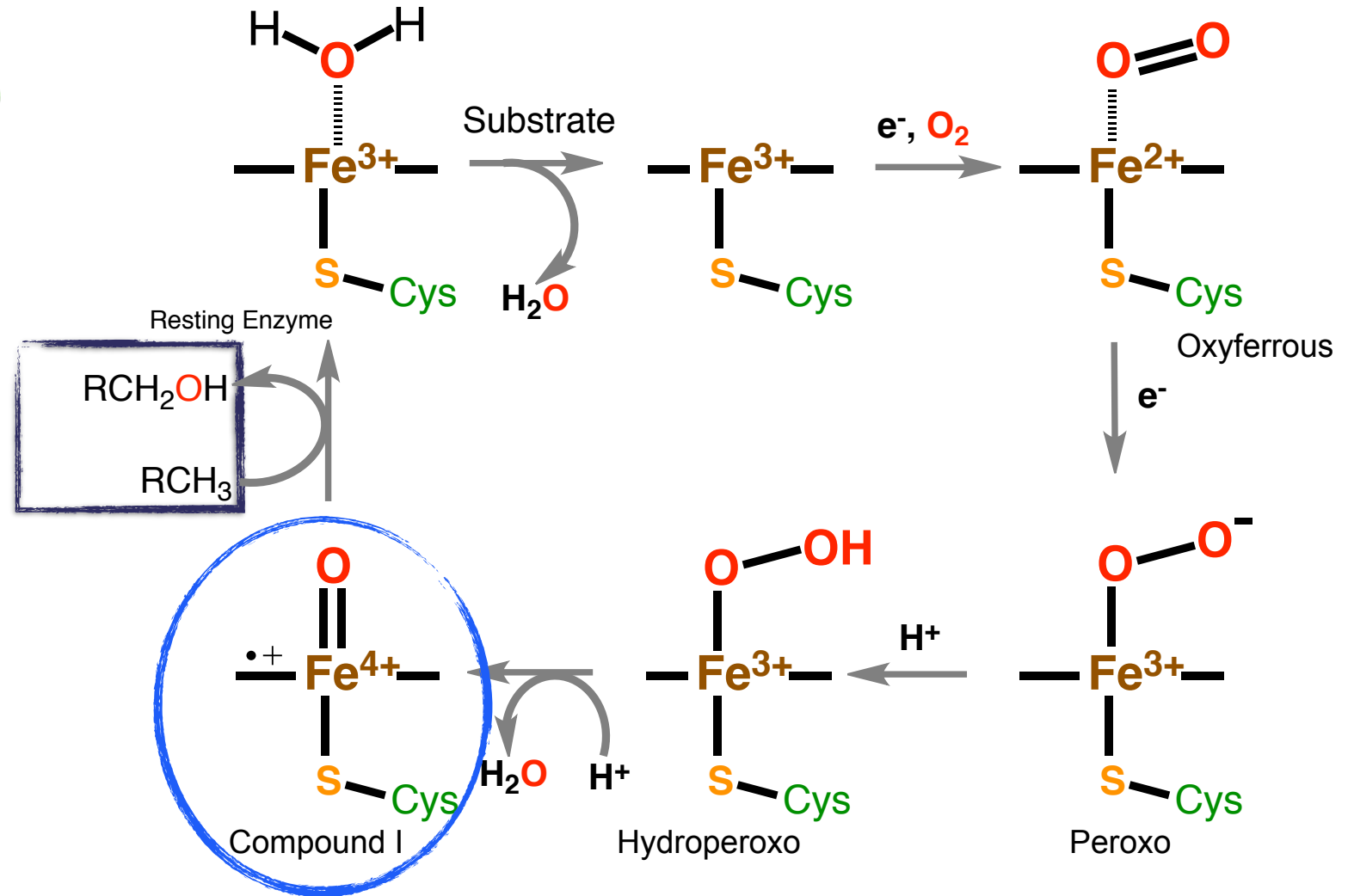
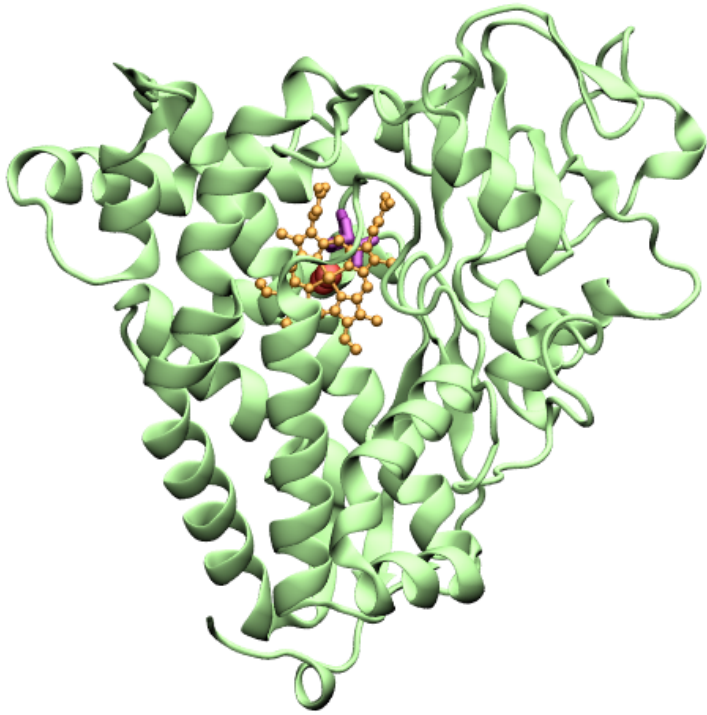
# P450 Catalysis

- Activation of the catalytic cycle
- Oxidation of substrate



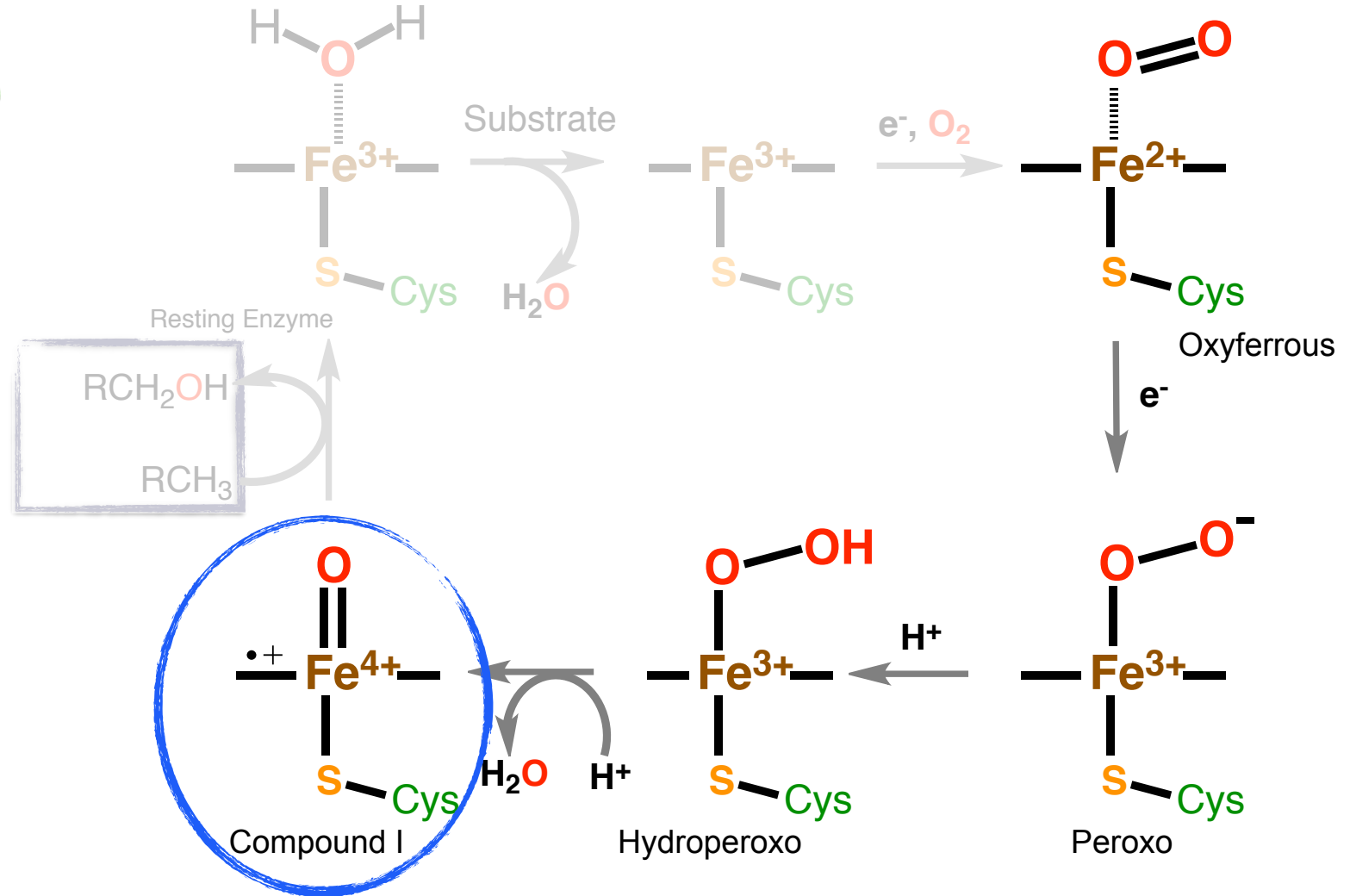
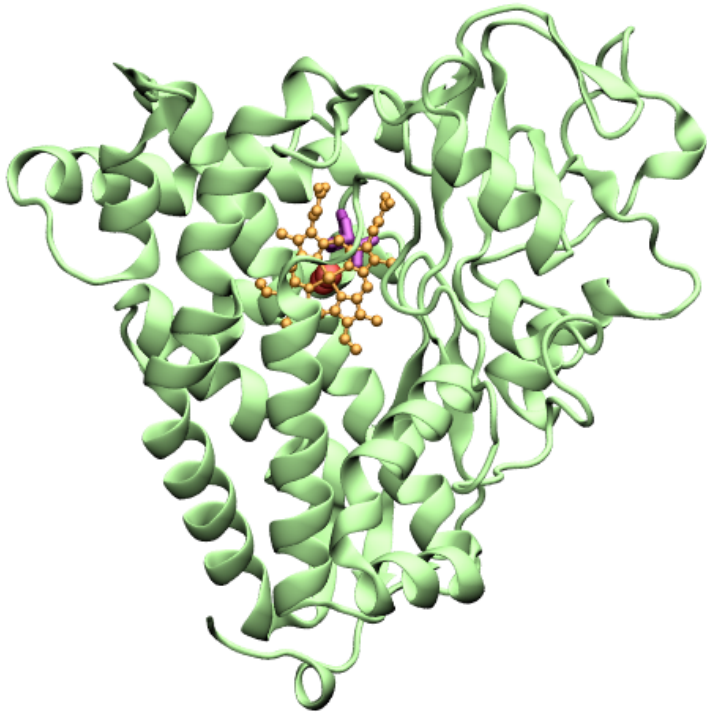
# P450 Catalysis

- Activation of the catalytic cycle
- Oxidation of substrate



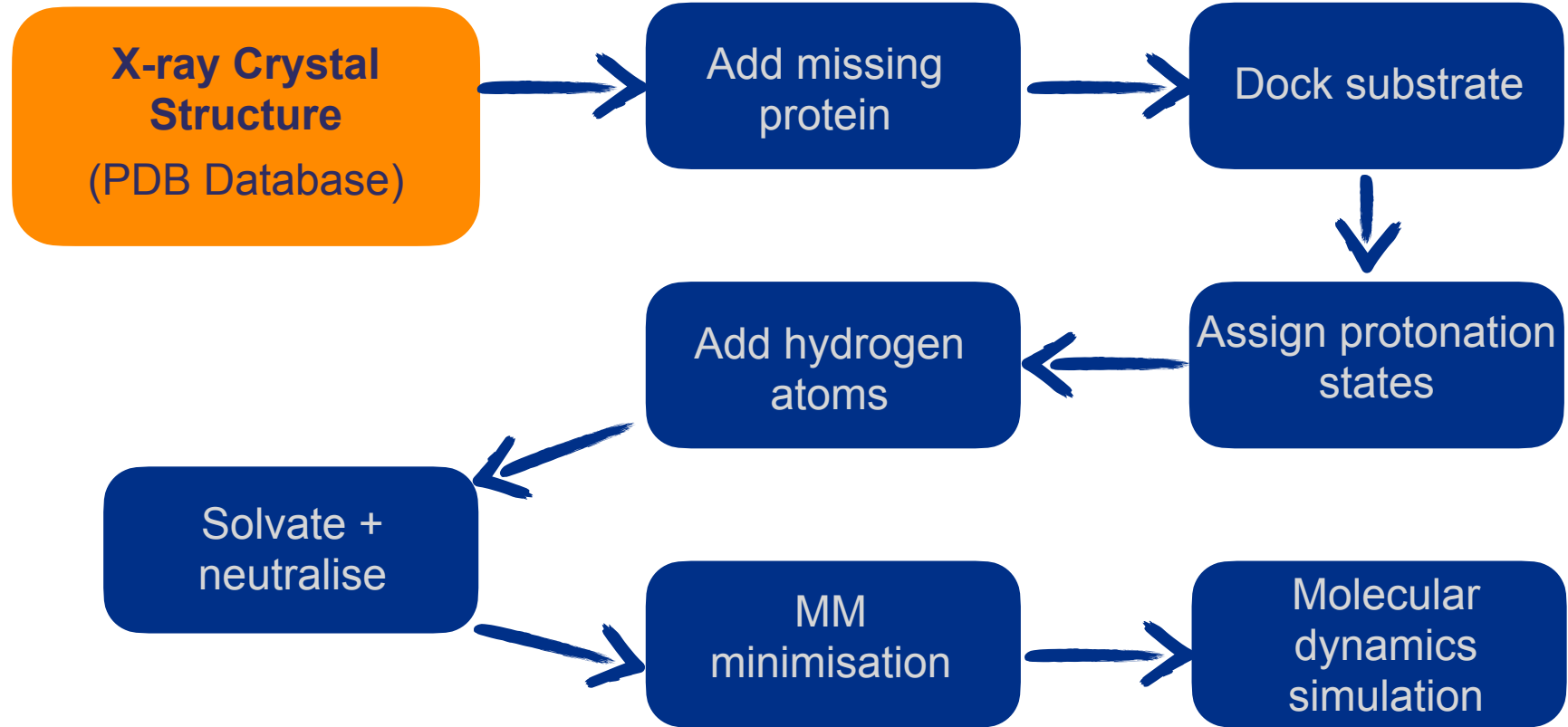
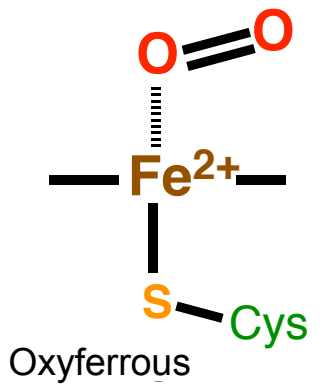
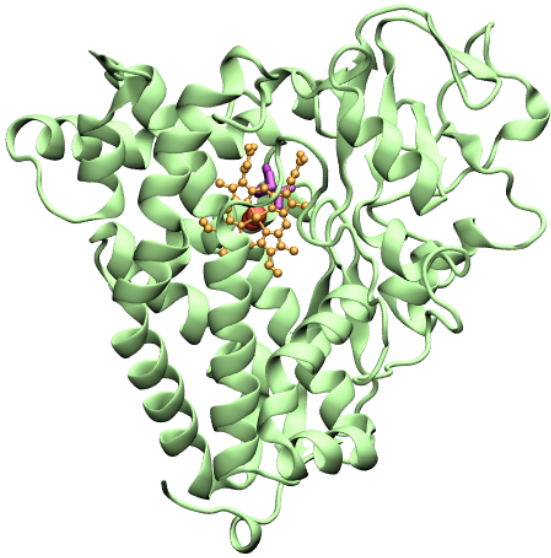
# P450 Catalysis

- Activation of the catalytic cycle
- Oxidation of substrate



# P450 Catalysis

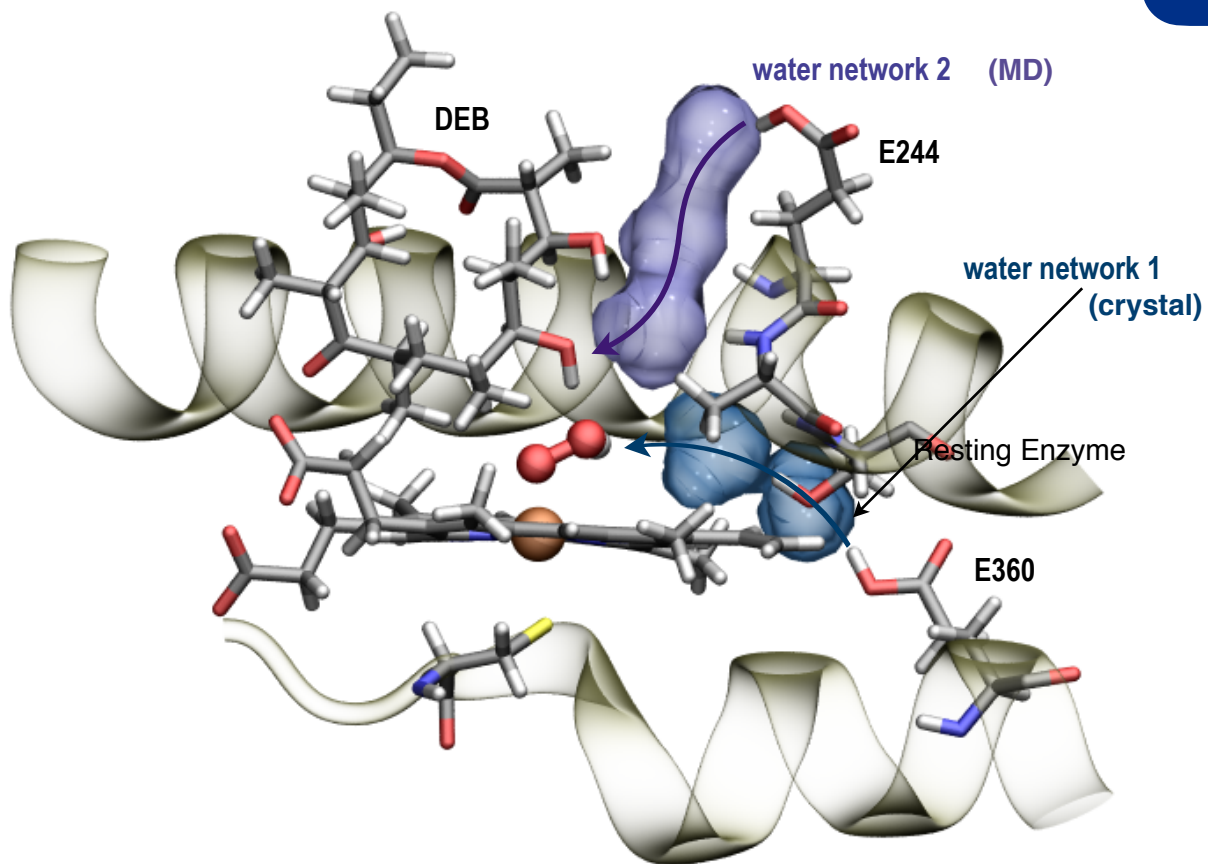
## Activation of the catalytic cycle: P450EryF



# P450 Catalysis

## Activation of the catalytic cycle: P450EryF

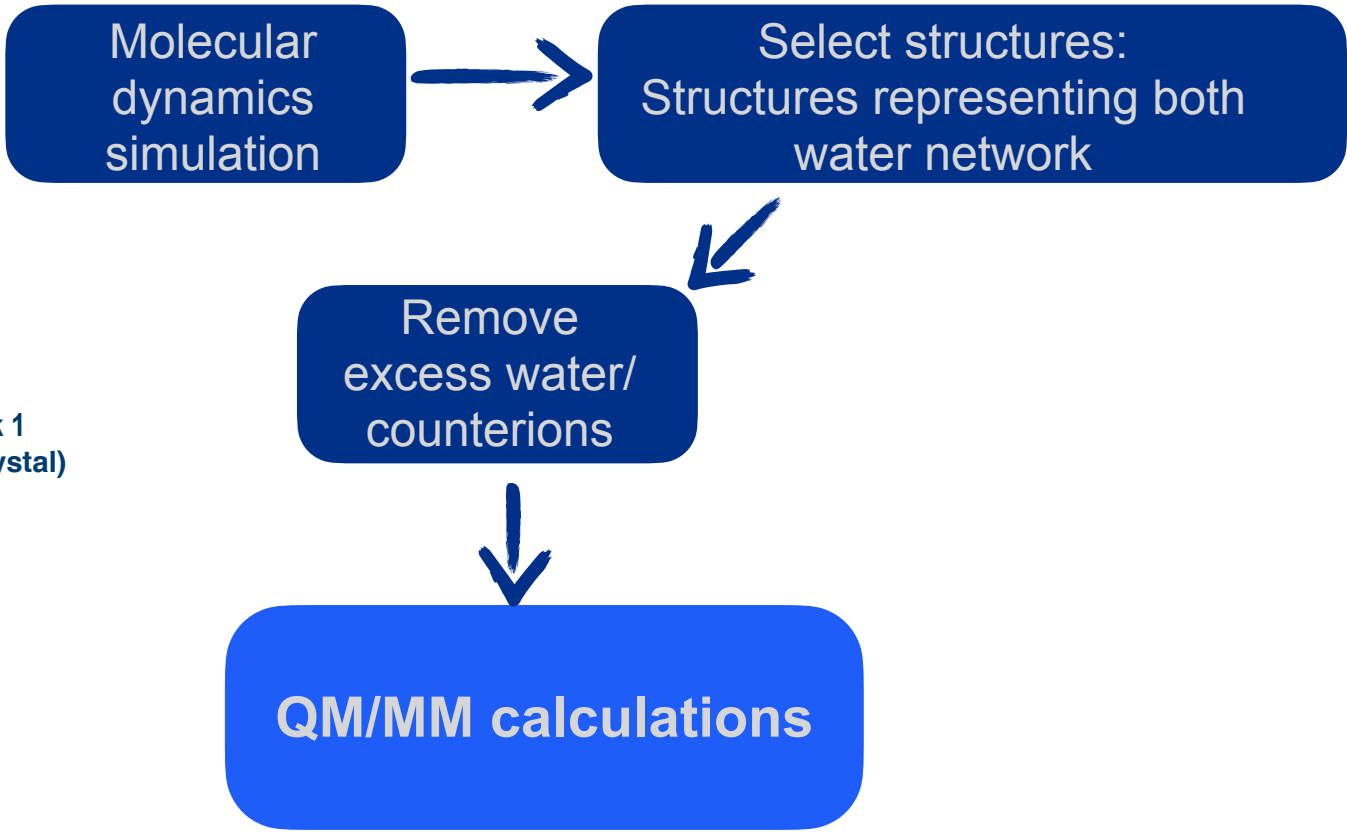
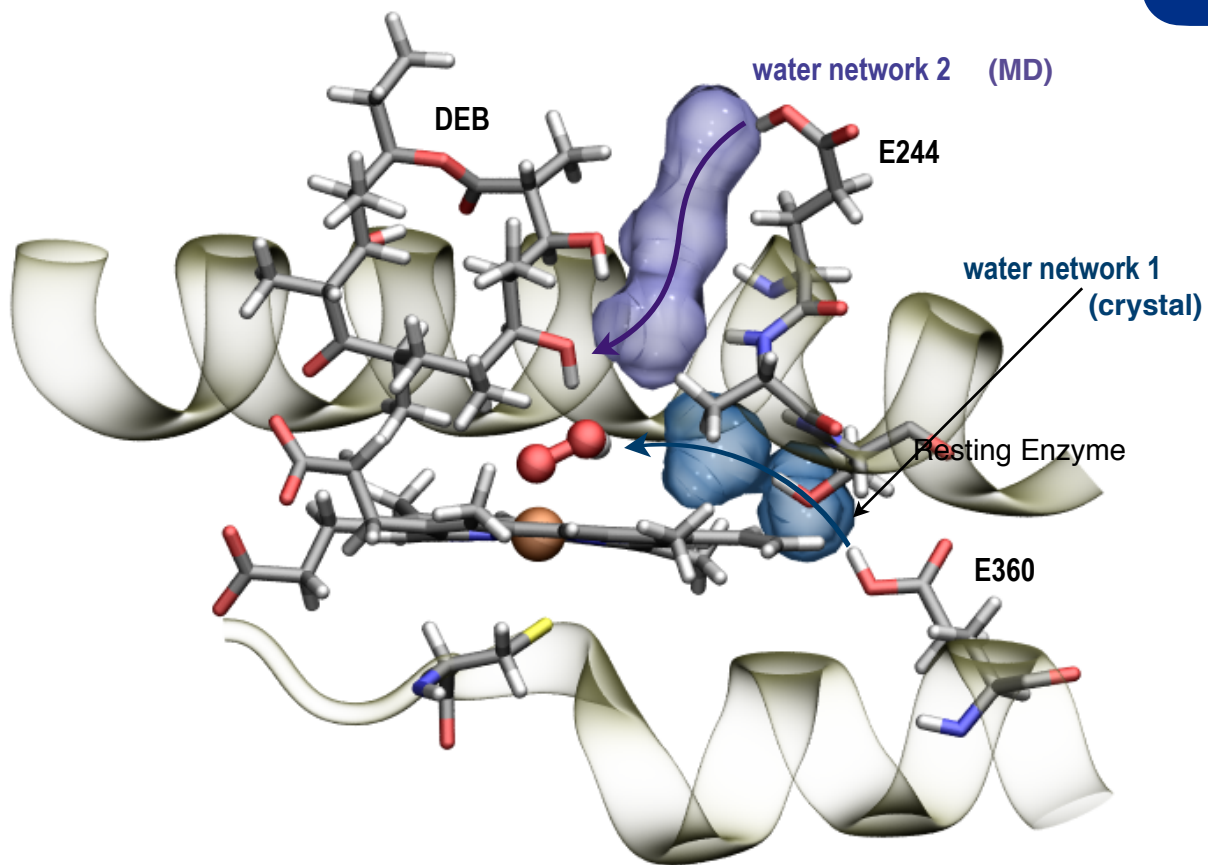
Molecular  
dynamics  
simulation



Sen & Thiel *JPCB*, 118, 2810 (2014)

# P450 Catalysis

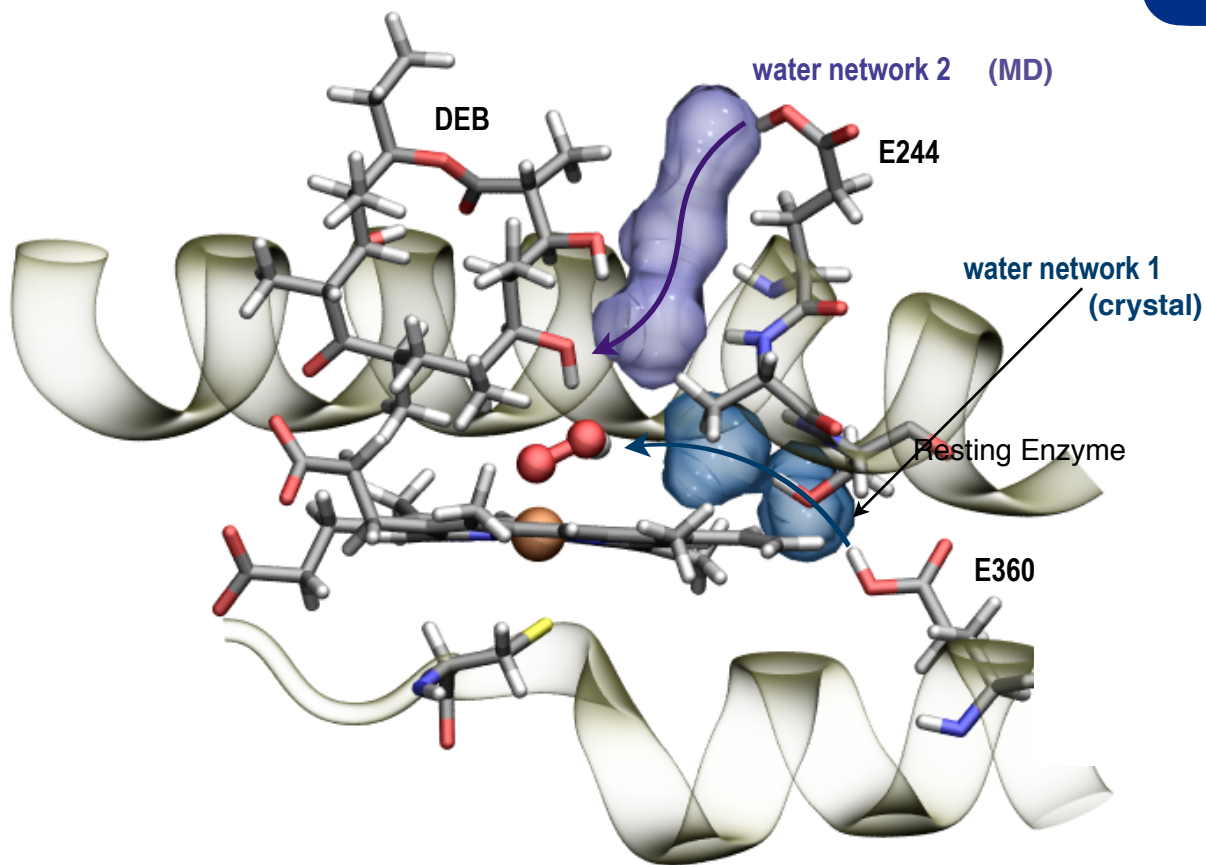
## Activation of the catalytic cycle: P450EryF



Sen & Thiel *JPCB*, 118, 2810 (2014)

# P450 Catalysis

## Activation of the catalytic cycle: P450EryF

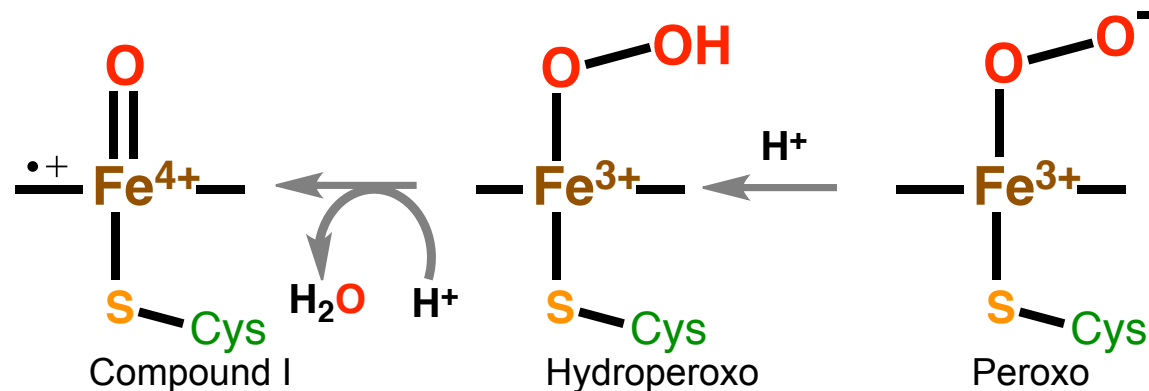
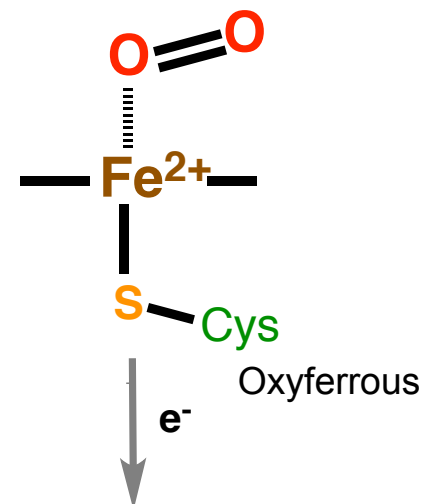


Sen & Thiel *JPCB*, 118, 2810 (2014)

Molecular dynamics simulation → Select structures: Structures representing both water network

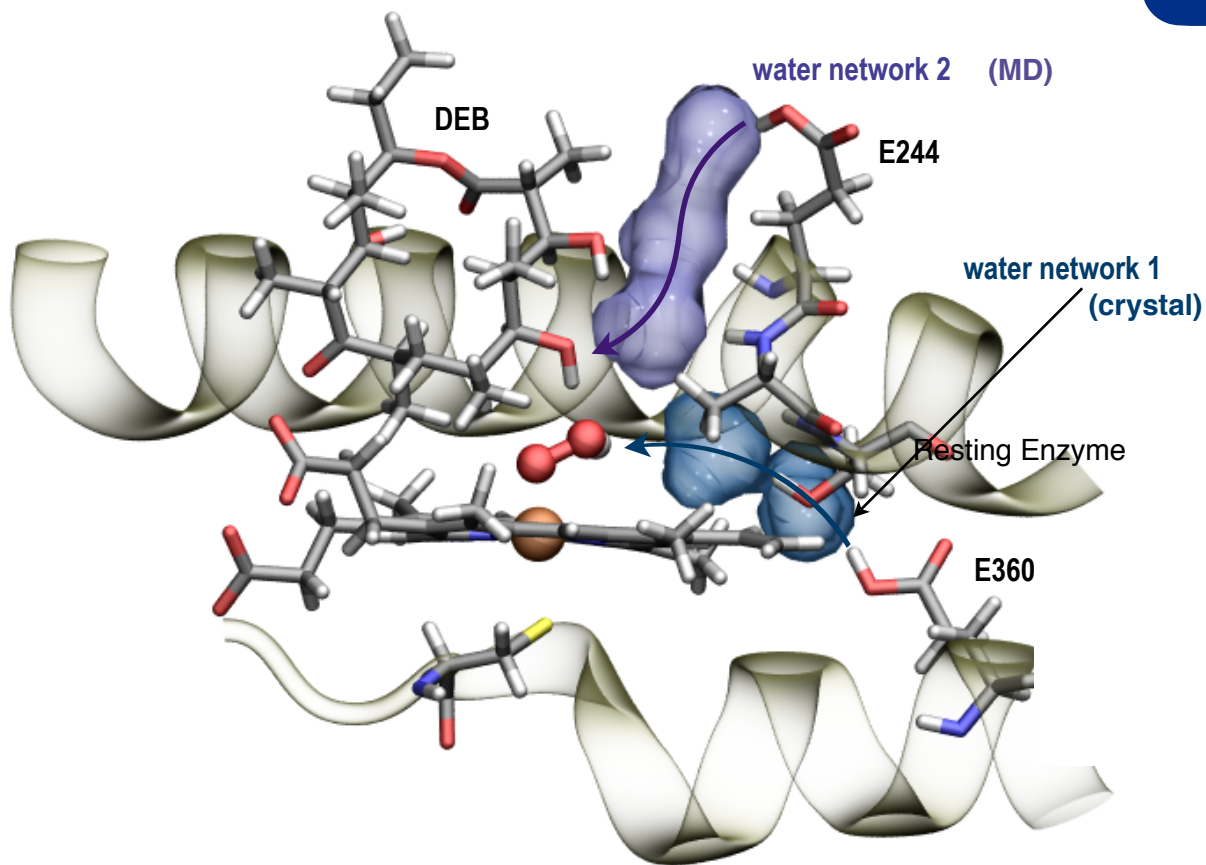
Remove excess water/counterions

QM/MM calculations

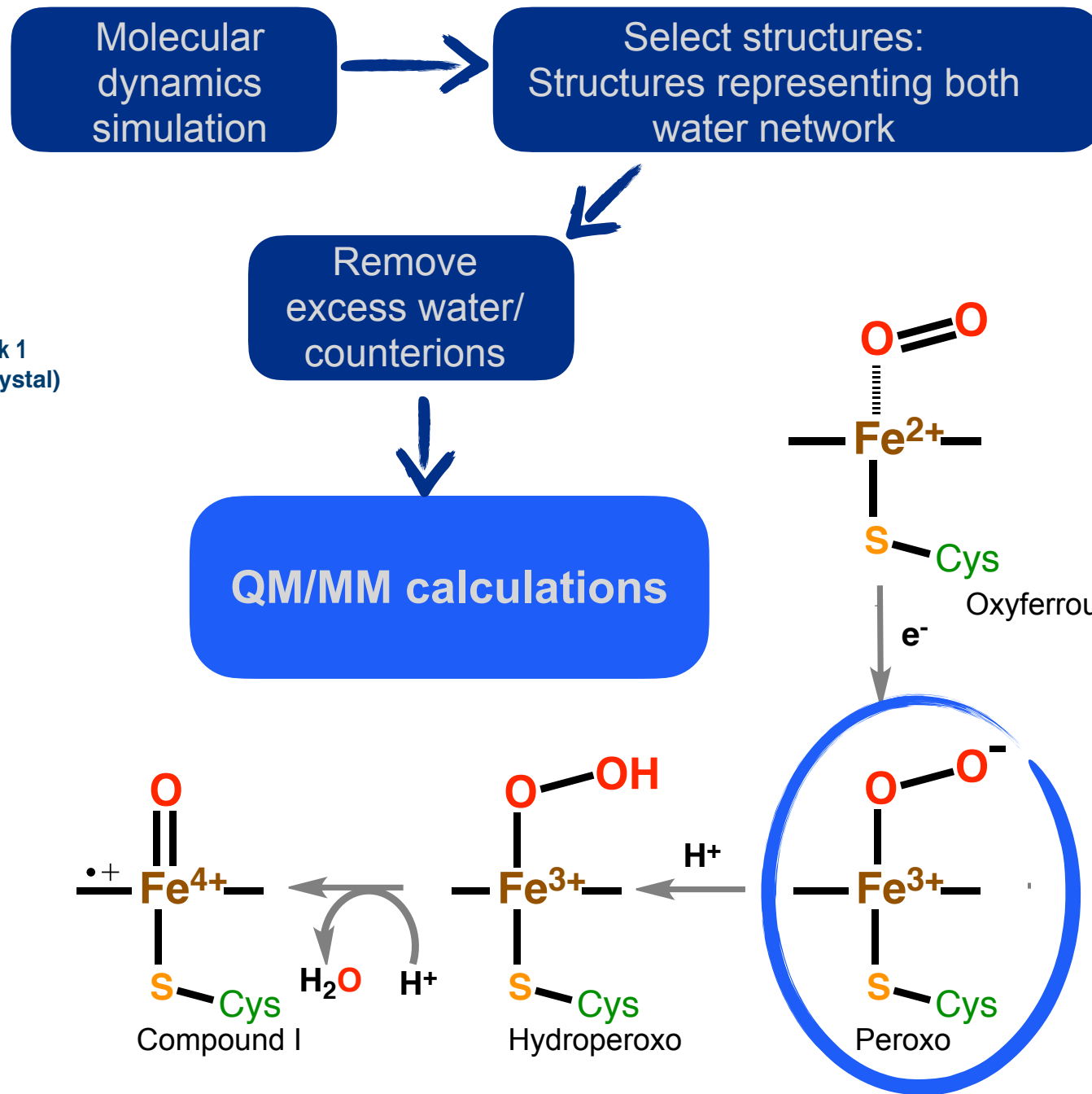


# P450 Catalysis

## Activation of the catalytic cycle: P450EryF



Sen & Thiel *JPCB*, 118, 2810 (2014)

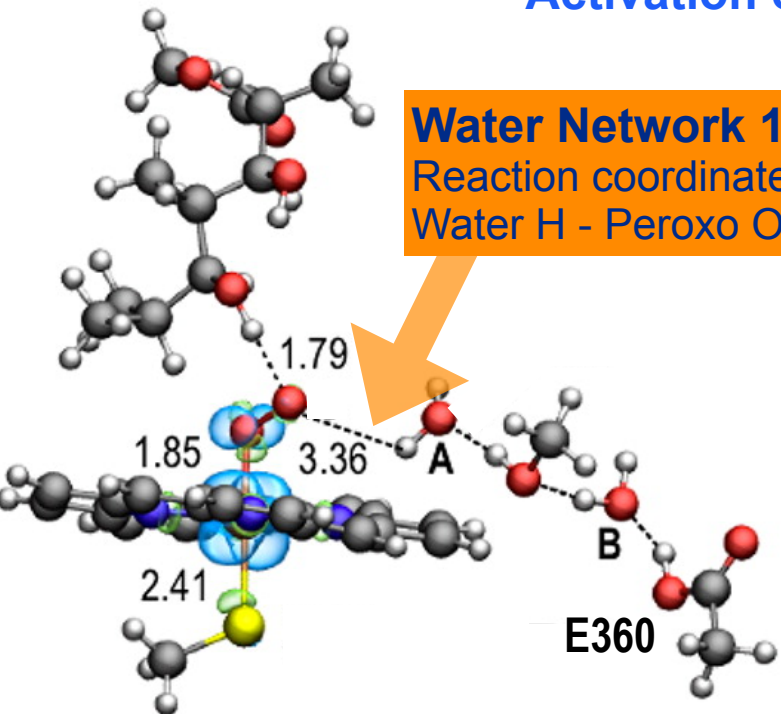


# P450 Catalysis

## Activation of the catalytic cycle: P450EryF

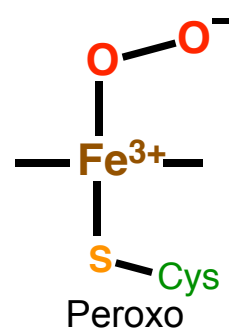
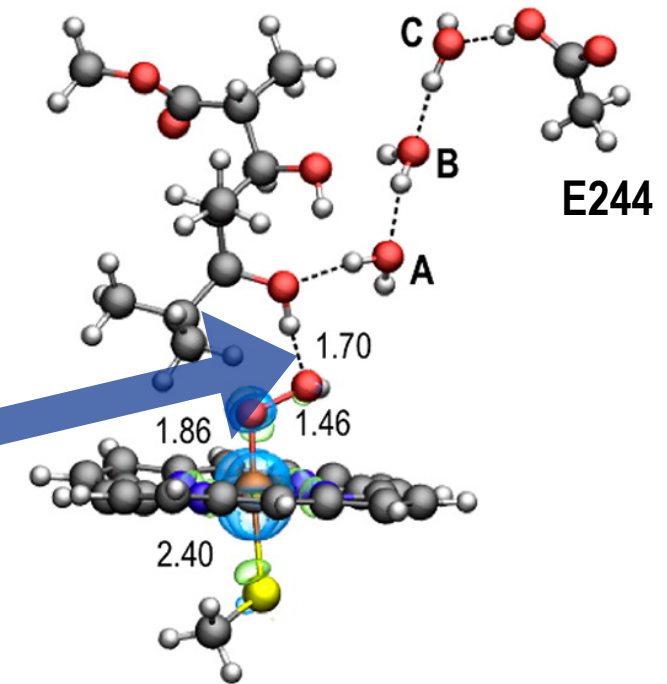
### Water Network 1

Reaction coordinate:  
Water H - Peroxo O distance

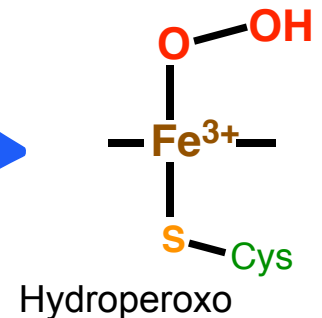


### Water Network 2

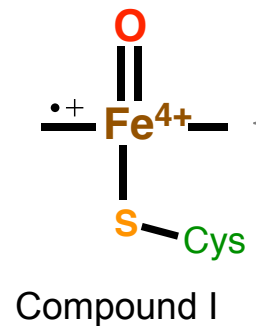
Reaction coordinate:  
Substrate H - peroxo O distance



Step 1



Step 2



Water Network 1 : Barrier 19.3 kcal/mol  
Water Network 2 : Barrier less

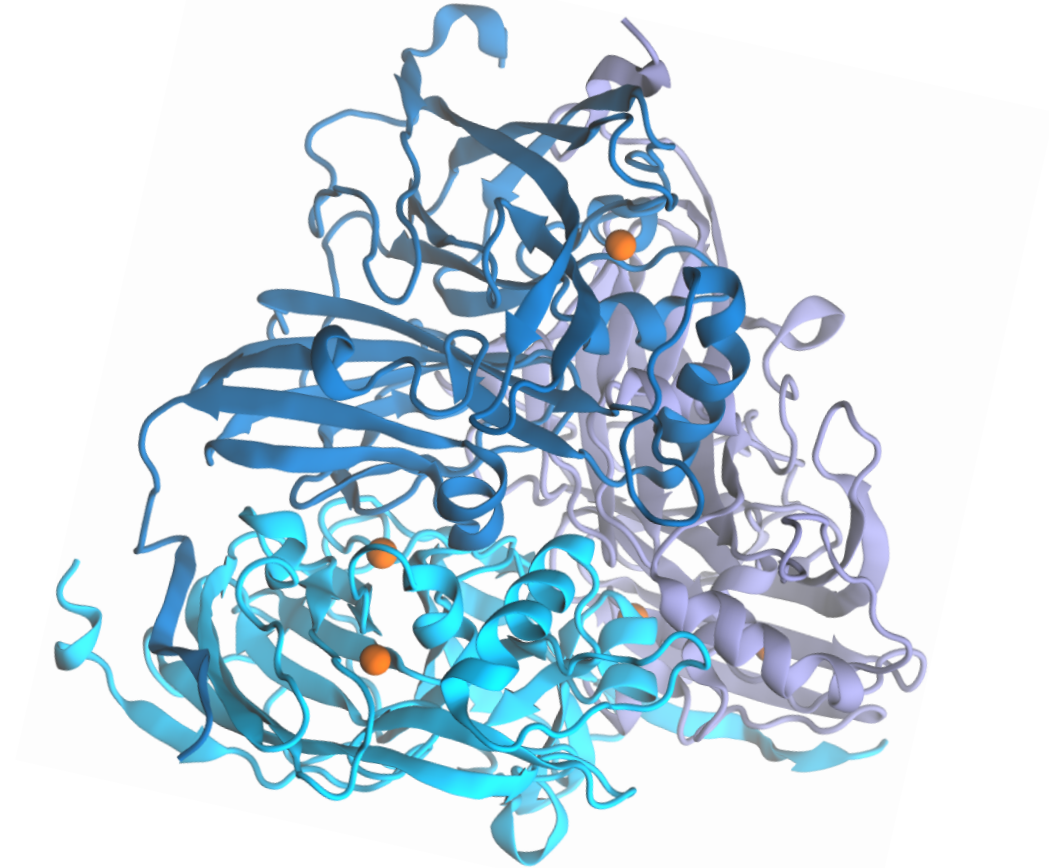
Water Network 1 : Barrier 17.6 kcal/mol  
Water Network 2 : Barrier 19.7 kcal/mol

**Water Network 2 : Preferred route**

# QM/MM for enzymes

- Characterise intermediates that are experimentally elusive; e.g., transition states
- Investigate mechanisms

## Copper nitrite reductases



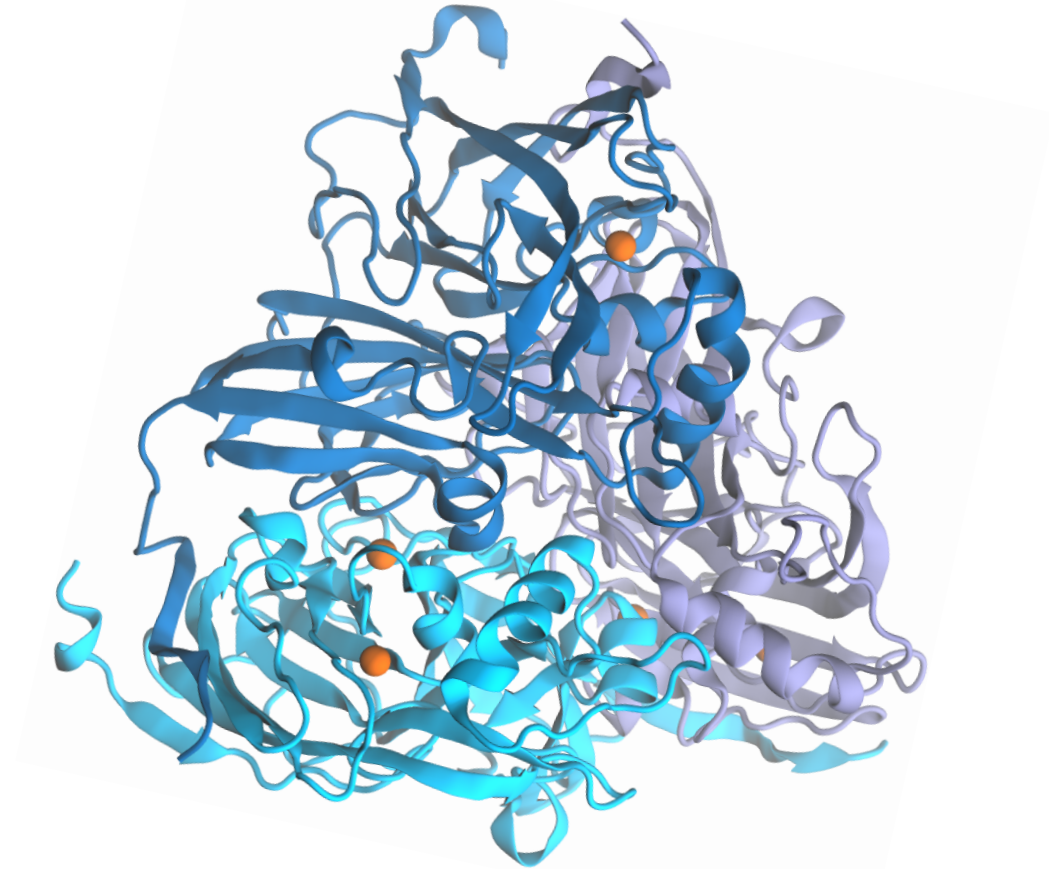
- CuNiRs are typically trimeric proteins containing two types of Cu sites:
- T1Cu: electron-donating Cu centre
  - T2Cu: the catalytic centre where the reduction of  $\text{NO}_2^-$  to NO occurs

# QM/MM for enzymes

- Characterise intermediates that are experimentally elusive; e.g., transition states
- Investigate mechanisms



## Copper nitrite reductases



- CuNiRs are typically trimeric proteins containing two types of Cu sites:
- T1Cu: electron-donating Cu centre
  - T2Cu: the catalytic centre where the reduction of  $\text{NO}_2^-$  to NO occurs

# QM/MM for enzymes

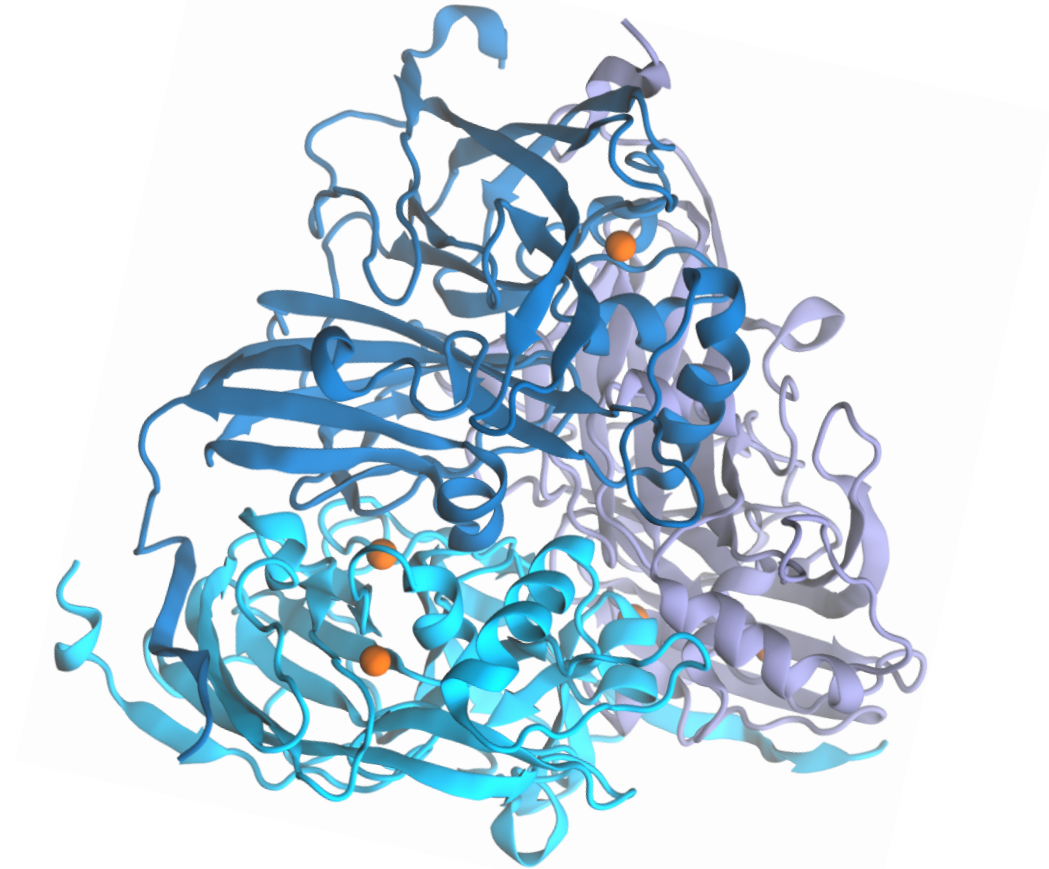
- Characterise intermediates that are experimentally elusive; e.g., transition states
- Investigate mechanisms



Reaction coordinate: Not obvious

- ▶ Nudged Elastic Band

## Copper nitrite reductases

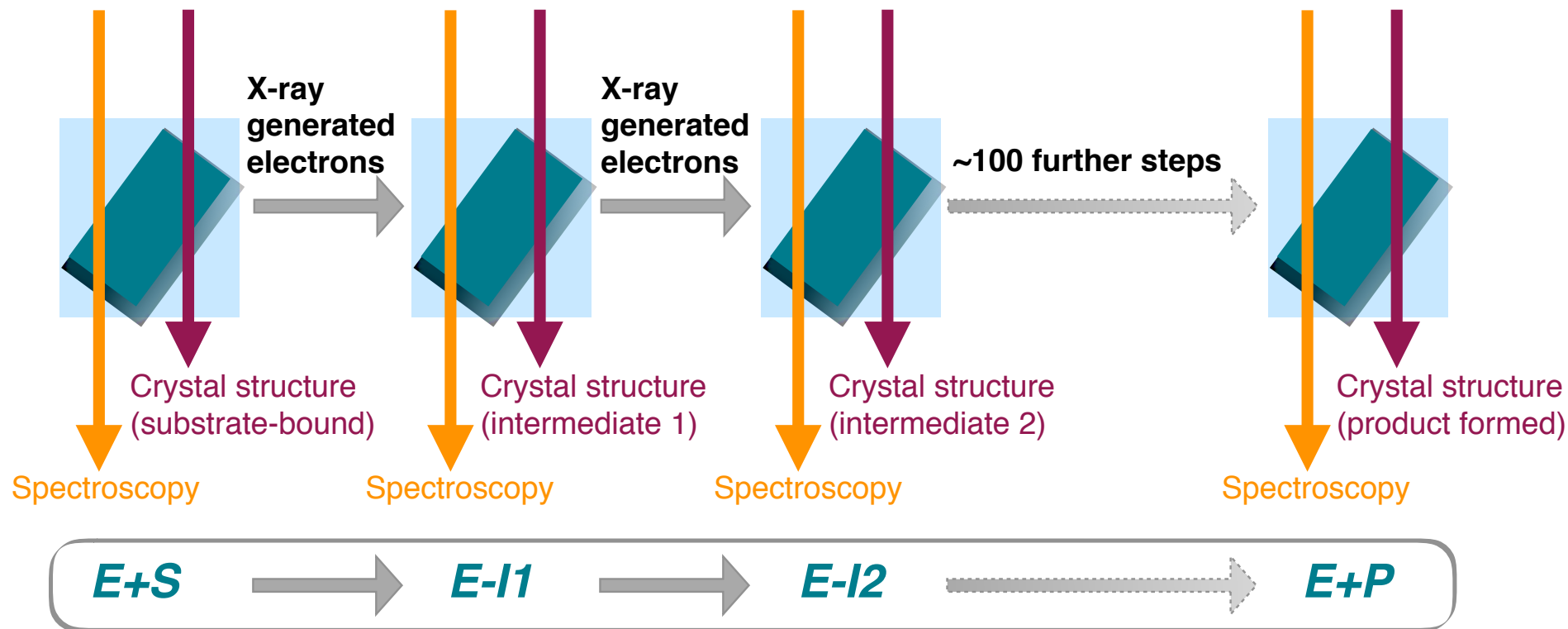


CuNiRs are typically trimeric proteins containing two types of Cu sites:

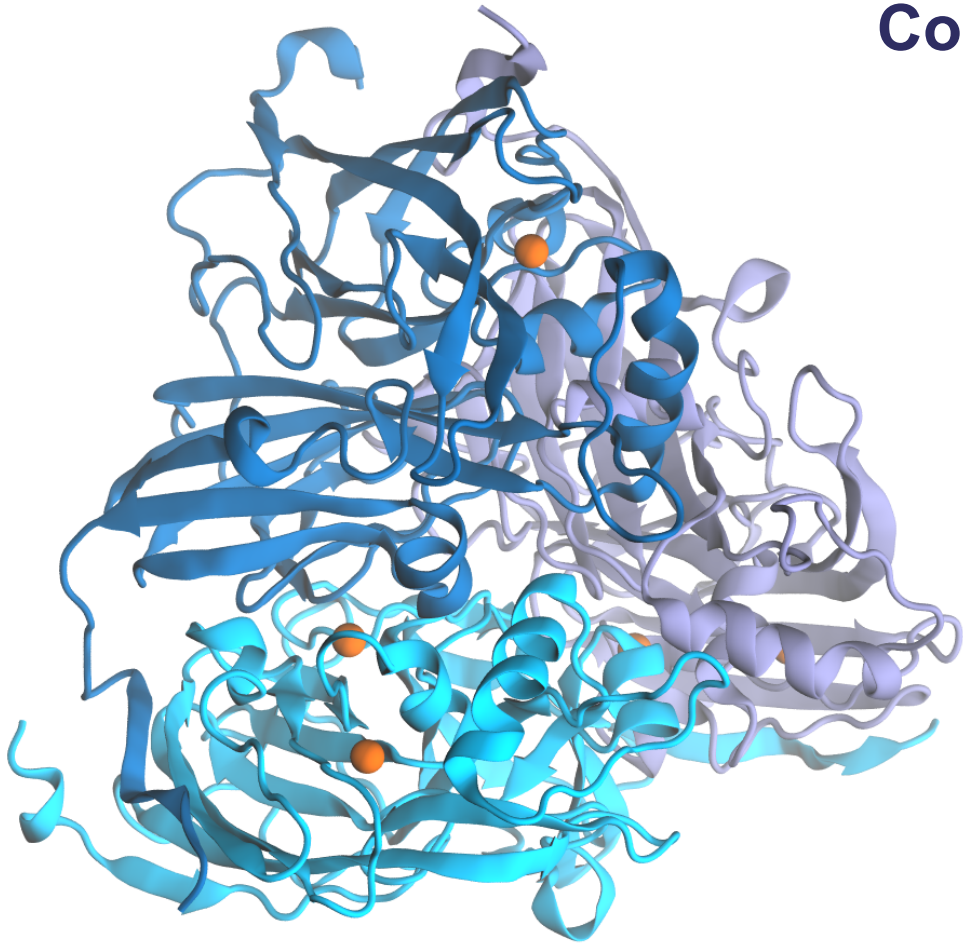
- T1Cu: electron-donating Cu centre
- T2Cu: the catalytic centre where the reduction of  $\text{NO}_2^-$  to NO occurs

# Multiple Structures from One Xtal (MSOX)

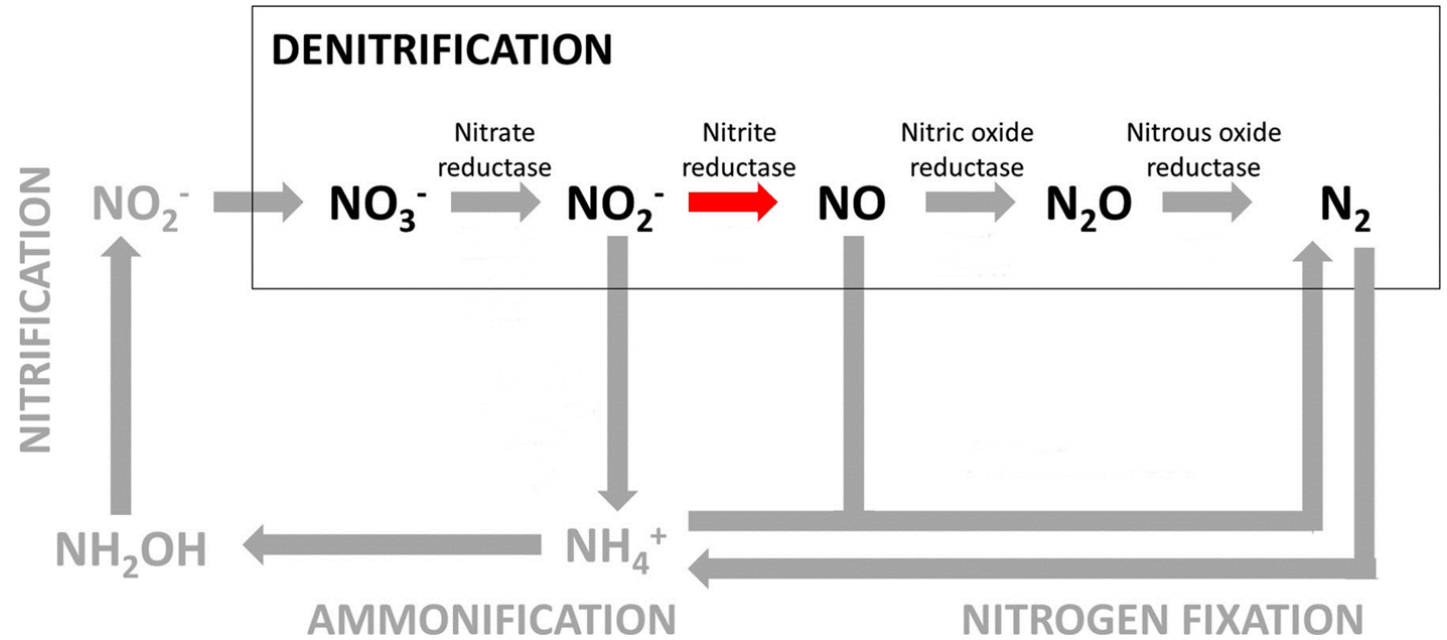
## Structural Movie of Catalysis



# Copper nitrite reductases

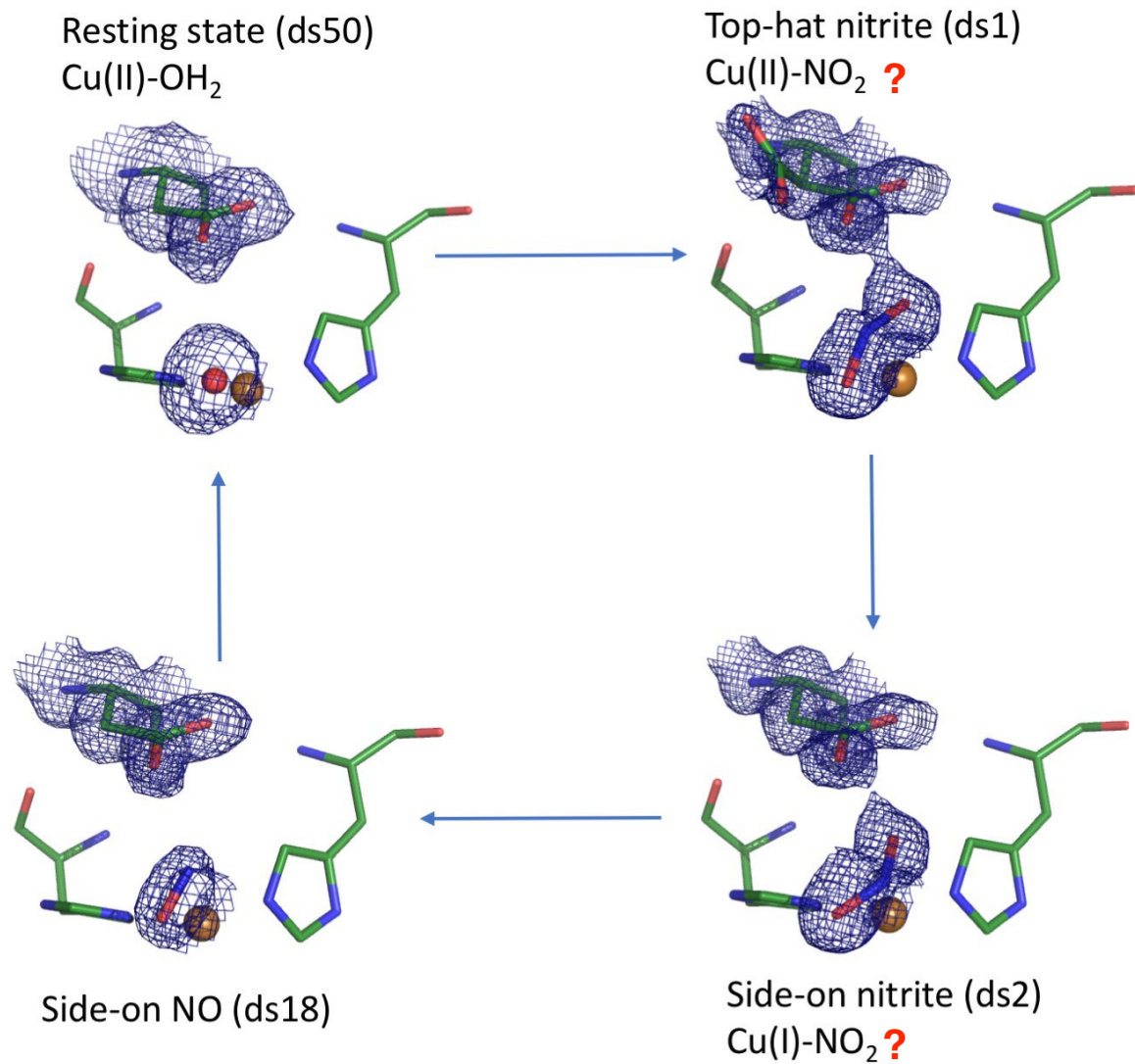


*Achromobacter cycloclastes* nitrite reductase (AcNiR) PDB ID: 2BW4



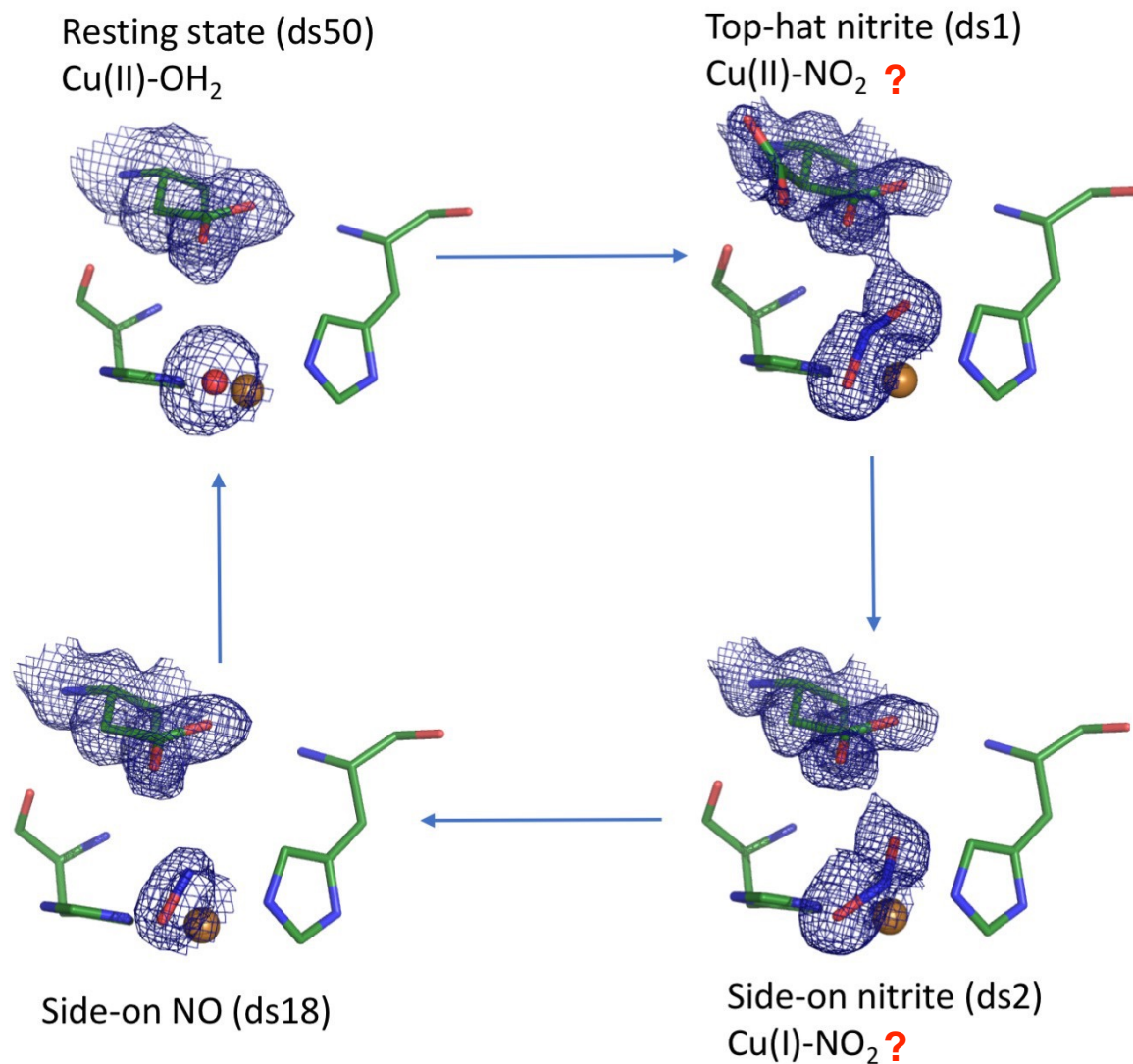
Reduction reaction: Ideal for MSOX

# Copper nitrite reductases



Assignment of structures to states  
along the structural movie

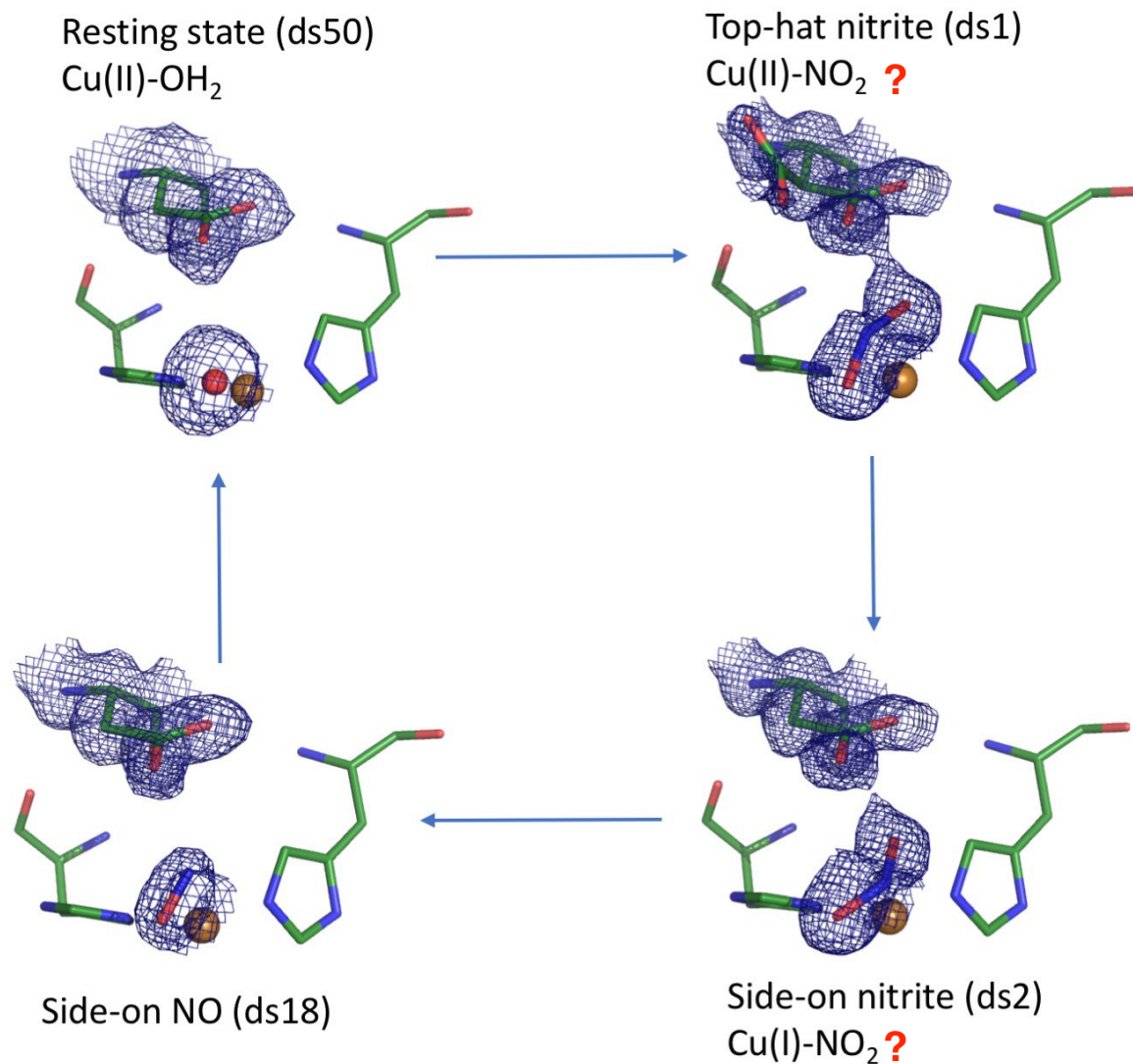
# Copper nitrite reductases



Assignment of structures to states along the structural movie

| Parameters                                    | <b>MSOX</b> | <b>QM/MM</b> |
|---|-------------|--------------|
| Metal redox state                             | No          | Yes          |
| Protonation states of ligands and side chains | No          | Yes          |

# Copper nitrite reductases

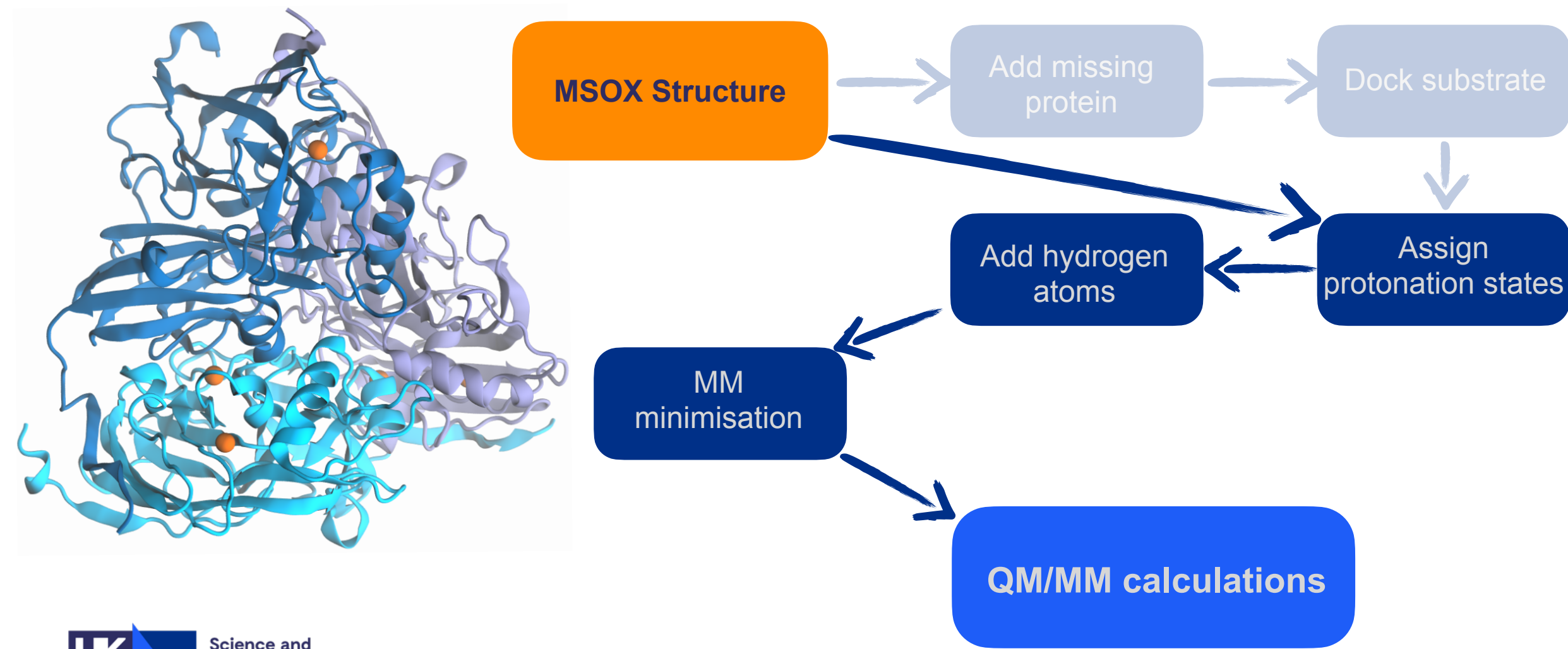


Assignment of structures to states along the structural movie

| Parameters                                    | <b>MSOX</b> | <b>QM/MM</b> |
|---|-------------|--------------|
| Metal redox state                             | No          | Yes          |
| Protonation states of ligands and side chains | No          | Yes          |

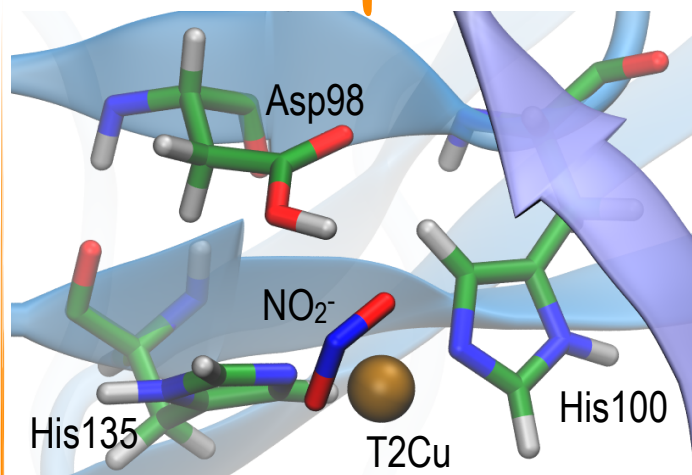
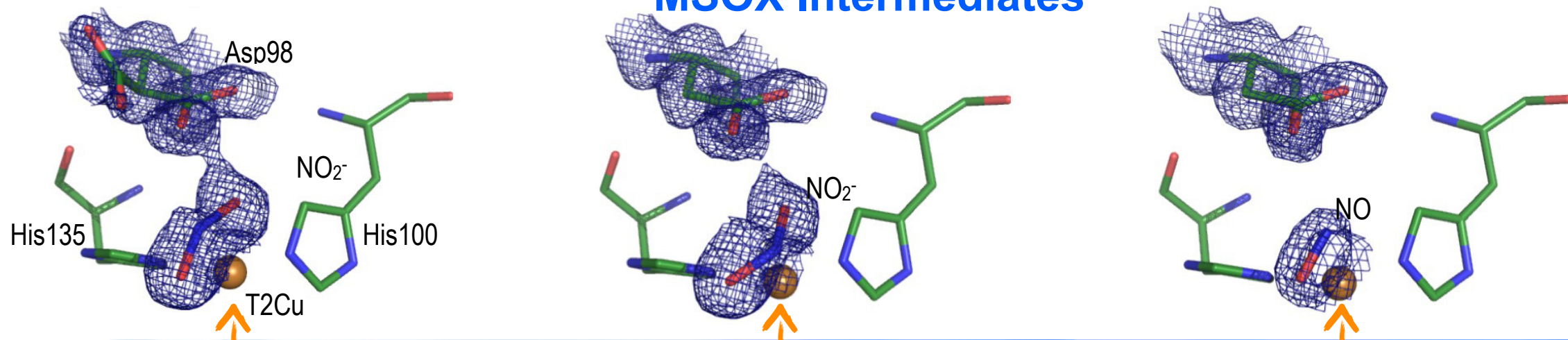
# Copper nitrite reductases

## WorkFlow

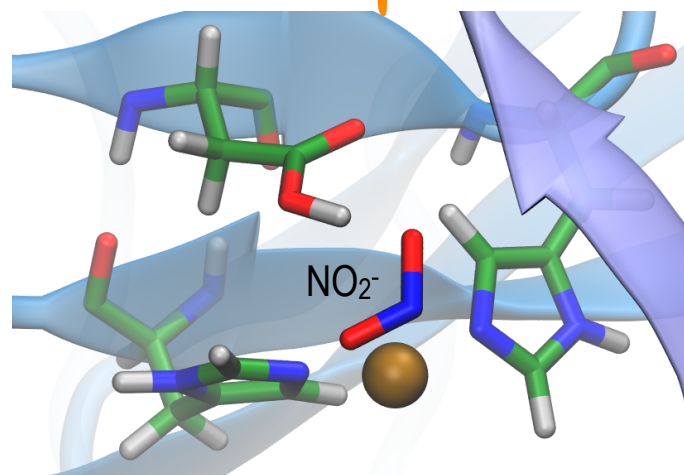


# QM/MM identification of the MSOX observed intermediates at the T2Cu site

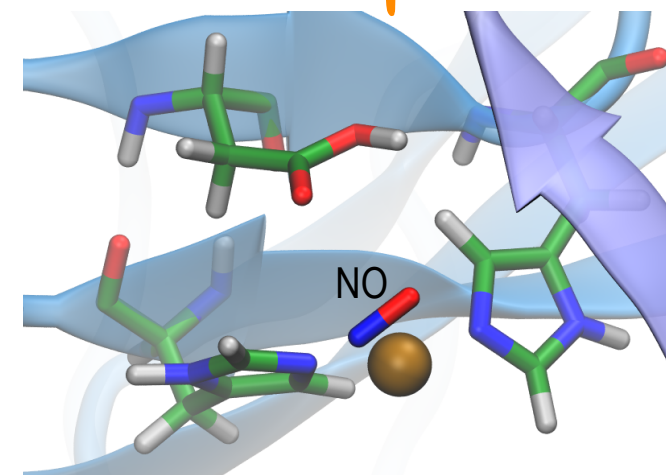
## MSOX Intermediates



**Cu(II)-NO<sub>2</sub><sup>-</sup> (HAT)**



**Cu(I)-NO<sub>2</sub><sup>-</sup> (SIDEON)**

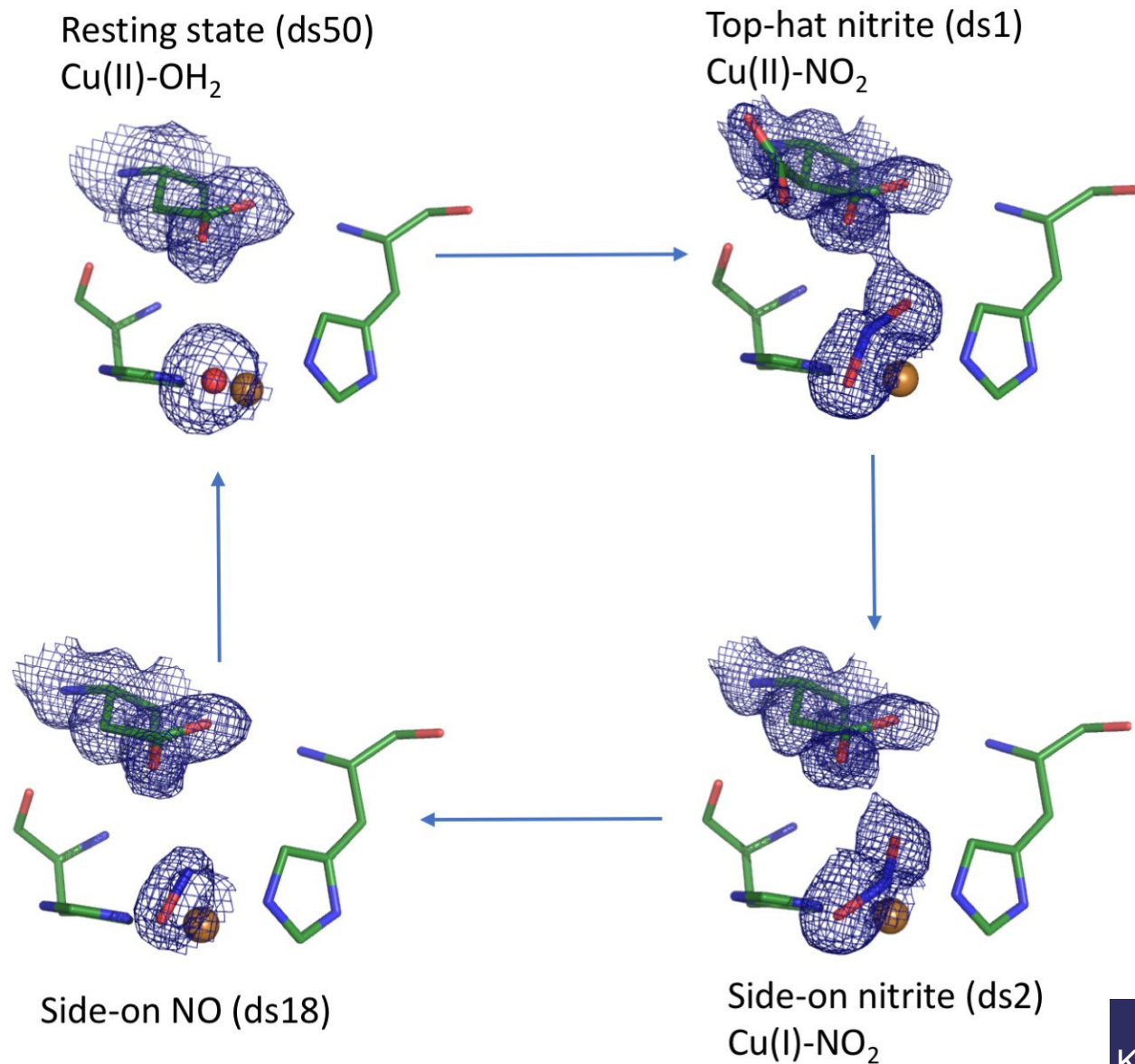


**Cu(I)-NO<sup>+</sup>**

**QM/MM Intermediates**

# Copper nitrite reductases

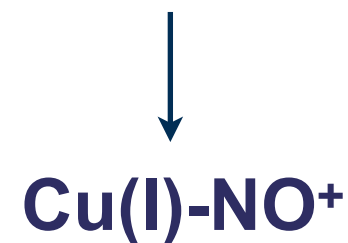
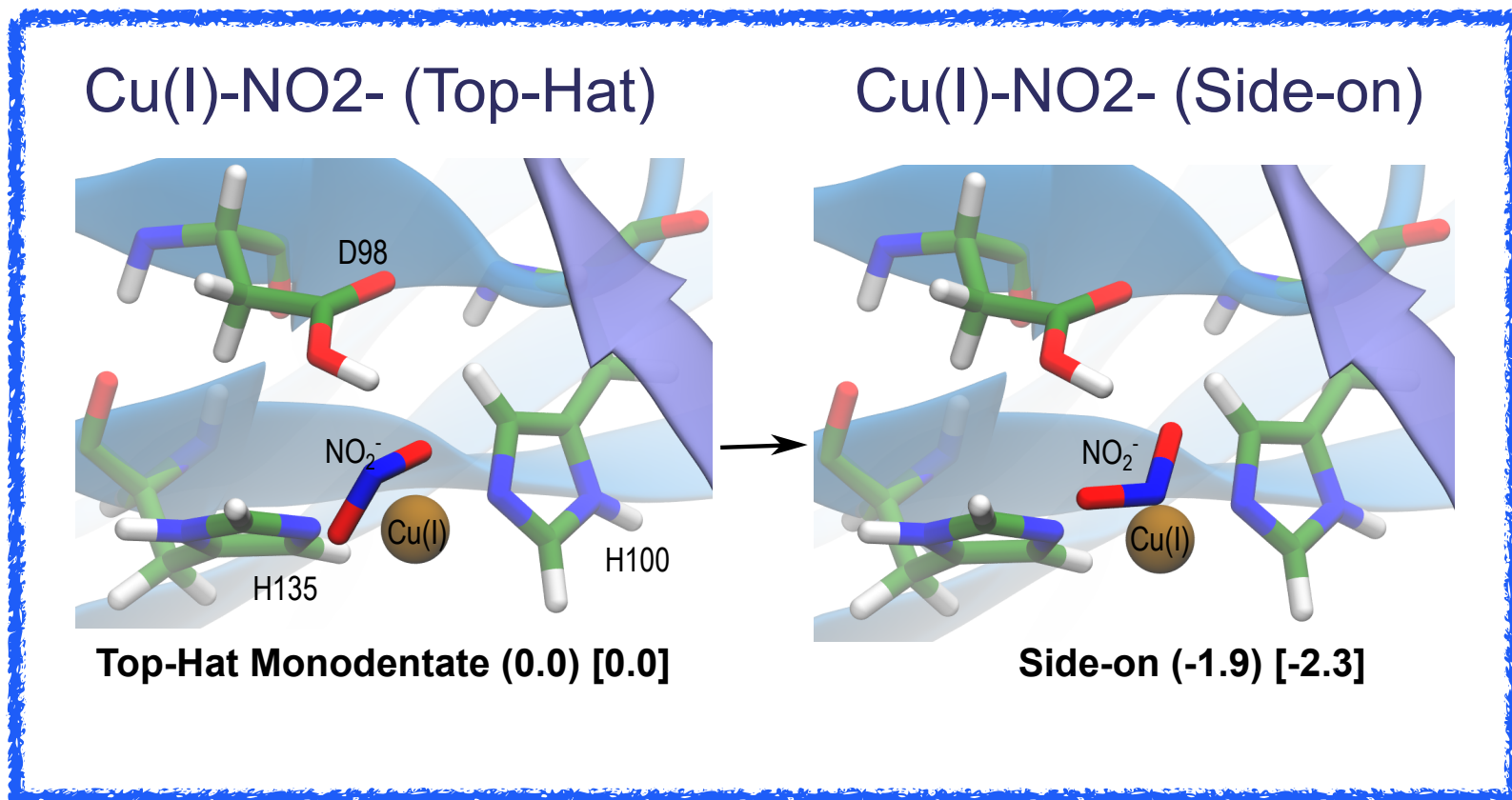
## Assignment of Intermediates



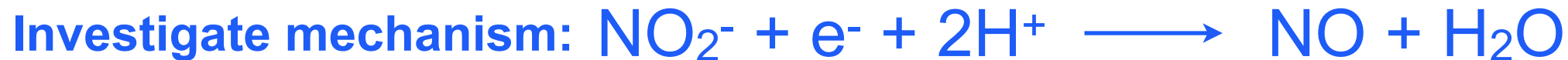
# Copper nitrite reductases



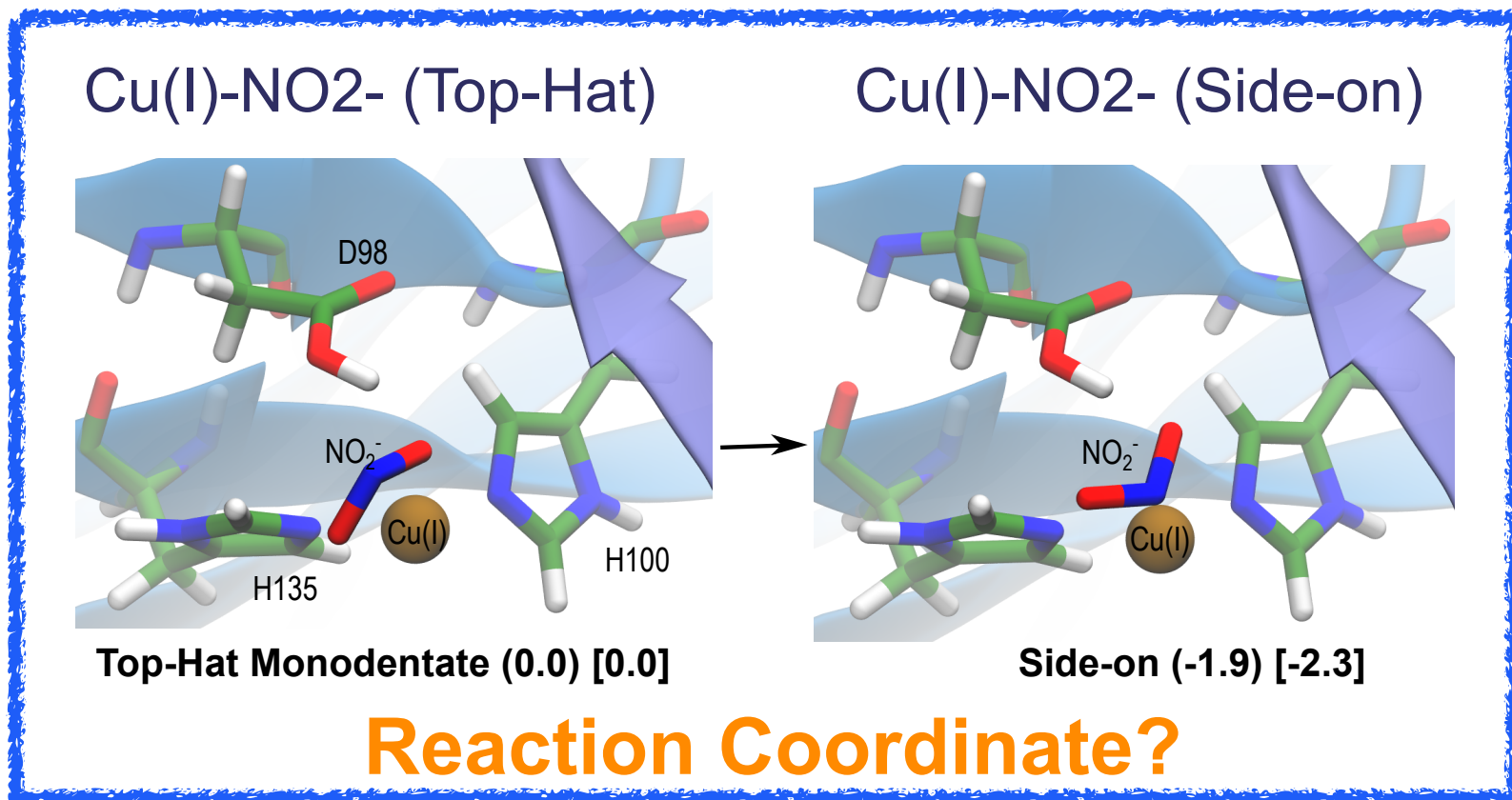
**Cu(II)-NO<sub>2</sub>- (Top-Hat)** →



# Copper nitrite reductases



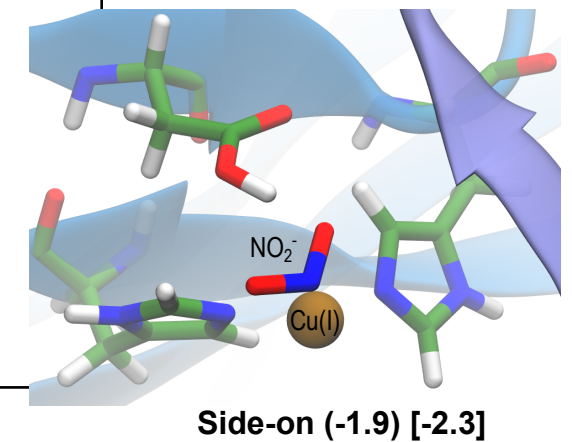
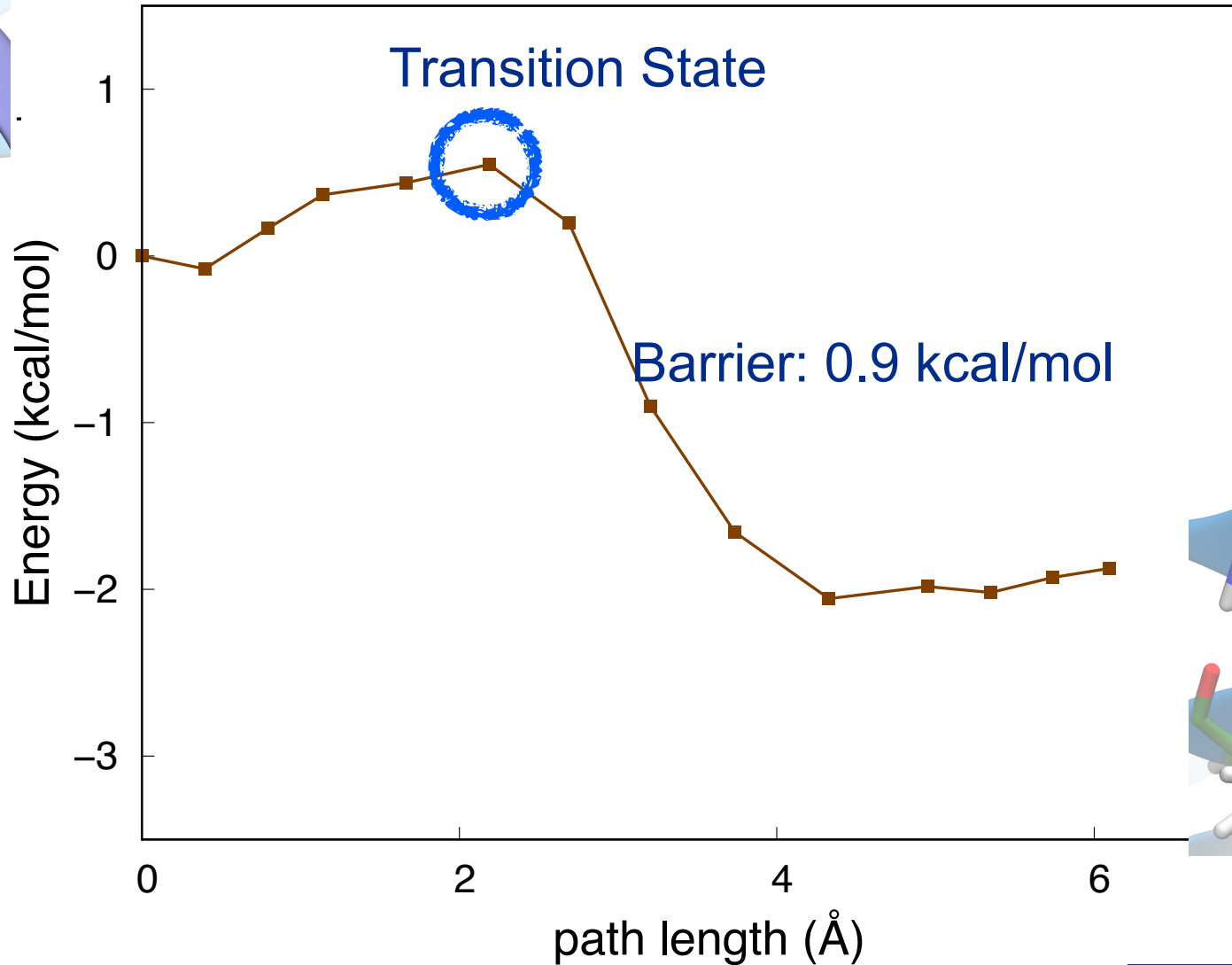
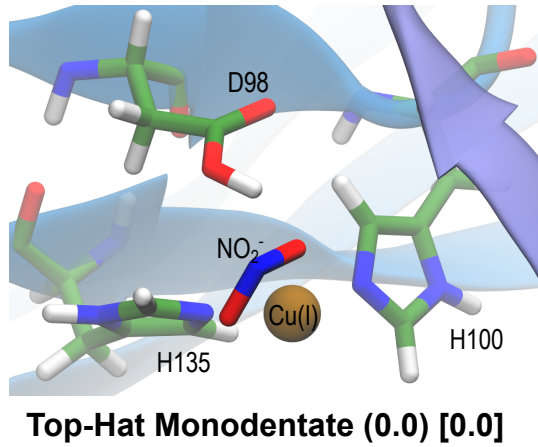
**Cu(II)-NO<sub>2</sub>- (Top-Hat)** →



↓  
**Cu(I)-NO<sup>+</sup>**

# Copper nitrite reductases

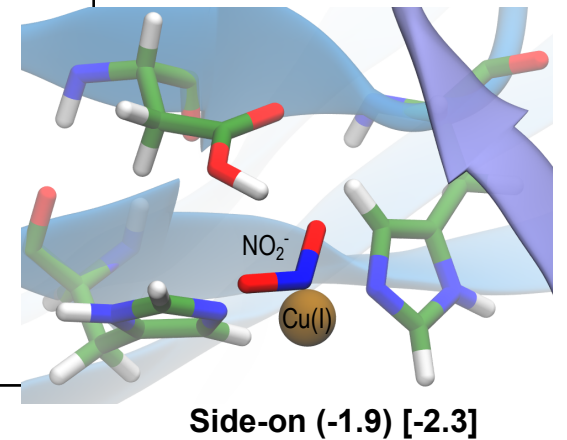
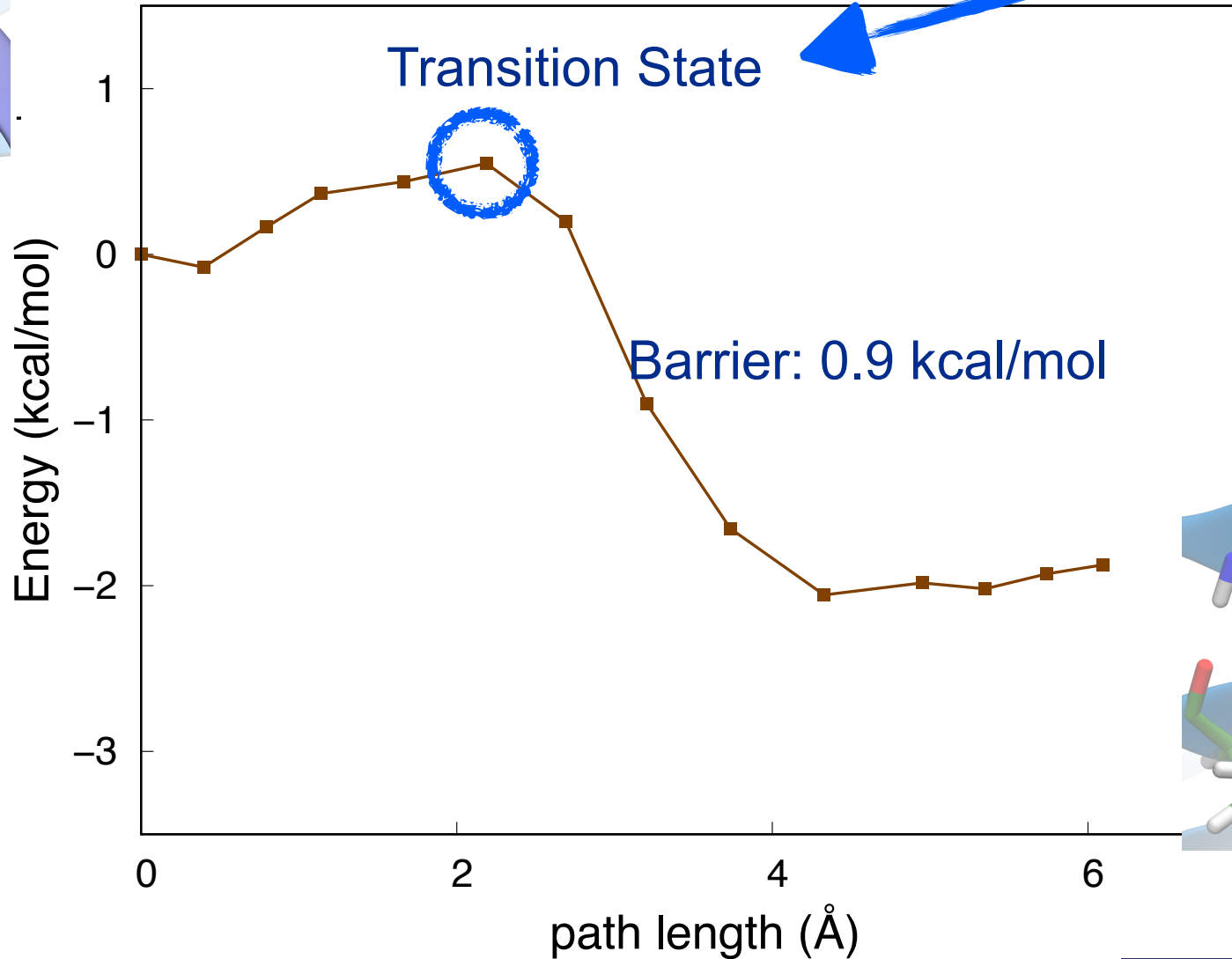
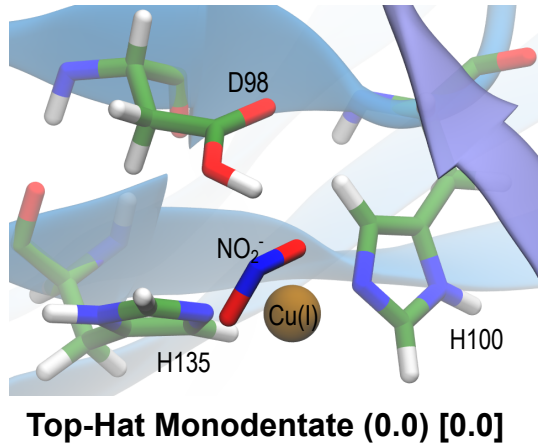
## Nudged Elastic Band



# Copper nitrite reductases

## Nudged Elastic Band

## Dimer Method

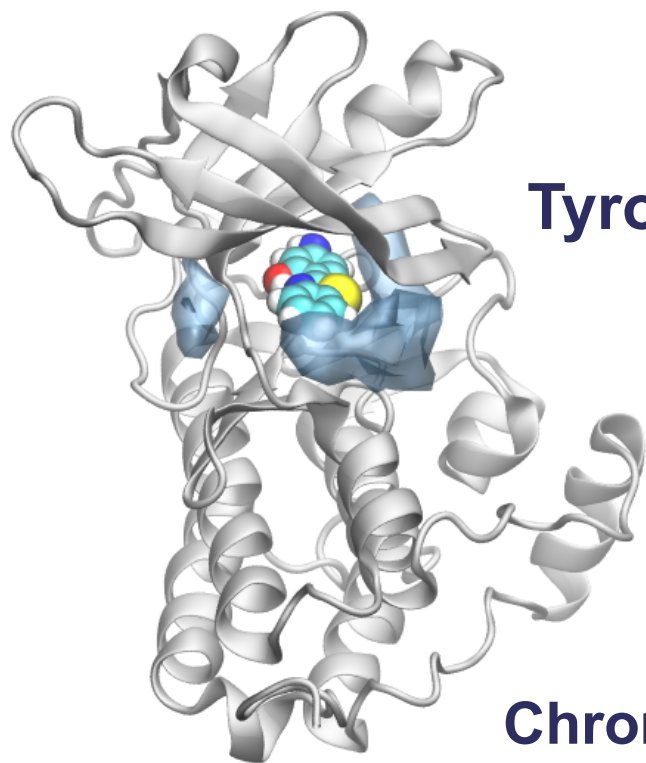


# QM/MM for enzymes

Emission property of chromophoric substrate

# QM/MM for enzymes

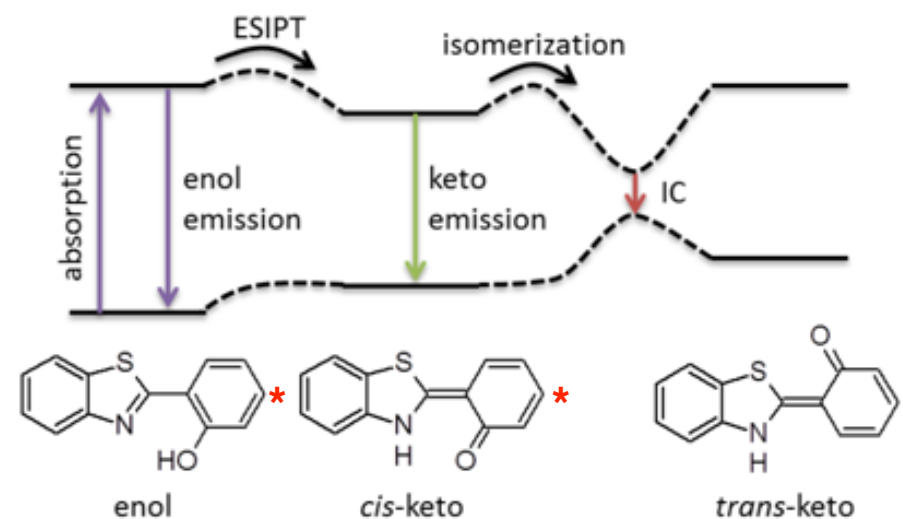
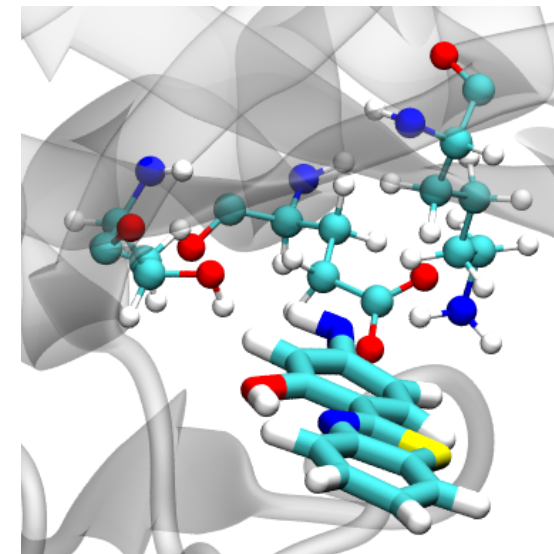
Emission property of chromophoric substrate



**Chromophore: HABT**

Phenylbenzothiazole (PBT) derivative  
-HN2 group at \* position

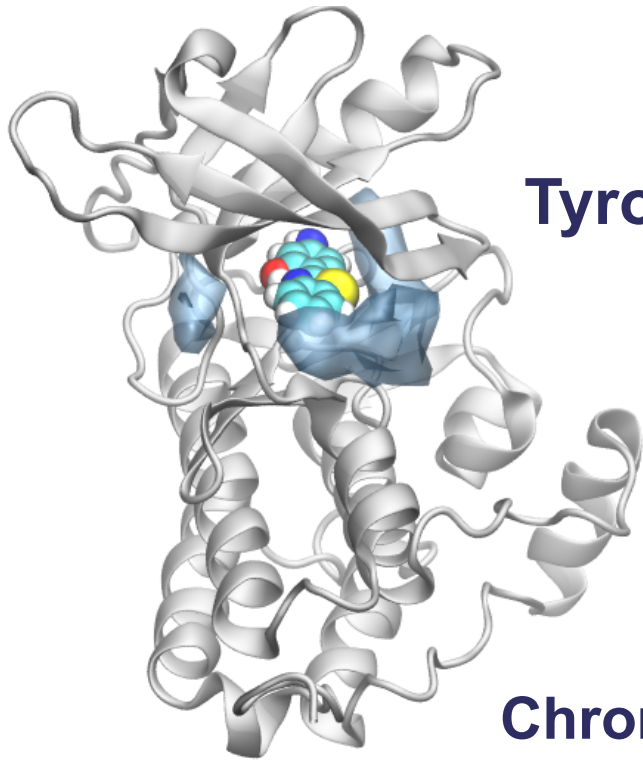
**Extending QM/MM  
with TD-DFT**



# QM/MM for enzymes

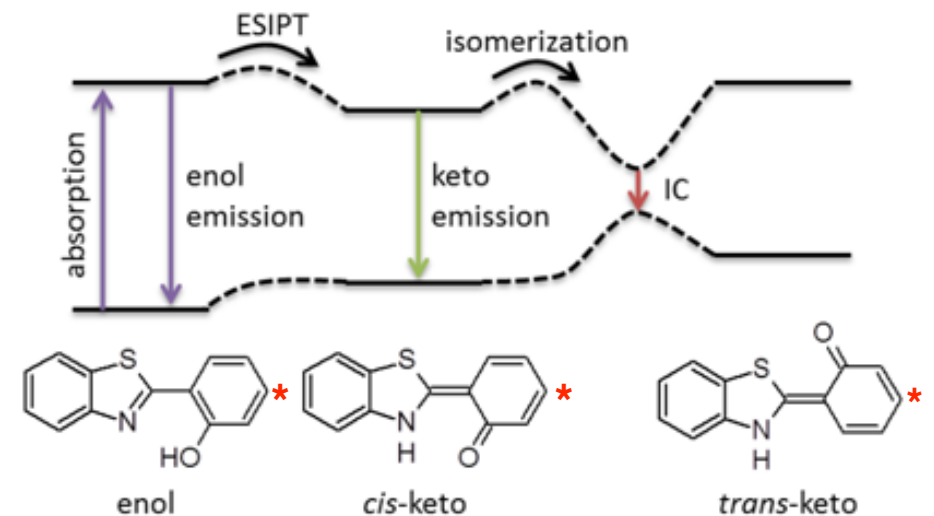
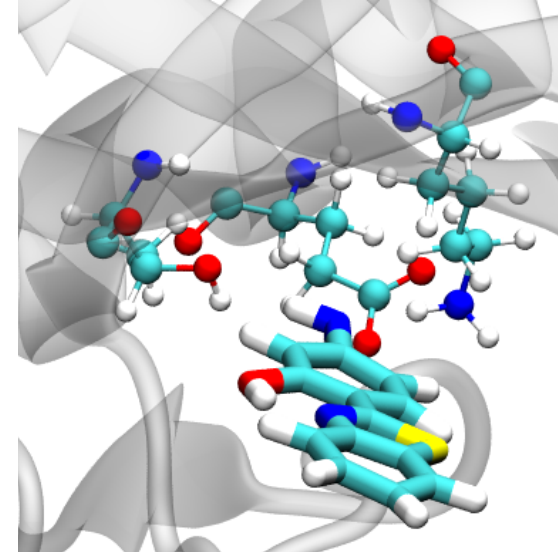
Emission property of chromophoric substrate

## Excited-State Proton Transfer Can Tune the Colour of Protein Fluorescent Markers



Phenylbenzothiazole (PBT) derivative  
-HN2 group at \* position

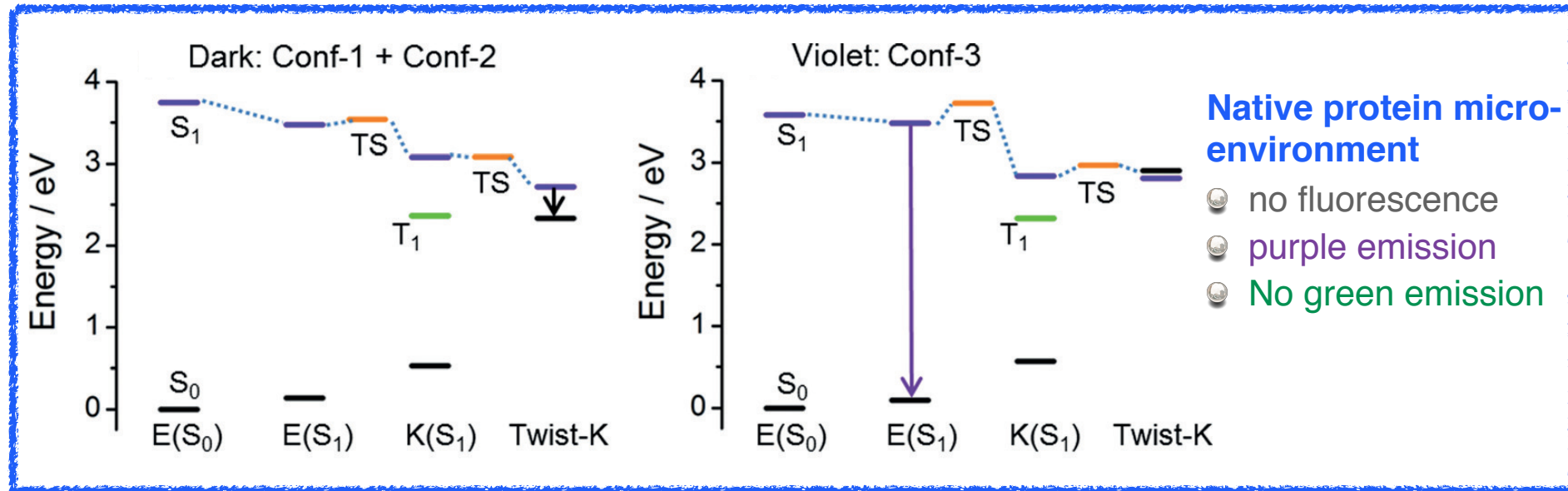
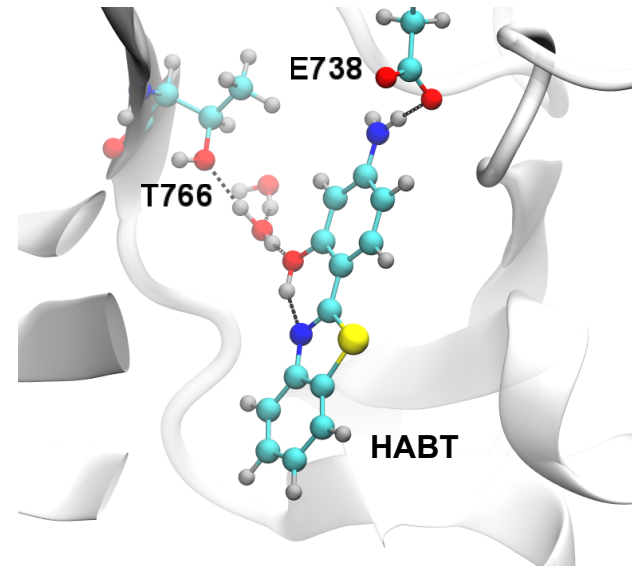
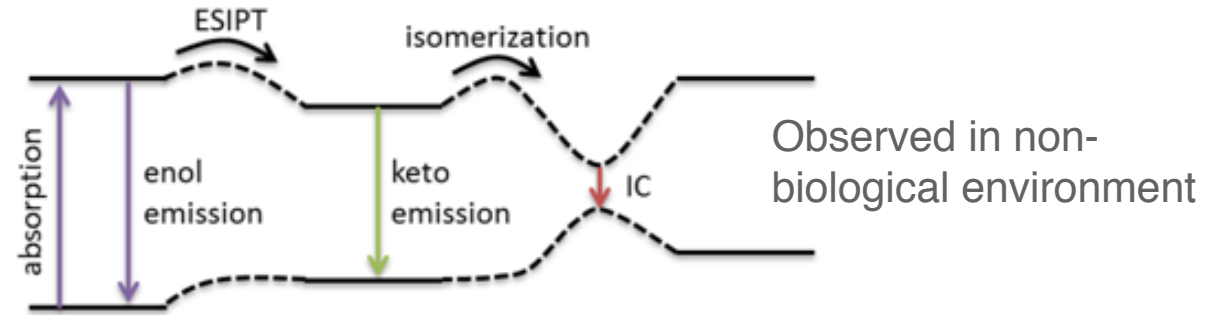
## Extending QM/MM with TD-DFT



# Tyrosine Kinase

## Emission properties of chromophore HABT

## Aim: Dual fluorescence



- Photoemission properties of a PBT compound bound to the protein is very sensitive to the active site conformation.
- The interaction of water and polar residues with the hydroxyl group of the PBT controls the rates of ES IPT and internal conversion.

# Conclusion

- Illustrated the wide range of application of QM/MM
- Very effective method to study systems where we need both electronic and structural information
- These calculations are not like black-box calculations and needs careful inspection at each step as it comes with it own pitfalls
  - ▶ Incorrect setup of model
    - Wrong protonation state of amino acids at active site
  - ▶ QM region too small
  - ▶ Basis set size
  - ▶ Not enough atoms in active region

# Acknowledgements

- Prof. Walter Thiel and Dr Mario Barbatti and their group members : MPI Kohlenforschung
- Dr Thomas Keal and Dr Chin Yong : STFC Daresbury Laboratory
- Dr Richard Strange, Dr Mike Hough : University of Essex
- Dr Svetlana Antonyuk, Dr Robert R. Eady and Prof. Samar Hasnain: University of Liverpool
- Prof. Adrian Mulholland and Dr Richard Lonsdale: University of Bristol



## Funding



## Supercomputing facilities:

- MPI Kohlenforschung and RZG supercomputing facilities
- Hartree Center
- SCARF Cluster@RAL



Science and  
Technology  
Facilities Council

# Thank you



Science and Technology Facilities Council



@STFC\_matters



Science and Technology Facilities Council



Science and  
Technology  
Facilities Council

# Questions?