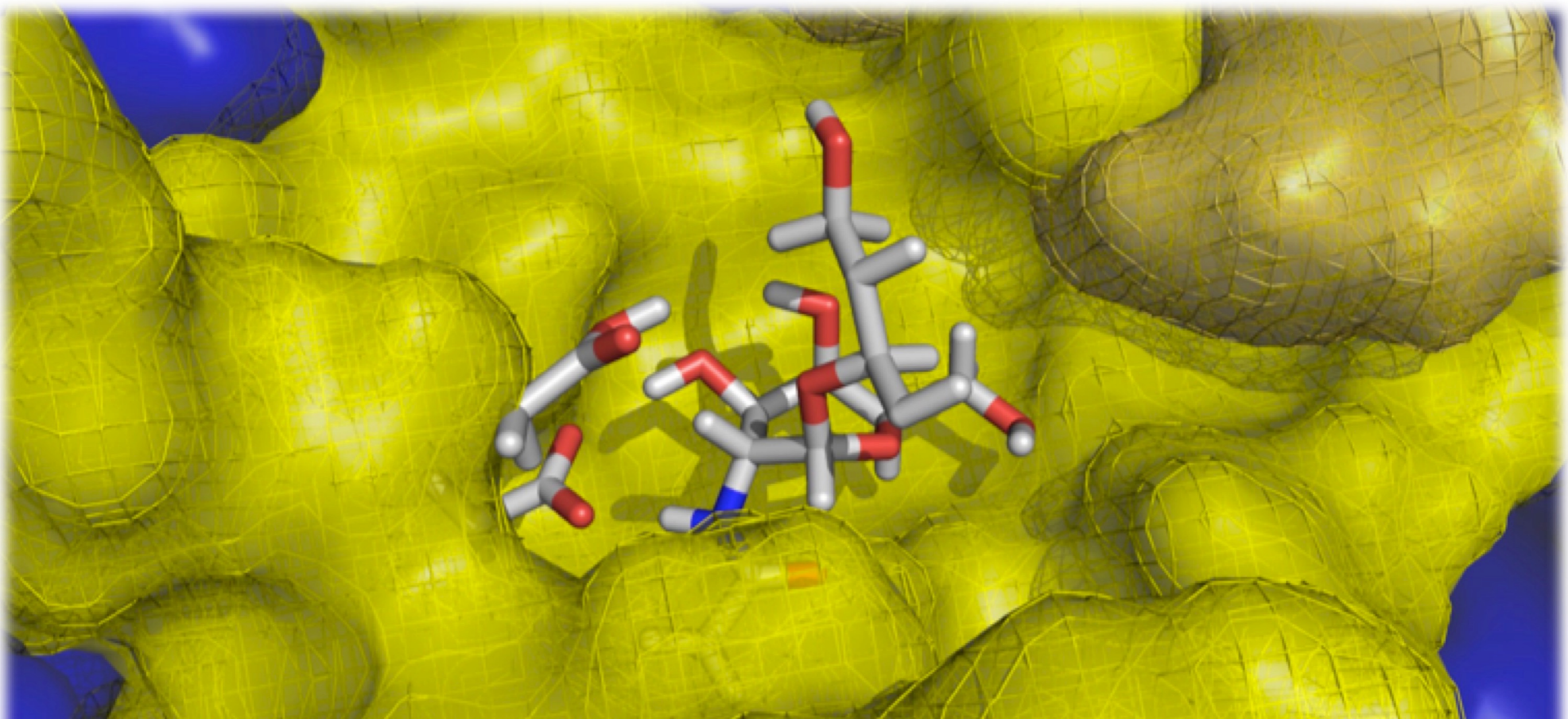


# The Catalytic Mechanism of an Enzymatic Reaction

**Pedro Alexandrino Fernandes,**

Dep. Chemistry & Biochemistry, University of Porto, Portugal

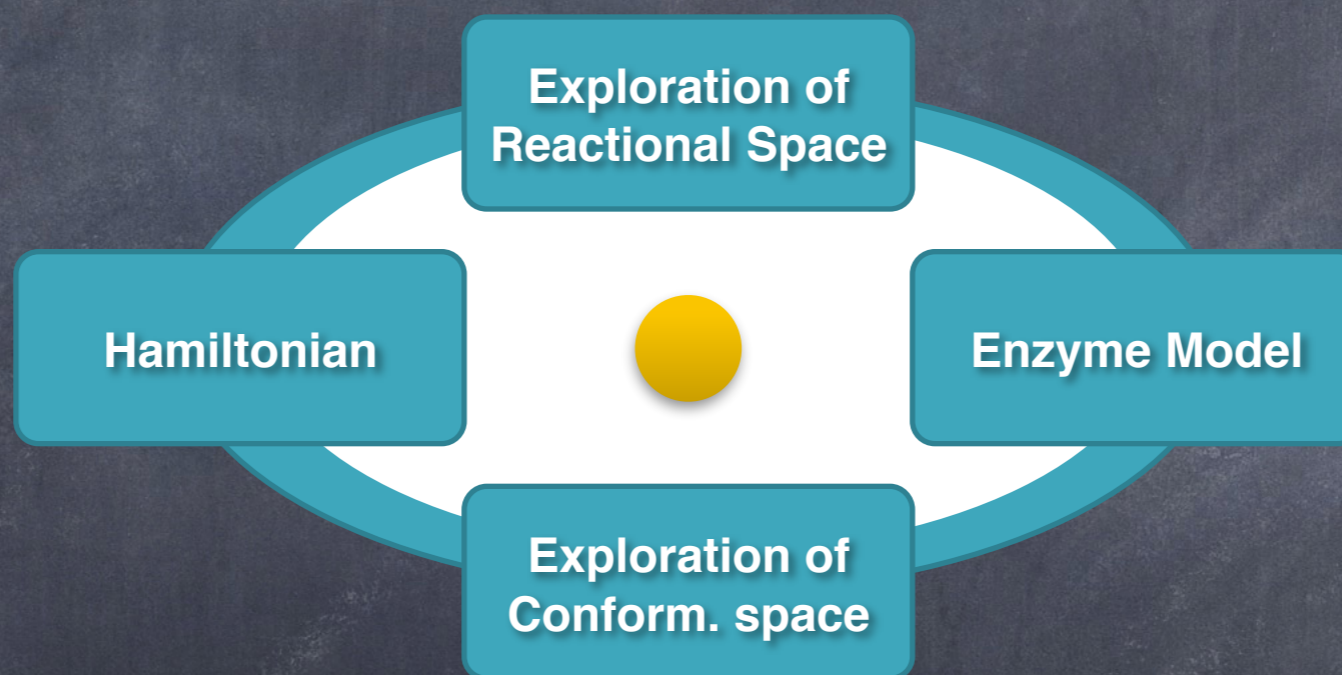
pedro.fernandes@fc.up.pt



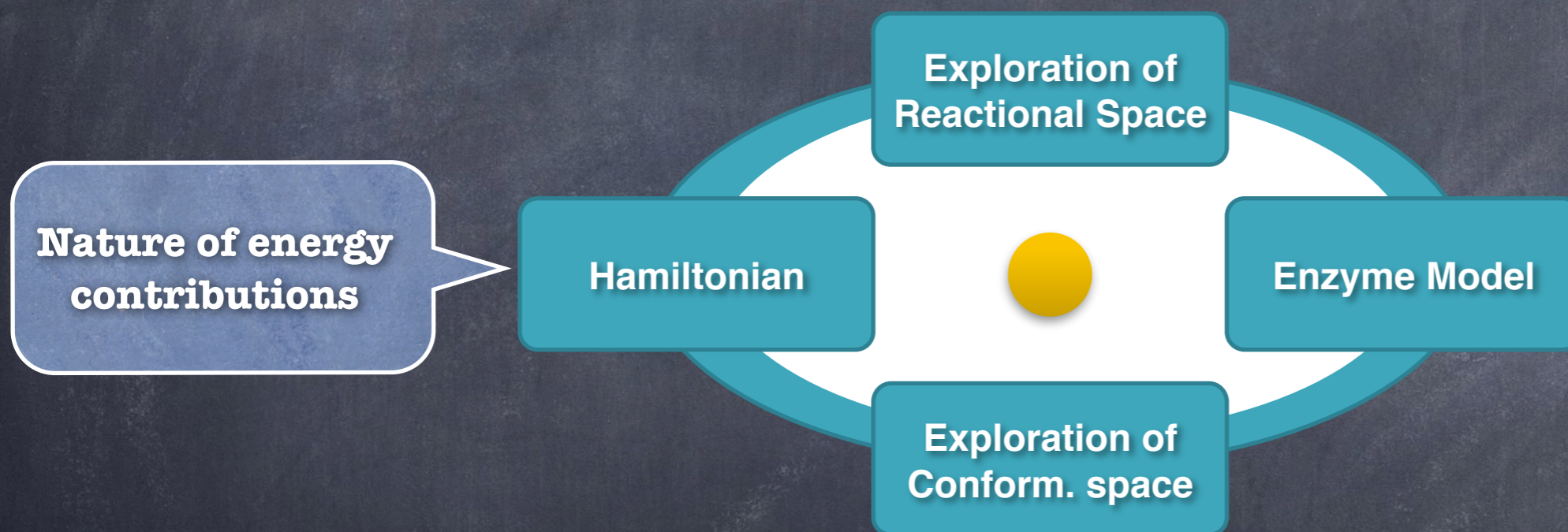
# Why enzymes are so difficult?

1. Nature of energy contributions
2. Spatial range of the atomic interactions
3. Unprecedented chemistry
4. Enzyme flexibility

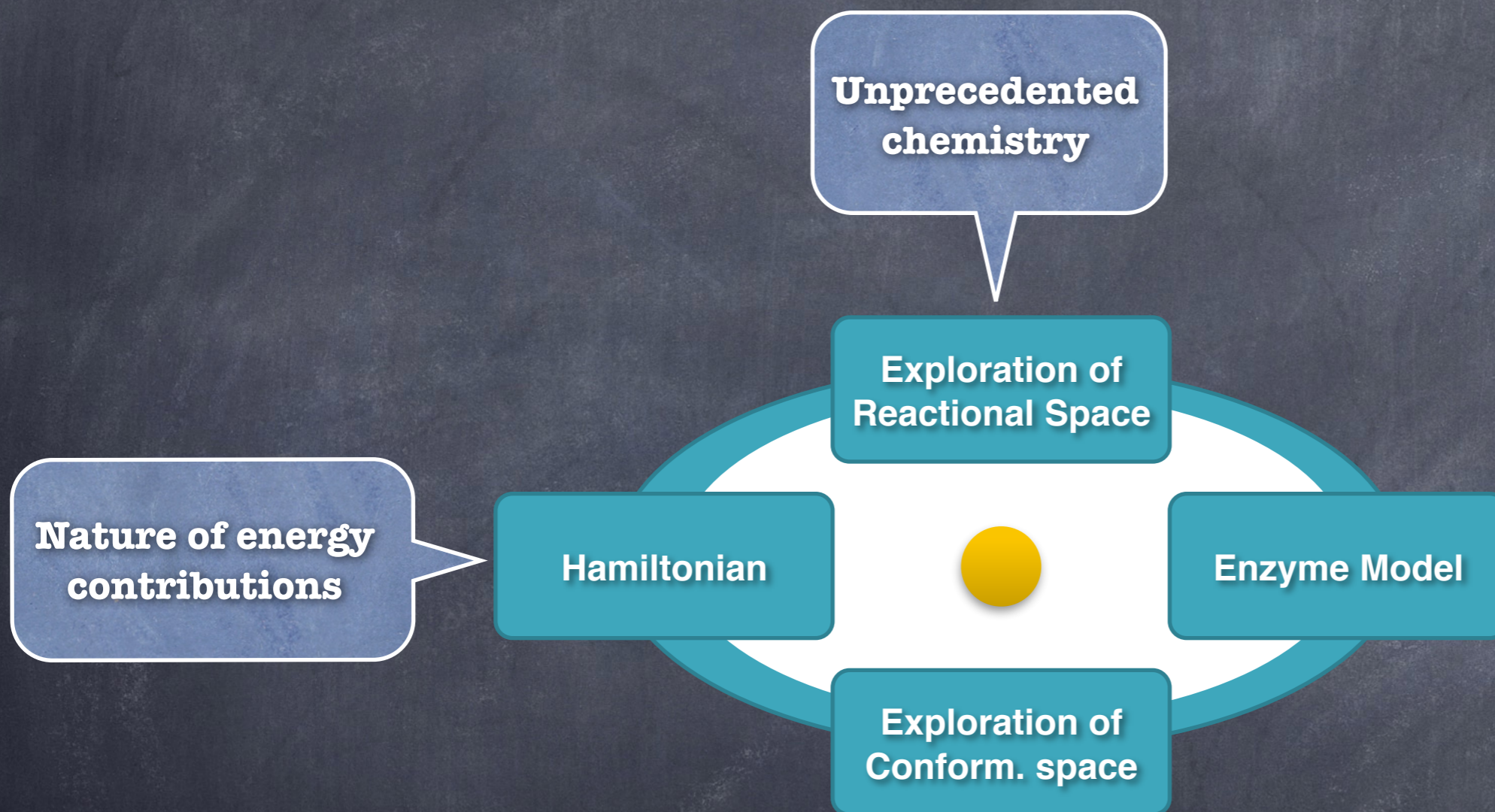
# Solutions for the problem



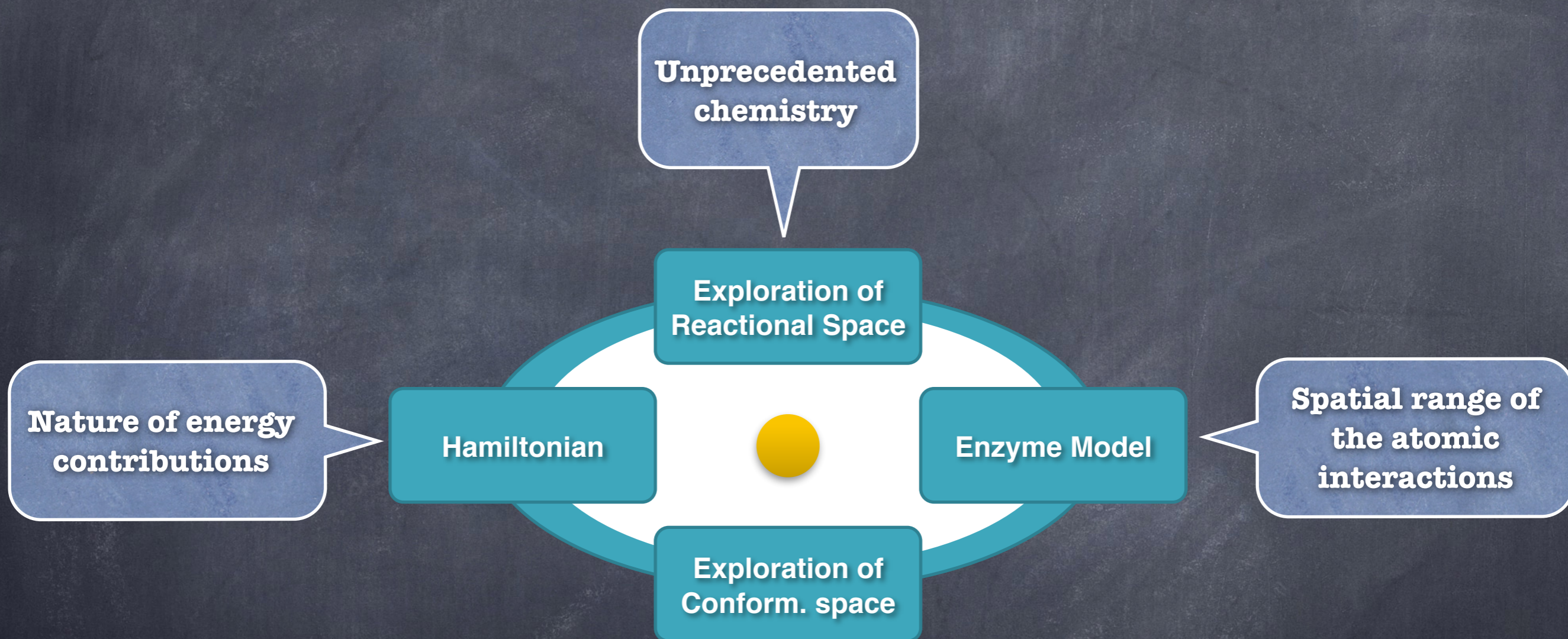
# Solutions for the problem



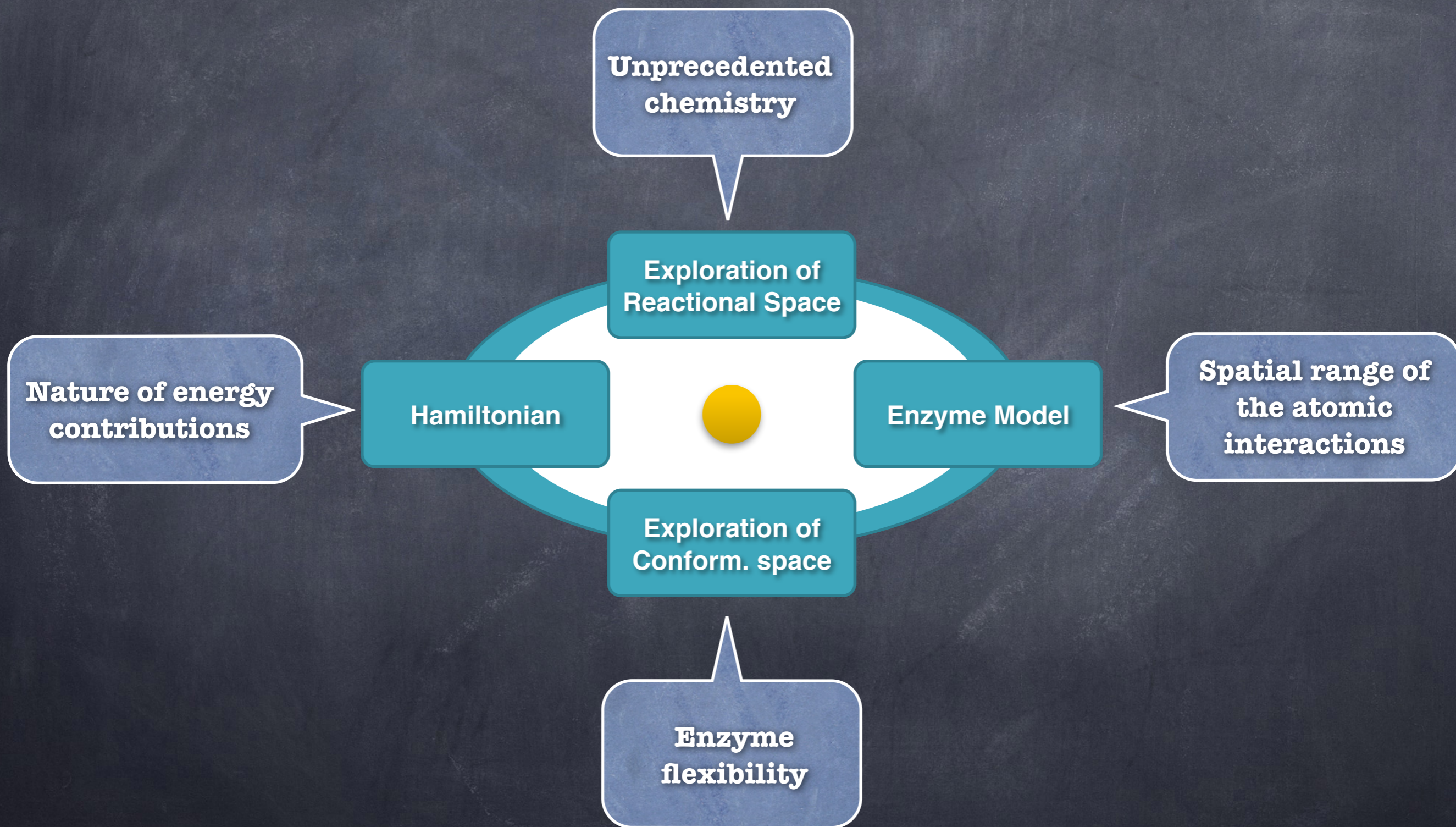
# Solutions for the problem



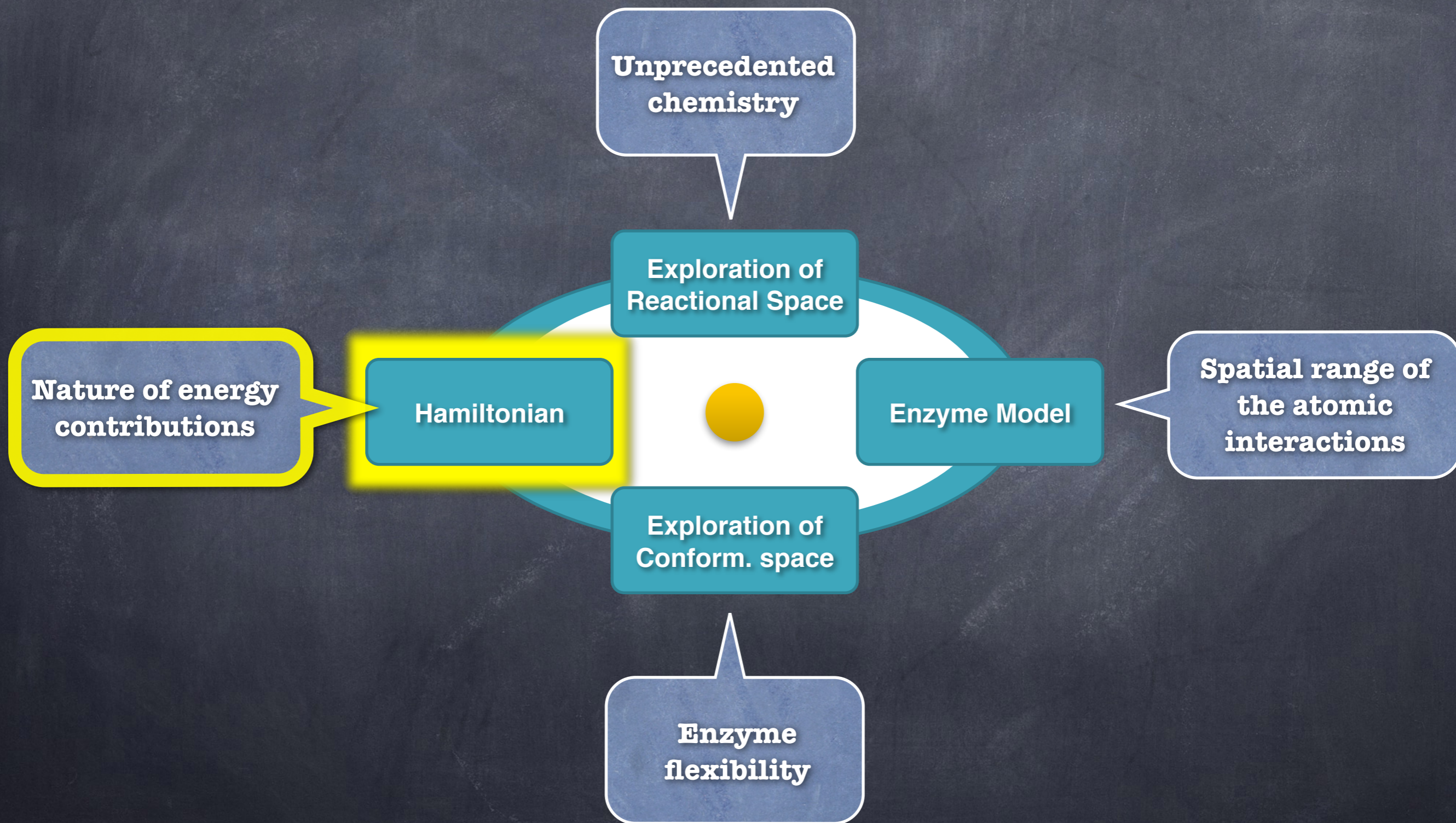
# Solutions for the problem



# Solutions for the problem



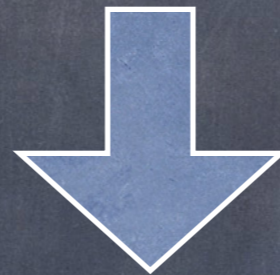
# Solutions for the problem





# The Hamiltonian

1. Nature of energy contributions



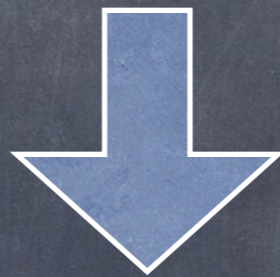
The **adequate Hamiltonian** must simultaneously describe well a wide range of interactions

# The Hamiltonian

Size of the QM region usually 100-500 atoms

# The Hamiltonian

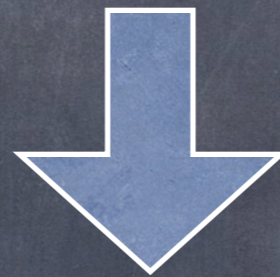
Size of the QM region usually 100-500 atoms



**DFT** is the only practical QM theoretical level

# The Hamiltonian

Size of the QM region usually 100-500 atoms



DFT is the only practical QM theoretical level



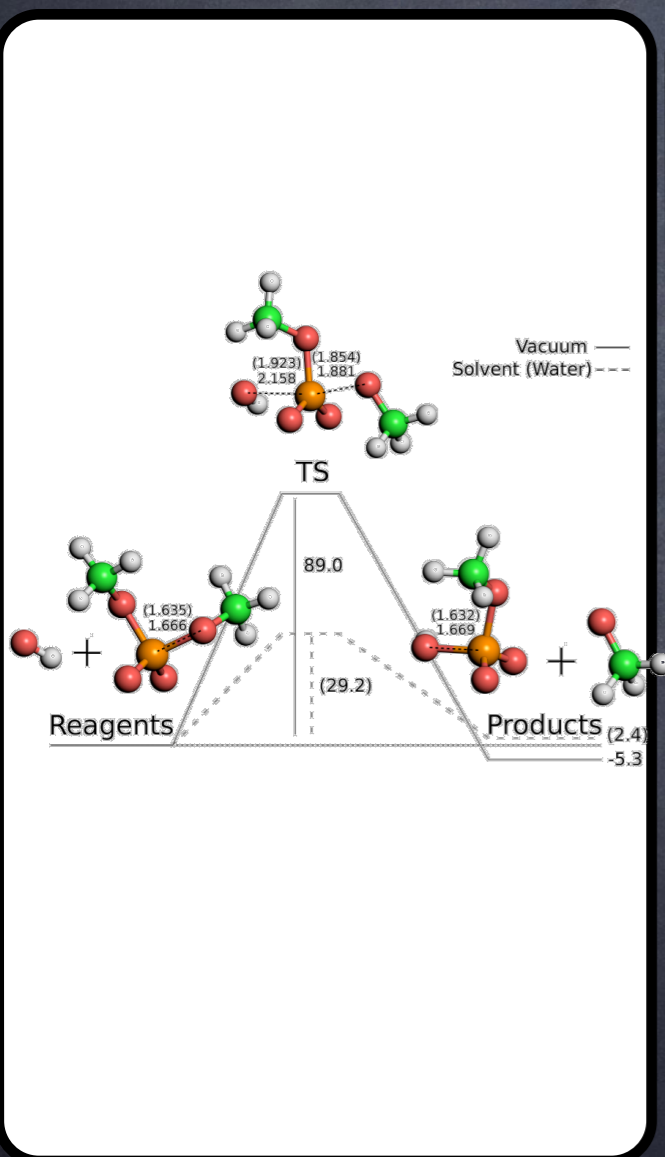
Problem: Functional performance is case-dependent

# Benchmarking Density Functionals

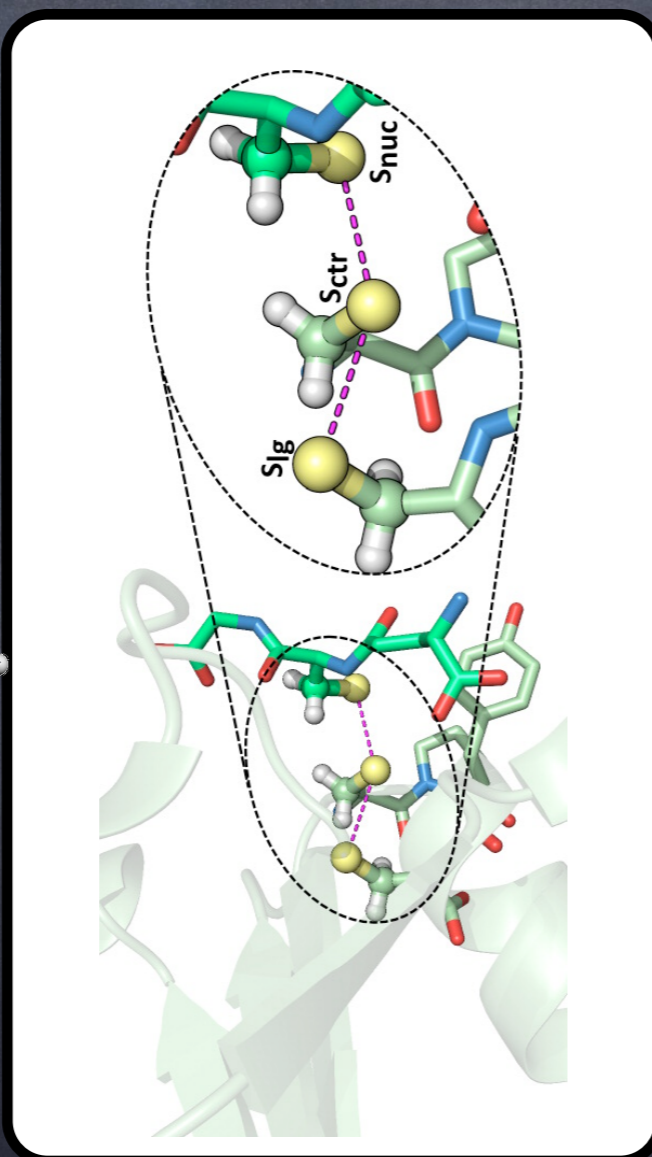
**Benchmarking of the DFT functionals  
before embarking in QM/MM  
calculations is highly advisable**

# Benchmarking Density Functionals

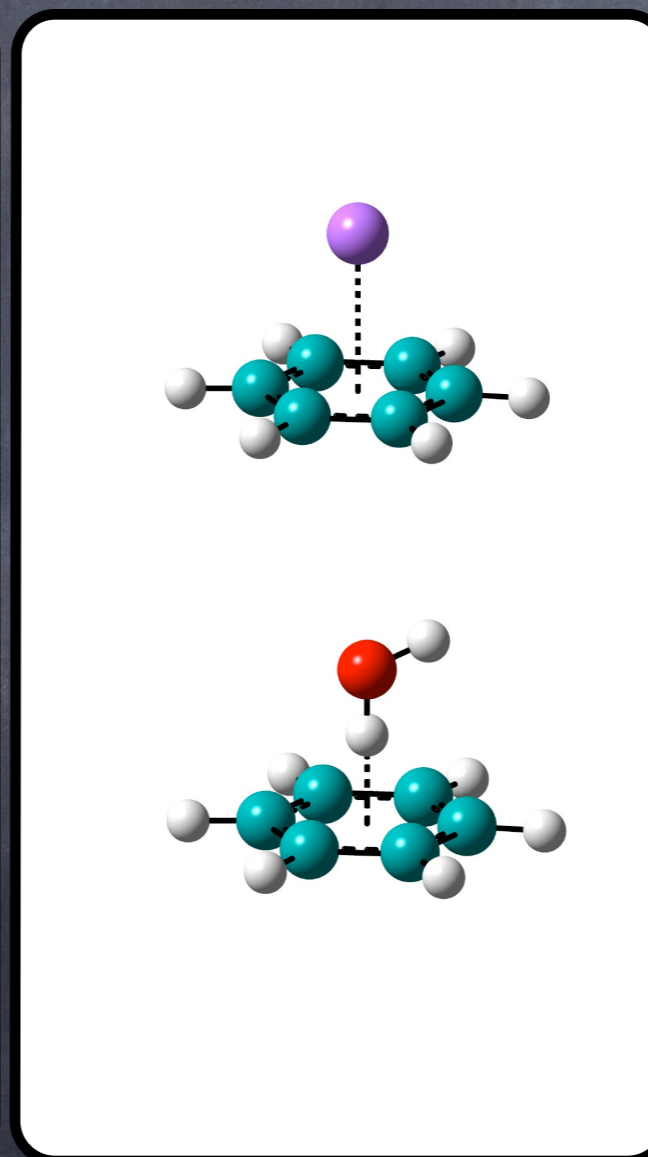
**Specific** heavy  
atom transfer



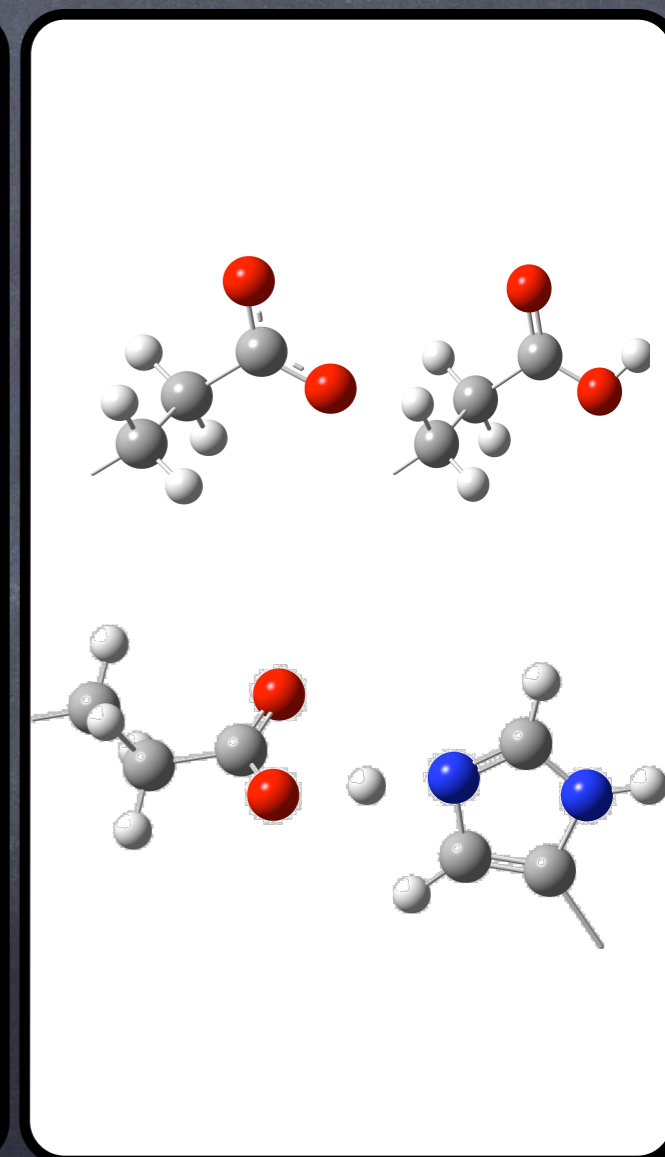
**Specific** heavy  
atom transfer



**Specific**  
interactions

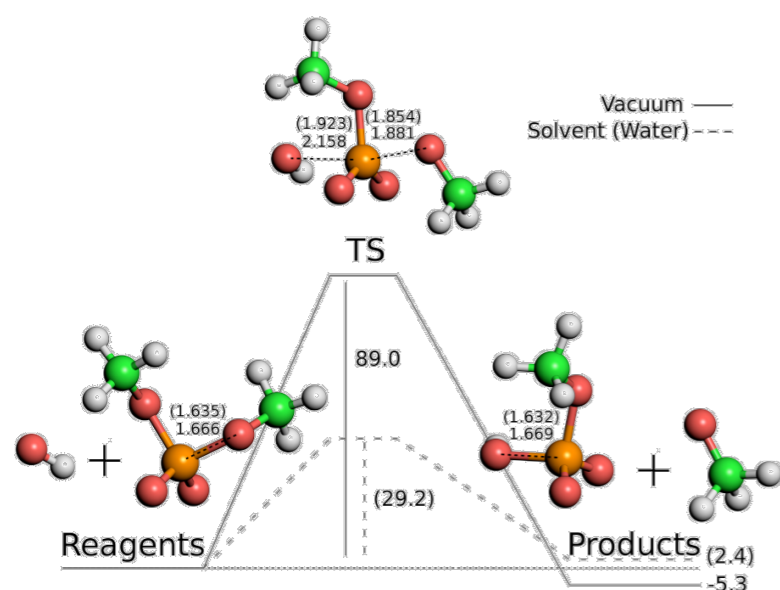


PA and PT of/between  
**residues**



# Benchmarking Density Functionals

**Specific** heavy atom  
transfer



**EXAMPLE:**

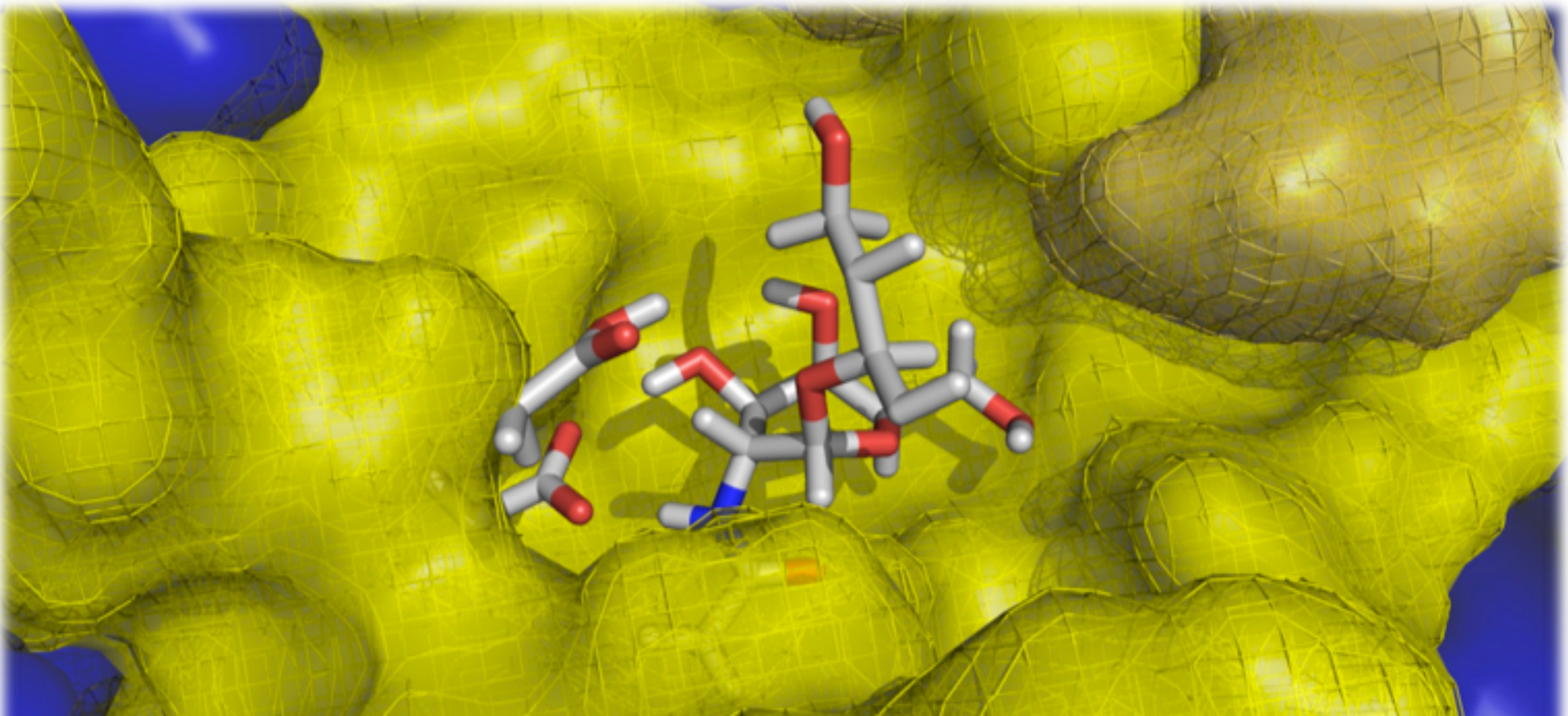
Hydrolysis of phosphodiester  
bonds.

# Benchmarking Density functionals

**Pedro Alexandrino Fernandes,**

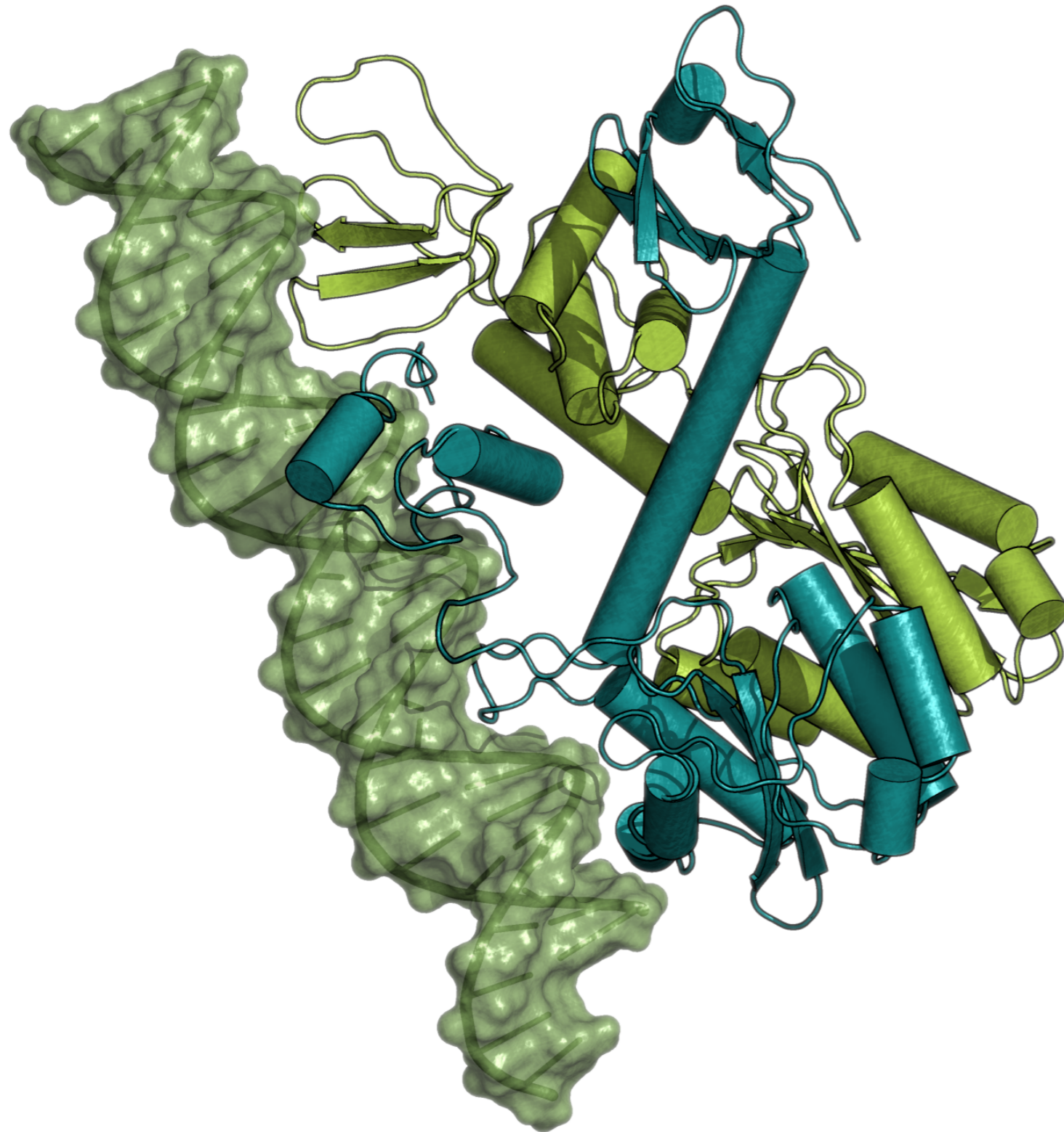
Dep. Chemistry & Biochemistry, University of Porto, Portugal

pedro.fernandes@fc.up.pt





# HIV-1 Integrase



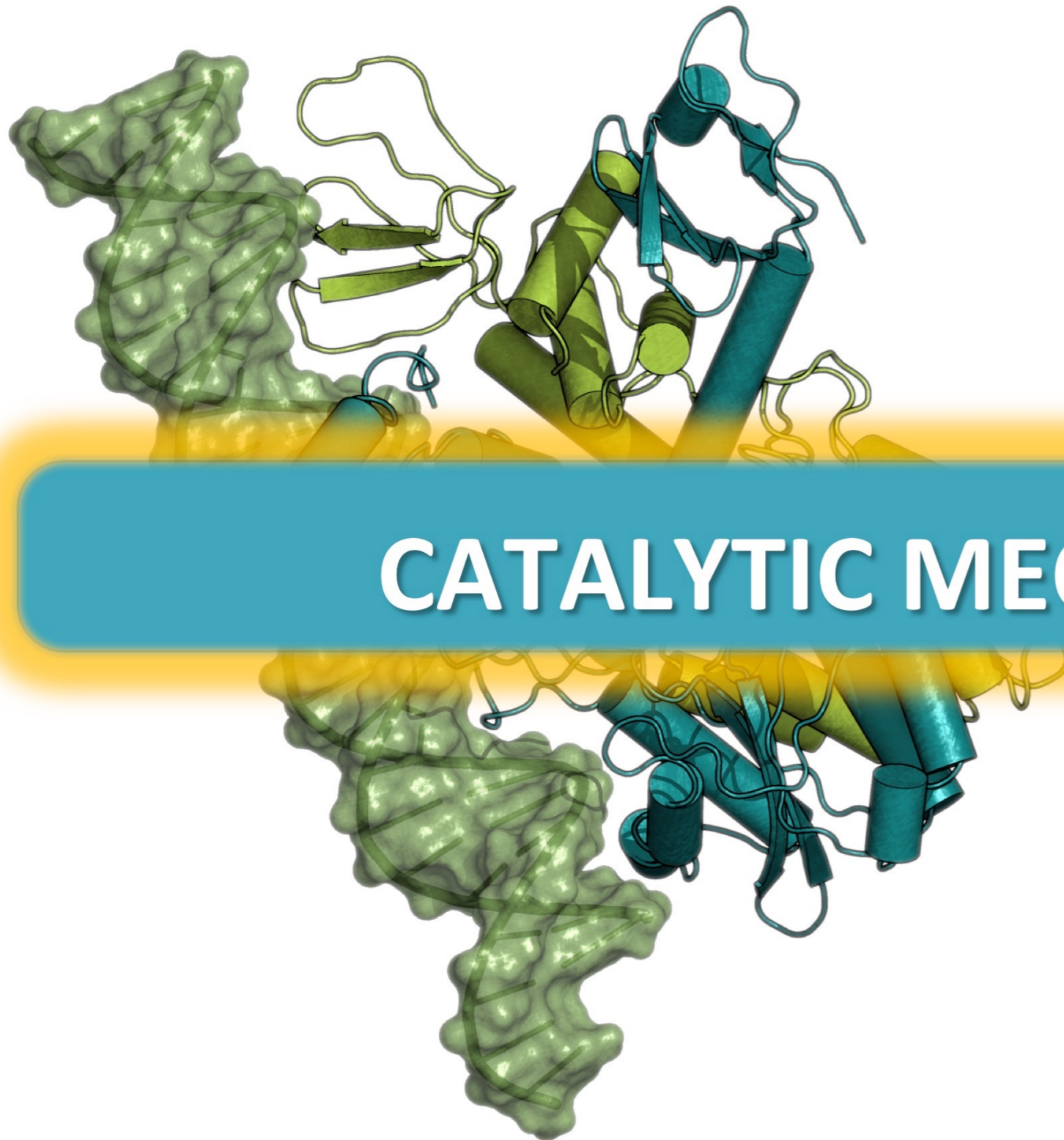
Integrates HIV-1 viral DNA into the human DNA.

Important for the treatment of HIV-1 Infection.

Drug discovery for Integrase extremely slow.

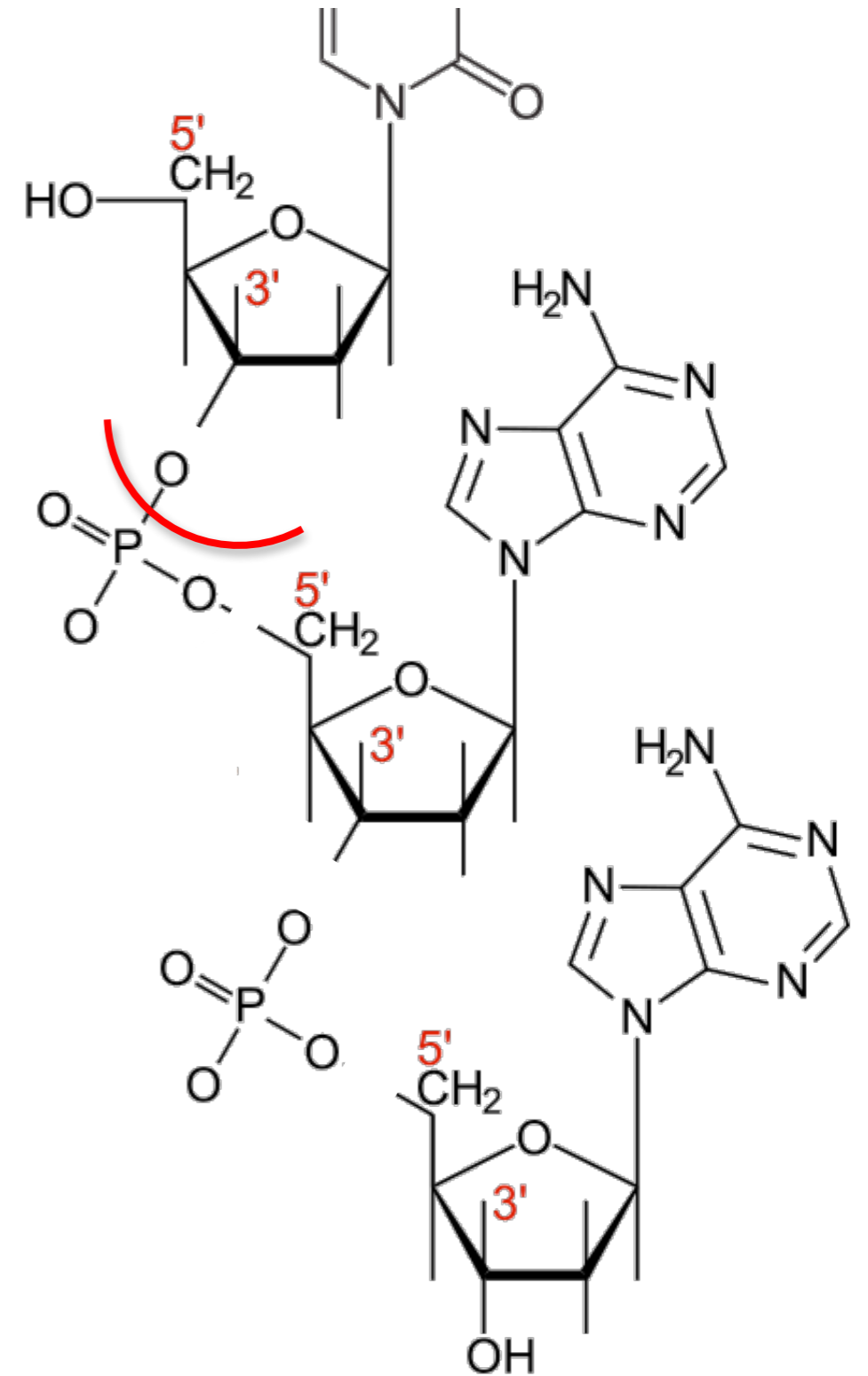
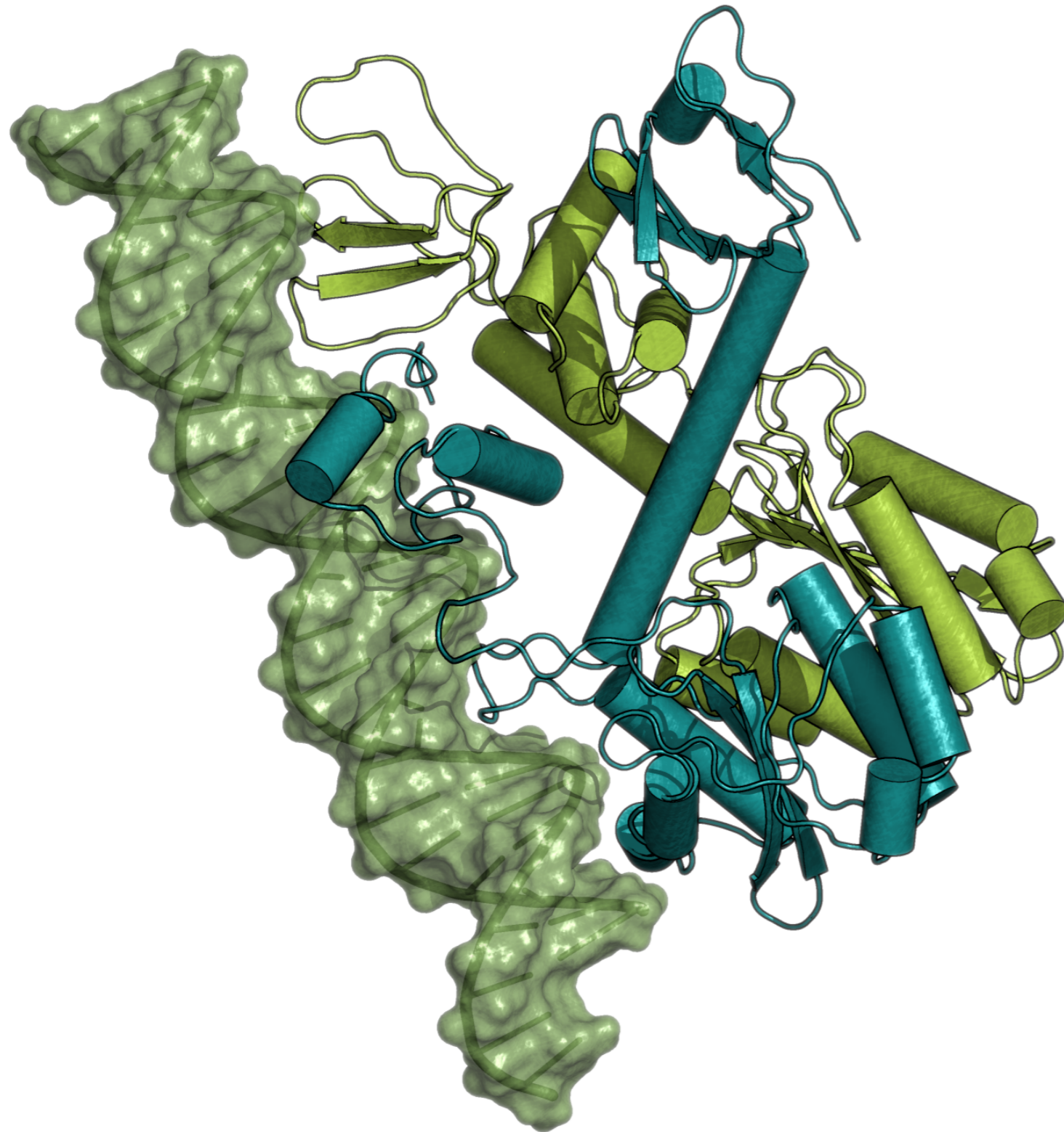
Lack of knowledge about structure and mechanism

# HIV-1 Integrase

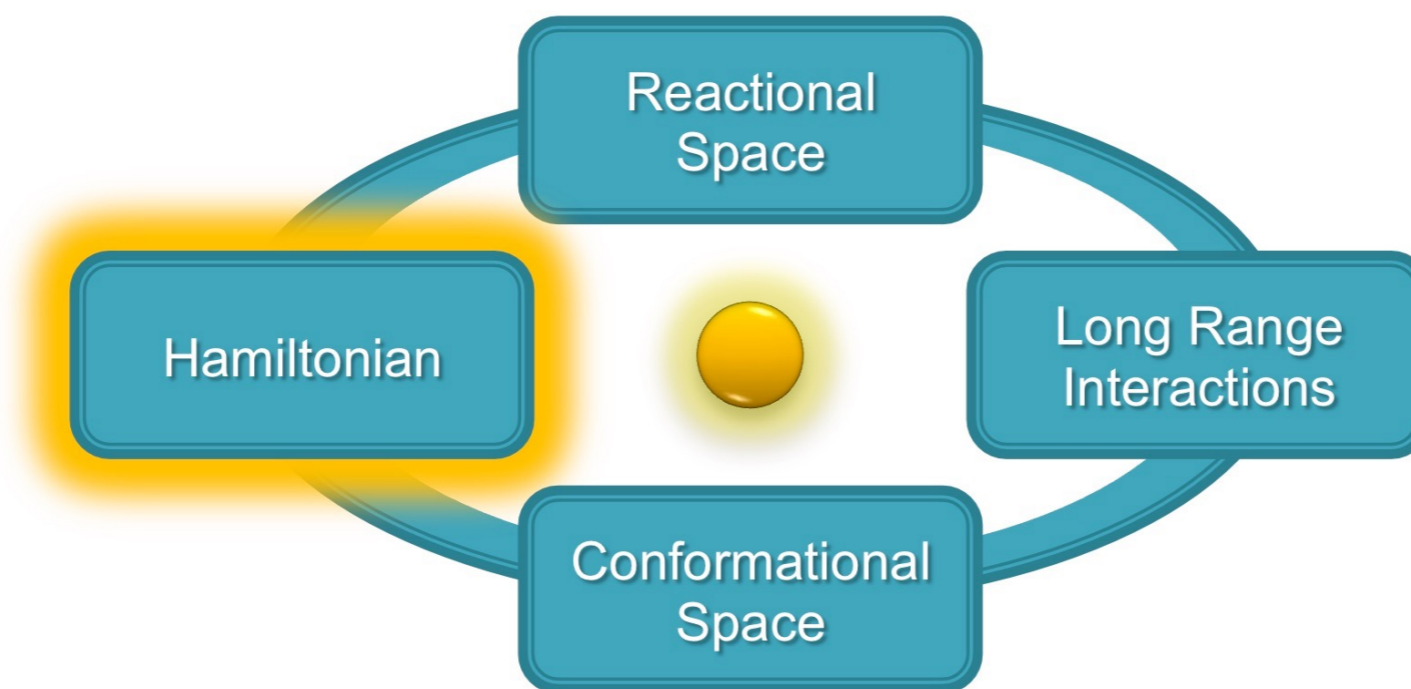


**CATALYTIC MECHANISM?**

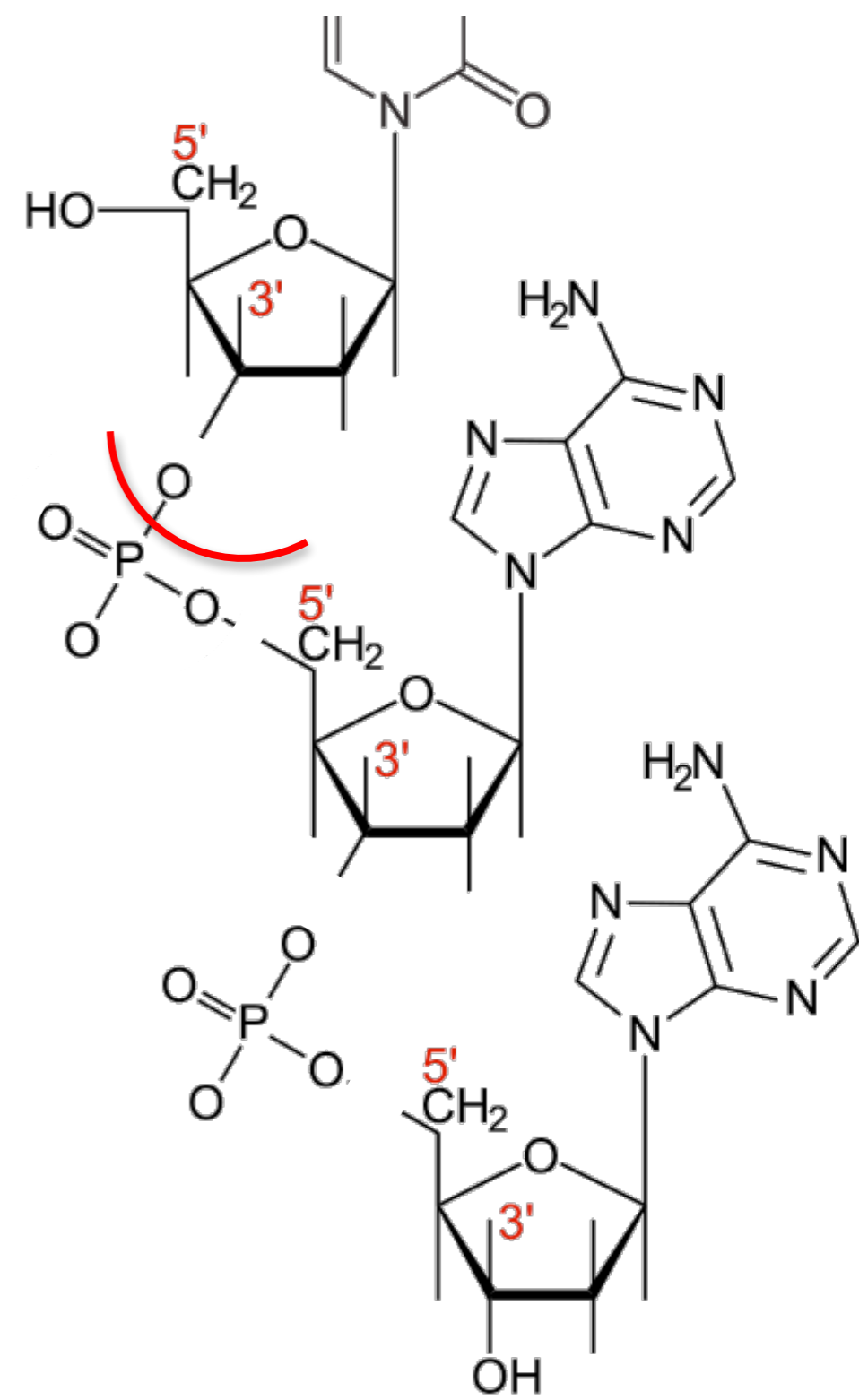
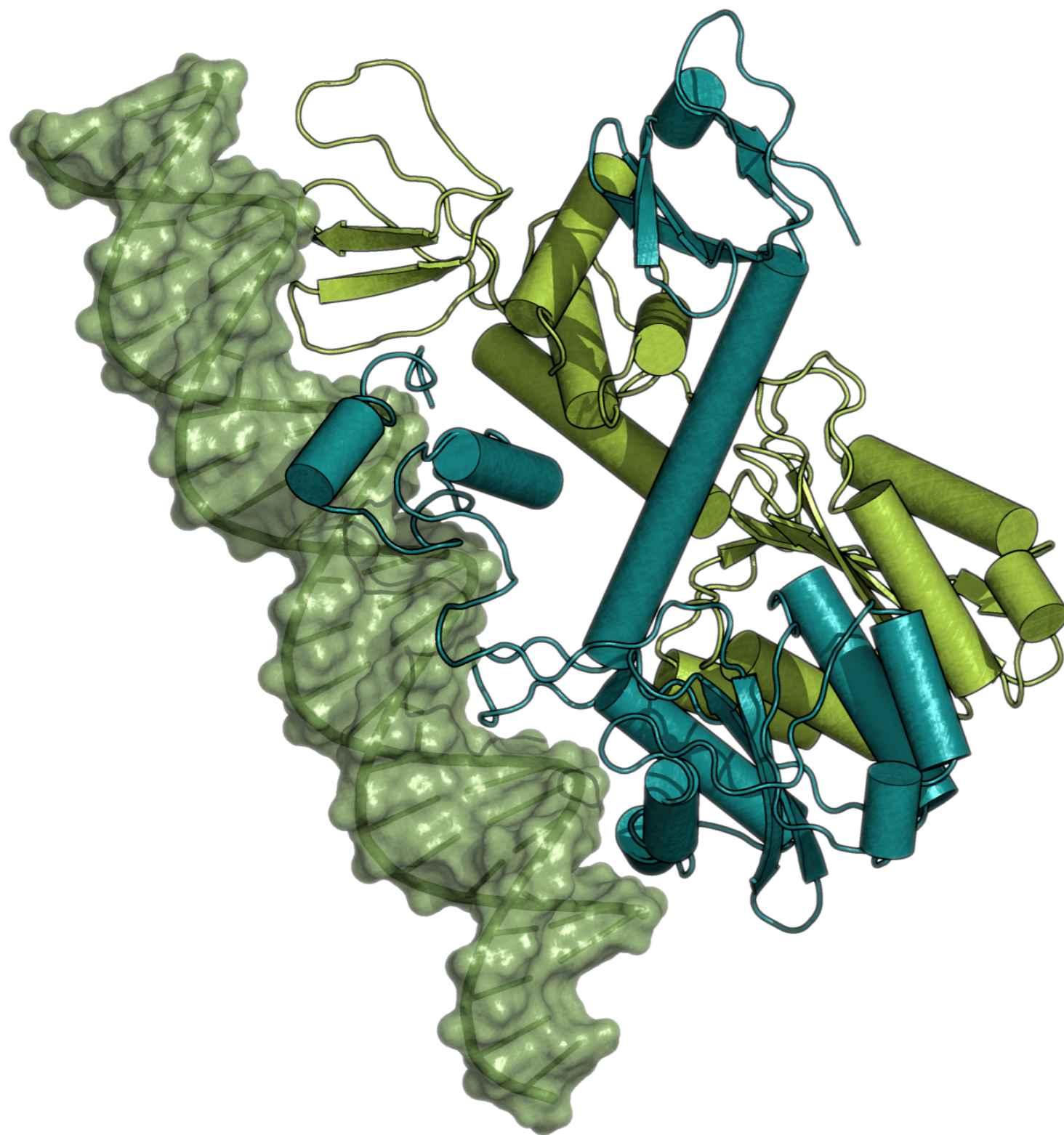
# HIV-1 Integrase



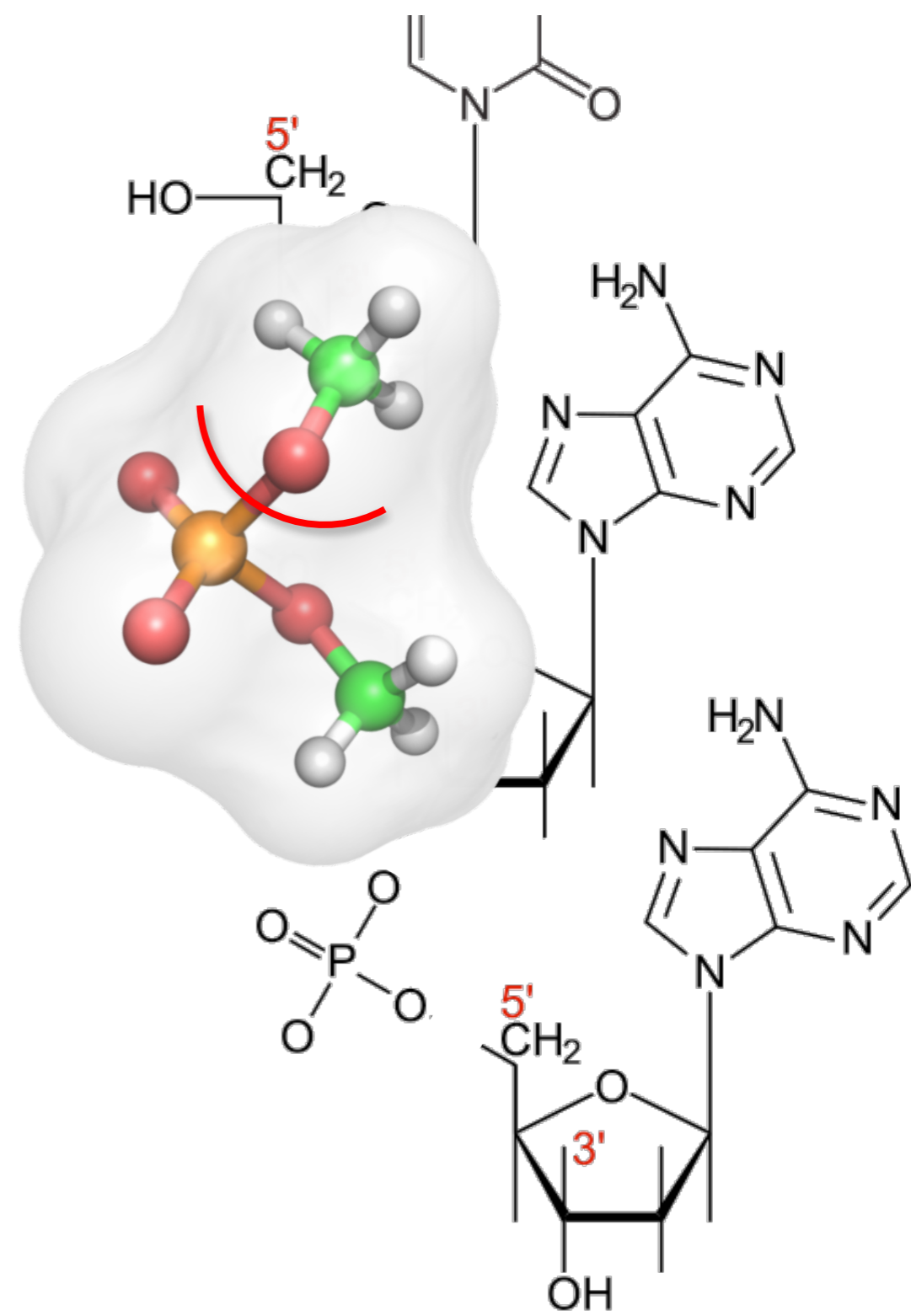
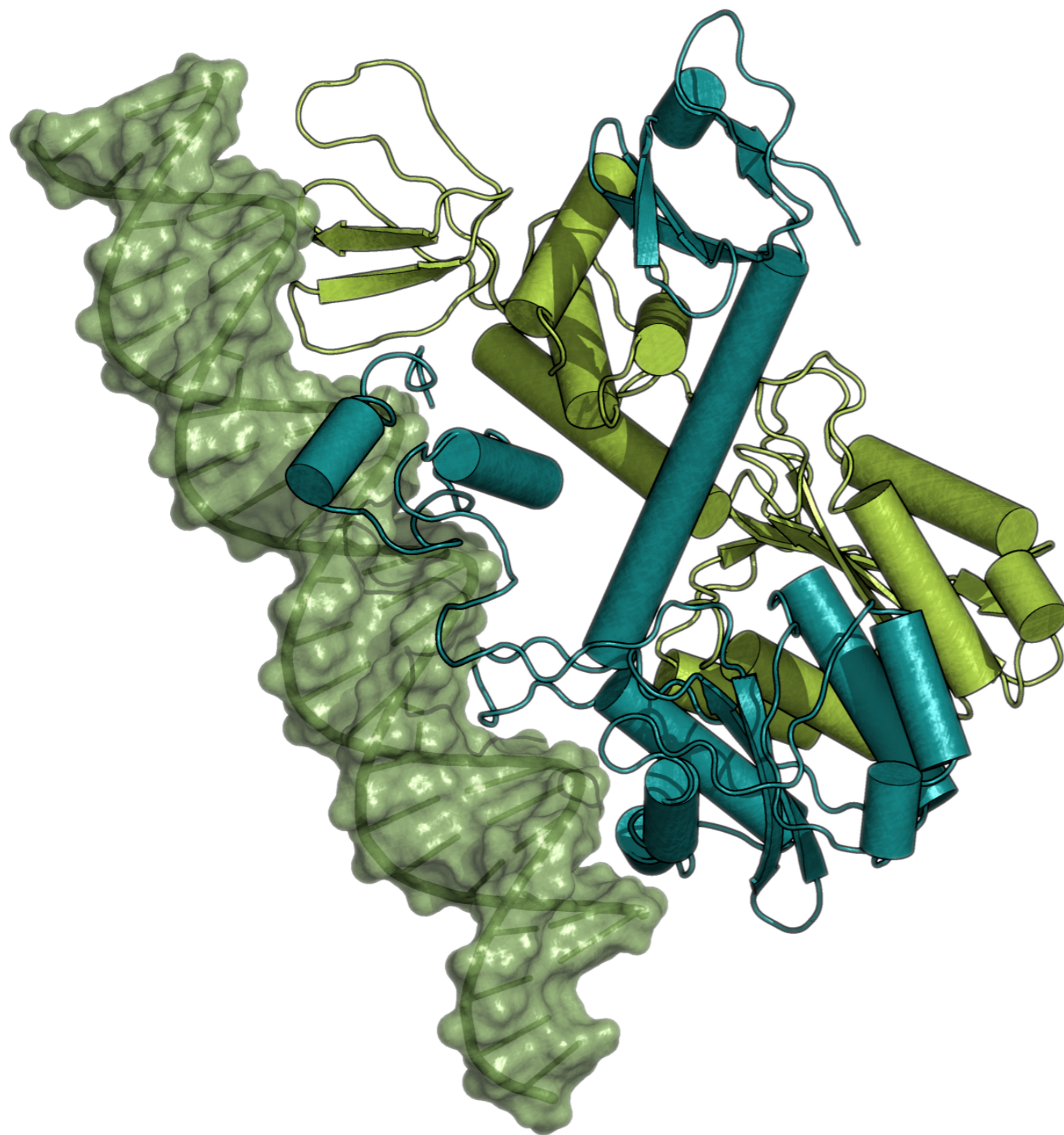
# The Hamiltonian



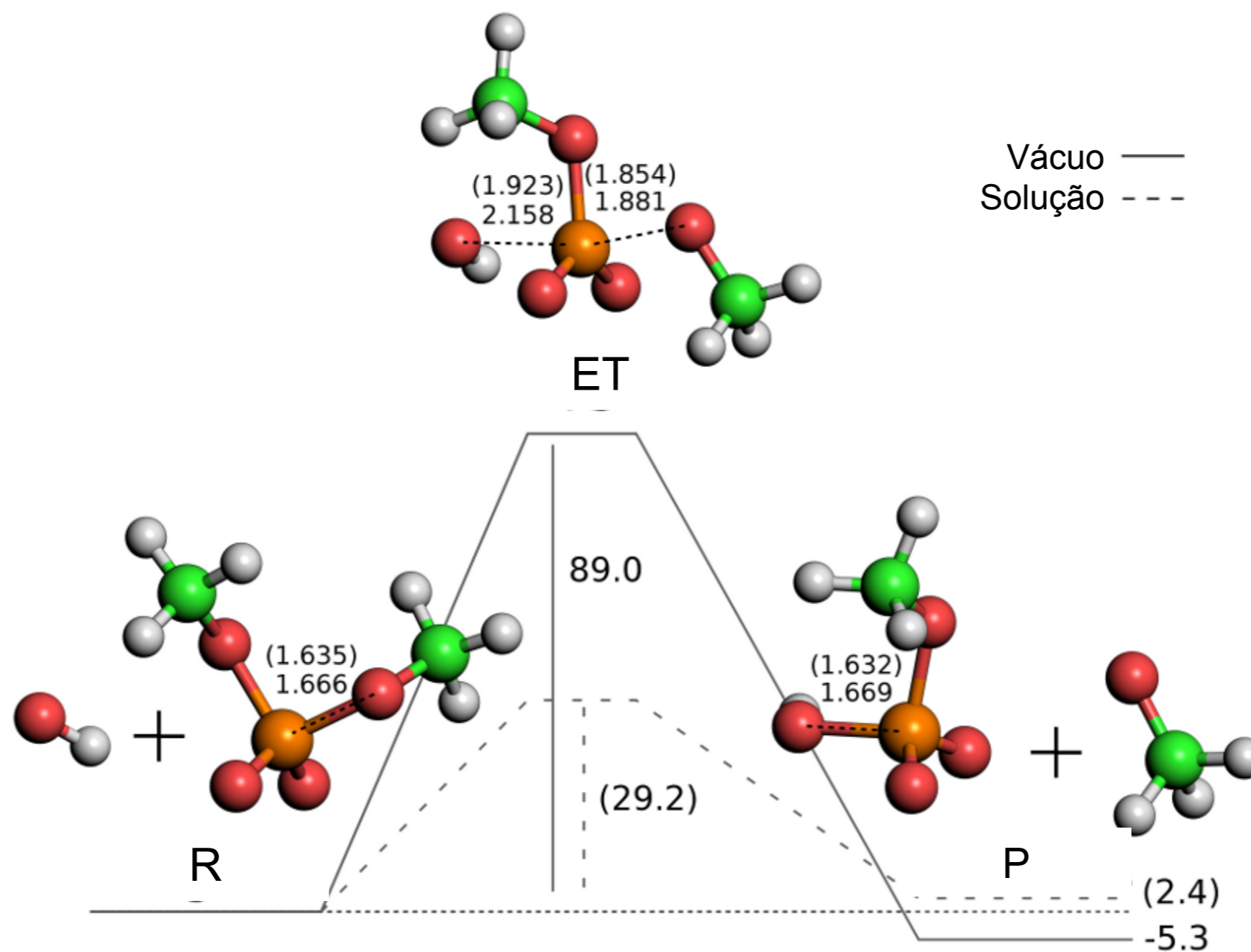
# The Hamiltonian



# Accuracy of Density Functionals?

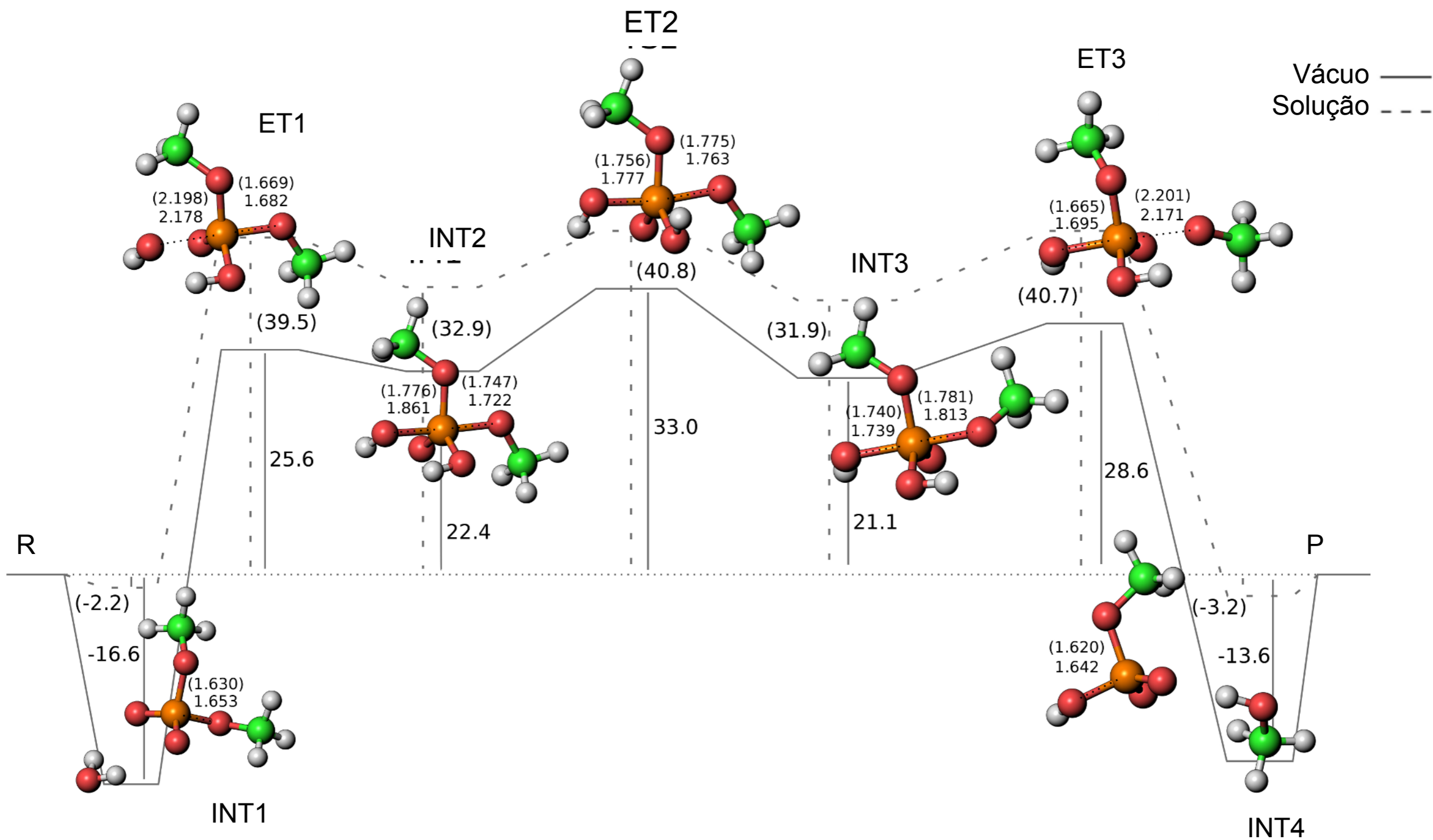


# Hydrolysis of Phosphodiester Bonds by $\text{HO}^-$ .



B3LYP/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p)

# Hydrolysis of Phosphodiester Bonds by H<sub>2</sub>O.



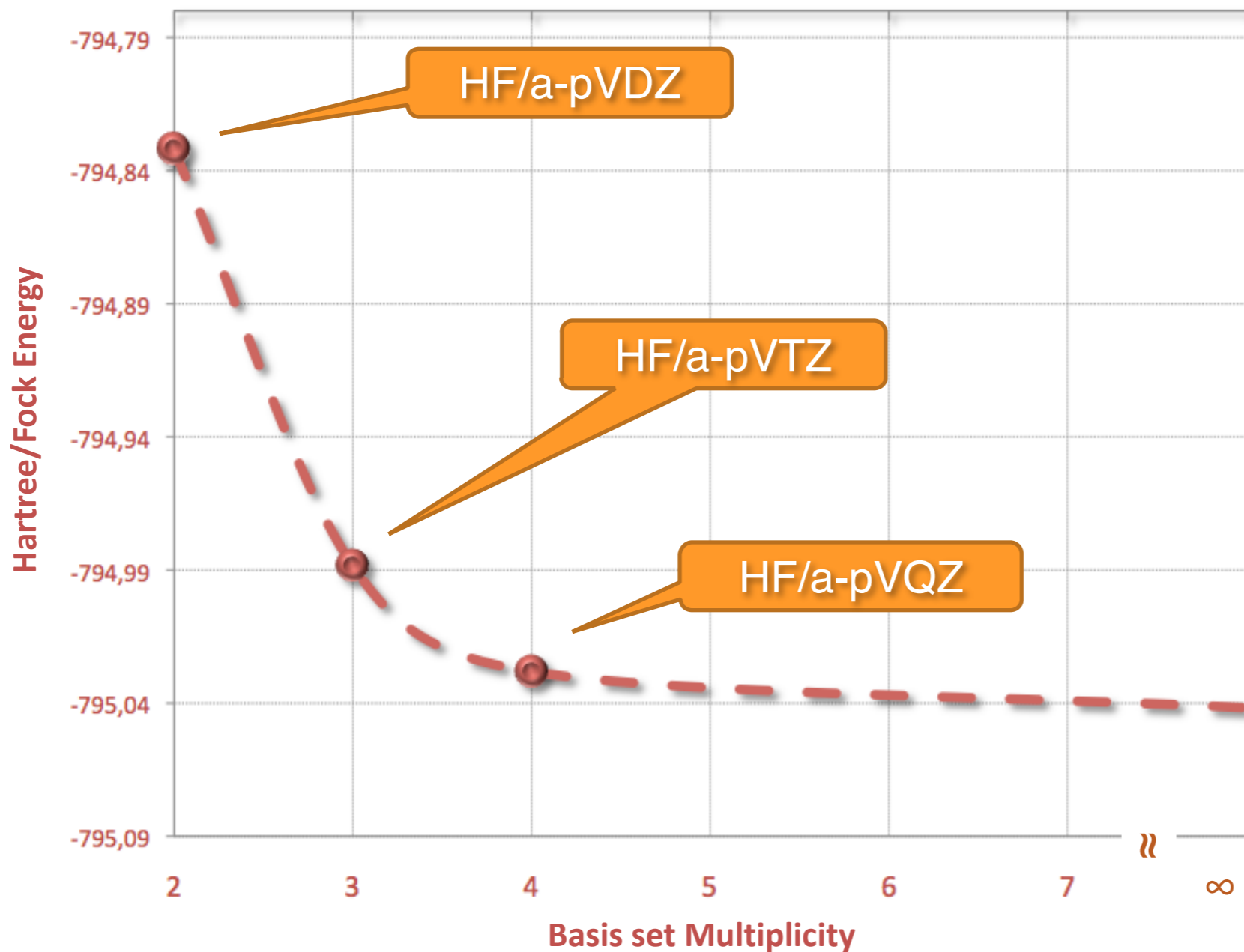
B3LYP/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p)



# Reference PES: CCSD(T)/CBS.



$$E_{\infty}^{\text{TOT}} = E_{\infty}^{\text{HF}} + E_{\infty}^{\text{corr}}, \quad E_{\infty}^{\text{HF}} = E_X^{\text{HF}} + a \cdot \exp(-bX), \quad E_{\infty}^{\text{corr}} = \frac{X^3}{X^3 - (X-1)^3} E_X^{\text{corr}} - \frac{X^3}{X^3 - (X-1)^3} E_{(X-1)}^{\text{corr}}$$



Halkier Method

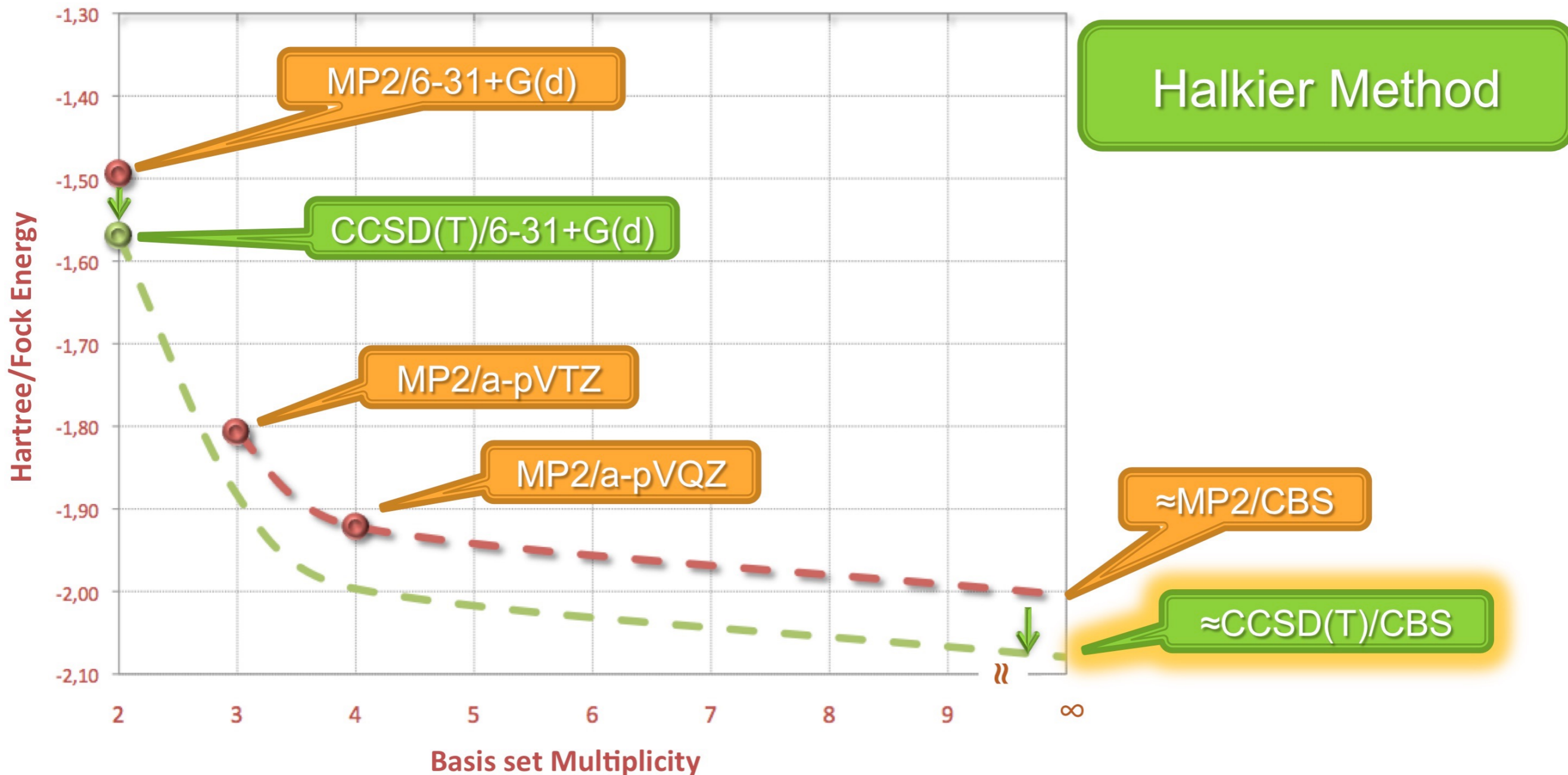
≈HF/CBS

# Reference PES: CCSD(T)/CBS.



$$E_{\infty}^{\text{TOT}} = E_{\infty}^{\text{HF}} + E_{\infty}^{\text{corr}}, \quad E_{\infty}^{\text{HF}} = E_X^{\text{HF}} + a \cdot \exp(-bX),$$

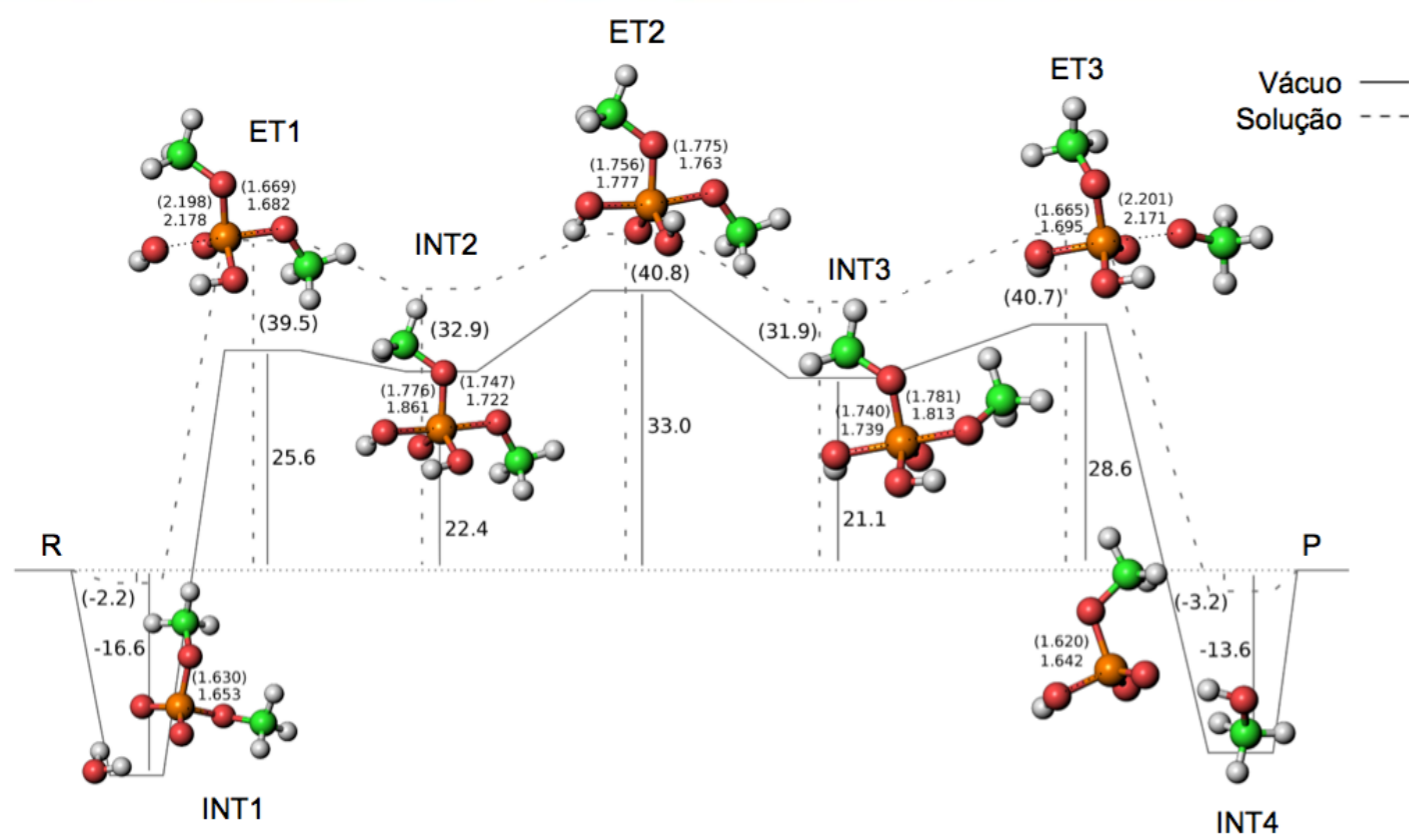
$$E_{\infty}^{\text{corr}} = \frac{X^3}{X^3 - (X-1)^3} E_X^{\text{corr}} - \frac{X^3}{X^3 - (X-1)^3} E_{(X-1)}^{\text{corr}}$$



# Reference PES: CCSD(T)/CBS.



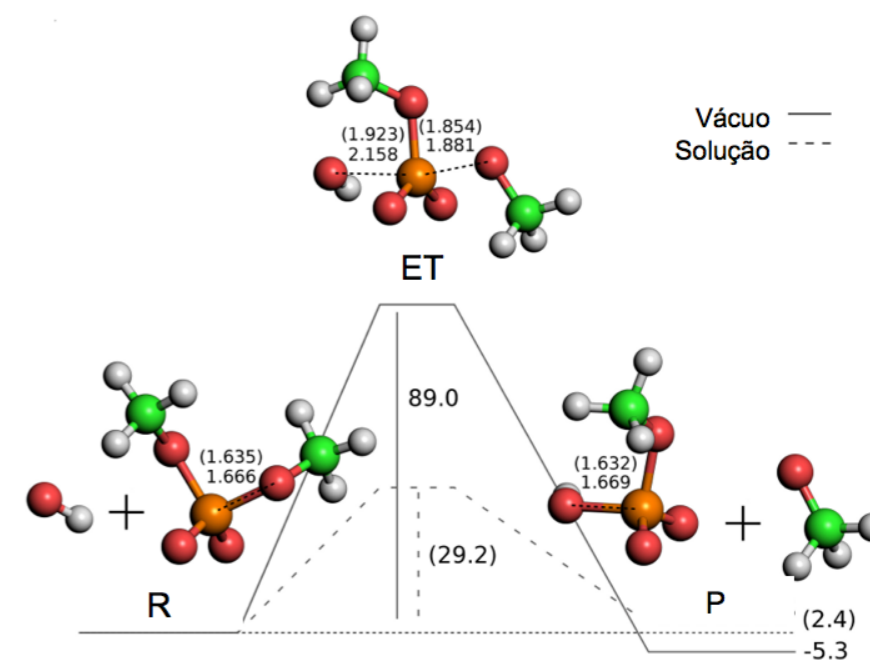
## Hydrolysis by H<sub>2</sub>O



$$\Delta G^*(\text{calculated}) = 38,0 \text{ kcal}\cdot\text{mol}^{-1}$$

$$\Delta G^*(\text{exp.}) \approx 38,5 \text{ kcal}\cdot\text{mol}^{-1}$$

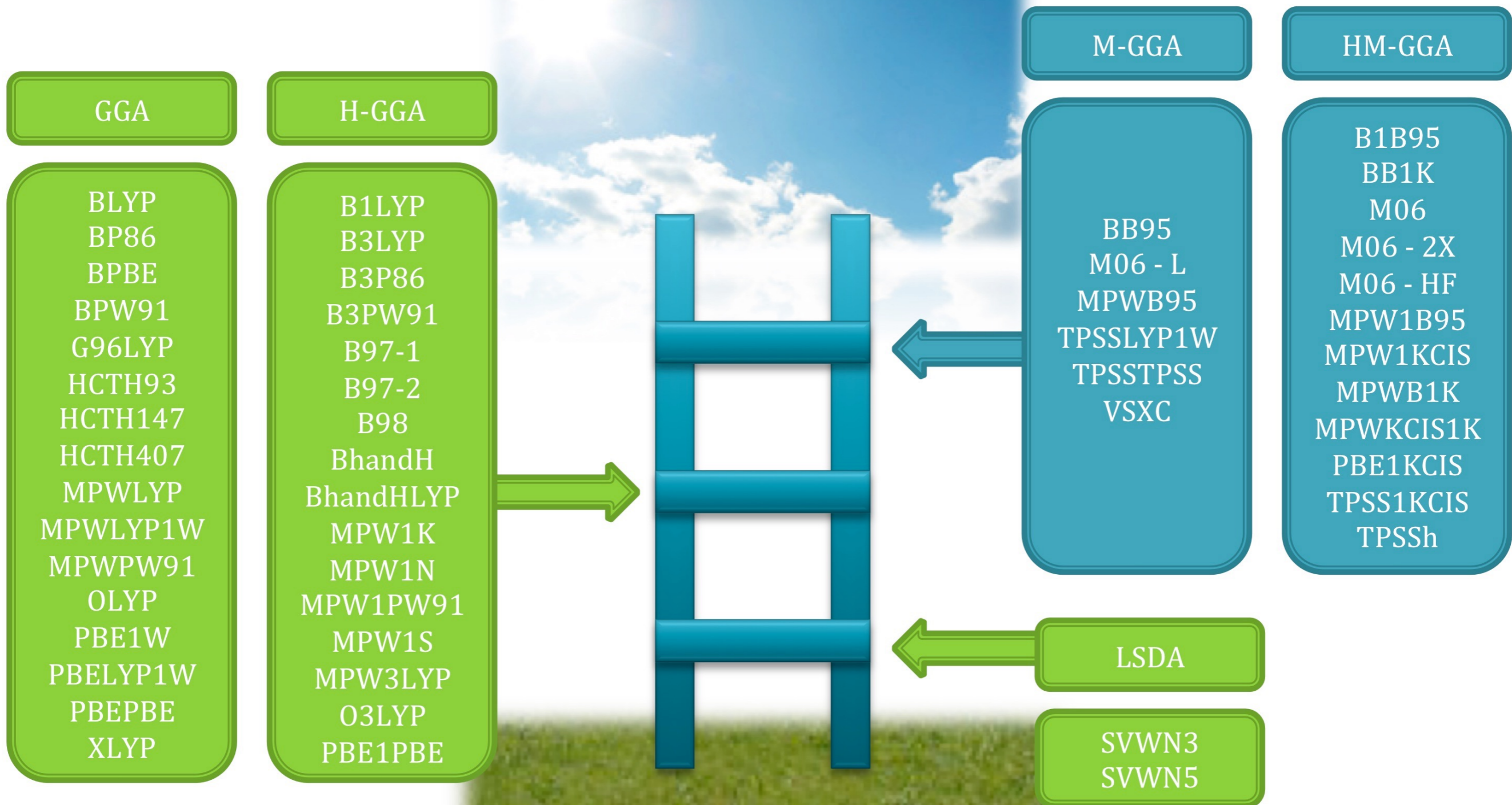
## Hydrolysis by HO<sup>-</sup>



$$\Delta G^*(\text{calculated}) = 34,7 \text{ kcal}\cdot\text{mol}^{-1}$$

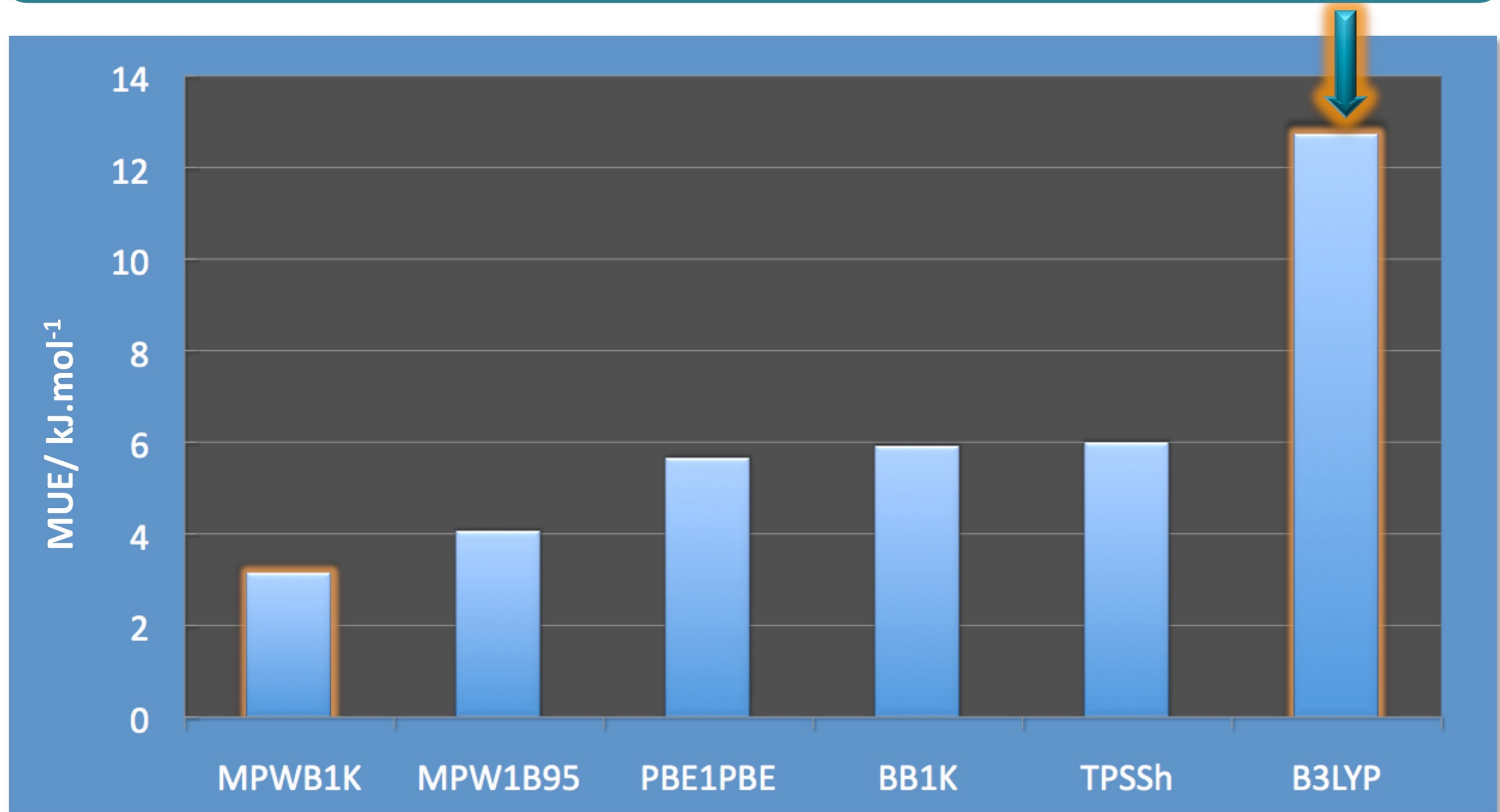
$$\Delta G^*(\text{exp.}) \approx 35 \text{ kcal}\cdot\text{mol}^{-1}$$

# Performance of 50 Density Functionals



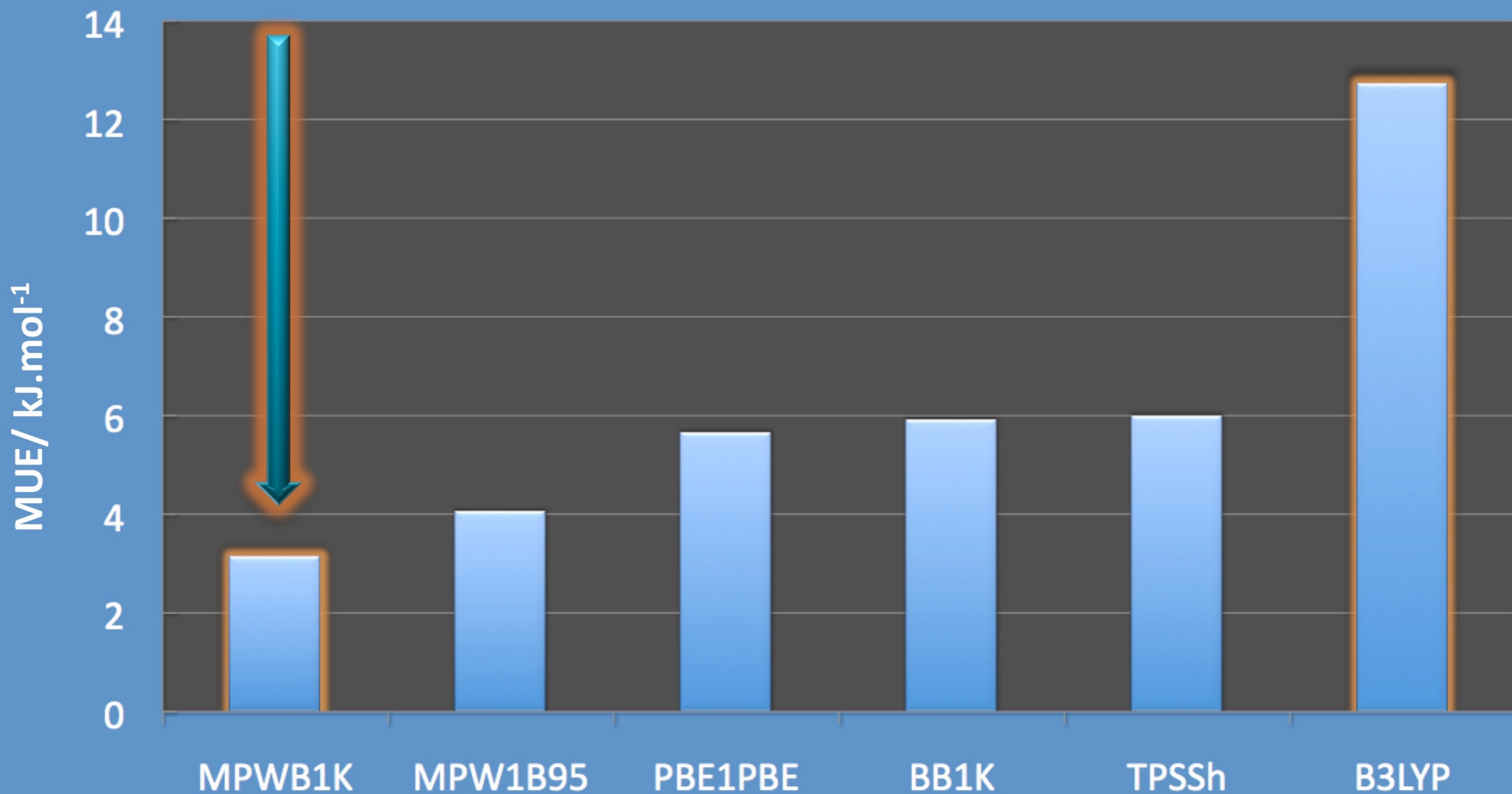
# Performance of 50 Density Functionals

B3LYP: TSs=2,9 kcal/mol; Minima=3,1 kcal/mol; **Average=3,0 kcal/mol**



# Performance of 50 Density Functionals

MPWB1K: TSs=0,3 kcal/mol; Minima=1,0 kcal/mol; Average=0,7 kcal/mol



# Conclusion



- At the end one knows the expectable accuracy of each density functional.

# Conclusion



- At the end one knows the expectable accuracy of each density functional.
- One may choose to use **B3LYP** knowing that the barriers and reaction energies will be overestimated by about 3kcal/mol.



# Conclusion



- At the end one knows the expectable accuracy of each density functional.
- One may choose to use **B3LYP** knowing that the barriers and reaction energies will be overestimated by about 3kcal/mol.
- One may choose to use MPWB1K knowing that the barriers and reaction energies will be very close to the ones of CCSD(T)/CBS



## Further Reading

JPCA, 2007, 111, 10439

JCTC, 2010, 6, 2281

JCTC, 2011, 7, 2059

JCTC, 2011, 7, 3898

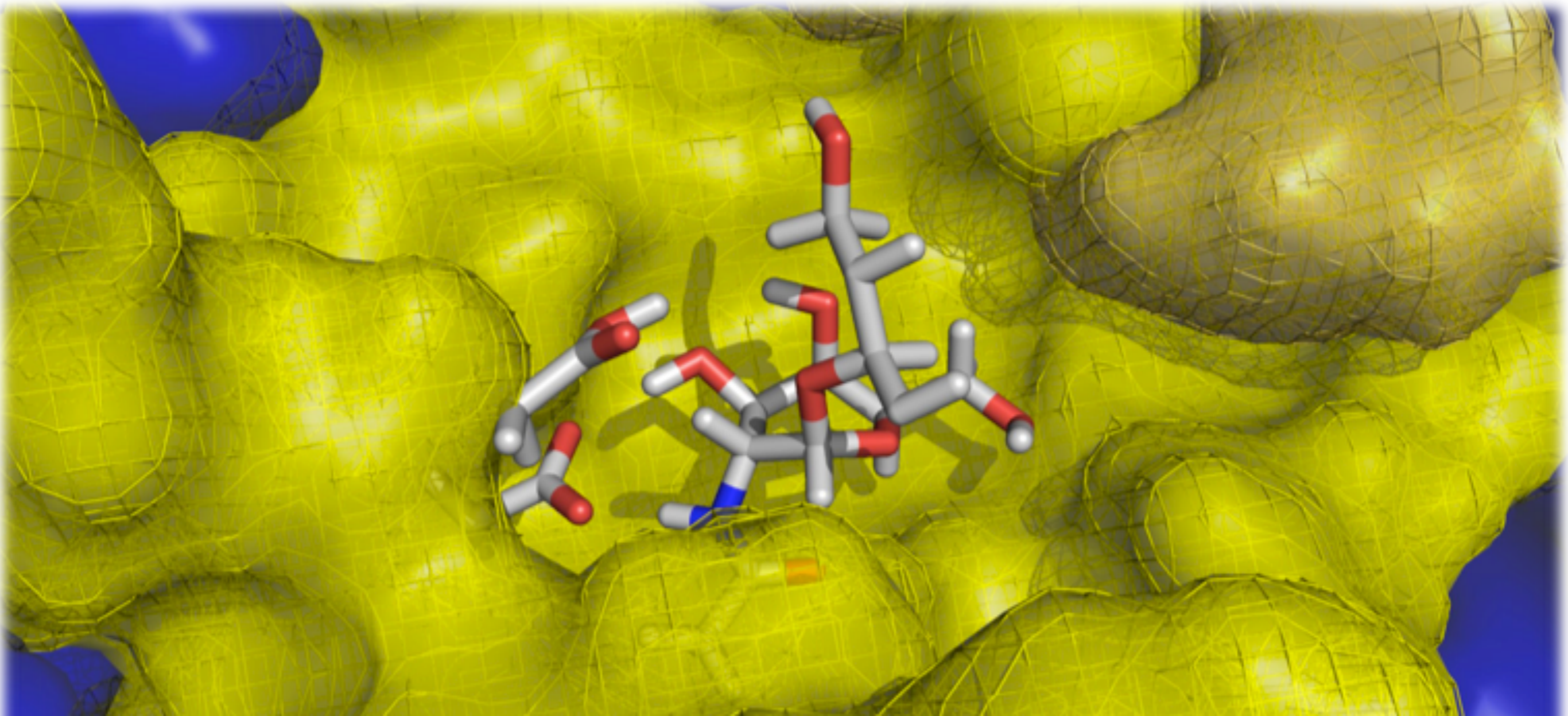
JCTC, 10, 4842

# Benchmarking Density functionals

**Pedro Alexandrino Fernandes,**

Dep. Chemistry & Biochemistry, University of Porto, Portugal

pedro.fernandes@fc.up.pt



# Benchmarking Density Functionals

What about the QM/MM interaction?

# Benchmarking Density Functionals

What about the QM/MM interaction?

Mechanical  
Embedding

Electrostatic  
Embedding

# Benchmarking Density Functionals

What about the QM/MM interaction?

Mechanical  
Embedding

Electrostatic  
Embedding



LJ Dispersion  
Classical electrostatics

# Benchmarking Density Functionals

What about the QM/MM interaction?

Mechanical  
Embedding

Electrostatic  
Embedding



LJ Dispersion  
Classical electrostatics



How reliable are the  
“QM” point charges?

# Benchmarking Density Functionals

What about the QM/MM interaction?

Mechanical  
Embedding



LJ Dispersion  
Classical electrostatics



How reliable are the  
“QM” point charges?

Electrostatic  
Embedding



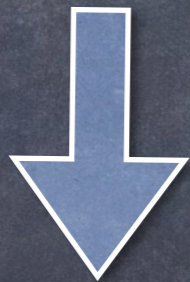
LJ Dispersion  
MM charges interact with  $\rho_e(\mathbf{r})$



# Benchmarking Density Functionals

What about the QM/MM interaction?

Mechanical  
Embedding



LJ Dispersion  
Classical electrostatics



How reliable are the  
“QM” point charges?

Electrostatic  
Embedding

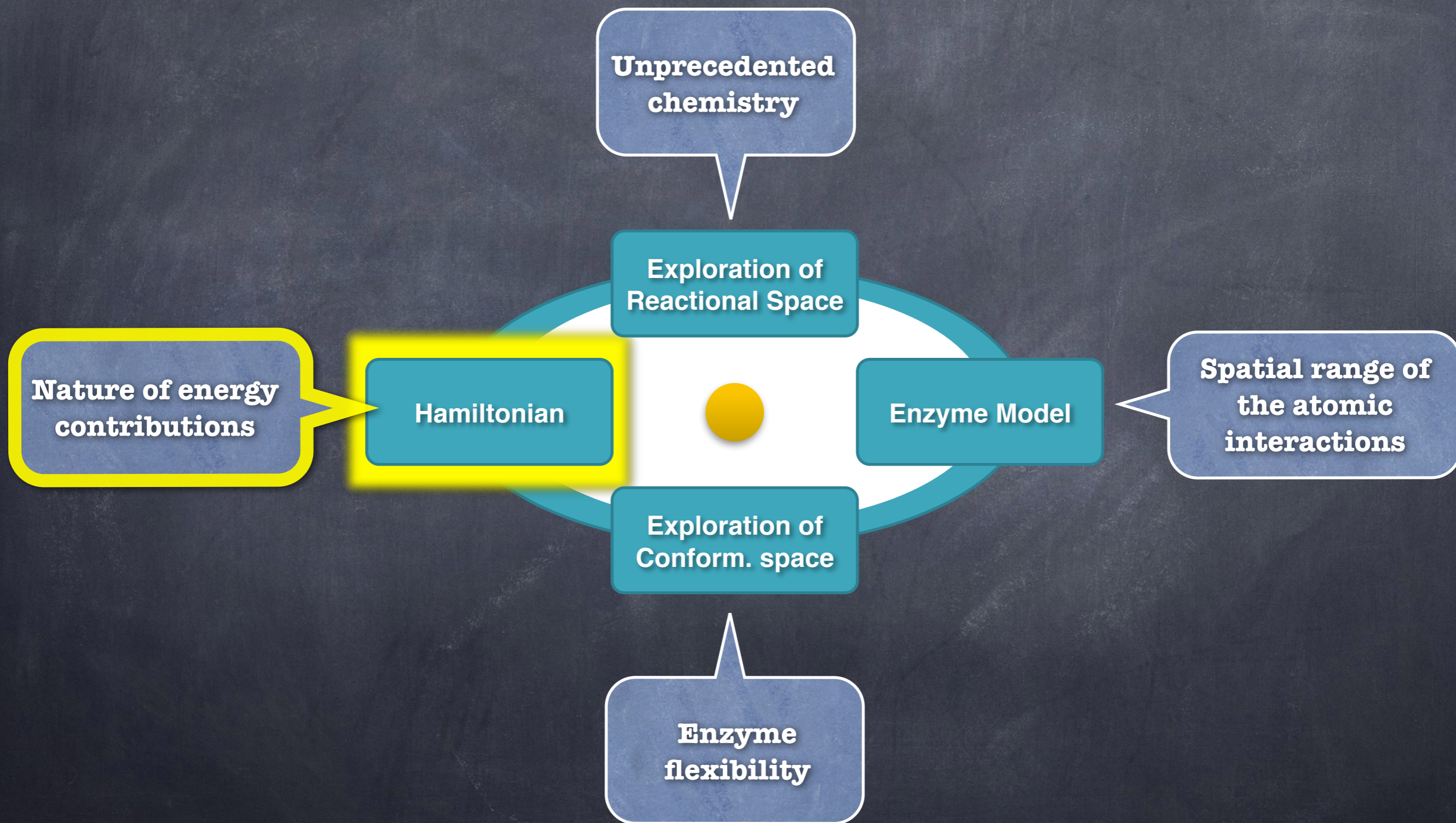


LJ Dispersion  
MM charges interact with  $\rho_e(\mathbf{r})$

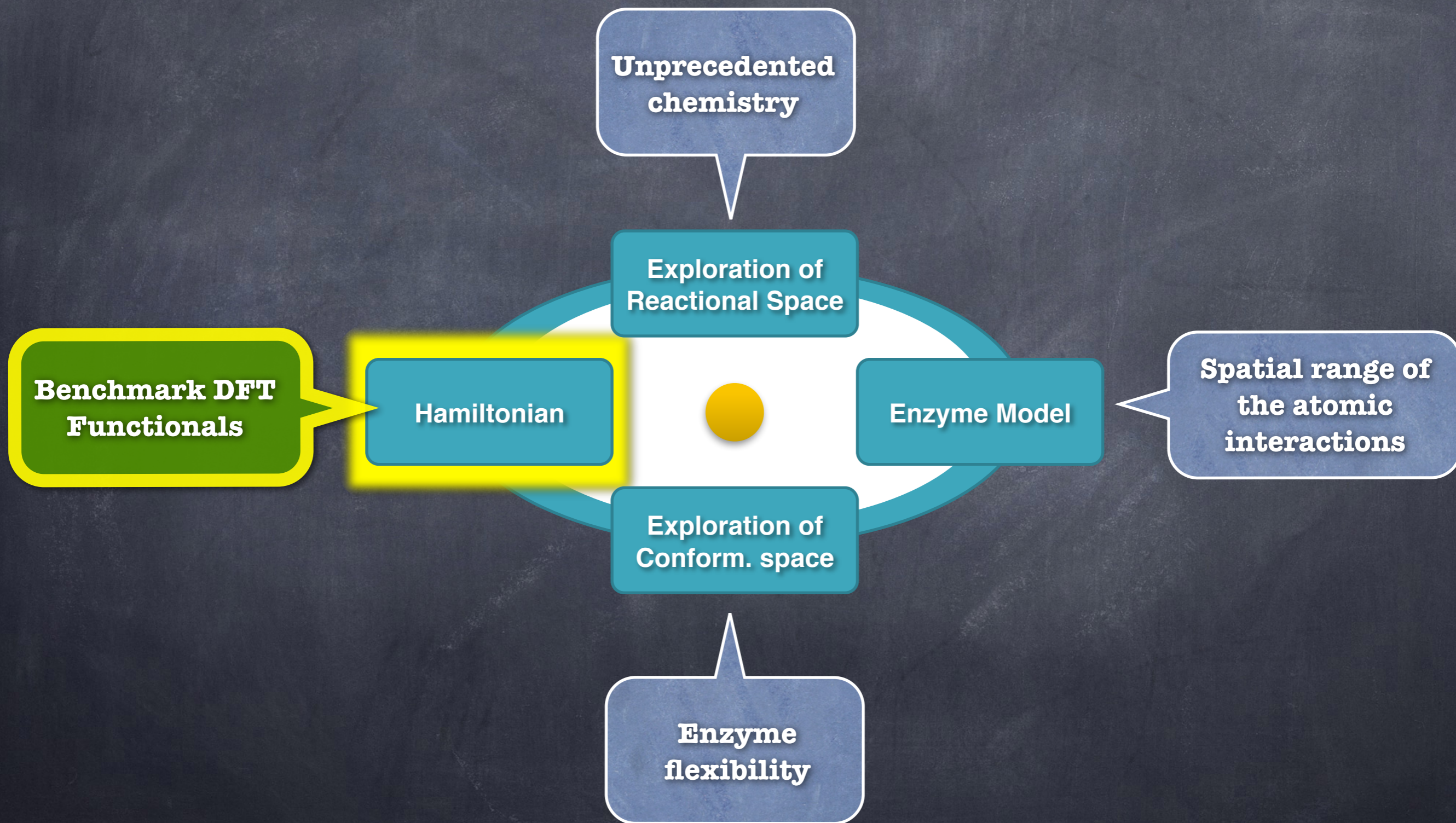


>>Time consuming!

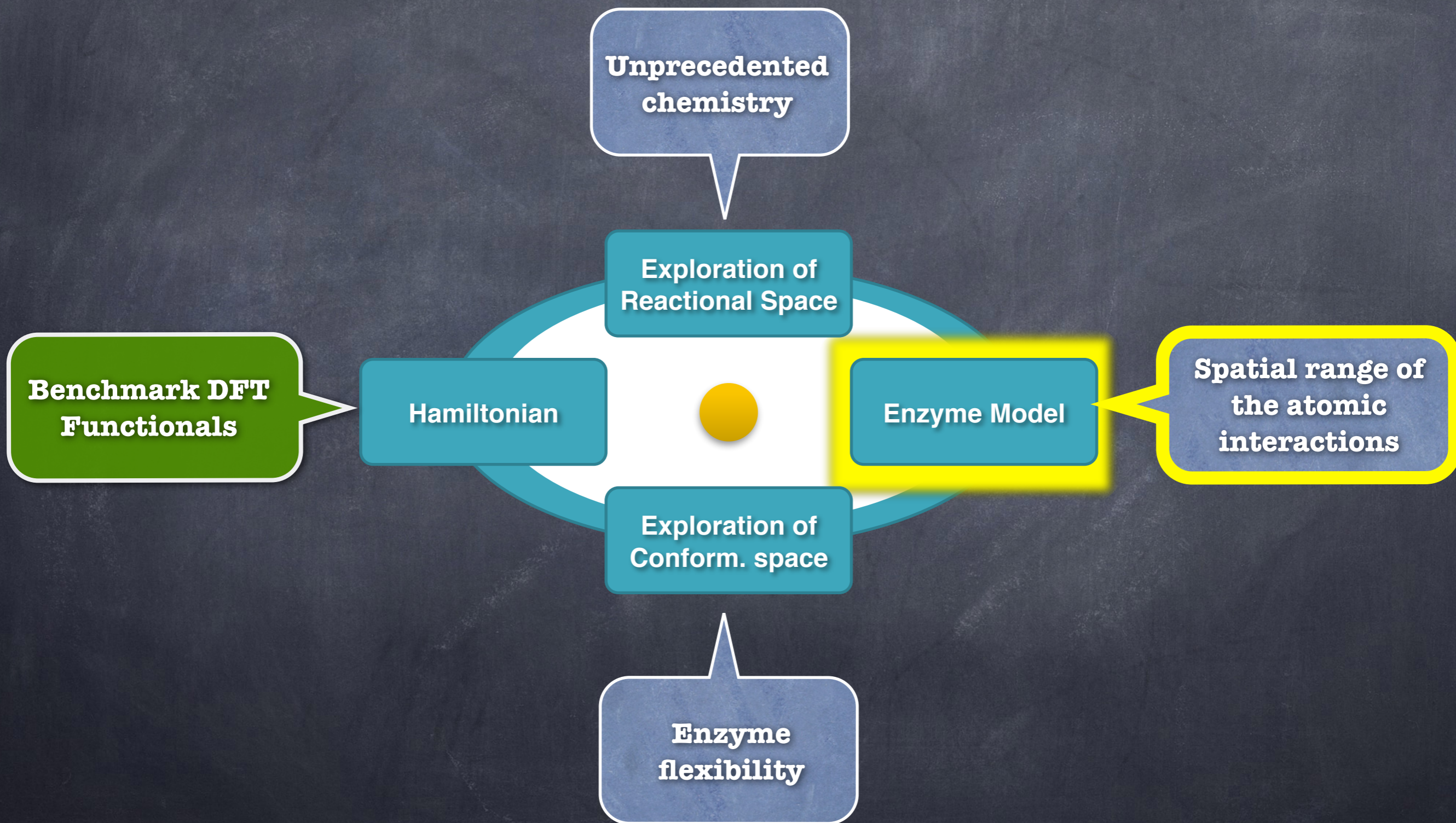
# Solutions for the problem



# Solutions for the problem

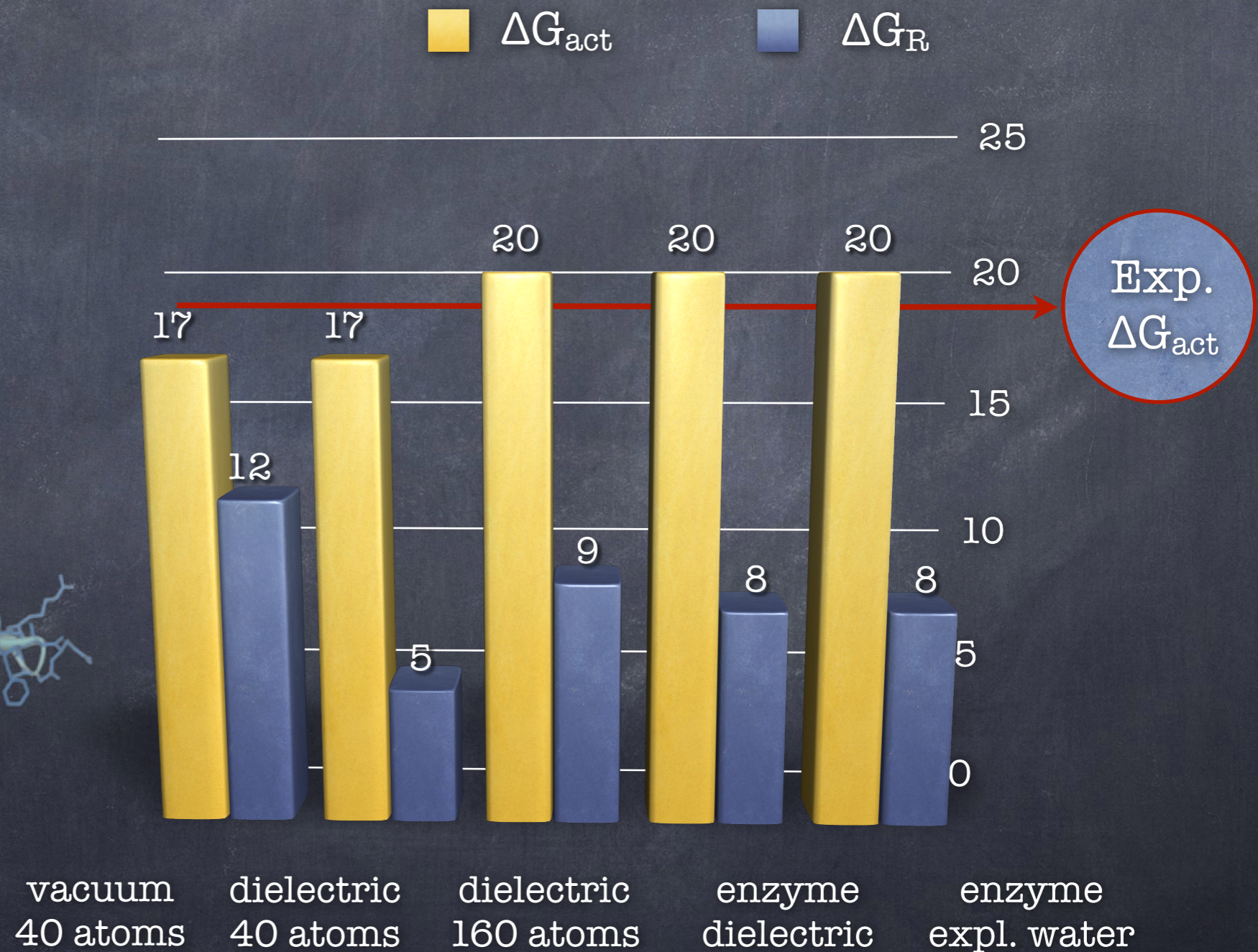
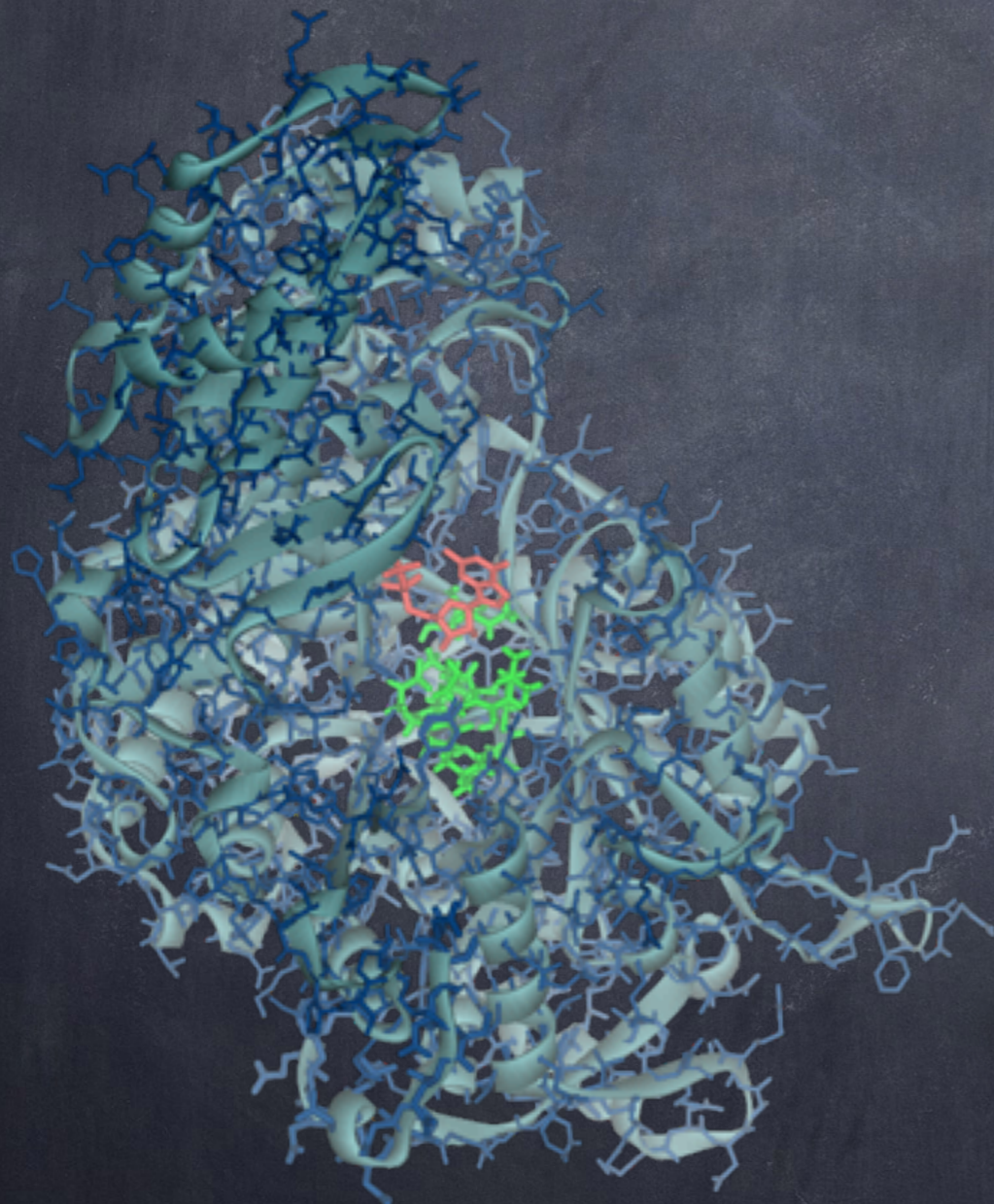


# Solutions for the problem



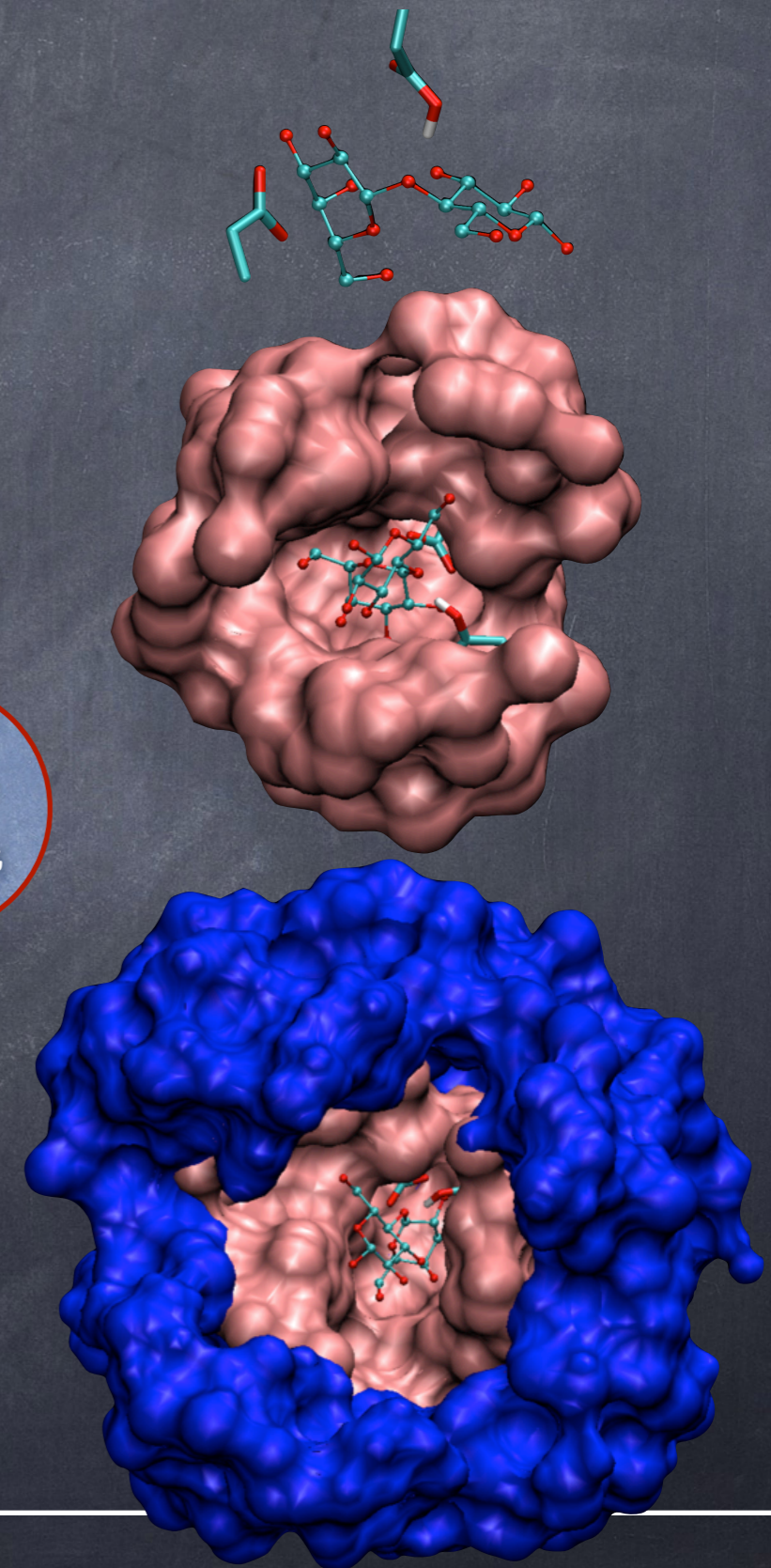
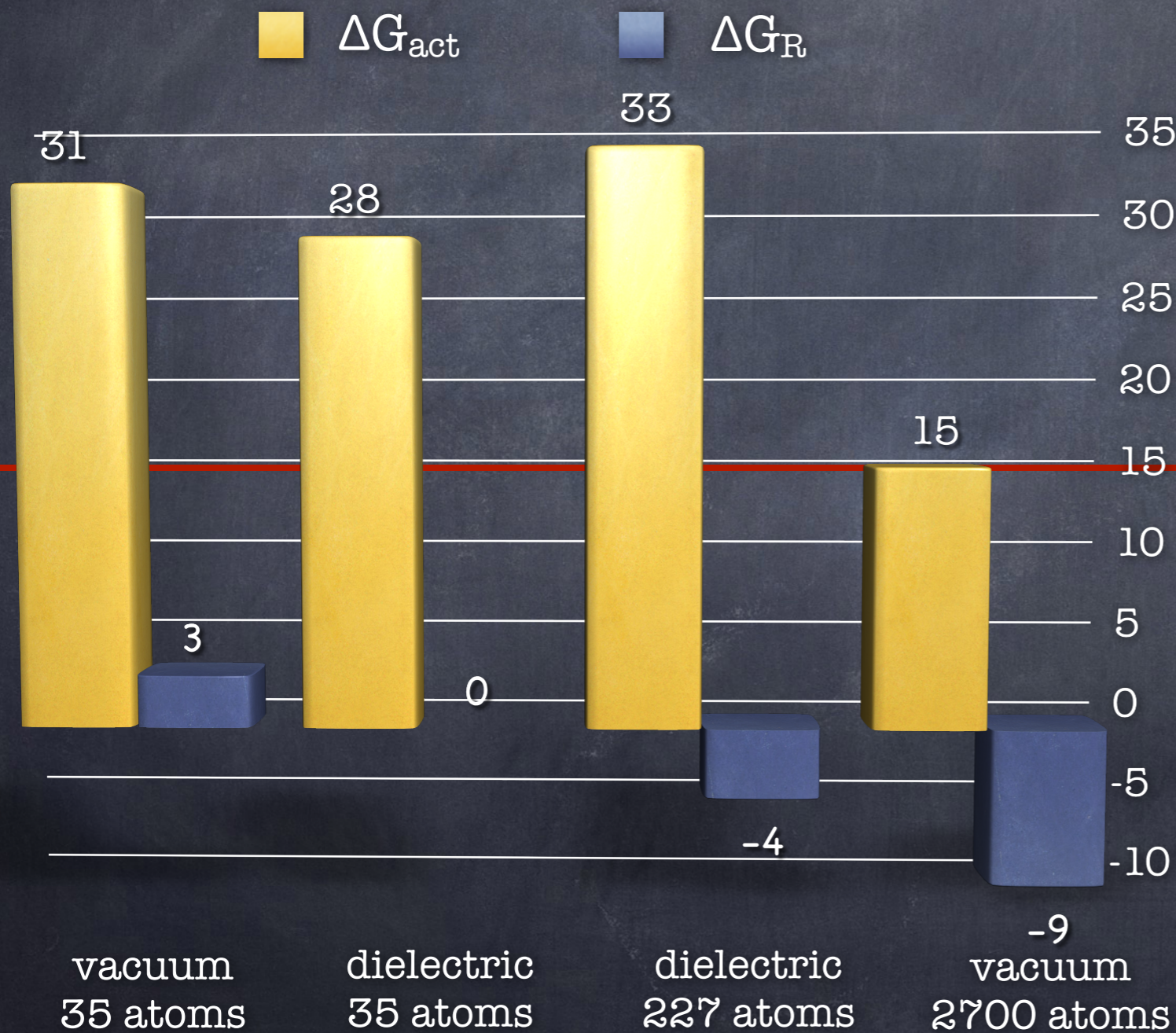
# The Molecular Models of Enzymes

## Ribonucleotide Reductase



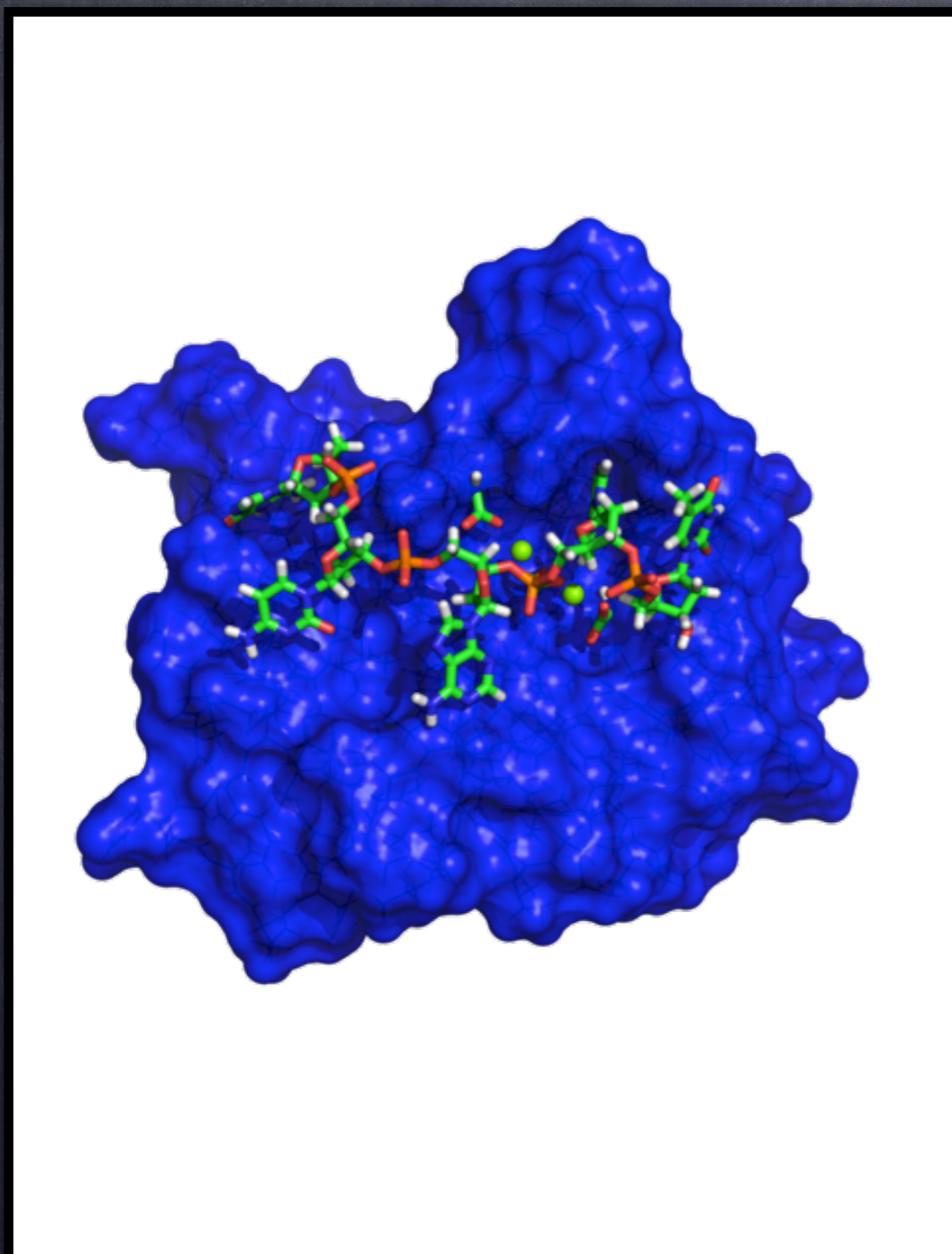
# The Molecular Models of Enzymes

## $\beta$ -Galactosidase



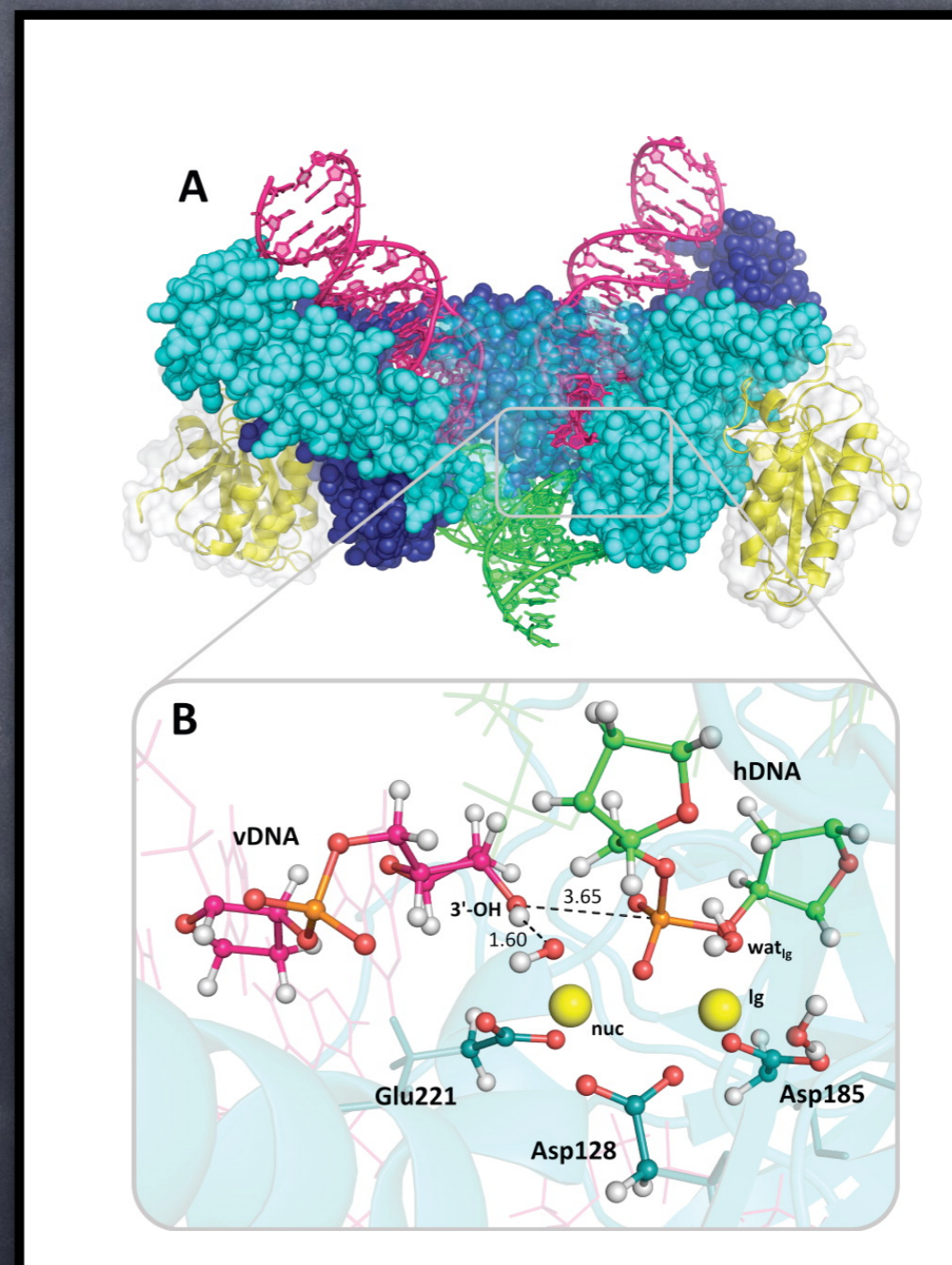
# The Molecular Models of Enzymes

3' End Processing  
>2.500 atoms



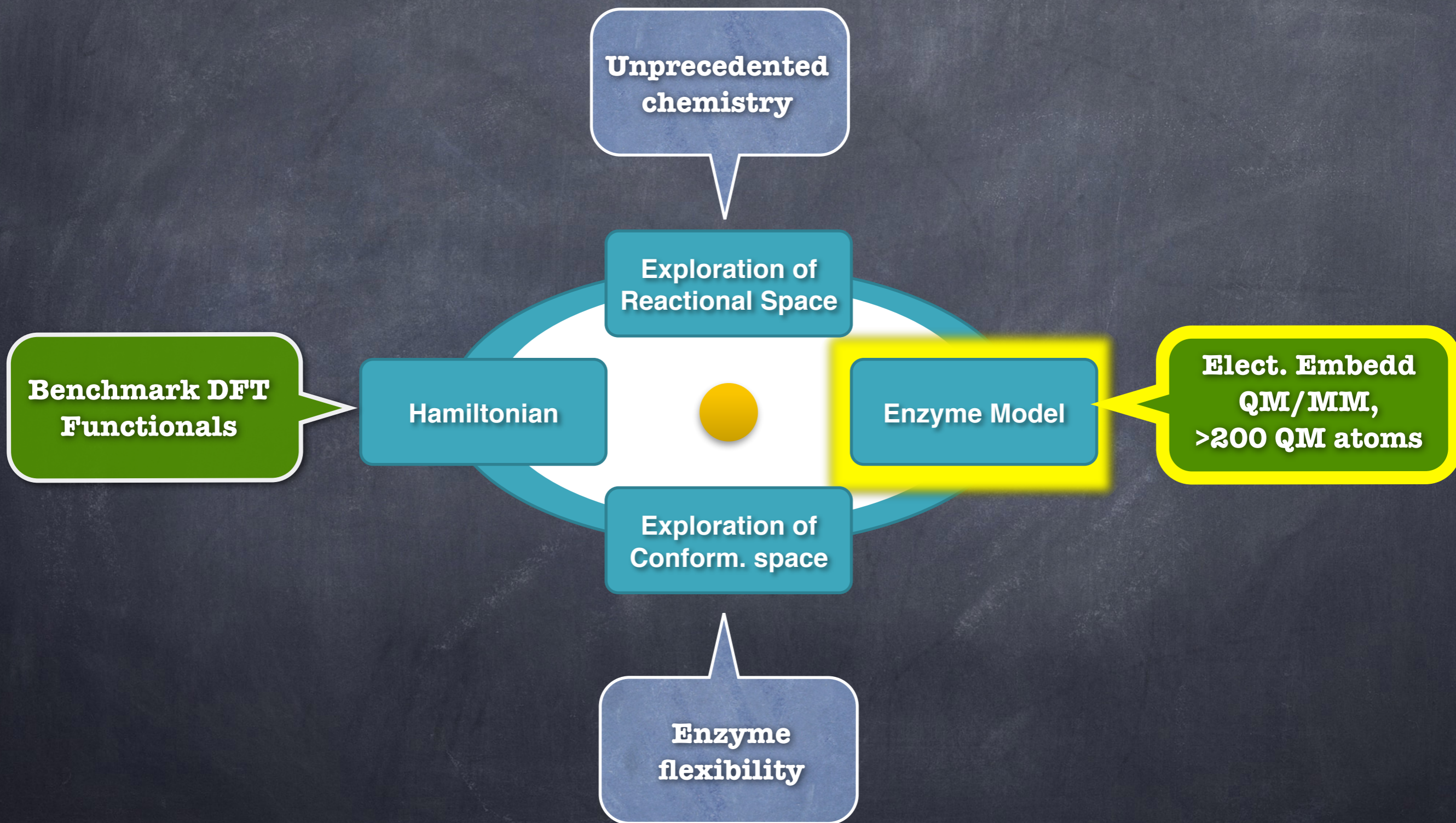
JACS, 134, 13436, 2012

Strand Transfer  
>22.000 atoms



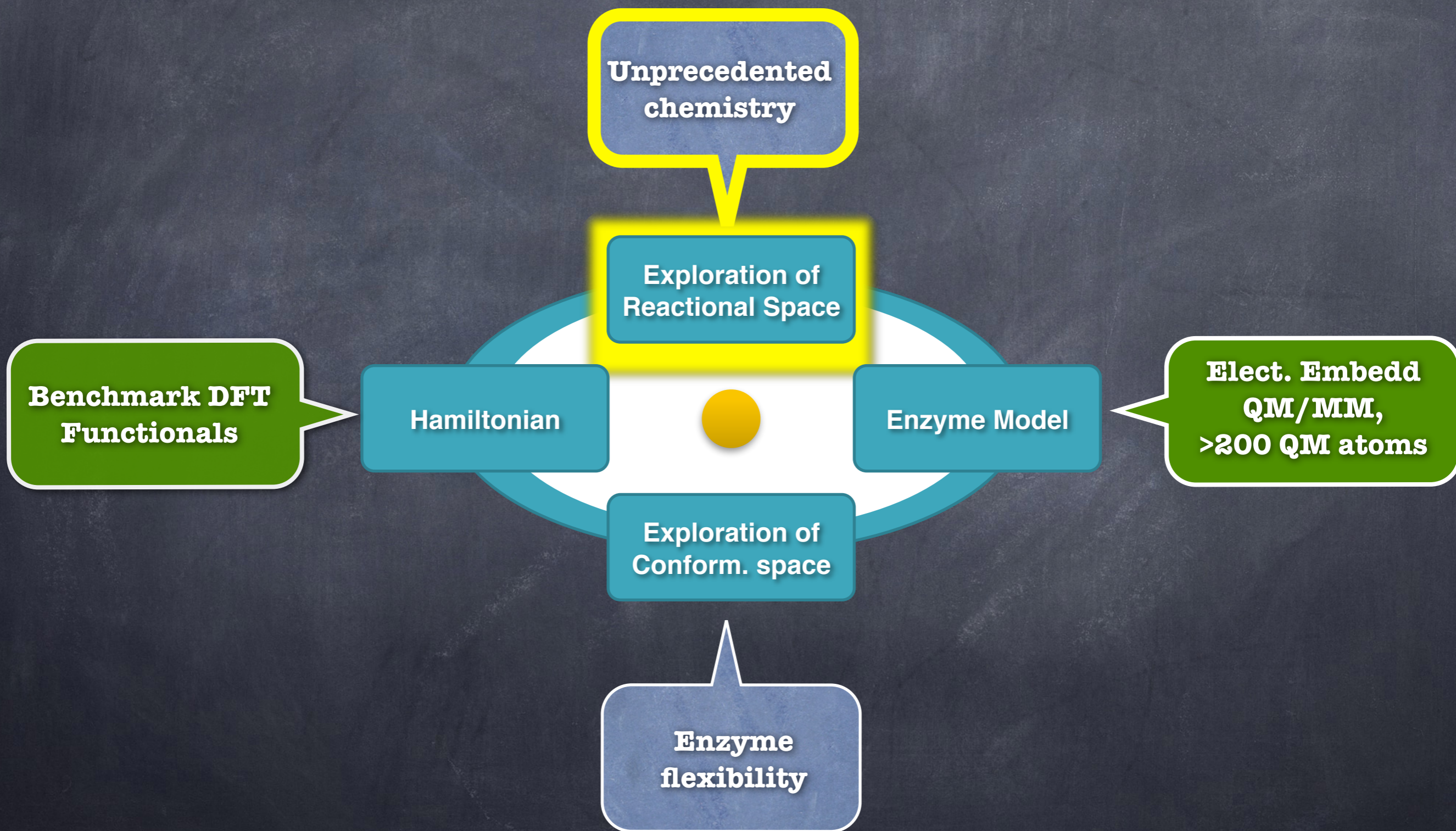
JCTC, 10, 5458, 2014

# Solutions for the problem



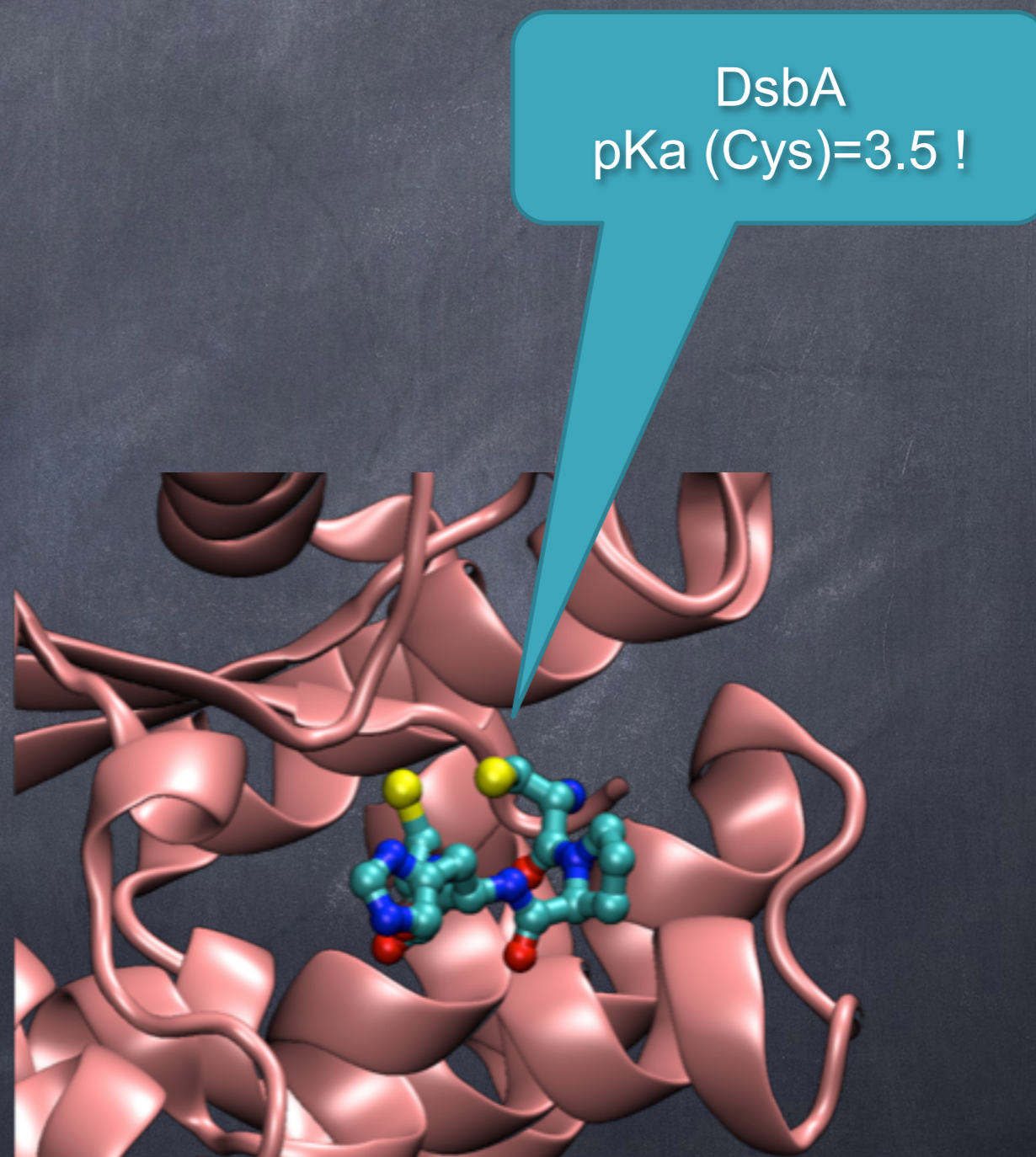


# Solutions for the problem



# Unprecedented Chemistry

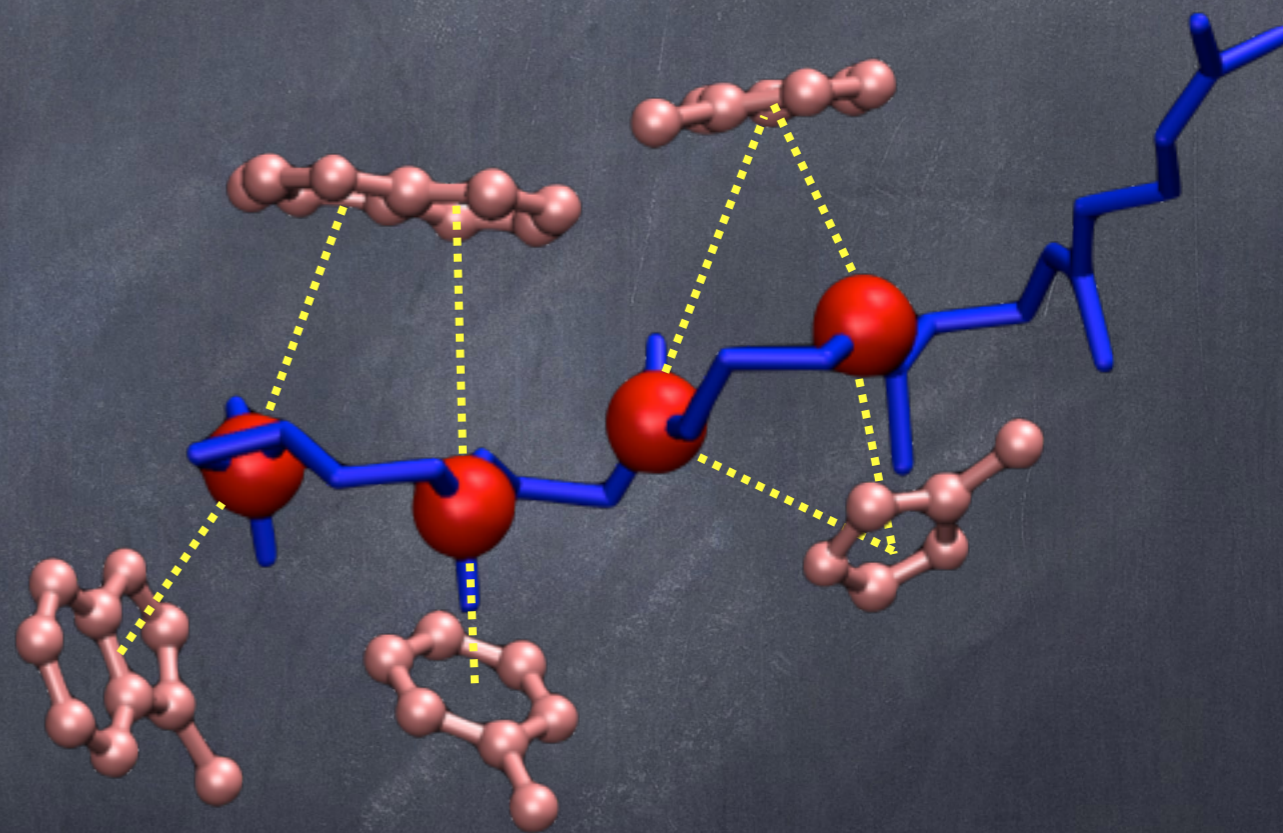
- Abnormal pKas



# Unprecedented Chemistry

## Oxidosqualene cyclase

- Abnormal pKas
- Stable carbocations

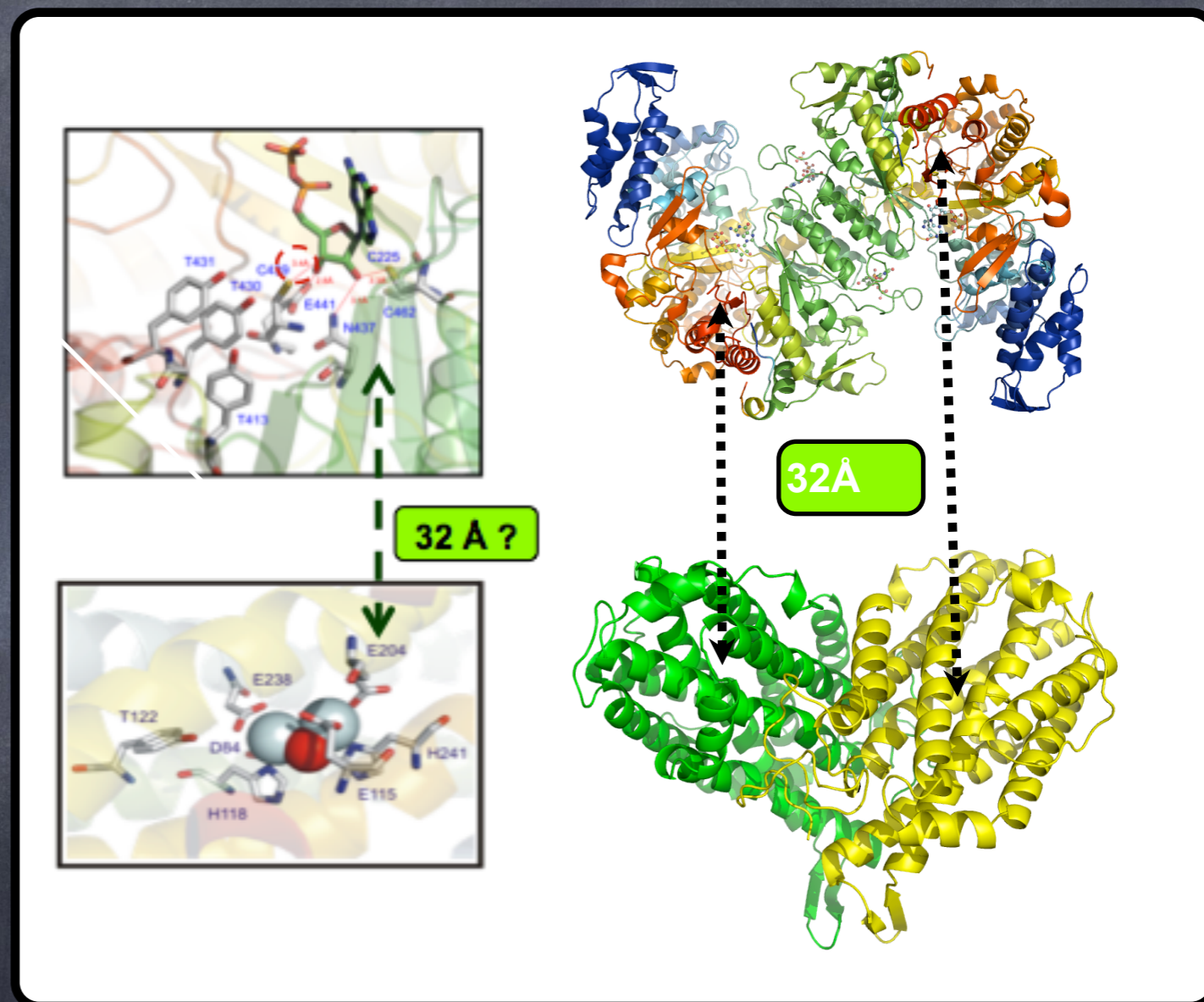


•  $\pi$ -cation interactions stabilize rate-limiting steps

# Unprecedented Chemistry

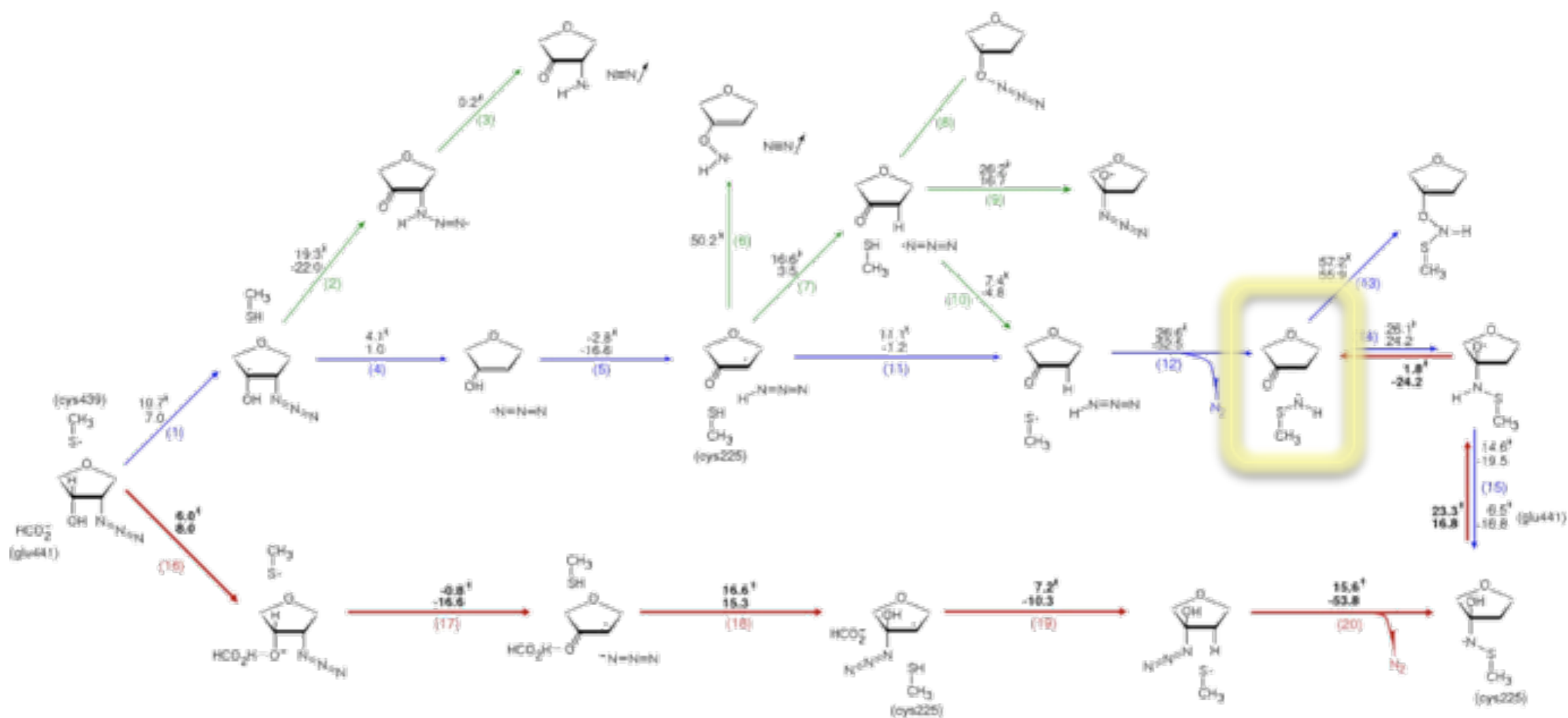
## Ribonucleotide Reductase

- Abnormal pKas
- Stable carbocations
- Long living aliphatic radicals

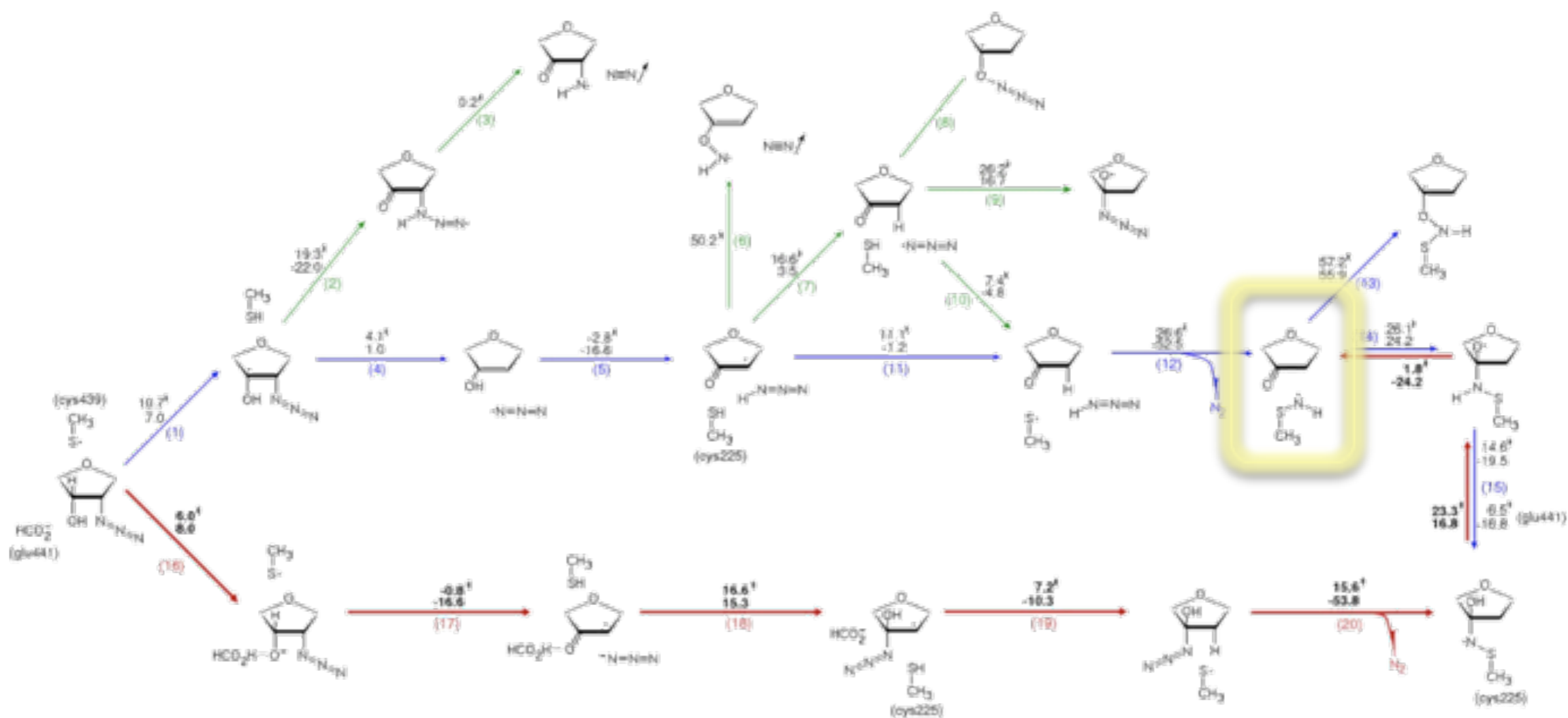


Dehydration through radical mechanism

# The diversity of the Chemical Space

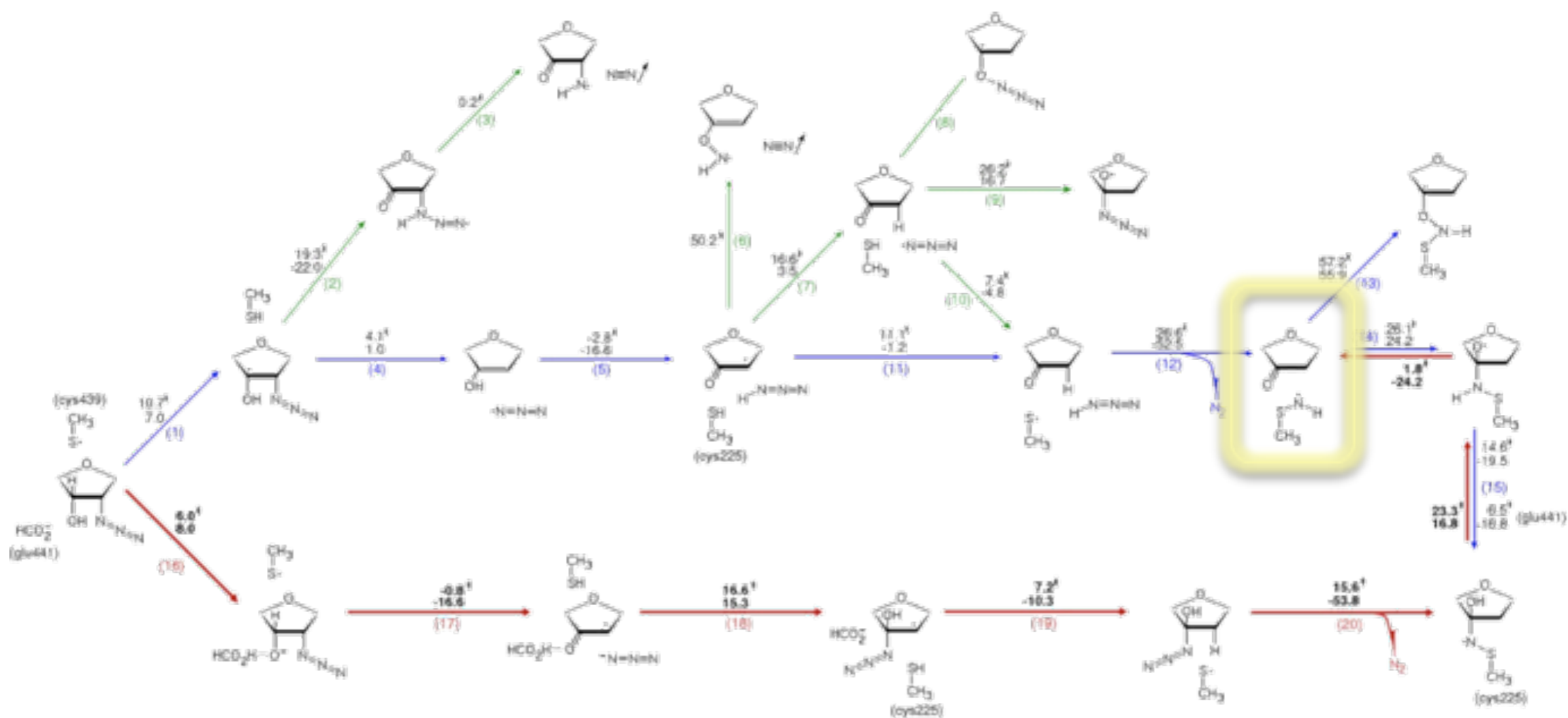


# The diversity of the Chemical Space



Systematic testing of mechanistic hypotheses

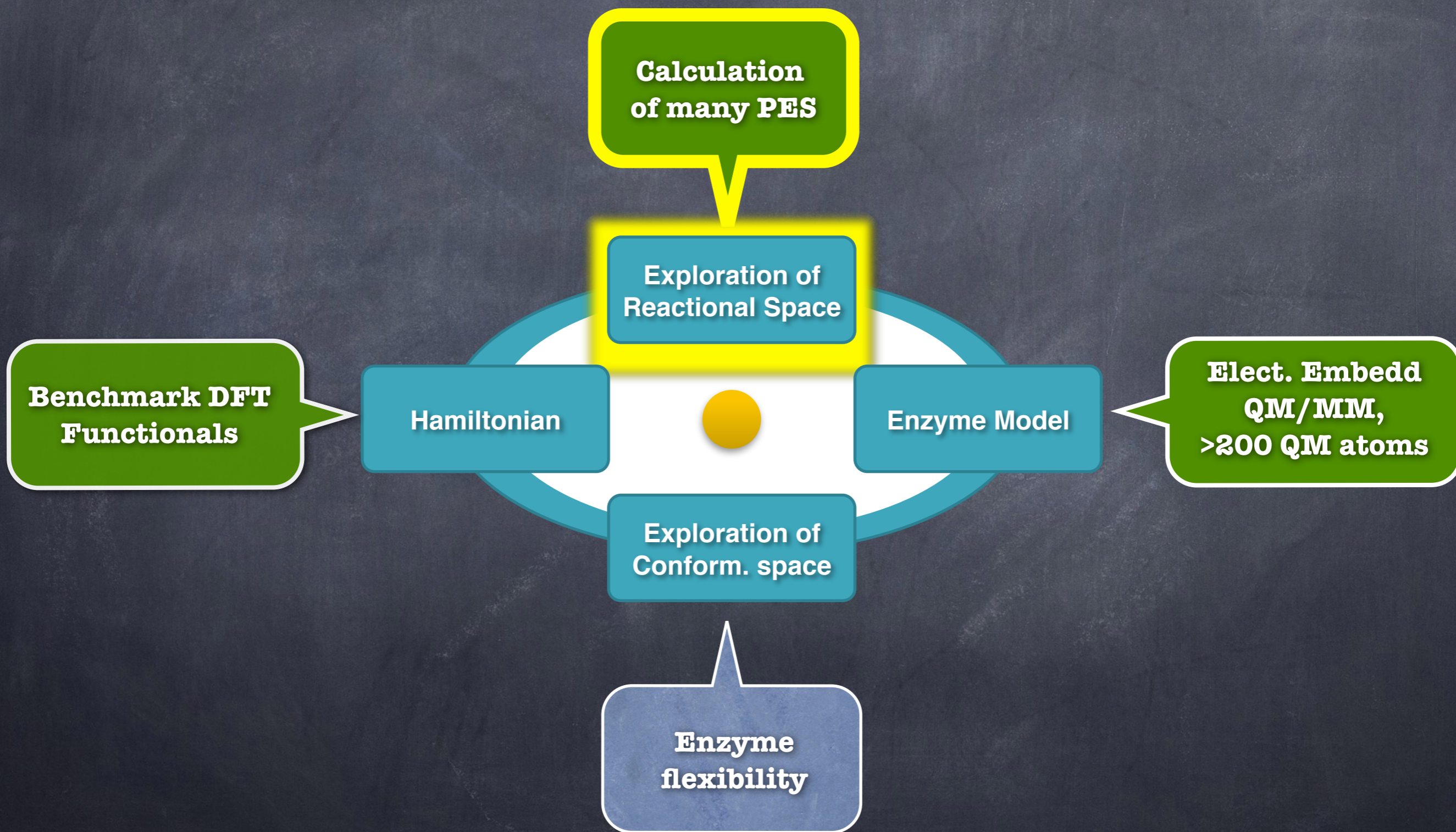
# The diversity of the Chemical Space



Systematic testing of mechanistic hypotheses

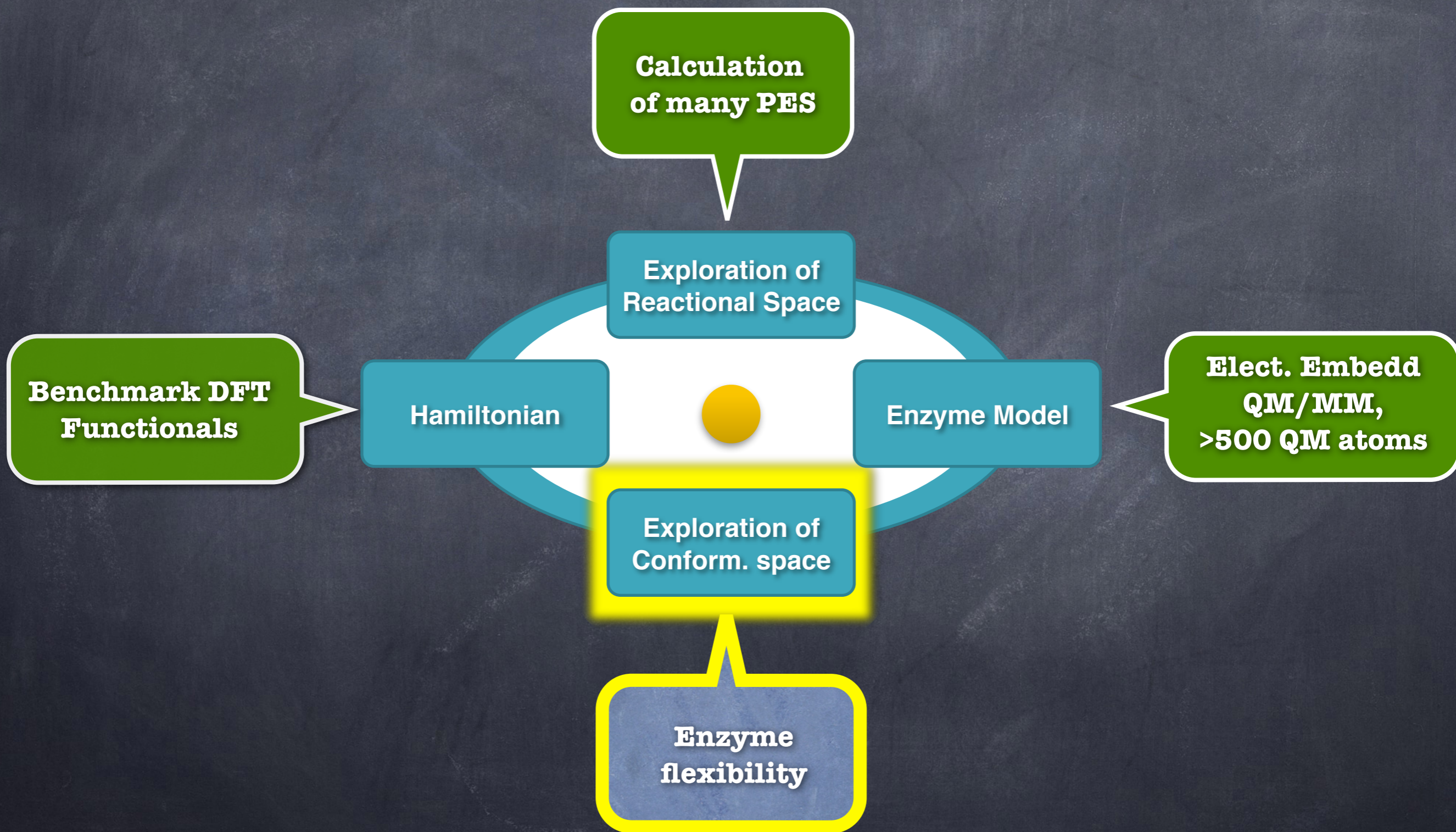
Cross-check with all available exp data

# Solutions for the problem





# Solutions for the problem



# Enzyme conformations

**Two  $\neq$  problems**

# Enzyme conformations

**Two  $\neq$  problems**



Generating relevant  
uncorrelated conformations

# Enzyme conformations

Two  $\neq$  problems



Opt=CPMD=QMMM MD  
CMD better

Generating relevant  
uncorrelated conformations

Other methods may  
unfold the enzyme

# Enzyme conformations

Two  $\neq$  problems



Generating relevant  
uncorrelated conformations

Opt=CPMD=QMMM MD  
CMD better

Other methods may  
unfold the enzyme



Averaging conformations  
or  
checking conformation effect

# Enzyme conformations

## Two ≠ problems



Generating relevant  
uncorrelated conformations

Opt=CPMD=QMMM MD  
CMD better

Other methods may  
unfold the enzyme



Averaging conformations  
or  
checking conformation effect

ab initio/DFT  
only checks effects

SE methods  
average within their  
accuracy

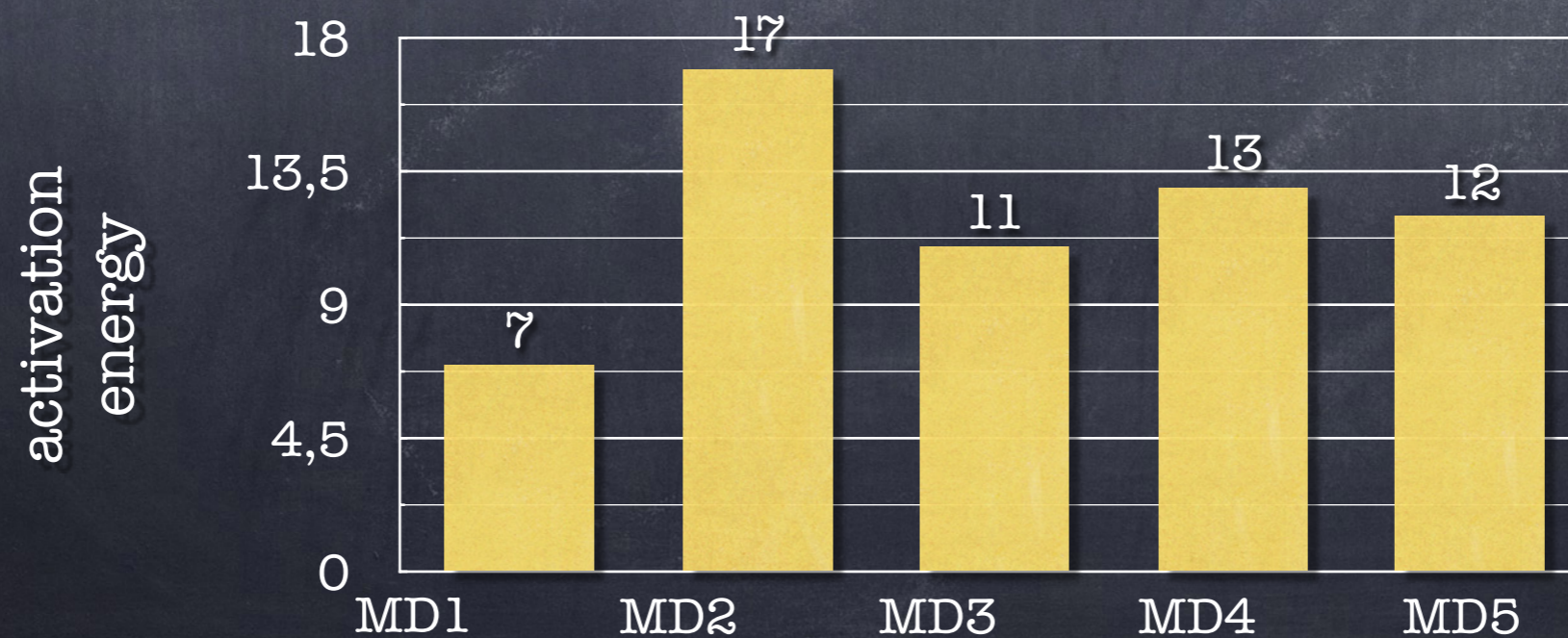
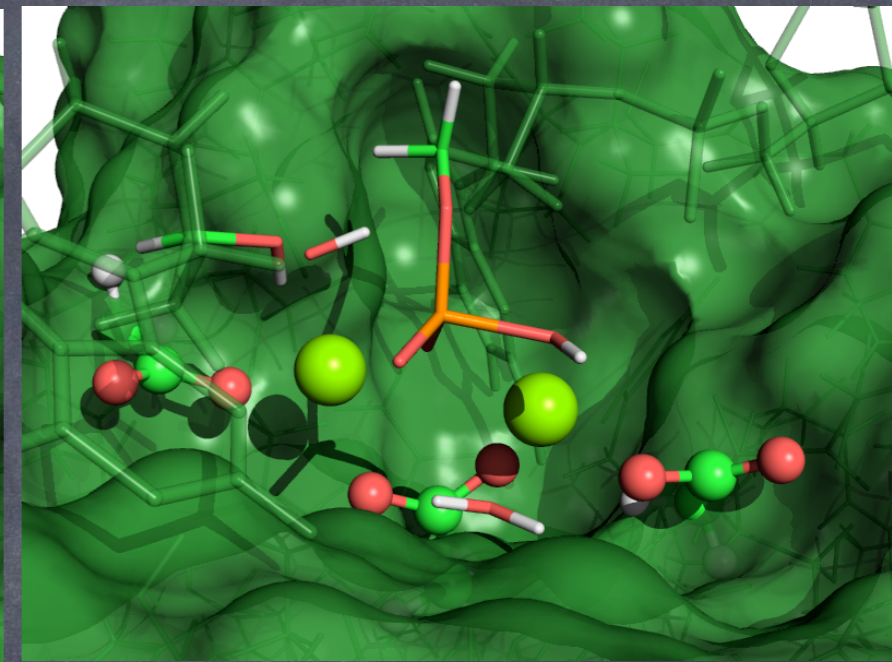
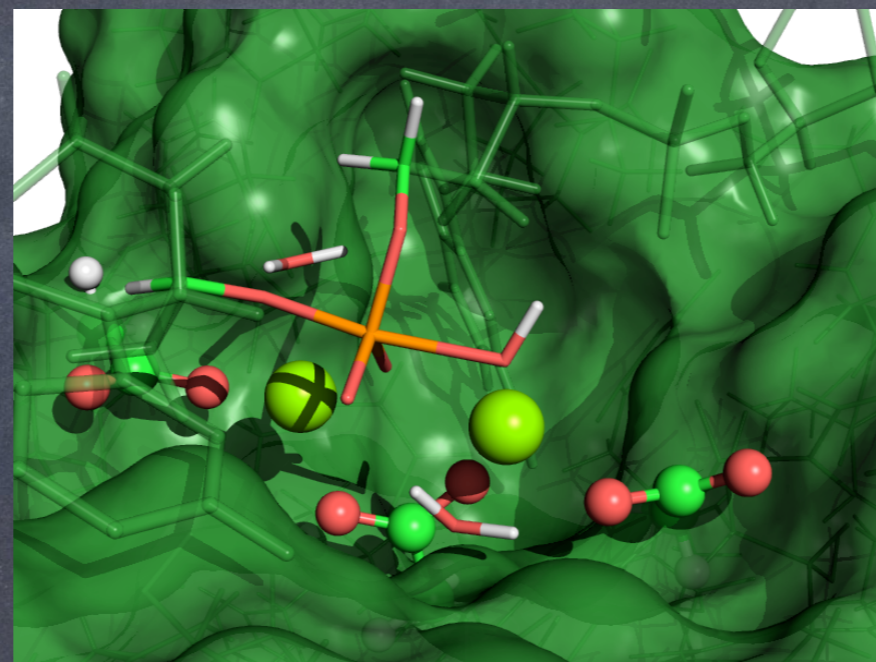
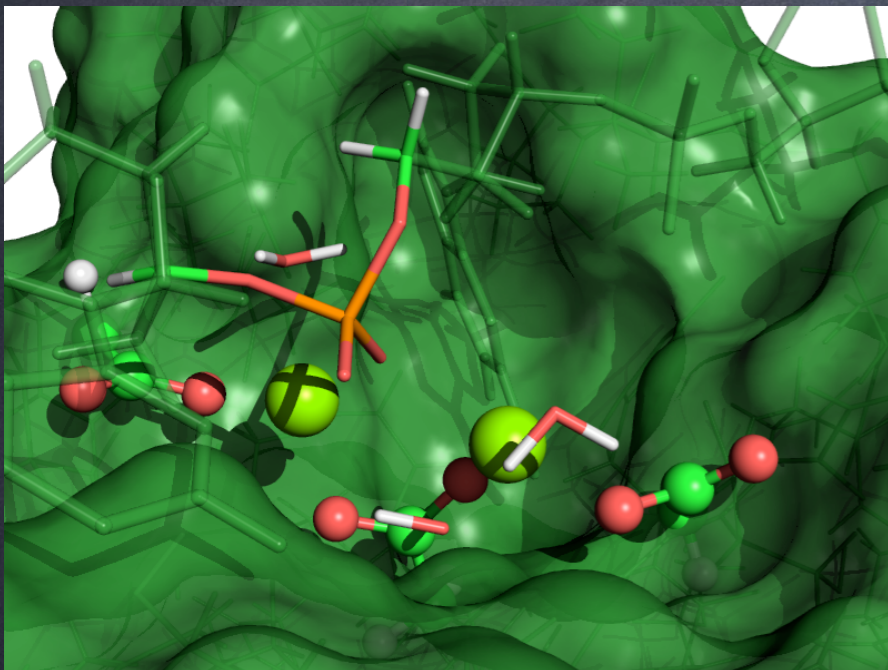
# Enzyme conformations

Hydrolysis of phosphodiester bonds

Reactants

TS

Products

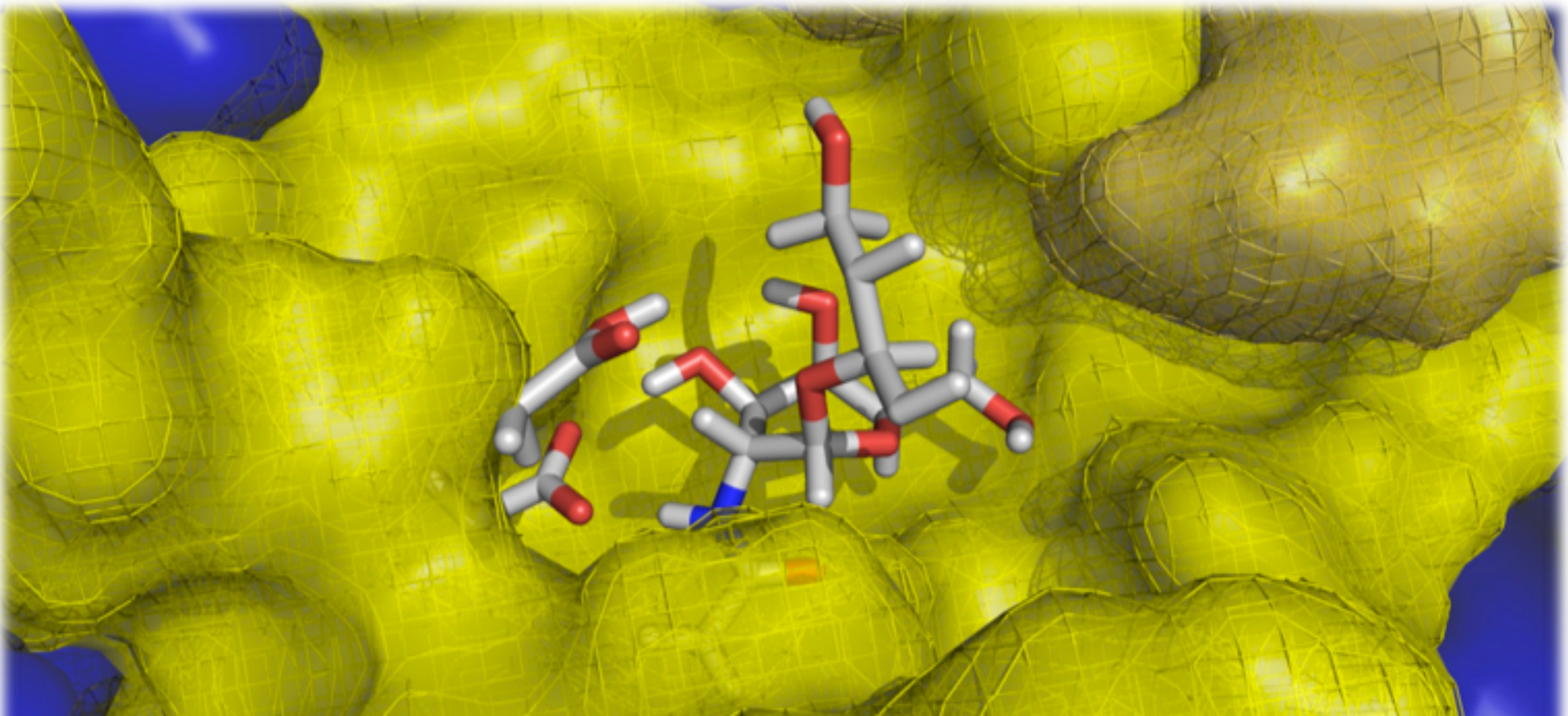


# Influence of Enzyme Conformation on Kinetics

**Pedro Alexandrino Fernandes,**

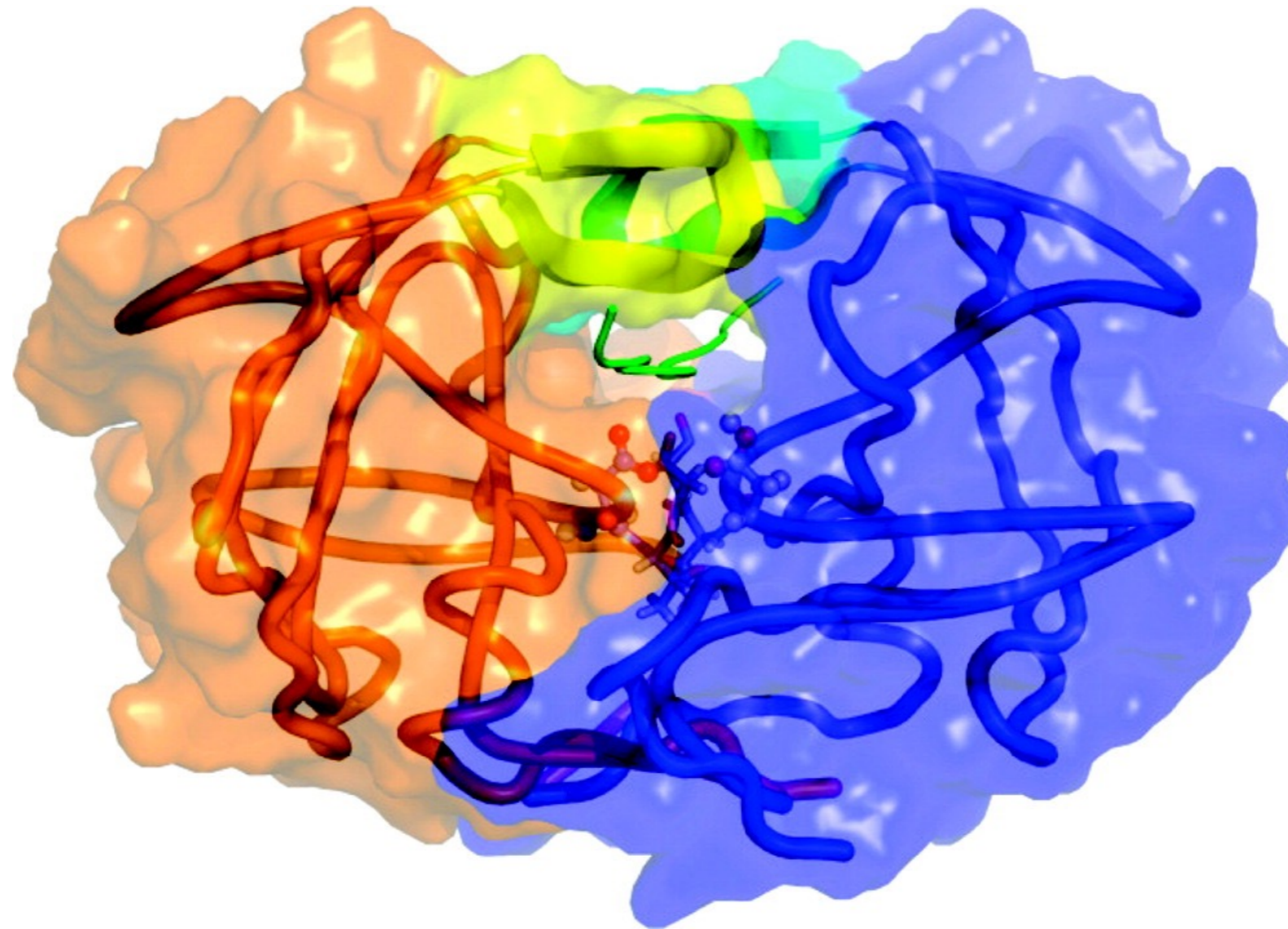
Dep. Chemistry & Biochemistry, University of Porto, Portugal

pedro.fernandes@fc.up.pt





# Example: HIV-1 Protease



# Example: HIV-1 Protease



QM/MM model.

QM layer: 30 atoms.

MM layer: 3232 atoms.

Electrostatic embedding in all calculations.

# Example: HIV-1 Protease

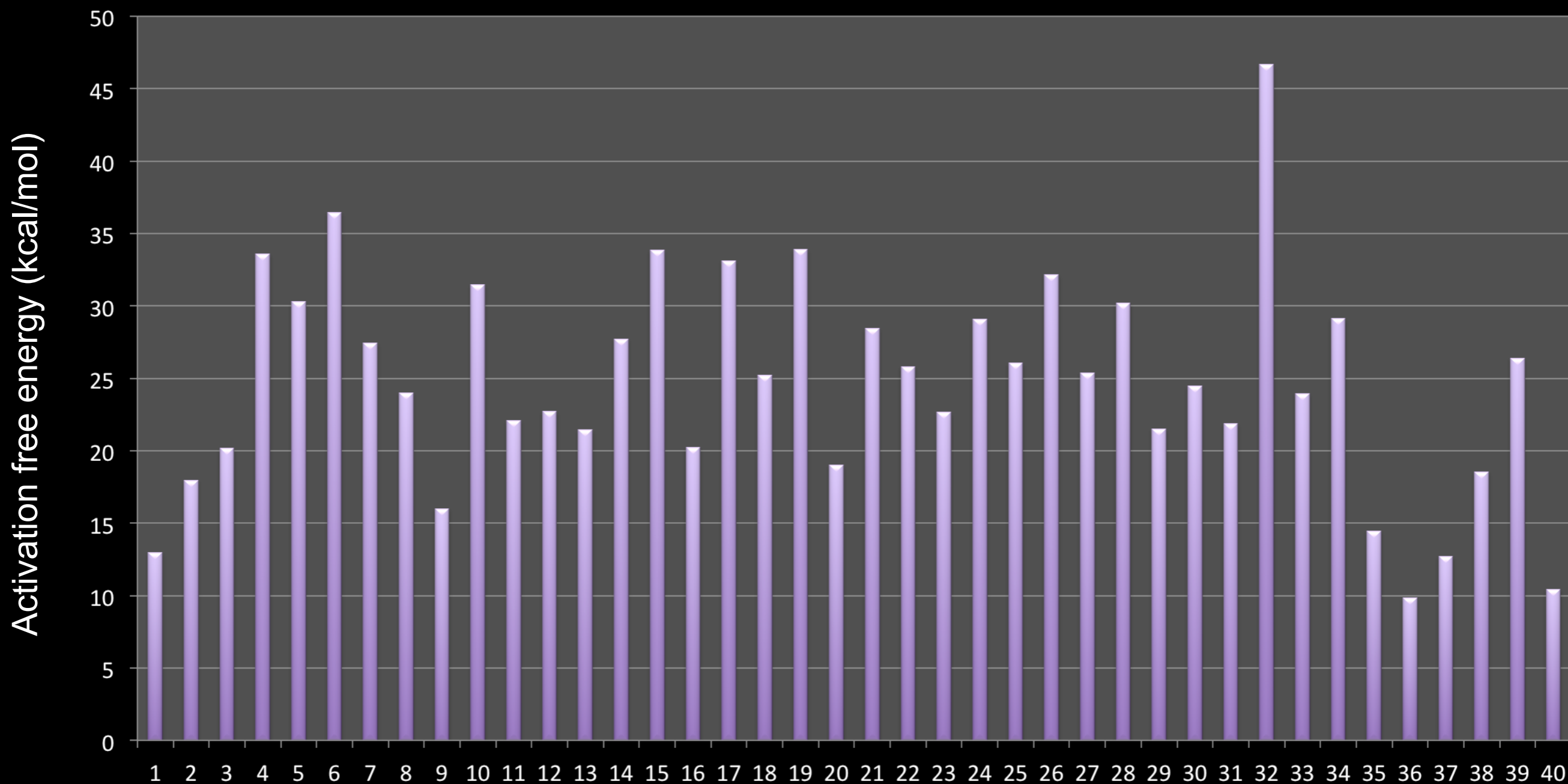


QM/MM model.  
QM layer: 30 atoms.  
MM layer: 3232 atoms.  
Electrostatic embedding in all calculations.

Extract coordinates every 2 ns of several MD simulations.  
Optimize the reactants state.  
Scan the distance between the nucleophile and the amide carbon.  
Optimize the maximum to a TS, without any constraints.  
Frequency calculation.  
IRC to identify the reactants.  
Optimize the reactants again.

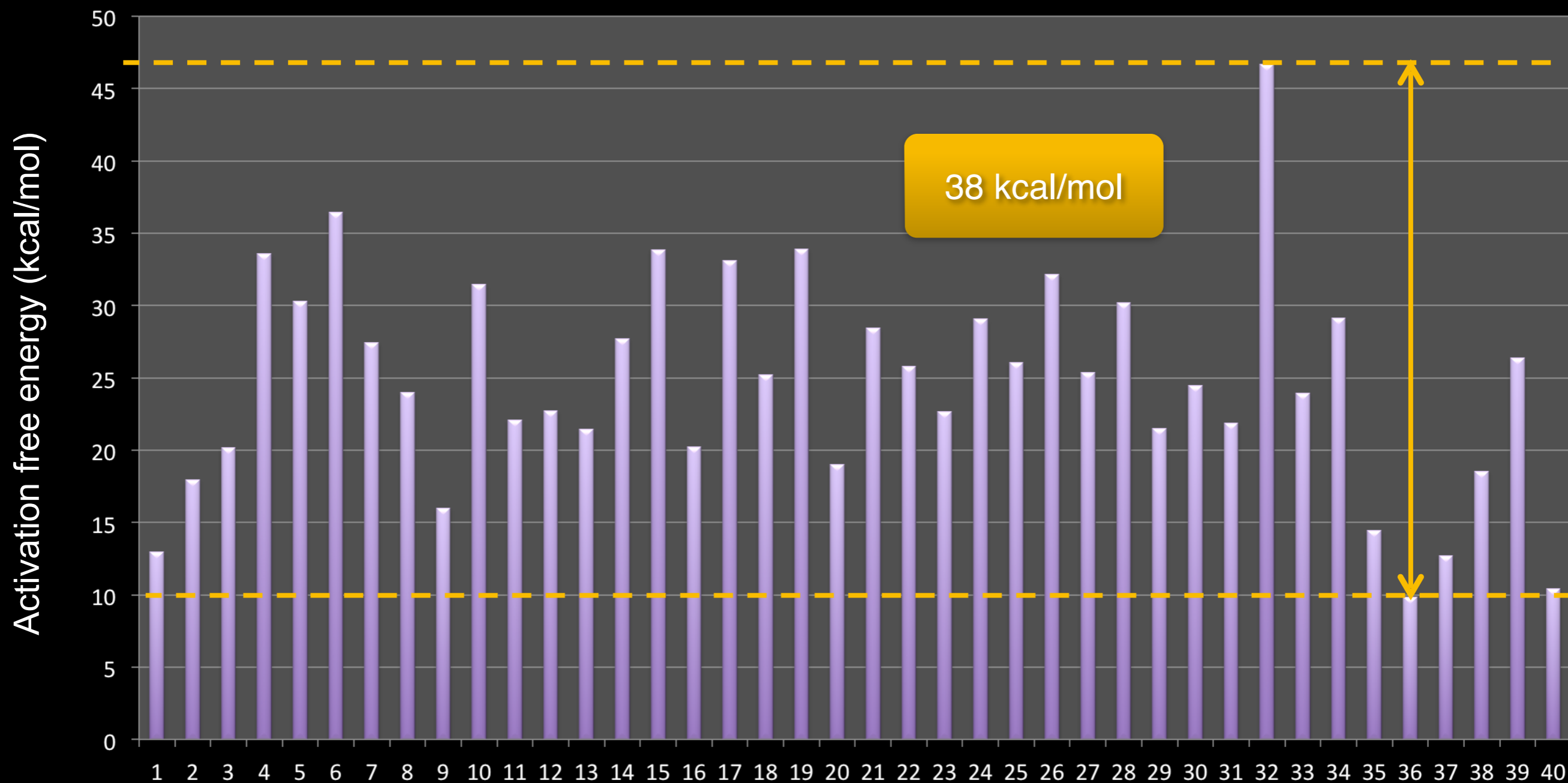
# Influence of Enzyme Conformation on Kinetics

Activation free energy for 40 different enzyme conformations



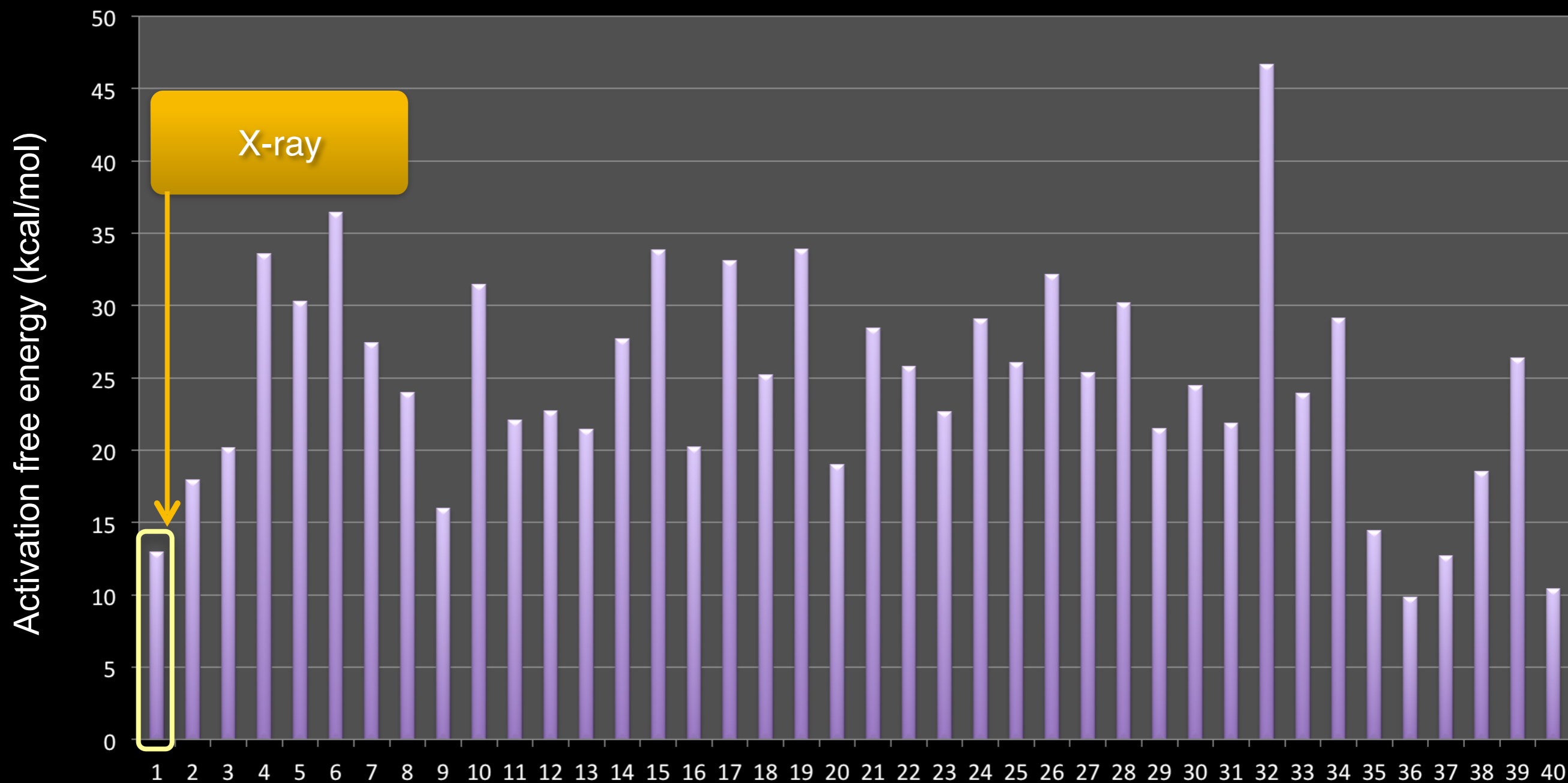
# Influence of Enzyme Conformation on Kinetics

Activation free energy for 40 different enzyme conformations



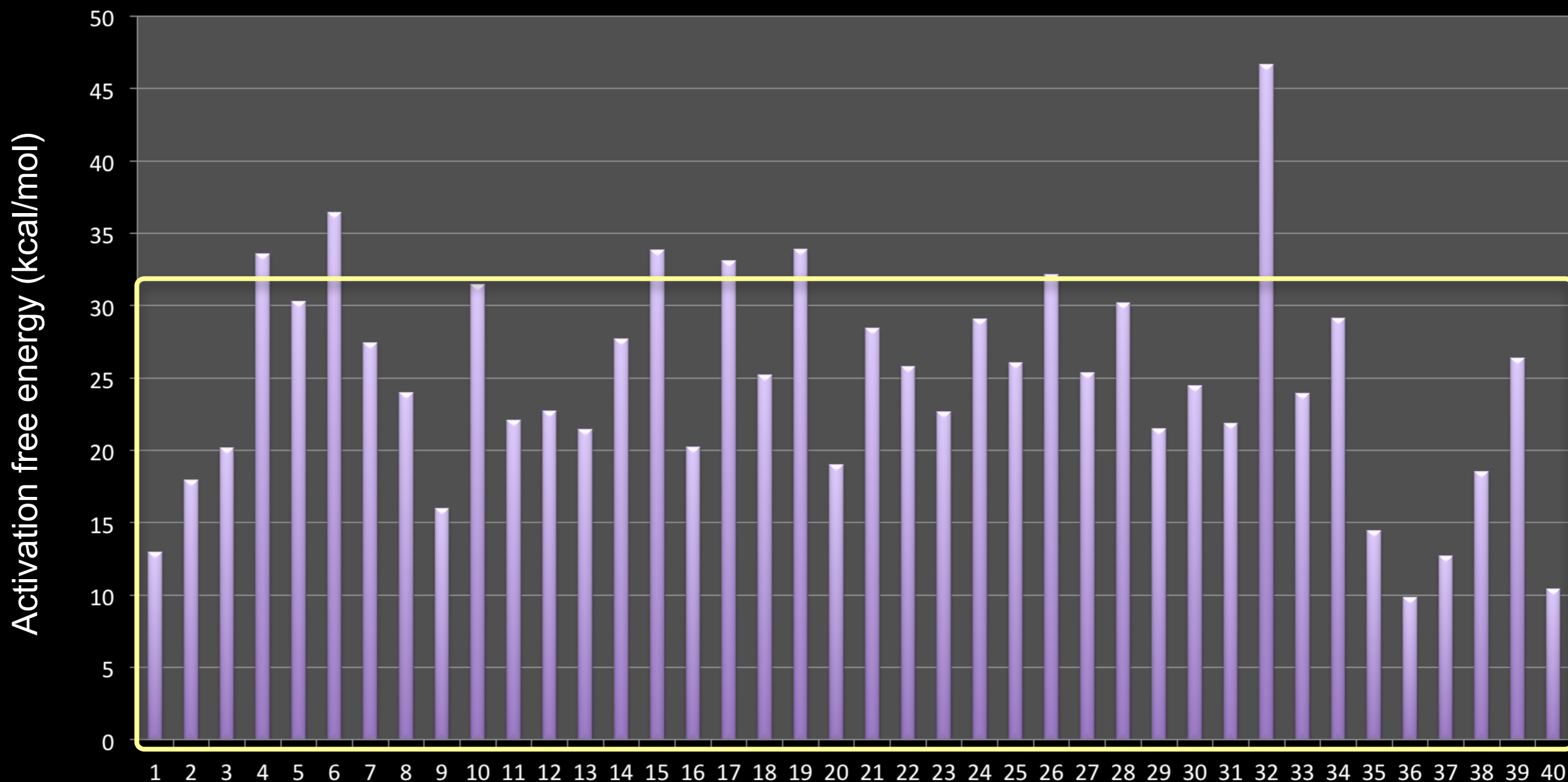
# Influence of Enzyme Conformation on Kinetics

$$\Delta G_{\text{X-ray}} = 13.4 \text{ kcal/mol} < \Delta G > = 9 \text{ kcal/mol}$$



# Influence of Enzyme Conformation on Kinetics

$\Delta G > \sim 34$  kcal/mol obtained with a different chemical mechanism



# Why $\Delta G_{\text{act}}$ fluctuates so much?



Origin of the fluctuations?



# Why $\Delta G_{\text{act}}$ fluctuates so much?



Origin of the fluctuations?



Differences at the active site ?

# Why $\Delta G_{\text{act}}$ fluctuates so much?



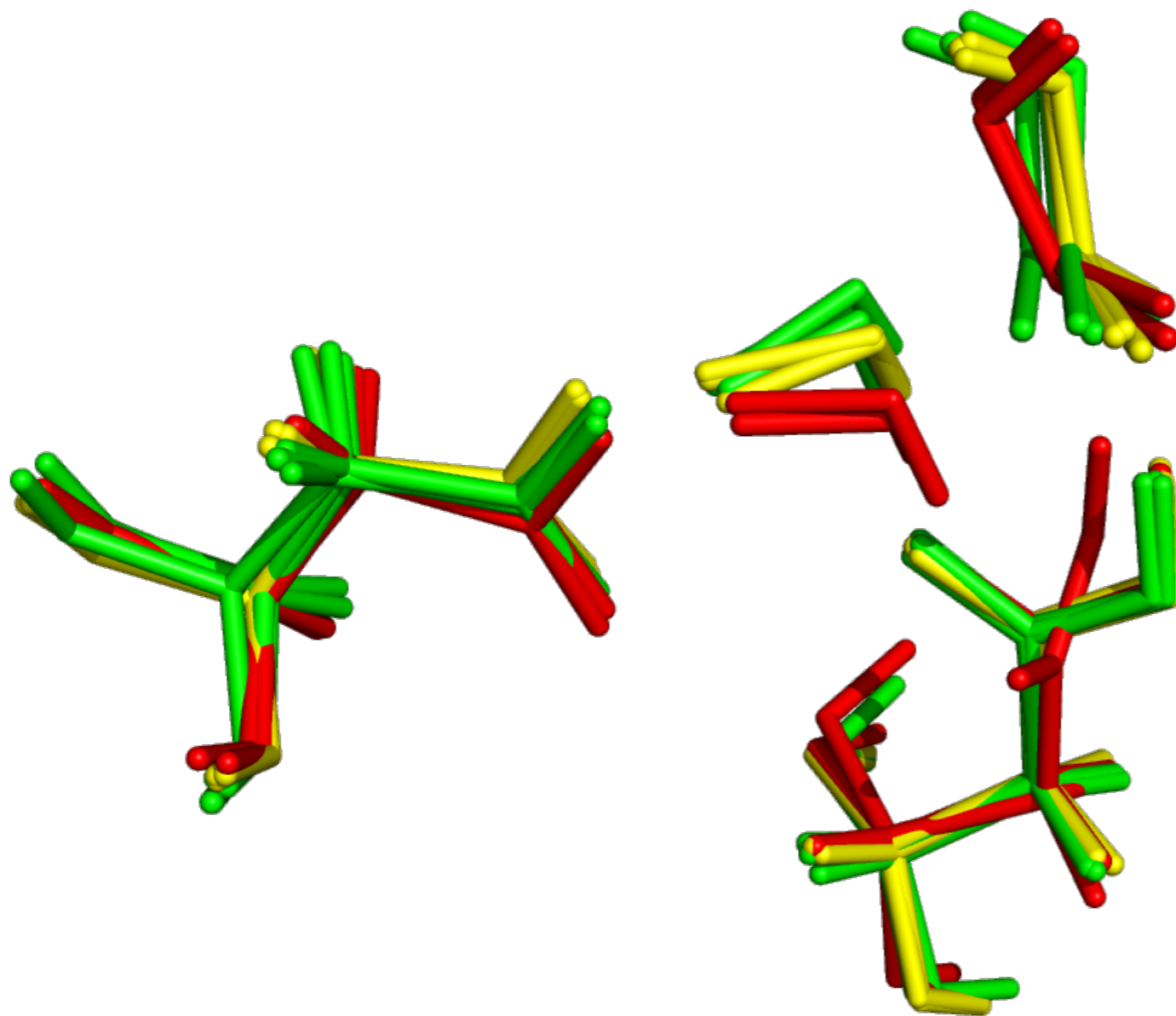
Origin of the fluctuations?



Differences at the active site ?

Differences in protein folding?

# Why $\Delta G_{\text{act}}$ fluctuates so much?



Active site  
 $\Delta G_{\text{act}} > 20$  kcal/mol

Active site  
 $\Delta G_{\text{act}} = 15-20$  kcal/mol

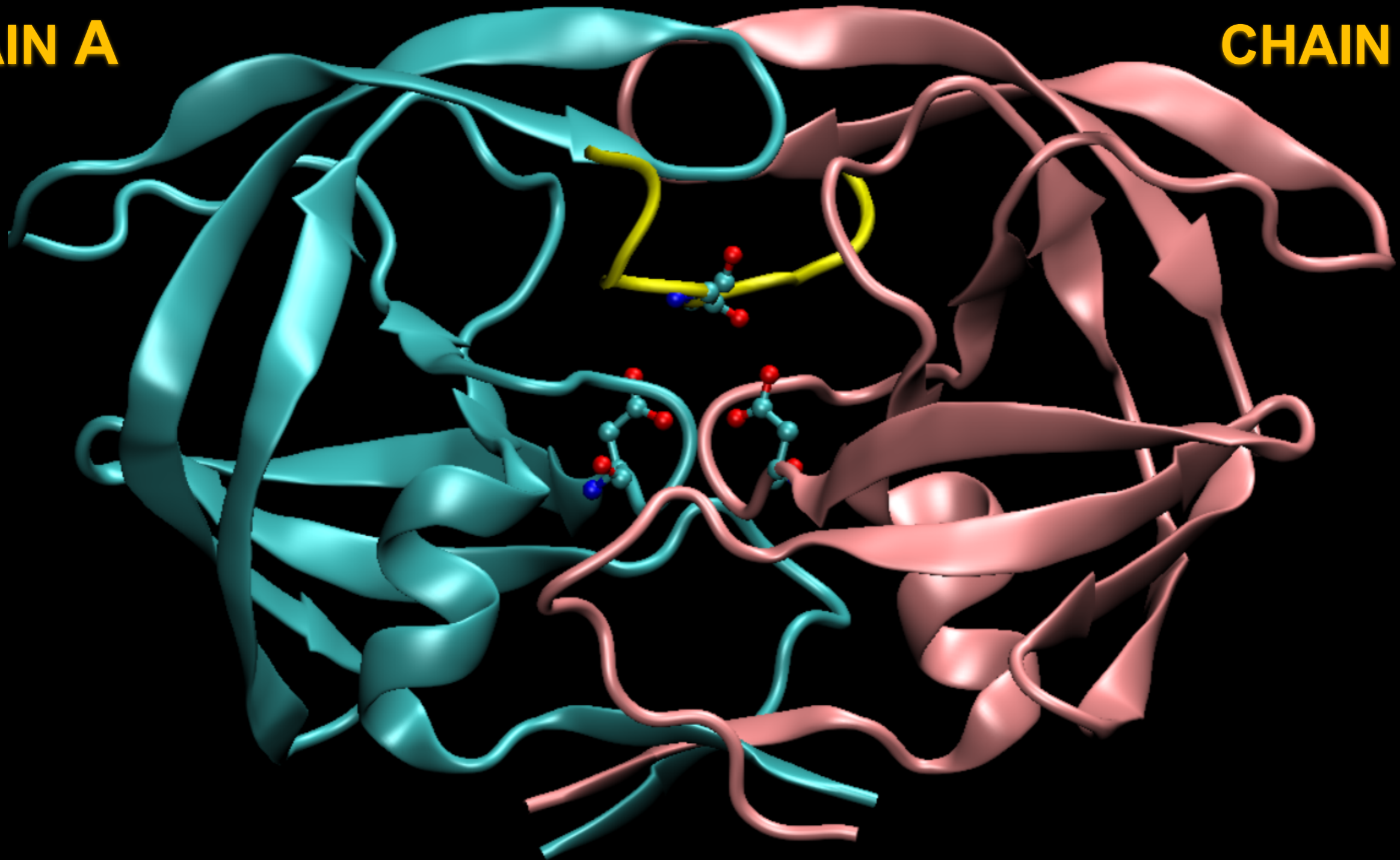
Active site  
 $\Delta G_{\text{act}} < 15$  kcal/mol

# Influence of long-range interactions



**CHAIN A**

**CHAIN B**

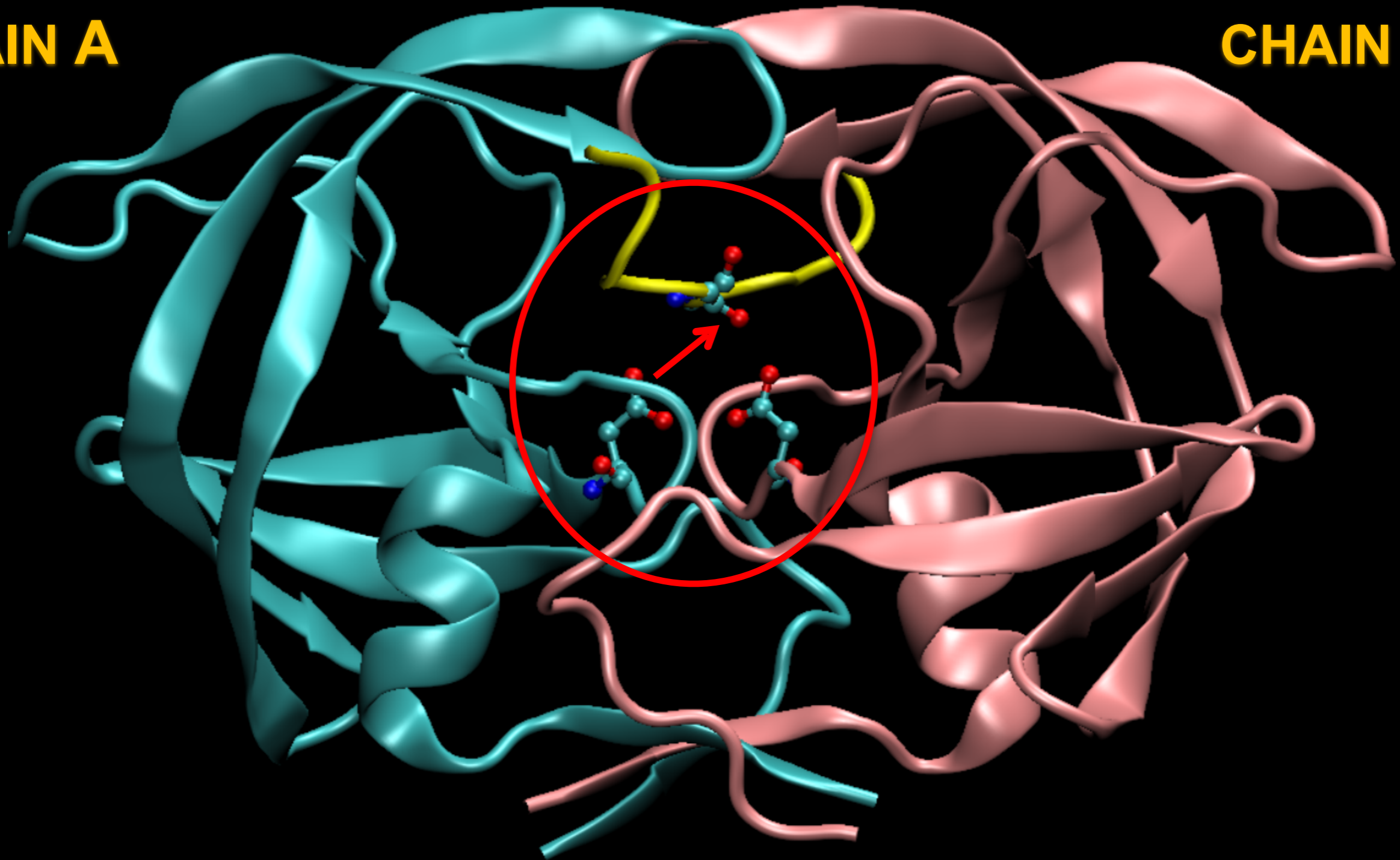


# Influence of long-range interactions



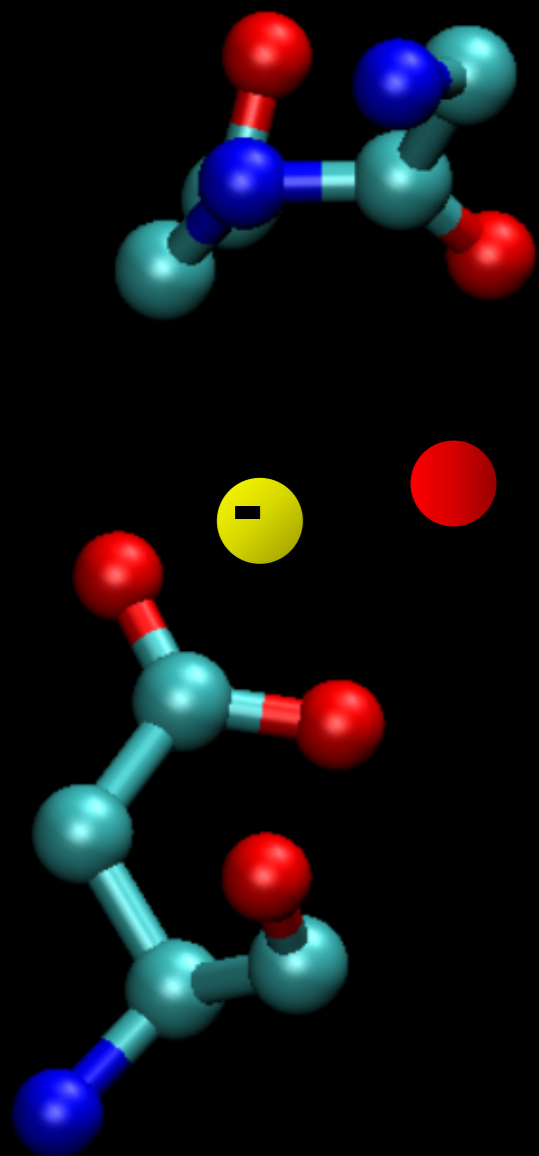
**CHAIN A**

**CHAIN B**

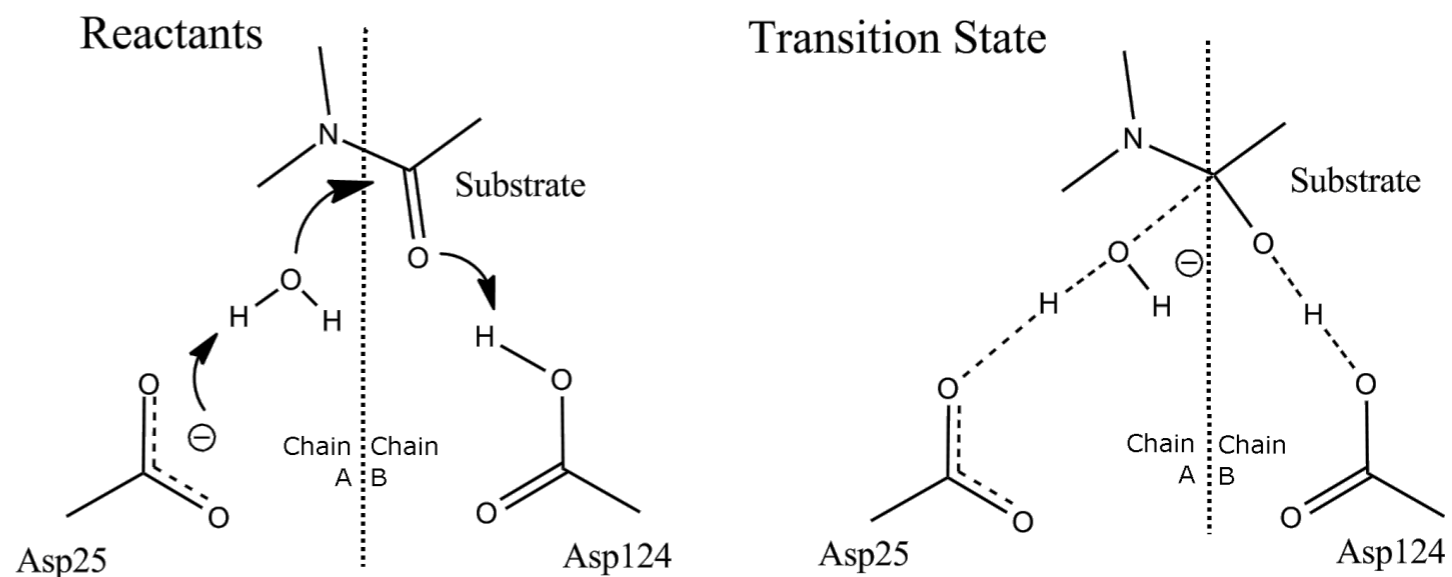
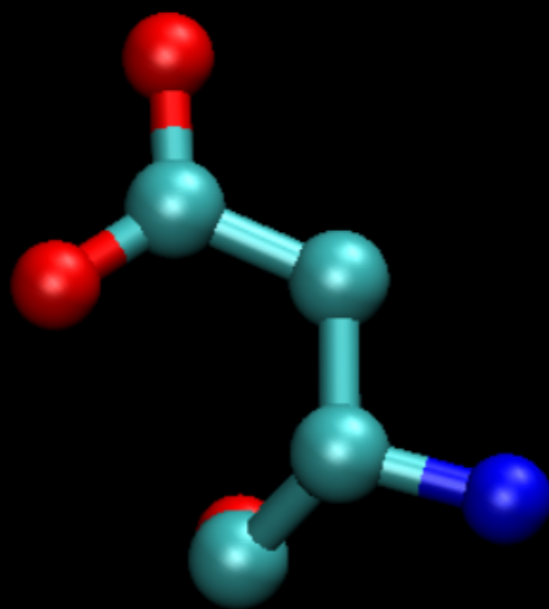


# Redistribution of electronic density between the R and the TS.

## CHAIN A



## CHAIN B



During the reaction mechanism significant charge density migrates from Asp25 of chain A to the substrate (direction A→B)

# Influence of each residue for $\Delta G_{\text{act}}$



A set of new calculations was done. In each one a residue was deleted.

## Influence of each residue for $\Delta G_{\text{act}}$



A set of new calculations was done. In each one a residue was deleted.

We generated 196 mutants, each with a different deletion.



## Influence of each residue for $\Delta G_{\text{act}}$



A set of new calculations was done. In each one a residue was deleted.

We generated 196 mutants, each with a different deletion.

The energy of the reactants and TS recalculated for each mutant.

## Influence of each residue for $\Delta G_{\text{act}}$



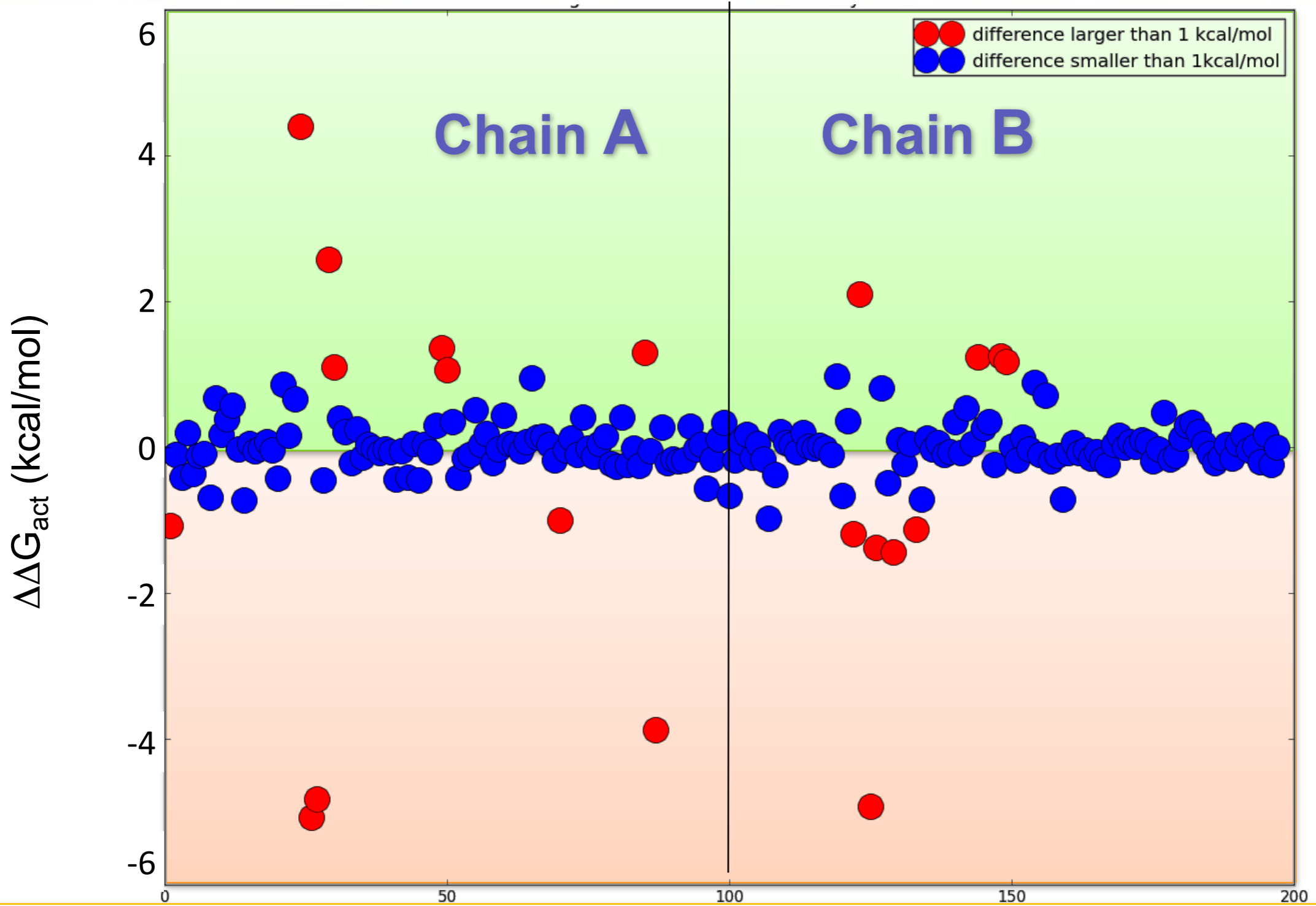
A set of new calculations was done. In each one a residue was deleted.

We generated 196 mutants, each with a different deletion.

The energy of the reactants and TS recalculated for each mutant.

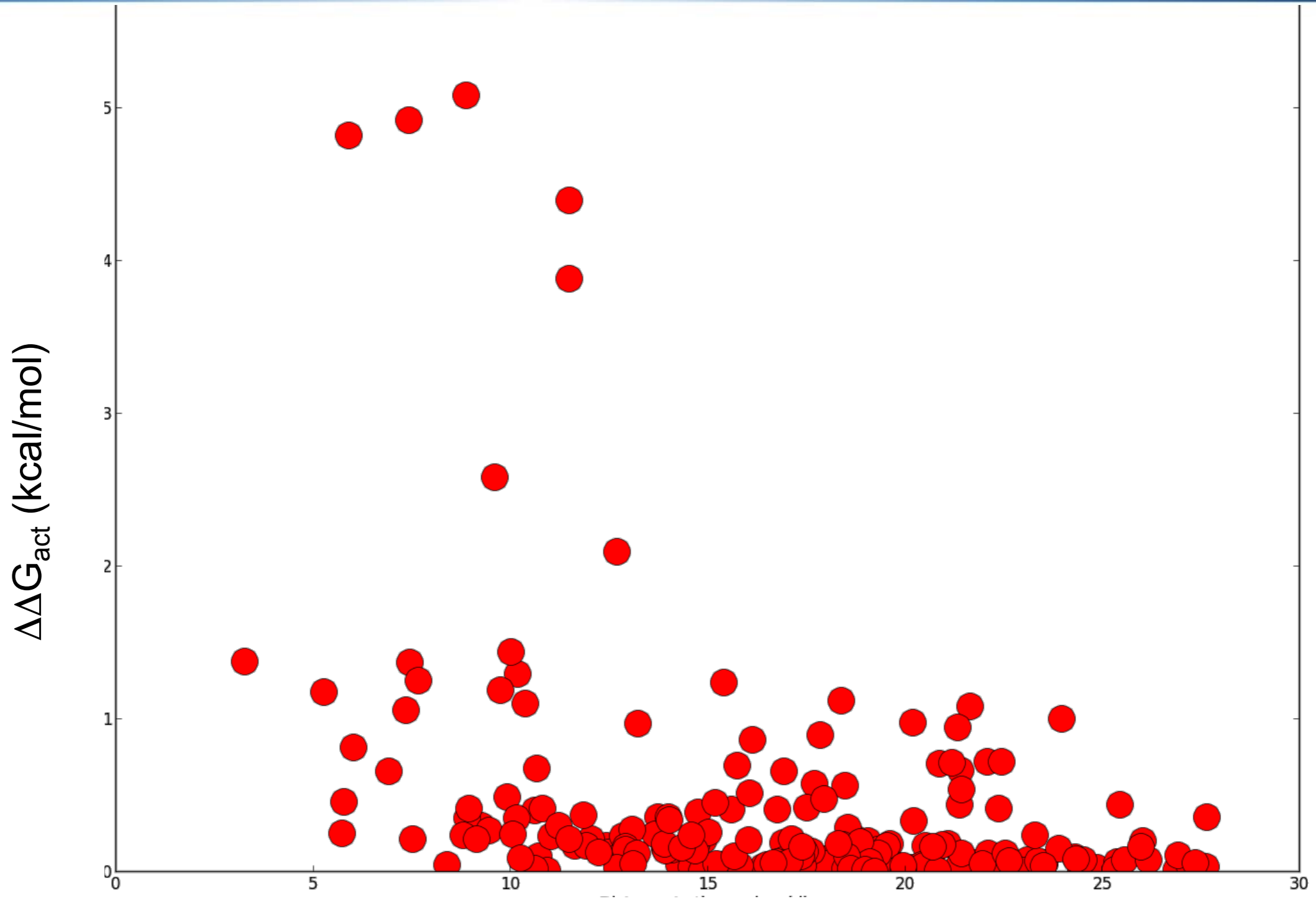
At the end we will have done  $196 \times 40 \times 2 = 15.680$  QM/MM calculations.

# Influence of each residue for $\Delta G_{act}$



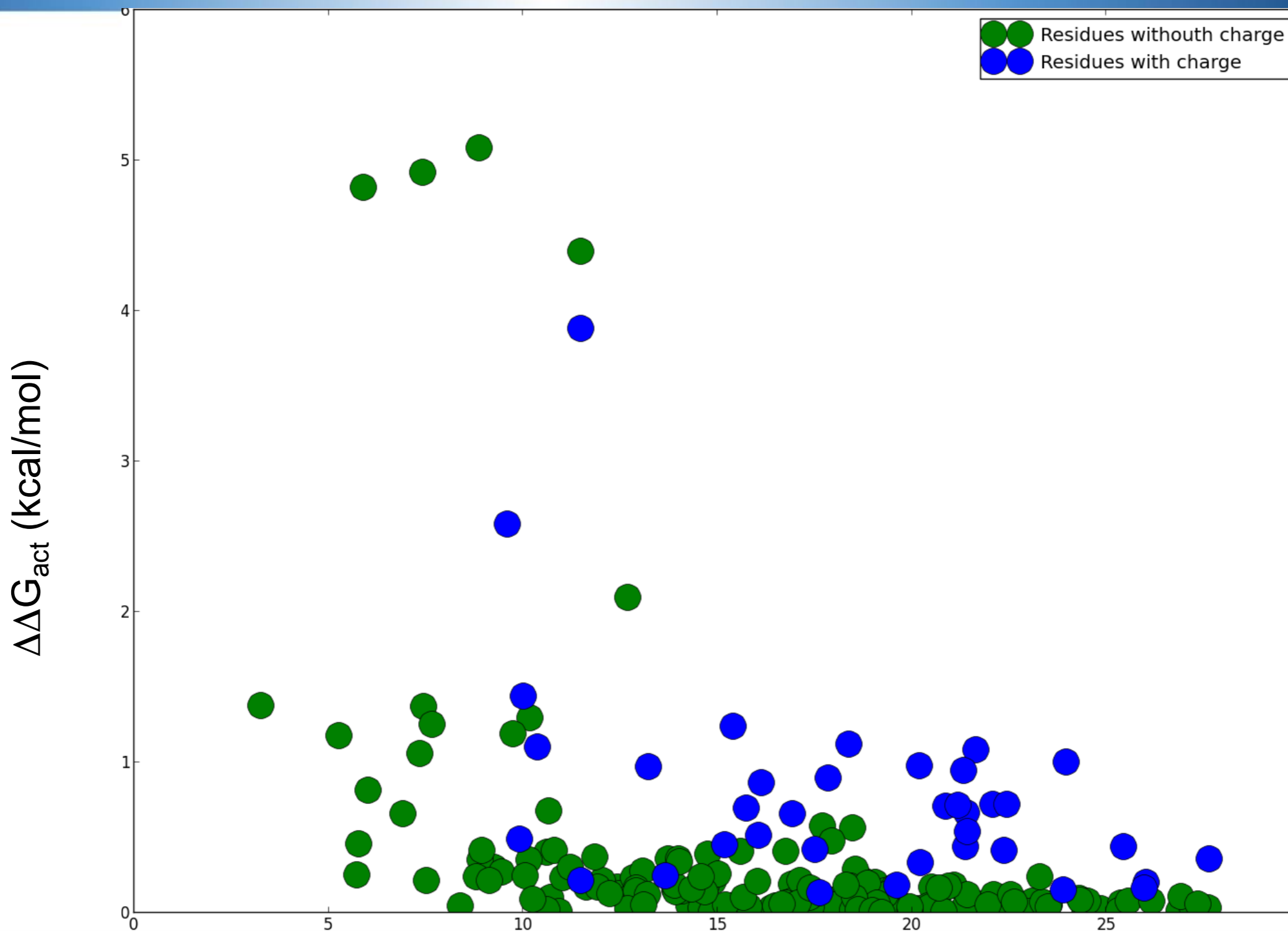
$\Delta\Delta G_{act}$  by residue (Larger than or smaller than 1 kcal/mol)

# Influence of each residue for $\Delta G_{act}$



$|\Delta\Delta G_{act}|$  vs. distance to the nucleophile

# Influence of each residue for $\Delta G_{\text{act}}$

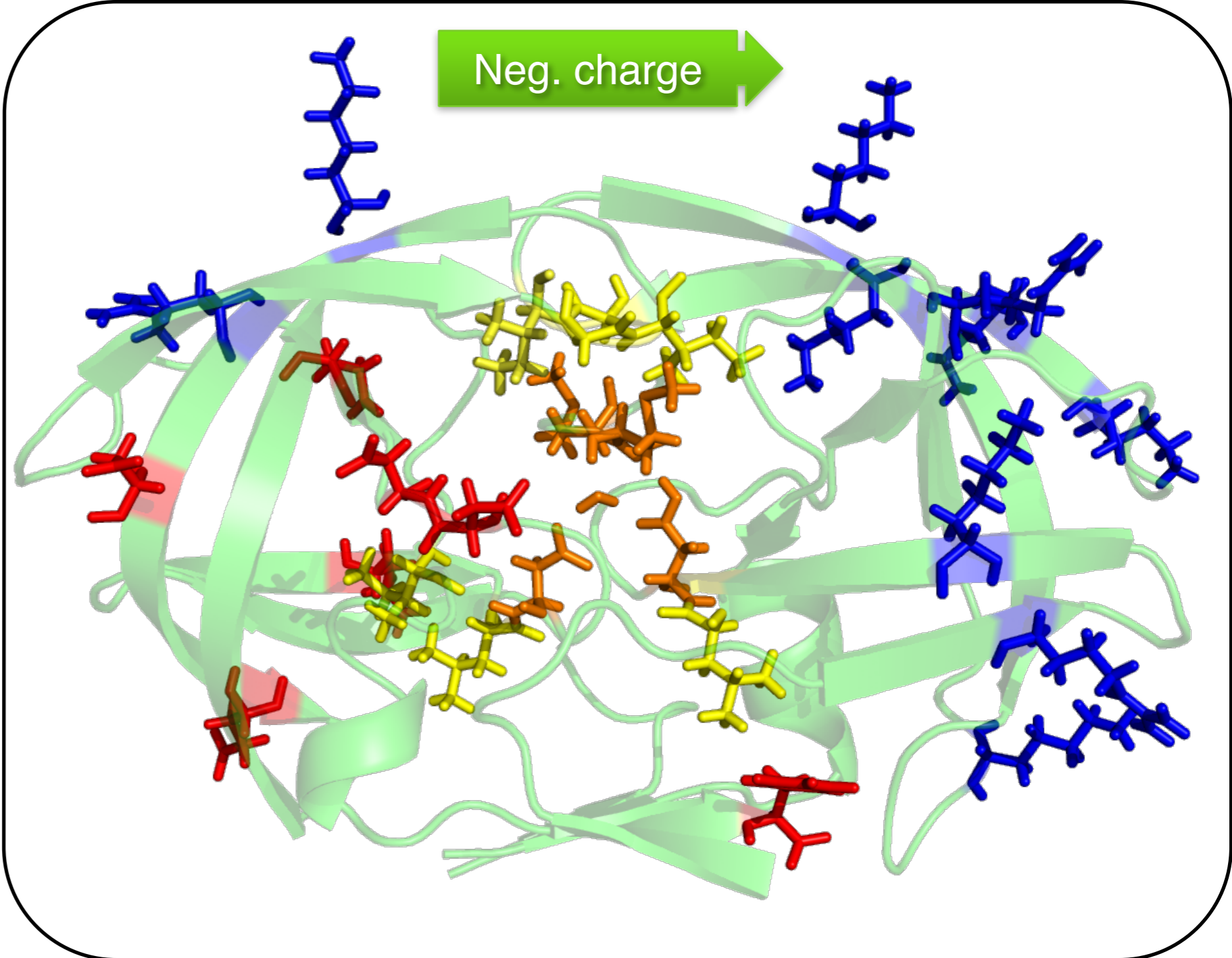


$|\Delta\Delta G_{\text{act}}|$  vs. distance to the nucleophile, discriminating charged residues

# Influence of each residue for $\Delta G_{act}$



- Positive residues
- Negative residues
- Neutral residues
- Asp25, H<sub>2</sub>O, S



Residues with  $\Delta\Delta G > 0.5$  kcal/mol.

# Influence of each residue for $\Delta G_{act}$

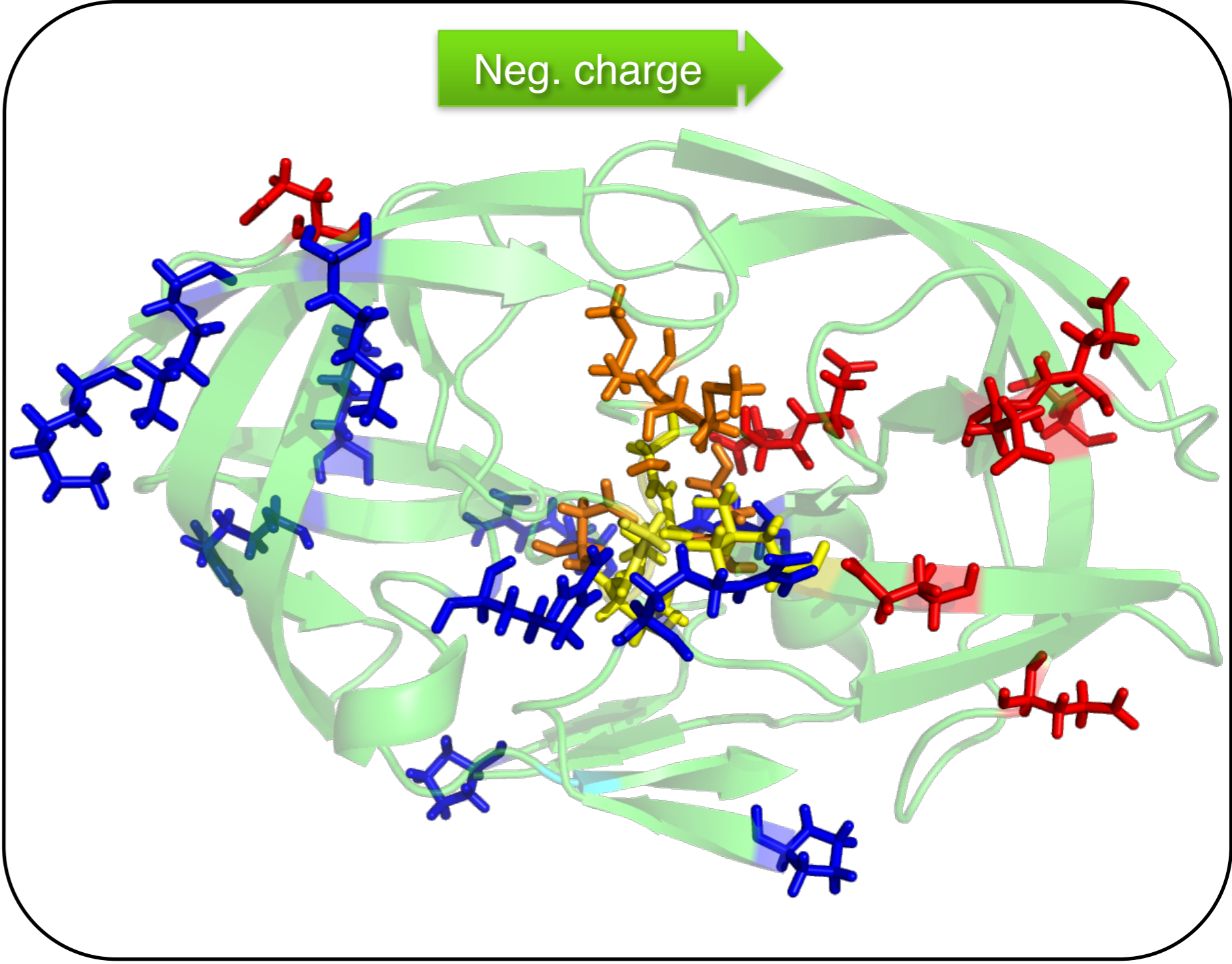


Positive residues

Negative residues

Neutral residues

Asp25, H<sub>2</sub>O, S



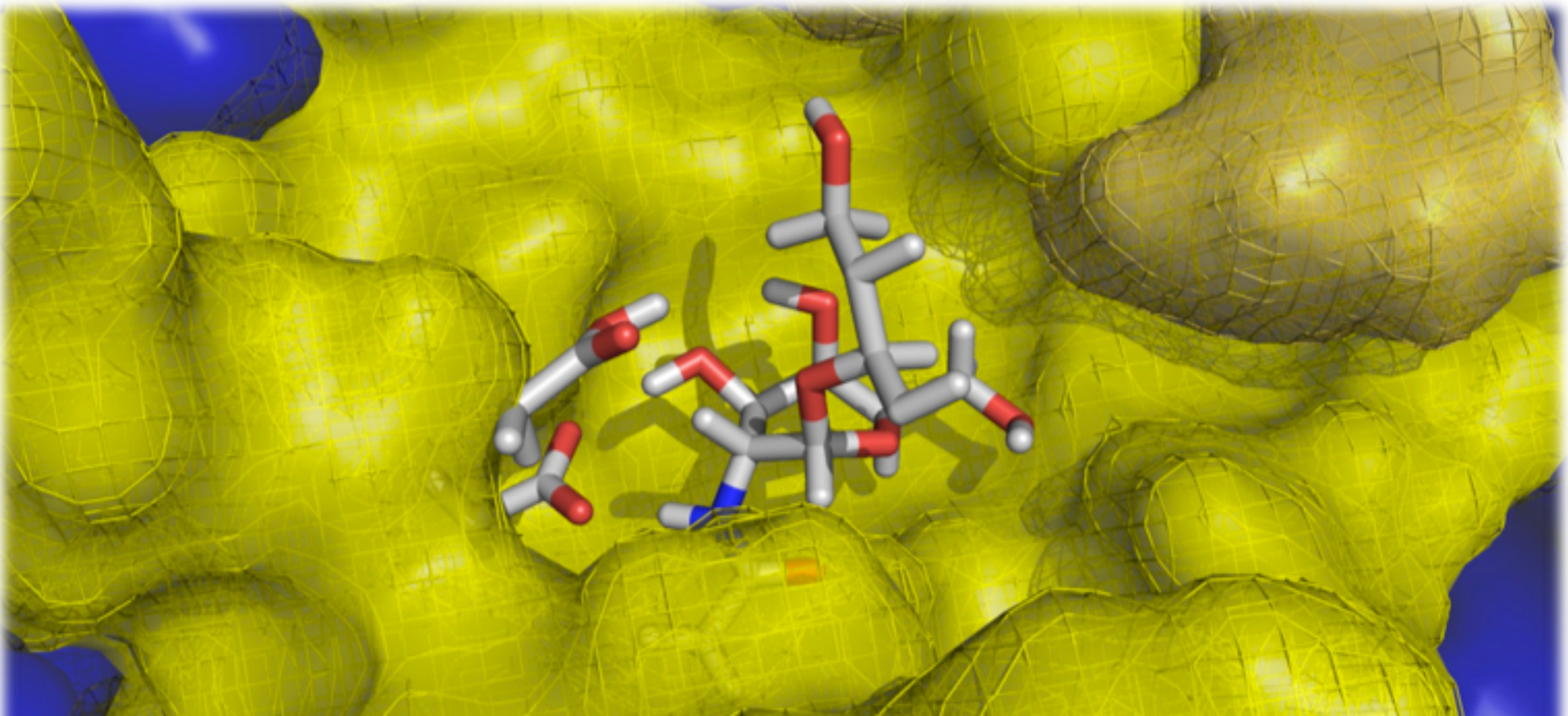
Residues with  $\Delta\Delta G < -0.5$  kcal/mol.

# Influence of Enzyme Conformation on Kinetics

**Pedro Alexandrino Fernandes,**

Dep. Chemistry & Biochemistry, University of Porto, Portugal

pedro.fernandes@fc.up.pt





# Conclusions

# Conclusions

1. **Enzyme kinetics** depends on a spectrum of **different kinds of interatomic interactions.**

The theoretical method should describe all these interactions with consistent accuracy.

# Conclusions

1. **Enzyme kinetics** depends on a spectrum of **different kinds of interatomic interactions.**

The theoretical method should describe all these interactions with consistent accuracy.

2. **Enzyme kinetics** may have meaningful contributions from **medium/long range interatomic interactions.**

The theoretical model should include all these interactions and treat them with consistent accuracy.

# Conclusions

1. **Enzyme kinetics** depends on a spectrum of **different kinds of interatomic interactions**.

The theoretical method should describe all these interactions with consistent accuracy.

2. **Enzyme kinetics** may have meaningful contributions from **medium/long range interatomic interactions**.

The theoretical model should include all these interactions and treat them with consistent accuracy.

3. **Enzyme kinetics** may have meaningful contributions from the **different conformations** the enzyme can adopt.

Methods to check and sample enzyme flexibility should be further developed. It is important to include conformational changes in the timescale of the enzyme kinetics, instead of just short scale fluctuations around the starting folding.

# Conclusions

1. **Enzyme kinetics** depends on a spectrum of **different kinds of interatomic interactions**.

The theoretical method should describe all these interactions with consistent accuracy.

2. **Enzyme kinetics** may have meaningful contributions from **medium/long range interatomic interactions**.

The theoretical model should include all these interactions and treat them with consistent accuracy.

3. **Enzyme kinetics** may have meaningful contributions from the **different conformations** the enzyme can adopt.

Methods to check and sample enzyme flexibility should be further developed. It is important to include conformational changes in the timescale of the enzyme kinetics, instead of just short scale fluctuations around the starting folding.

**In short, the CONSISTENCY in the correctness of the treatment of all these aspects is the key for accurate studies of enzyme catalysis.**

# Enzyme Catalysis

**Pedro Alexandrino Fernandes,**

Dep. Chemistry & Biochemistry, University of Porto, Portugal

pedro.fernandes@fc.up.pt

