

SBM CDT 2020

Computational Module

Day 1

Fernanda Duarte

Assoc. Prof Computational Organic Chemistry
University of Oxford, UK



Twitter @duarte_group

<http://fduartegroup.org>



Github: duartegroup

Timetable

	Tuesday	Thursday	Friday
9:00-10:00		Lecture 2	Lecture 3
10:30-12:00	Lecture 1	Project Work	Project Work
14:00- 15:00	Lab session	Project Work	Presentations (preparation)
15:30- 17:00	Lab session	Project Work	Presentations

Workplan

Lab session

- Basic tutorials:
 1. Warm up: calculation of water
 2. Conformational analysis (cyclohexane)
 3. Transition state optimisation
 - Work through them as you get familiar with the softwares

Project Work

- Work on the projects you have chosen (4 people)
- Read the literature (exp. & comp data) associated to it
- Reproduce some of the calculations
- Explore any novel aspect that you may find relevant

Presentations

- 15 min per group + 5 min questions
- Background/ rational behind the technique used/ main findings /conclusions
- Any novel aspect explore will be a +

Presentations

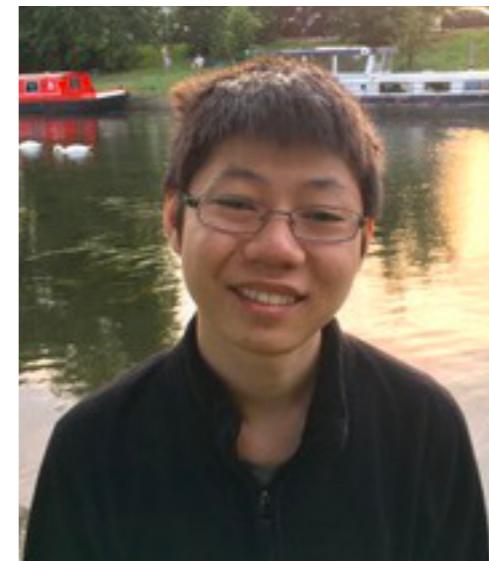
- 15 min per group + 5 min questions
- Background/ rational behind the technique used/ main findings /conclusions
- Any novel aspect explore will be a +

Demonstrators



Tom Young

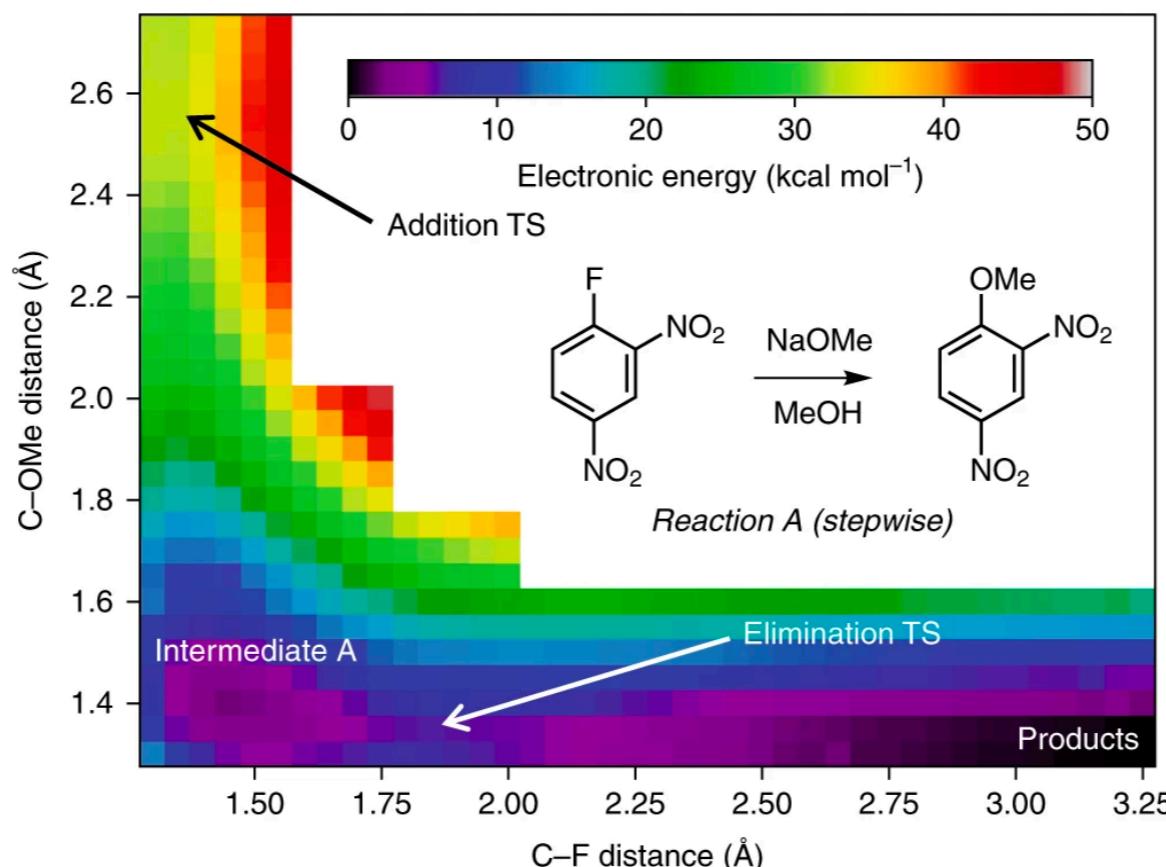
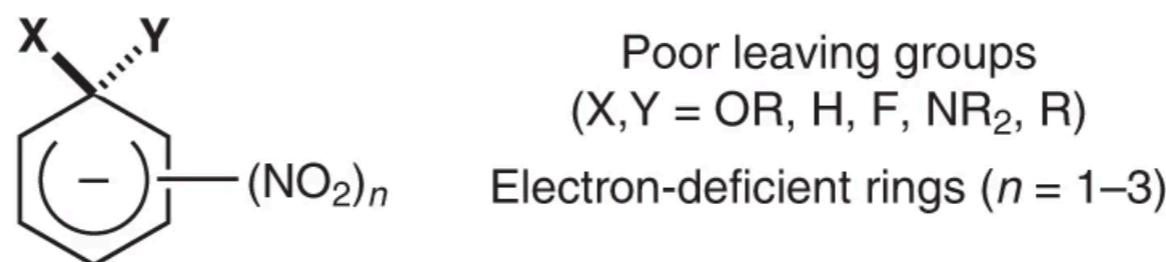
Theory and Modelling in Chemical Sciences (TMCS) CDT



Jonathan Yong

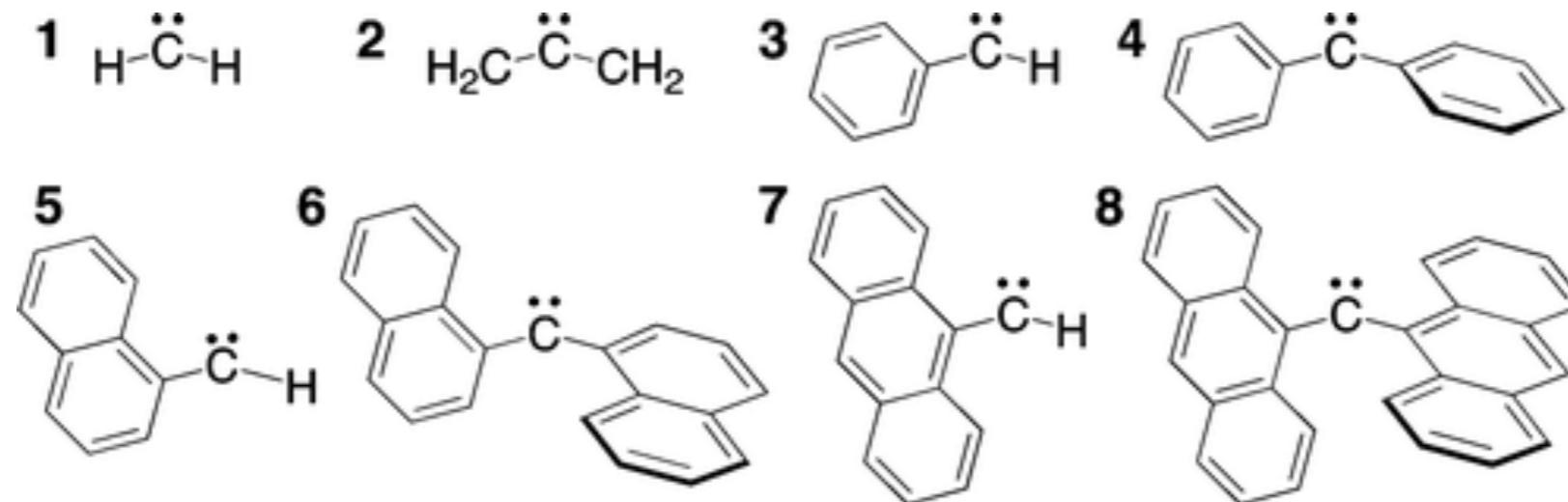
Synthesis for Biology and Medicine CDT

Project 1: Concerted nucleophilic aromatic substitutions



Technique: Density Functional Theory; **Software:** ORCA

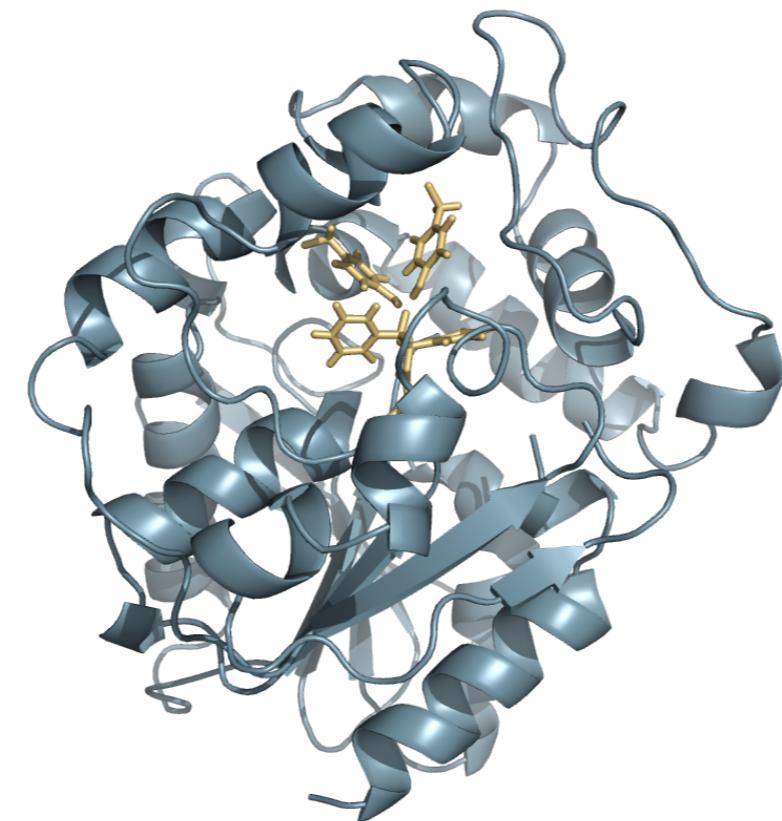
Project 2: Carbene Stabilization by Aryl Substituents. Is Bigger Better?



J. Am. Chem. Soc., **2007**, 129, 3763, DOI: 10.1021/ja068899t.

Technique: Density Functional Theory; **Software:** ORCA

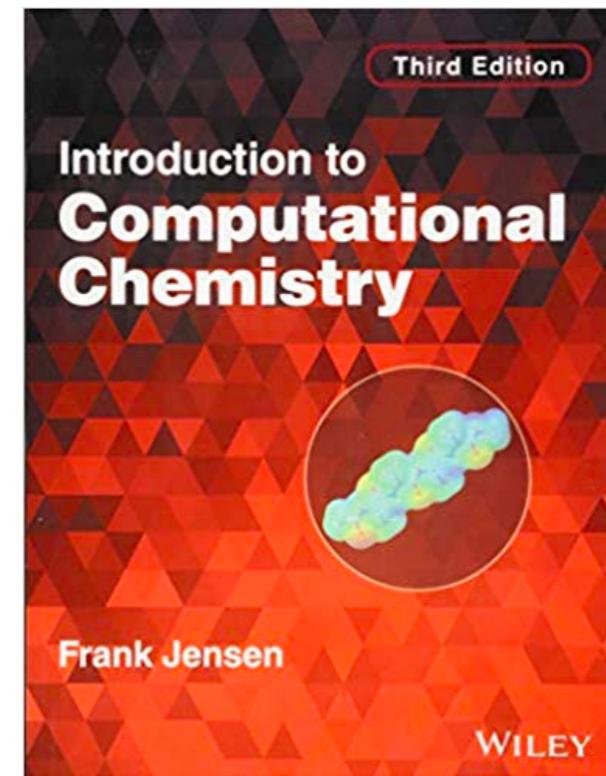
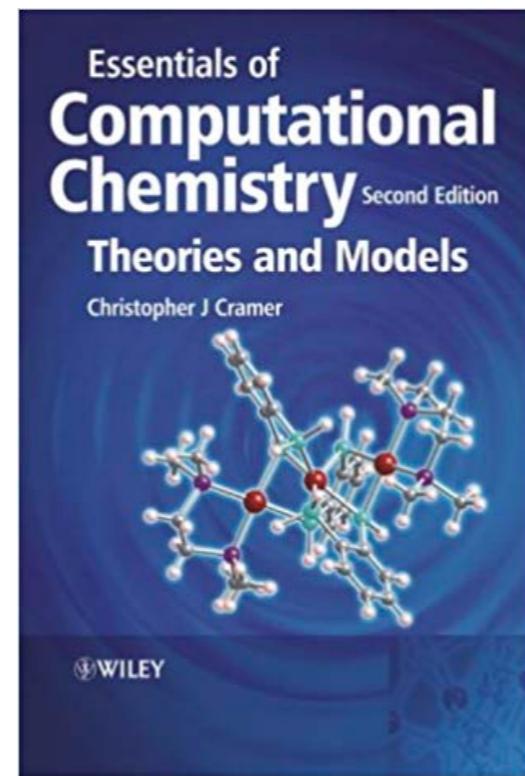
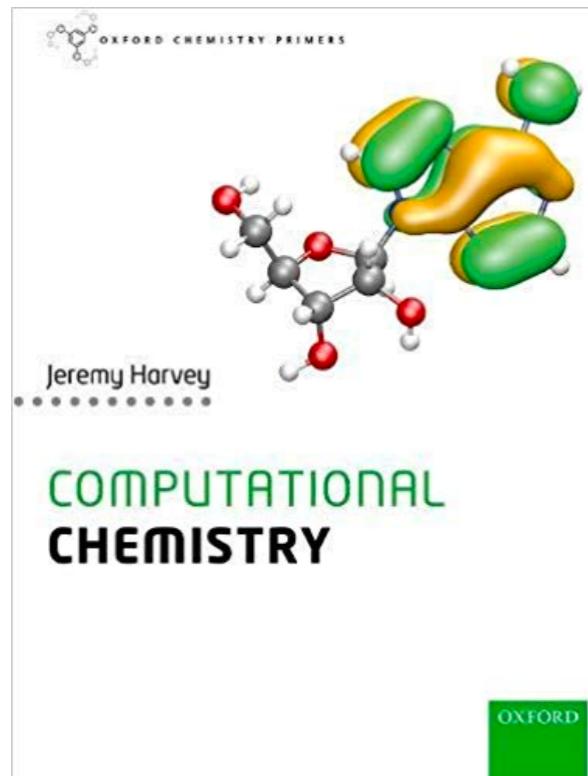
Project 3: Expanding the Catalytic Triad in Epoxide Hydrolases and Related Enzymes



ACS Catal. **2015**, 5, 5702–5713. DOI: 10.1021/acscatal.5b01639.

Technique: Molecular Dynamics; **Software:** GROMACS

Useful References



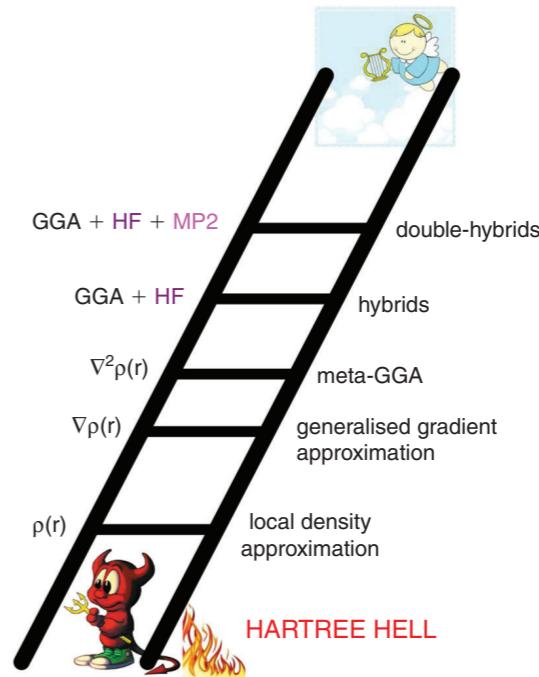
Useful References

A Trip to the Density Functional Theory Zoo: Warnings and Recommendations for the User

Aust. J. Chem. **2019**, 72, 563

<https://doi.org/10.1071/CH19023>

HEAVEN OF CHEMICAL ACCURACY



Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals

Mol. Phys. **2017**, 115, 2315

<https://doi.org/10.1080/00268976.2017.1333644>



Lecture 1

Outline

- What is - and why is it relevant?
- A bit of history
- How do we use it?

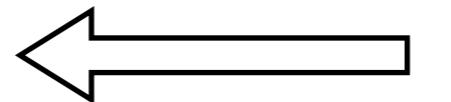
What is - and why is it relevant?

Computational chemistry is a branch of chemistry that uses computer simulation to assist in solving chemical problems (Wikipedia)

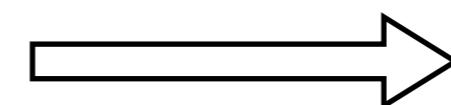
Macroscopic Observables



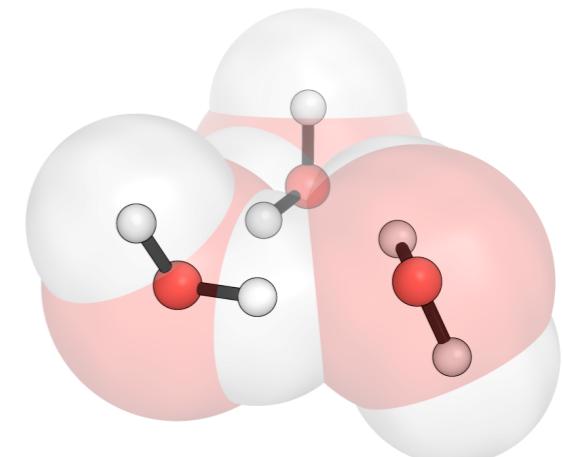
Predict



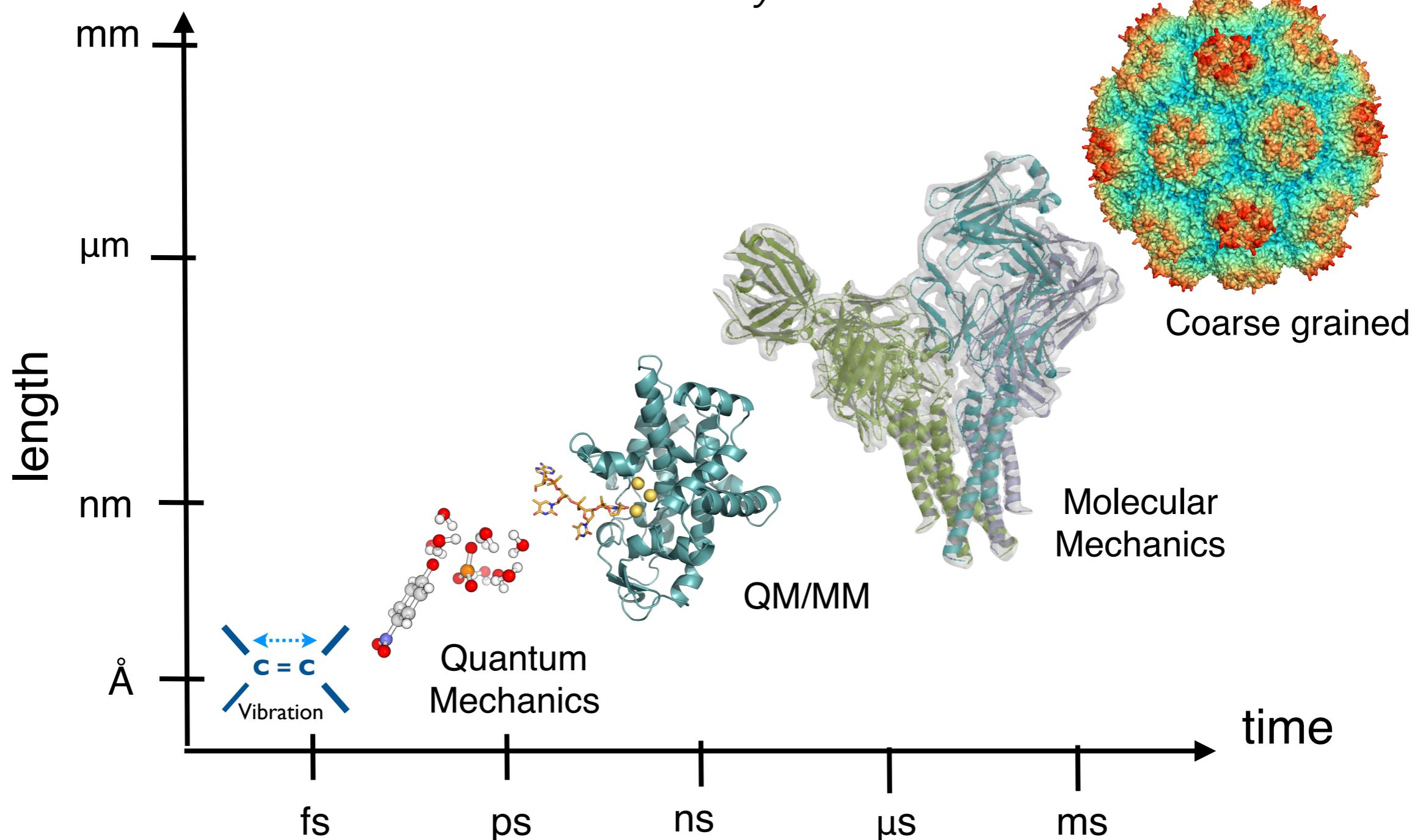
Electronic/Molecular
Structure



Interpret

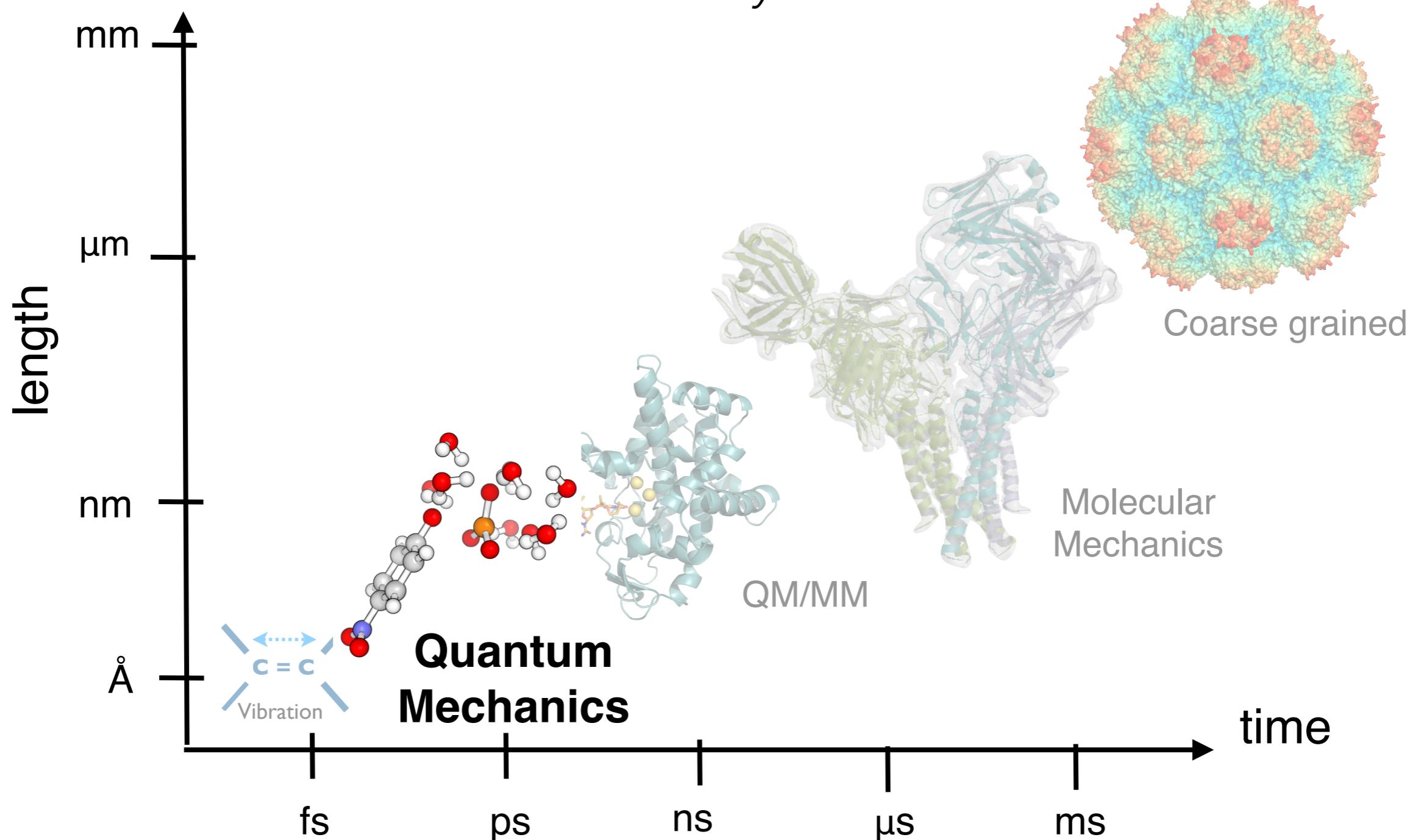


What is - and why is it relevant?



“Everything should be made as simple as possible, but not simpler.”

What is - and why is it relevant?



“Everything should be made as simple as possible, but not simpler.”

Computational Chemistry

A bit of history

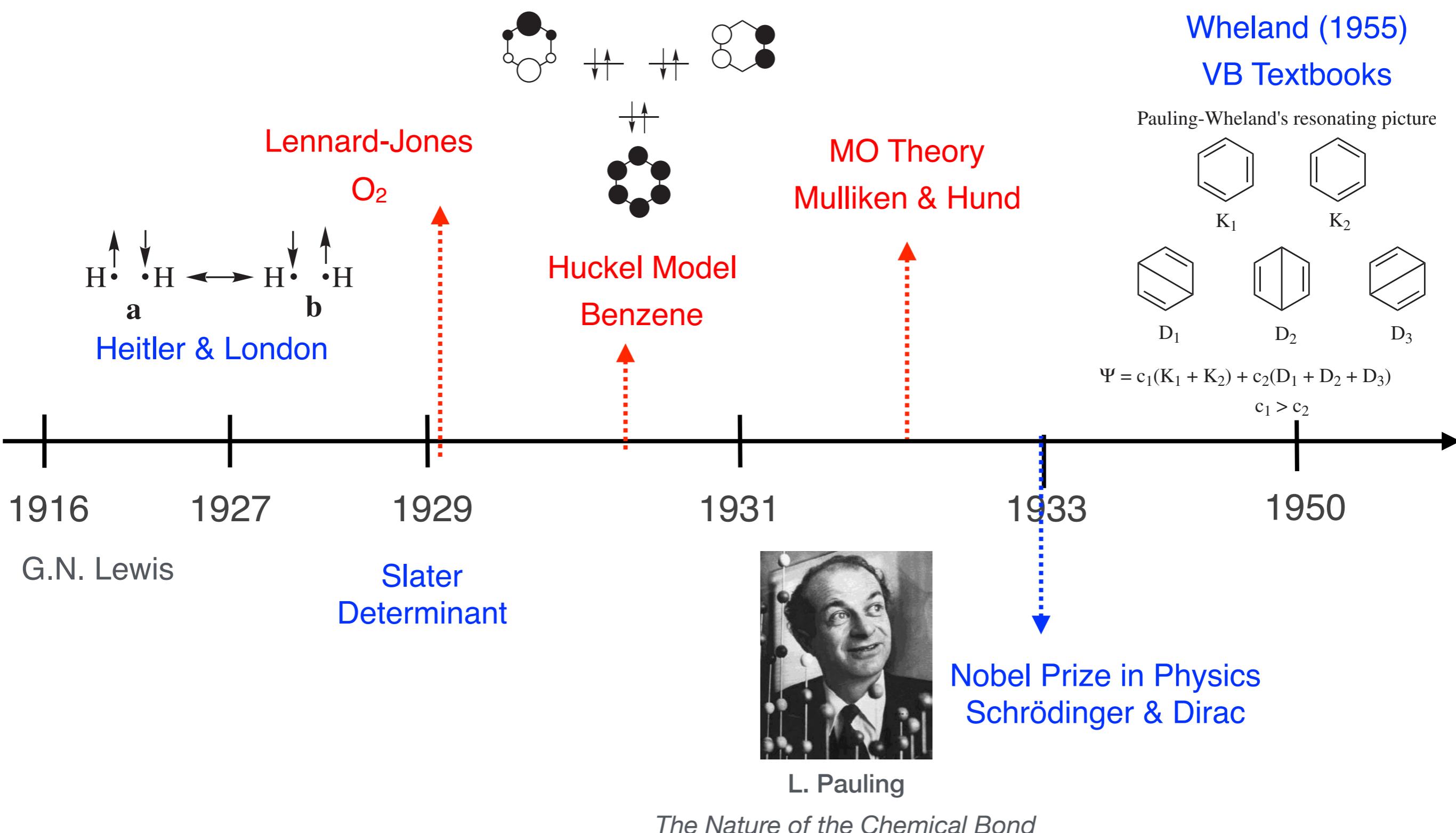
The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that the exact application of these laws leads to equations much too complicated to be soluble.

It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation

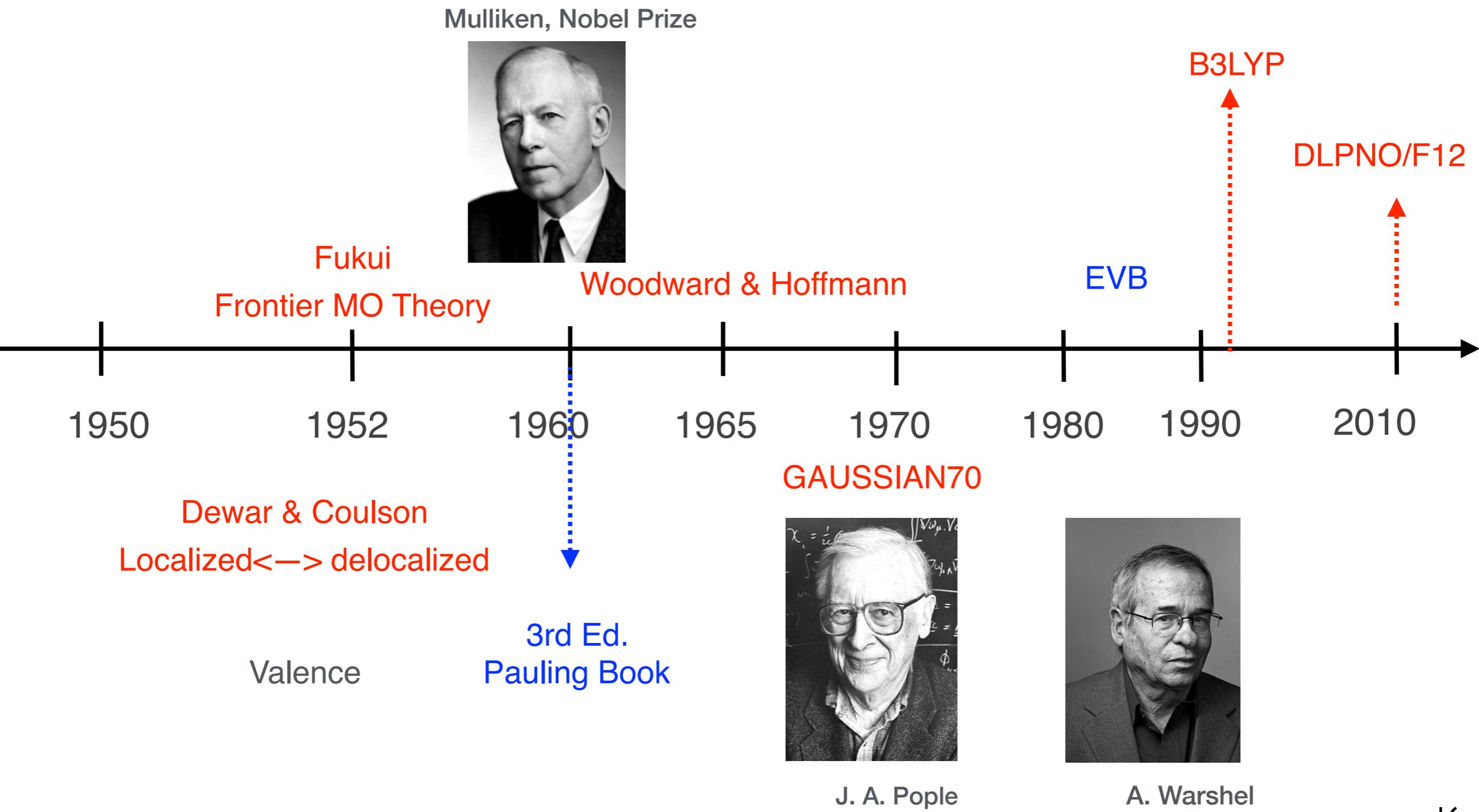
Paul A. M. Dirac

Proc. Royal Soc. Lond. A, 1929, 123, 792

A bit of history

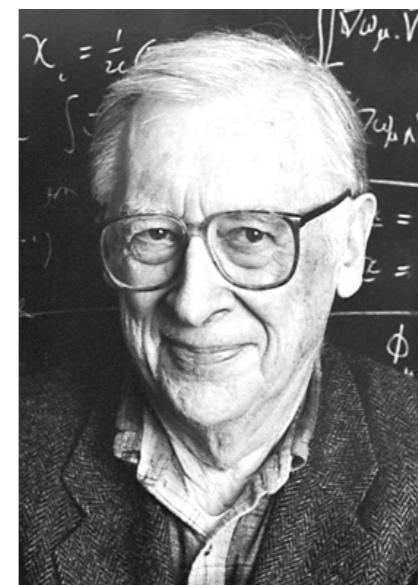
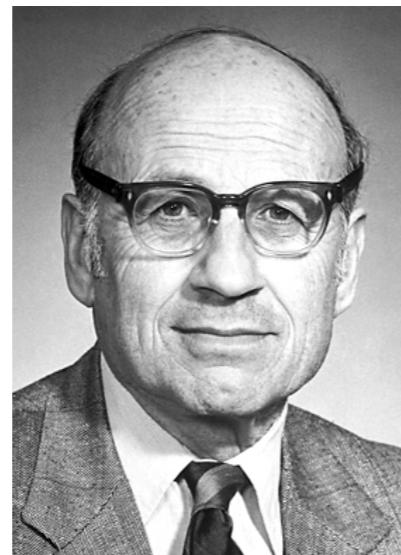


A bit of history



A bit of history

65 years later.....



The Nobel Prize in Chemistry 1998

Walter Kohn

"for his development of the density-functional theory"

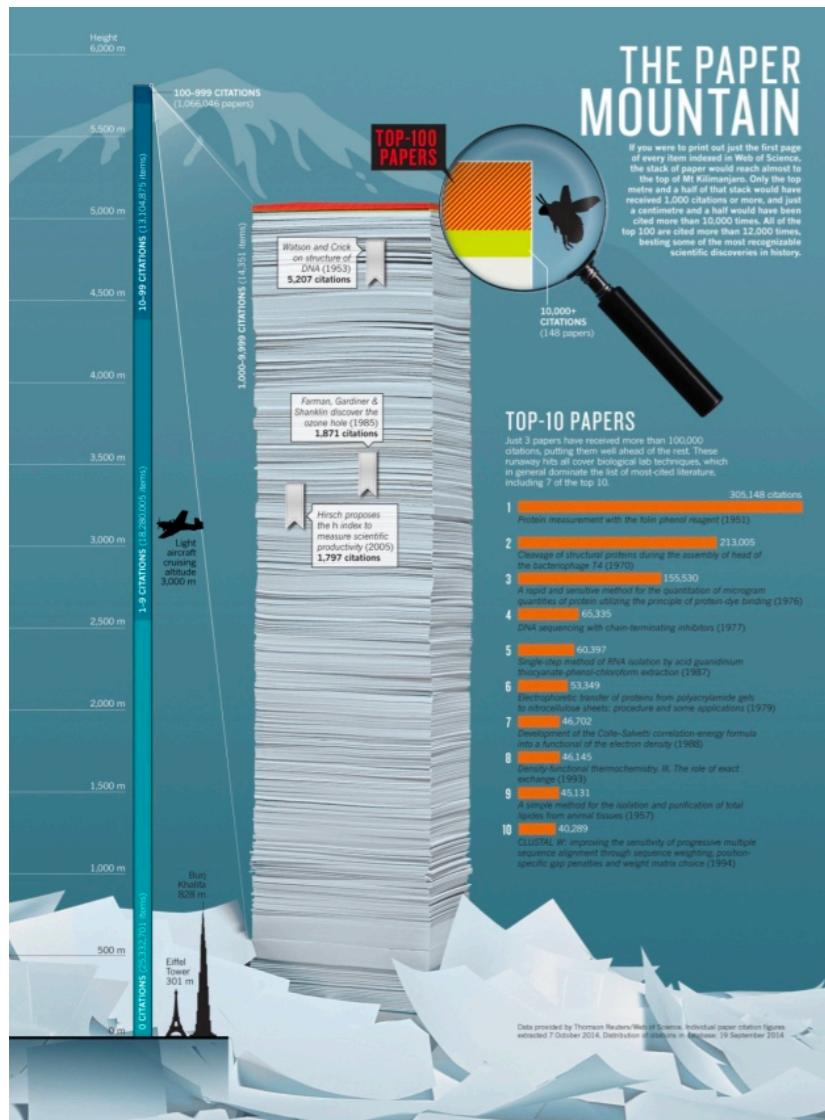
and

John A. Pople

"for his development of computational methods in quantum chemistry".

A bit of history

The 100 most-cited research of all time



1. Protein measurement with the folin phenol reagent.
J. Biol. Chem. **1951**, *193*, 265.
Lowry O. H., Rosebrough N. J. Farr A. L. & Randall, R. J.
7. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.
Phys. Rev. B **1988**, *37*, 785.
Lee C., Yang W.* & Parr R. G. (*also 93)
8. Density-functional thermochemistry. III. The role of exact exchange
J. Chem. Phys. **1993**, *98*, 5648.
Becke, A. D. (also 25)
35. Self-consistent equations including exchange and correlation effects.
Phys. Rev. **1965**, *140*, A1133.
Kohn, W. & Sham, L. J. (also 39)

A bit of history

Then everything is solved.....

“Overall, it is not clear to us that any significant accurate information that was not already apparent from experiment either has been, or could have been, reliably garnered purely from computations.”

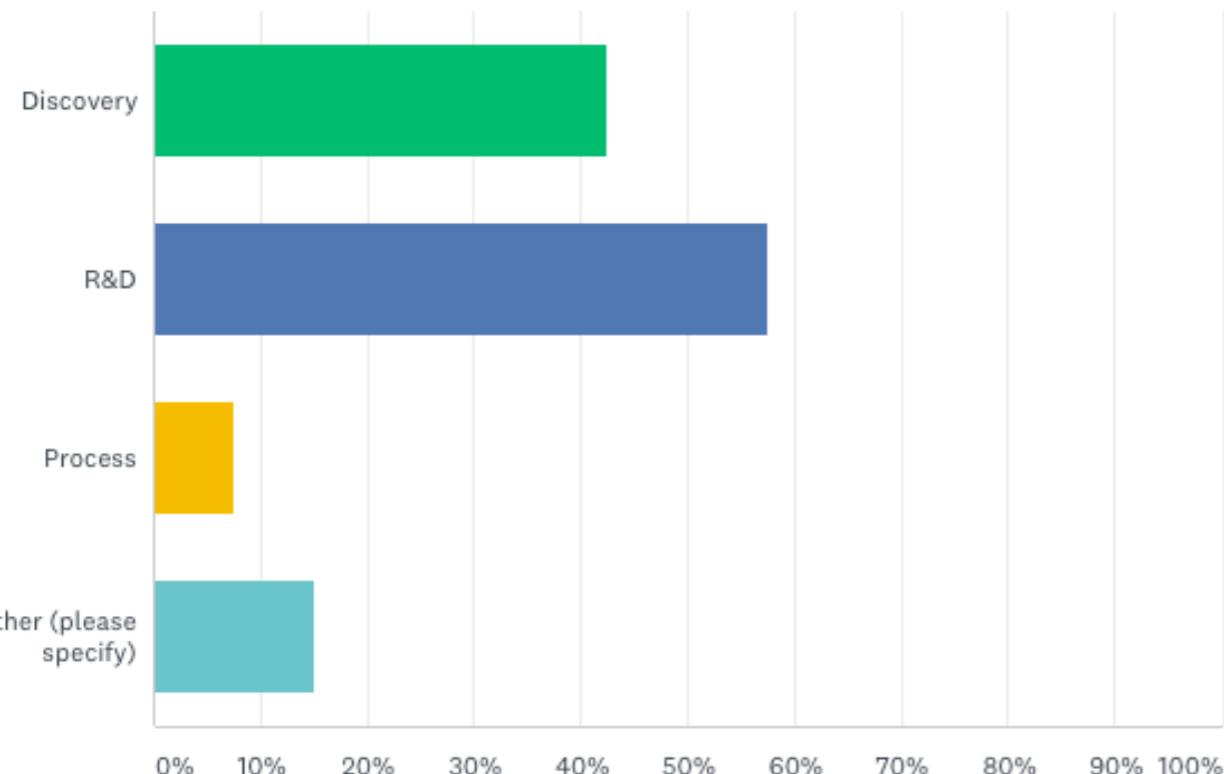
D. A. Singleton, J. Am. Chem. Soc., 2015, 137, 3811

Can Computational Chemistry Predict, or Just Explain?

A bit of history

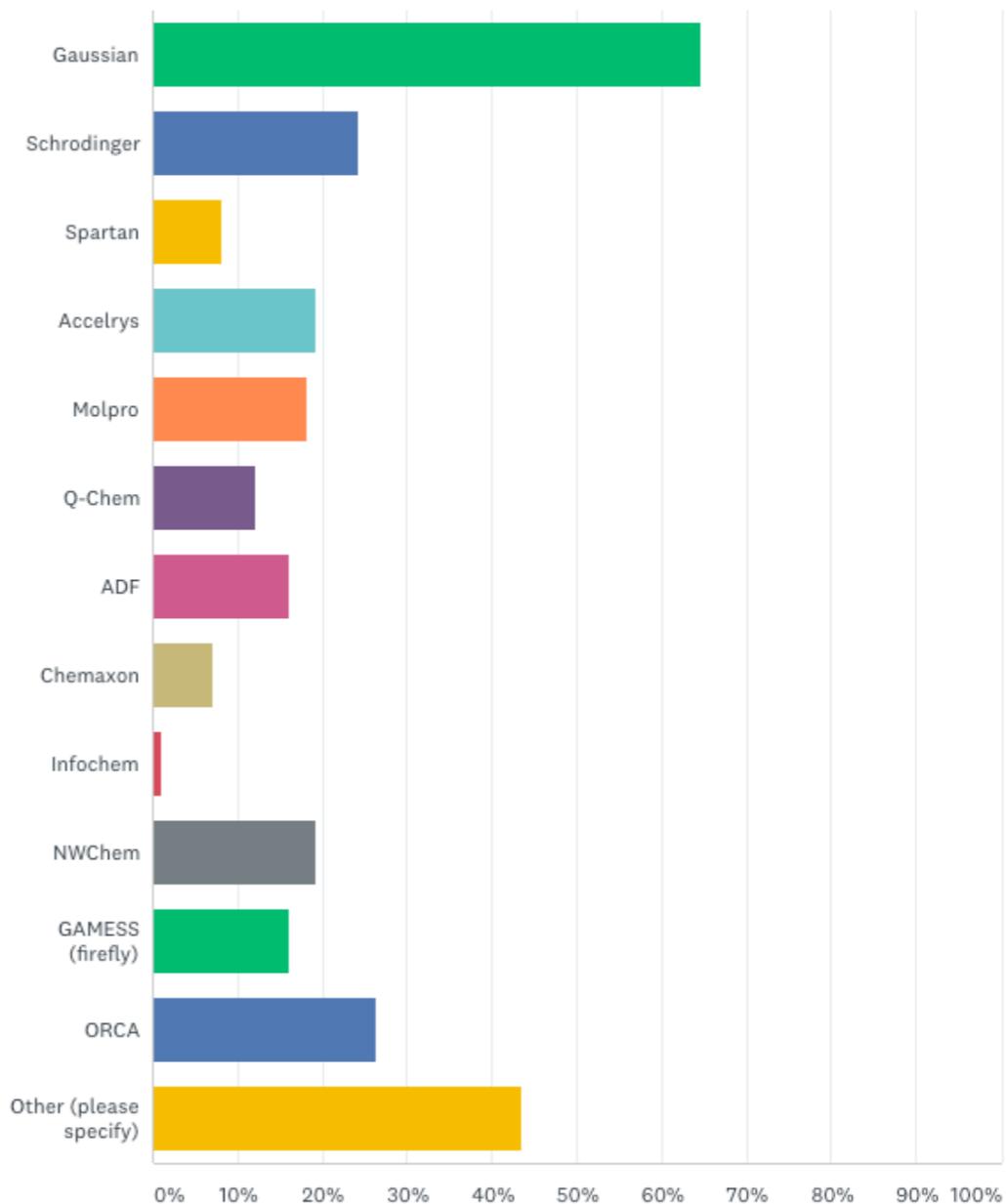
If you are in Industry, what is your division?

Answered: 40 Skipped: 60



What commercial or academic software do you use?

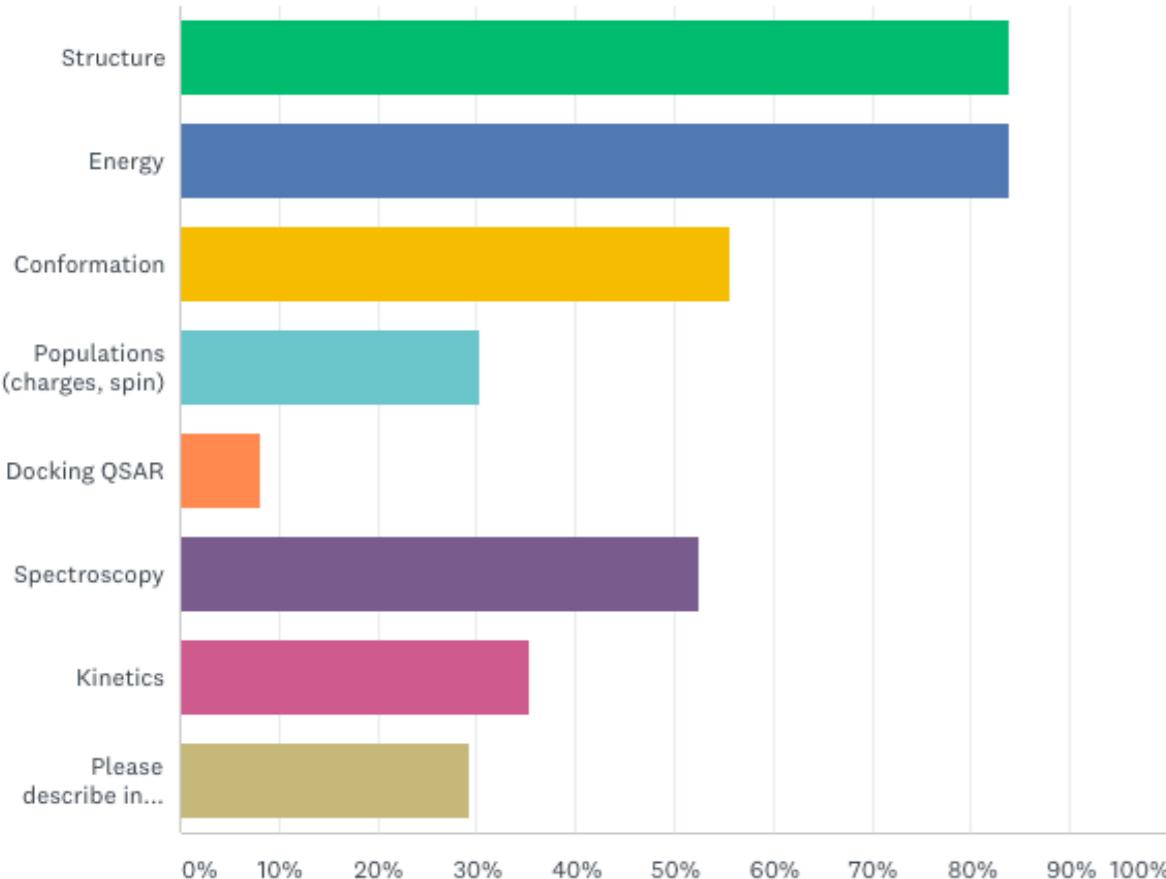
Answered: 99 Skipped: 1



A bit of history

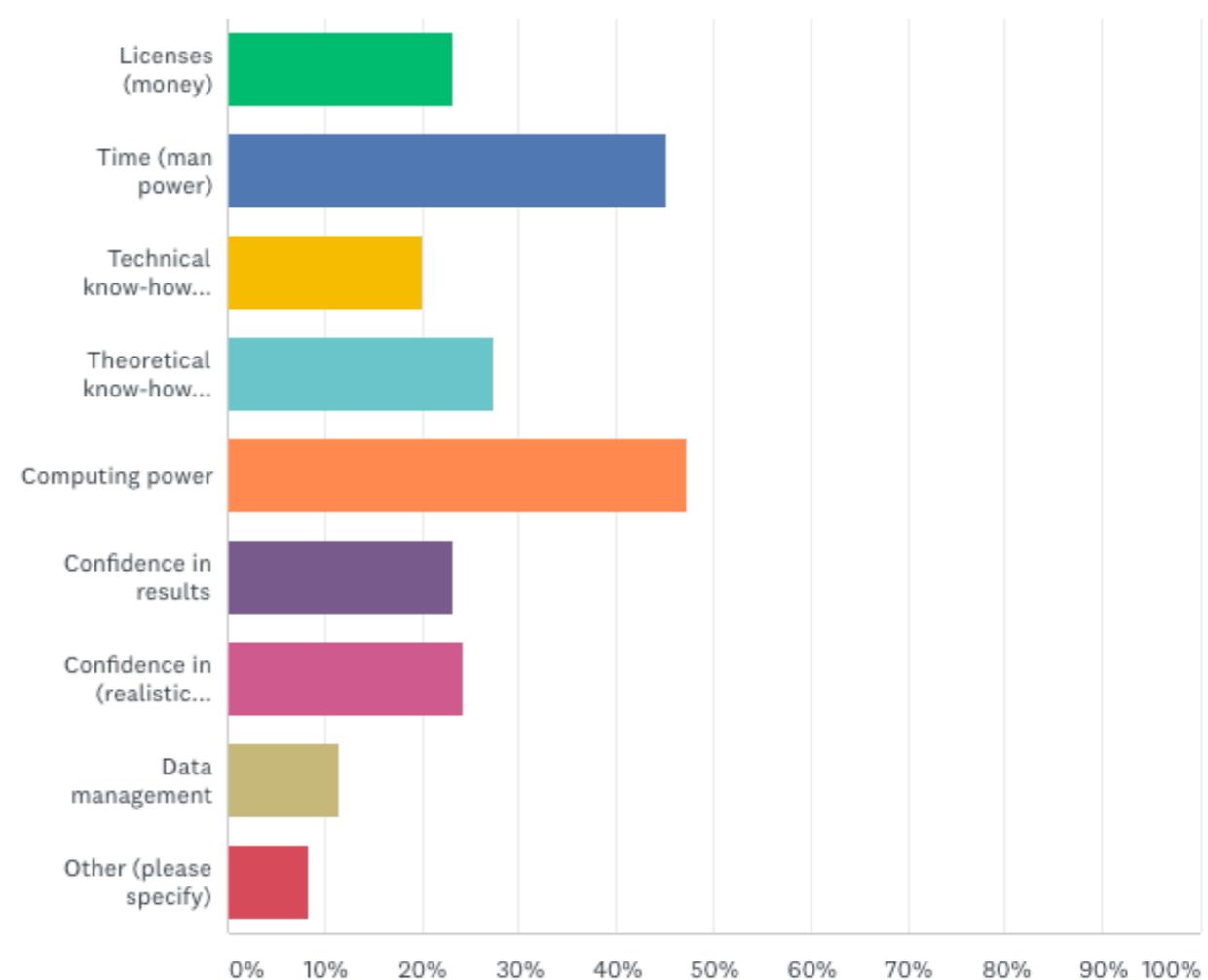
Please describe your most common implementation of ab initio computational chemistry. What question are you looking to answer.

Answered: 99 Skipped: 1



If you use QM regularly, what is the most significant bottleneck?

Answered: 95 Skipped: 5



How do we use it?

Shorter Lead-time to new drugs or materials

Prediction of new properties

Enhanced knowledge of molecular properties

Scoring and refinement of molecular libraries

Enhanced visual and quantitative communication of research topics

Rationalisation of experimental outcomes



Optimisation of existing processes

Help in the characterisation of novel molecules

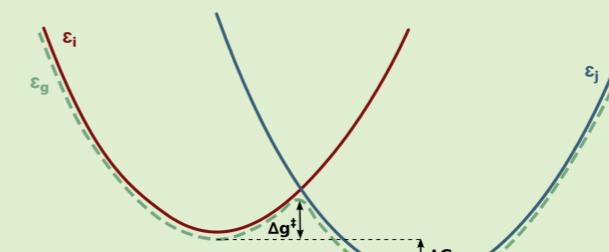
How do we use it?

Theory

$$\hat{H}_e \psi_e = E_e \psi_e$$

$$\hat{H}_e = \frac{-\hbar^2}{2m} \sum_i^{\text{electrons}} \nabla_i^2 - \sum_i^{\text{electrons}} \sum_A^{\text{nuclei}} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i < j}^{\text{electrons}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

Modelling



Experiments

Synthesis

Kinetics

spectroscopy



“Artificial Intelligence will not replace chemists. But chemists who don’t use (AI) will be replaced by those who do” Willem Van Hoorn

Computational Chemistry

How do we use it?

Which System Do I Have?



What Do I Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse

Computational Chemistry

How do we use it?

Which System Do I Have?



What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse

25

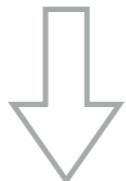
Computational Chemistry

How do we use it?

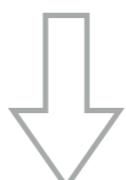
Which System Do I Have?



What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse

Computational Chemistry

How do we use it?

Which System Do I Have?



What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)

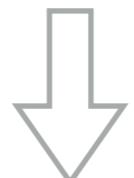


Interpret/Analyse

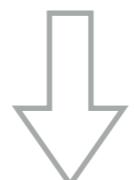
Computational Chemistry

How do we use it?

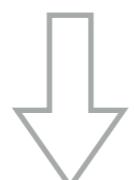
Which System Do I Have?



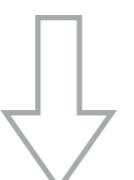
What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse



Often the most interesting result is when the “calculation gets it wrong”

Computational Chemistry

How do we use it?

Which System Do I Have?



What Do I Want to Compute (and Why)?



Which Model /Method Should I Choose?



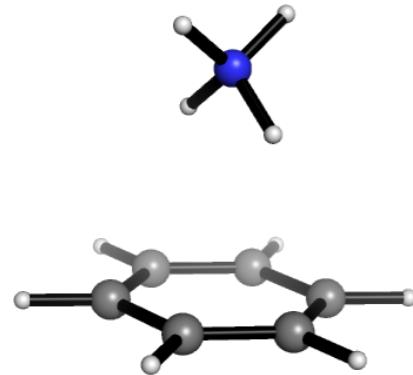
Verify Approach (vs. Experiment)



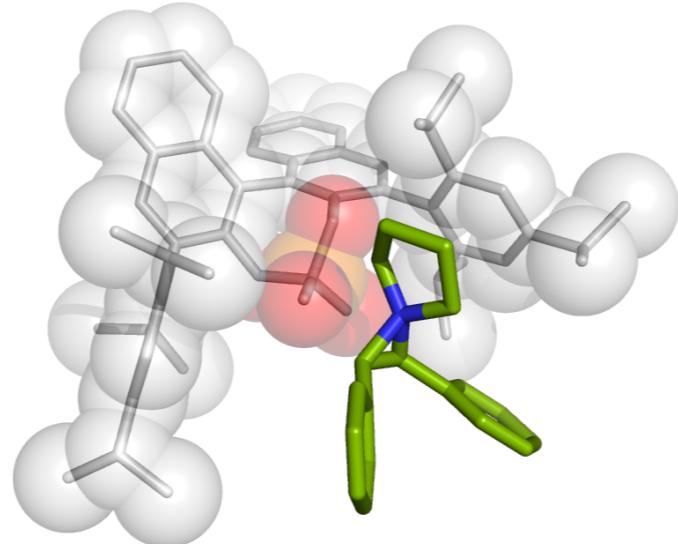
Interpret/Analyse

1. Which System Do I Have?

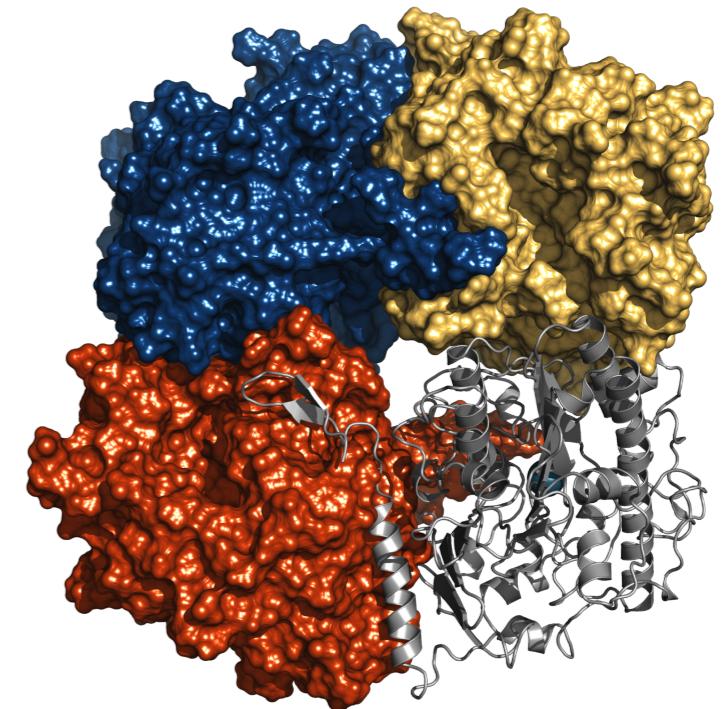
Size



17 atoms

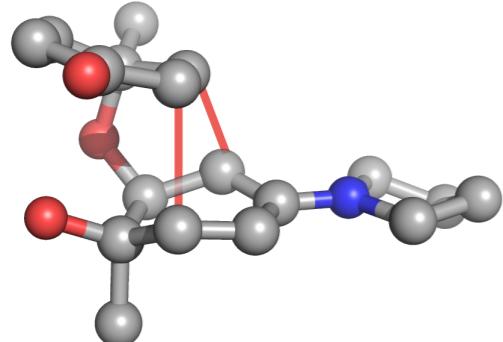


100 atoms

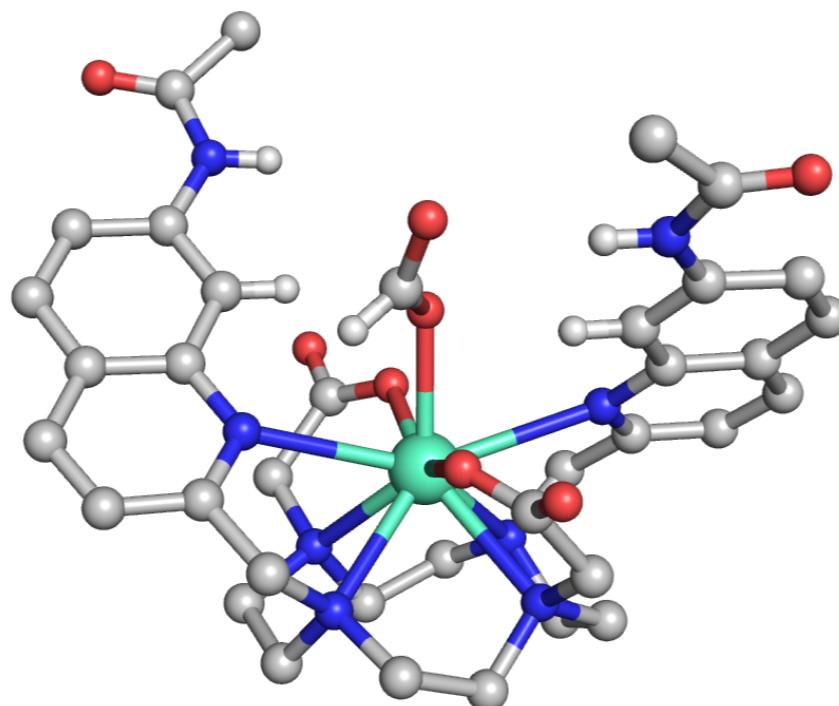


32000 atoms

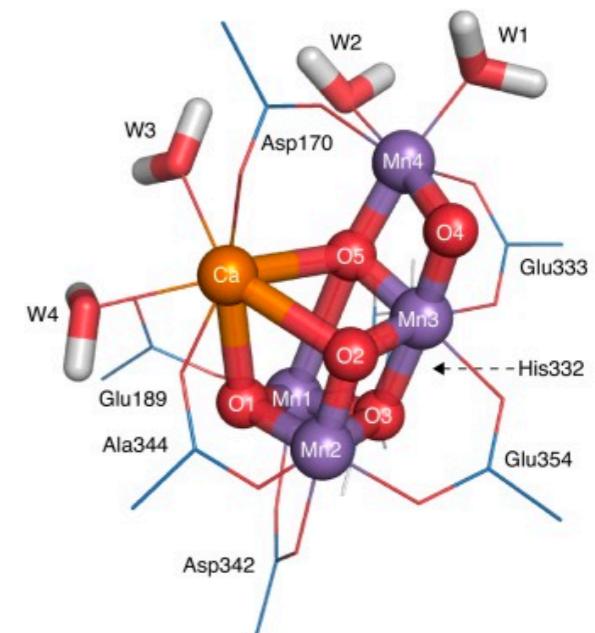
Complexity



Organic molecule



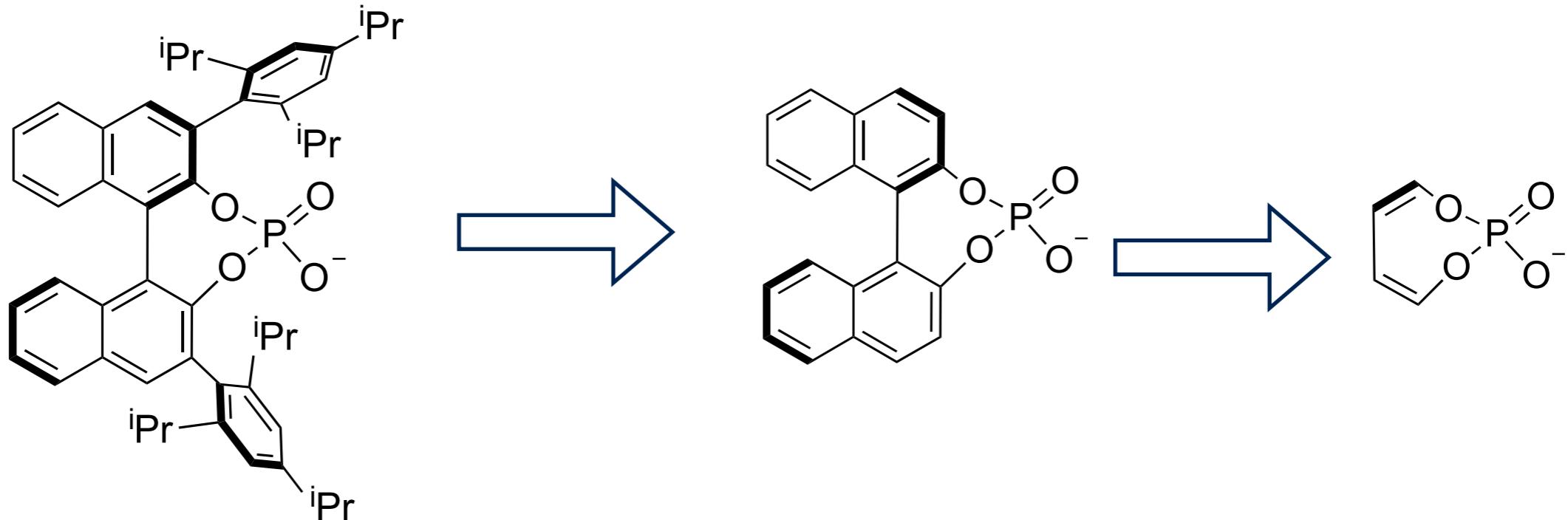
Eu-complex



Photosystem II
Multiple spin states

1. Which System Do I Have?

Model System



- Choice of model will depend on the problem and the available resources.
- Are the substituents required to understand the problem?
- A small model may be a good starting point.
- Factors out steric vs. electronic effects.

Computational Chemistry

How do we use it?

Which System Do I Have?



What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse

2. What Do You Want to Compute (and Why)?

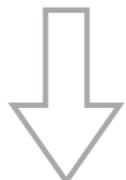
Geometry optimisation

IR and Raman intensities	Bonding properties	Reaction mechanisms
Kinetic Parameters		
Atomic charges & Multipoles		Optical spectroscopy
LFER	Magnetic coupling constants	
NMR chemical shifts		Ionization energies & Electron affinities
	EPR spectra	

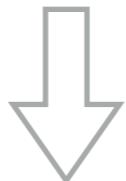
Computational Chemistry

How do we use it?

Which System Do I Have?



What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse

3. Which Method Should I Choose?

Electronic Schrödinger Equation

$$\hat{H}_e \psi_e = E_e \psi_e$$

$$\hat{H}_e = \frac{-\hbar^2}{2m} \sum_i^{electrons} \nabla_i^2 - \sum_i^{electrons} \sum_A^{nuclei} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i < j}^{electrons} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

3. Which Method Should I Choose?

Quantum Chemical Methods

Closed vs Open Shell

Single Reference vs Multireference (MR)

HF (SCF)

CASCF

CISD

MRCISD

CCSD

MRCCSD

MP2

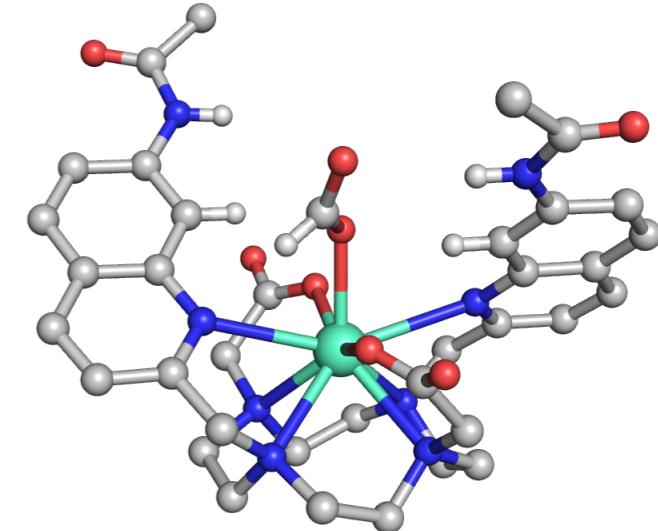
CASPT2

DFT ... works surprisingly well

3. Which Method Should I Choose?

Not easy to Answer

- TM systems are difficult to benchmark.
 - Size, electronic structure, spin-orbit relativistic effects..
 - Low amount and often poor quality of data ...
- Radicals
 - Spin contamination, multideterminantal character.
 - Result usually very sensitive to the level of theory chosen.



Often different approaches/levels of theory are required for different properties

3. Which Method Should I Choose?

Non-Empirical

John Perdew

PBE, TPSS, etc.



Semi-Empirical

Axel Becke

B3LYP, B97, etc.

- Functional form satisfies exact constraints
- Free parameters can be determined by fitting to appropriate norms

- Flexible functional form
- Fitting the coefficients to CCSD(T)/CBS limit



Combined

wB97: semi-empirical approach/ zeroth-order parameters constrained to be exact for infinite UEG.

MS1 mostly non-empirical, one of its parameters fitting to atomisation energies and barrier heights.

3. Which Method Should I Choose?

Non-Empirical

John Perdew

PBE, TPSS, etc.



Semi-Empirical

Axel Becke

B3LYP, B97, etc.

Popularity does not imply Accuracy

A Trip to the Density Functional Theory Zoo: Warnings and Recommendations for the User
Aust. J. Chem. 2019, 72, 563 <https://doi.org/10.1071/CH19023>

3. Which Method Should I Choose?

The Jacob's Ladder Classification of DFT (Perdew 2000)

Heaven of Chemical Accuracy

1 kcal/mol for reaction energies or barrier heights

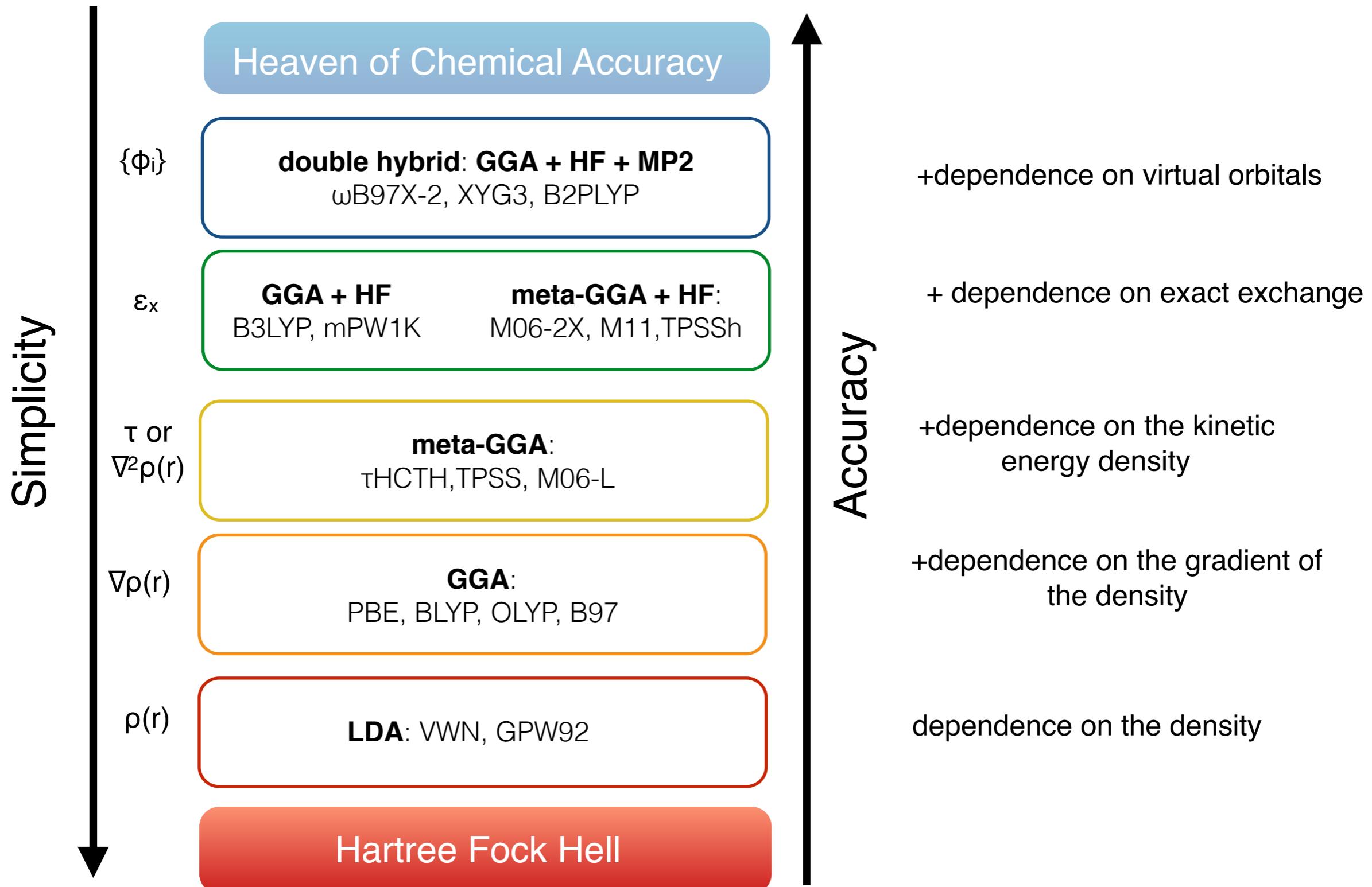
0.1 kcal/mol for noncovalent interaction energies

0.1 eV for electronic excitation energies.

Hartree Fock Hell

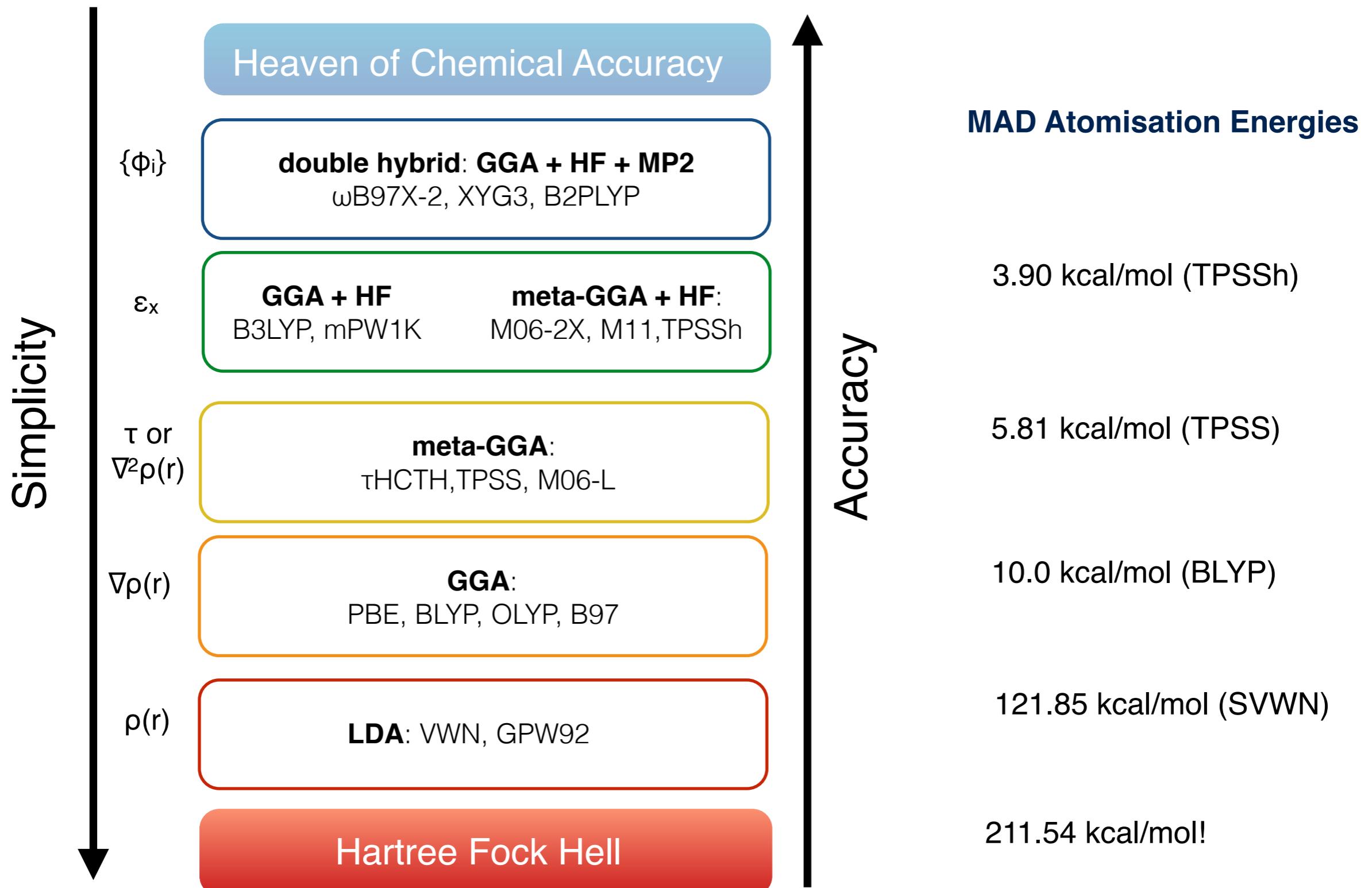
3. Which Method Should I Choose?

The Jacob's Ladder Classification of DFT (Perdew 2000)



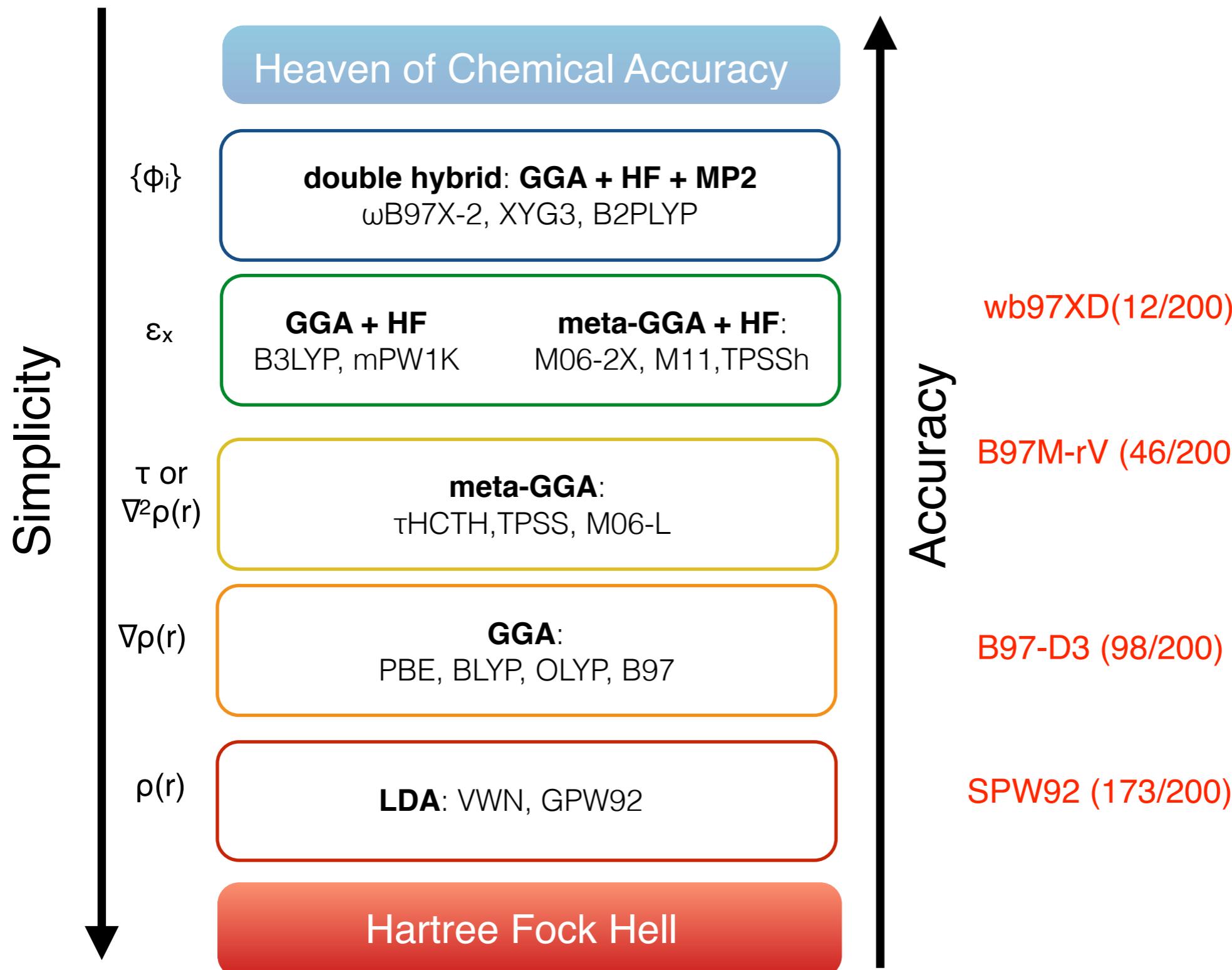
3. Which Method Should I Choose?

The Jacob's Ladder Classification of DFT (Perdew 2000)



3. Which Method Should I Choose?

The Jacob's Ladder Classification of DFT (Perdew 2000)



Functional/Basis Sets

During this afternoon tutorial you will use:

h2o_bad_geometry.inp

```
! BP86 def2-SVP SP

*xyz 0 1
H -1.0 0.0 0.0
O 0.0 0.0 0.0
H 1.5 0.0 0.0
*
```

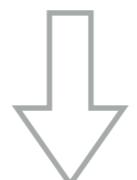
BP86 is a GGA functional composed of:
Becke 1988 exchange functional & Perdew 1986 correlation
functional

def2-svp: Split valence polarization basis set

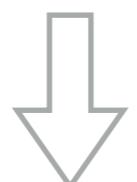
Computational Chemistry

How do we use it?

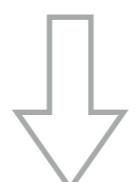
Which System Do I Have?



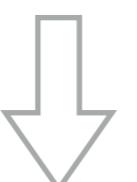
What Do You Want to Compute (and Why)?



Which Model /Method Should I Choose?



Verify Approach (vs. Experiment)



Interpret/Analyse



Often the most interesting result is when the “calculation gets it wrong”

Interpret/Analyse

Understand what you are comparing to ...

- Have ready your model/hypothesis
- Do you understand the experiment (theory, setup, sources of error)?
- Numerical precision (algorithm) vs Physical precision (comparable with experiment)
- Are there other interpretations I should consider?
- Test something that work and something that does not (false positive)

If the exp. method has a known error of 0.01 Å for a certain bond length, then there is no point in reporting geometries to 0.0000001 Å)