

Infraestrutura Nacional de **Computação Distribuída**



Course on Computational Biosciences using HPC systems

> 12-29 September 2022 - Online UCIBIO / INCD







FMUP FACULDADE DE MEDICINA

i4HB

Computational Biosciences Using HPC Systems – Module 6

Hybrid Quantum-Mechanics/Molecular Mechanics Methods

Sérgio F. Sousa & Arménio Barbosa

Email: sergiofsousa@med.up.pt UCIBIO/REQUIMTE, BioSIM – Departamento de Biomedicina Faculdade de Medicina da Universidade do Porto, Portugal

Email: aj.Barbosa@fct.unl.pt UCIBIO-i4HB, Biomolecular Engineering Lab – Departamento de Química School of Science and Technology, NOVA University Lisbon, Portugal

www.biosim.pt

https://sites.fct.unl.pt/biomolecular eng/

September 29, 2022



UCIBIO is formed by the collaborative efforts of researchers from the **University of Porto** and **University NOVA of Lisbon**.

UCIBIO's research activities occur mainly at the campus of Faculty of Sciences and Technology of the NOVA University of Lisbon (FCT-NOVA), and at the Faculty of Sciences (FCUP), the Faculty of Pharmacy (FFUP), the Faculty of Medicine (FMUP) and Instituto de Ciências Biomédicas Abel Salazar (ICBAS-UP) from the University of Porto.

UCIBIO combines key expertise in Chemistry and Biological Sciences with an ambitious strategic plan to maximize its national and international impact in terms of scientific productivity, advanced training and translation to society. In the national context, UCIBIO's key strength lies on its **broad scope of fundamental and applied research, standing at the interface of Chemistry, Biology and Engineering** to address pertinent questions at **atomic, molecular, sub-cellular and cellular levels, including cell-to-cell interactions and population evolutionary dynamics**.





BioSIM Research Group

Location

BioSIM – Biomolecular SIMulations Research Group UCIBIO/REQUIMTE - Departamento de Biomedicina Faculdade de Medicina da Universidade do Porto Portugal www.biosim.pt













BioSIM Research Group

Research

BioSIM – Biomolecular SIMulations Research Group

Our research team bridges the gap between theory and experiment applying and developing state-of-the-art computational tools focusing on Enzymatic Catalysis, Drug Discovery and Molecular Recognition

For that we combine: QM/MM Methods, Quantum Mechanics, Molecular Dynamics, Docking, Virtual Screening, and Free Energy Perturbation methods, always in close linking with experiment.

Several Software Applications and Databases have also been developed and made available to the scientific community.

Biomolecular Engineering Lab

Location

Biomolecular Engineering Lab UCIBIO-i4HB, – Departamento de Química School of Science and Technology, NOVA University Lisbon, Caparica, Portugal https://sites.fct.unl.pt/biomolecular_eng/











NOVA SCHOOL OF SCIENCE & TECHNOLOGY



Biomolecular Engineering Lab

Research

Biomolecular Engineering Lab

Composed by a multidisciplinary team with expertise in Applied and Computational Chemistry, Biotechnology, Biomedical Engineering, Chemical Engineering and Physics of Materials.

The group is dedicated to MINIMAL BIOMIMETIC SYSTEMS, combining designed molecular recognition agents with functional materials, for Bioseparation, Biocatalysis, Sensing & Diagnostics, and Nanomedicine.

The main modeling tools to acheive these goals are: Molecular Dynamics, Docking, Virtual Screening, Protein Structure Prediction, Biomaterials Simulations.

We have outstanding benefit for the *interconnected in silico and experimental* in impactful research.

Website: https://sites.fct.unl.pt/biomolecular_eng/home

Quantum Mechanics

Quantum Mechanics is a fundamental theory in physics that provides a description of the physical properties of nature at the scale of atoms and subatomic particles.

	Smallest Particle	Parameters	Bond-Breaking	Number of Atoms
QM	Electrons	Not Required	Possible	100-300 (with luck)
MM	Nuclei	Required	Not Directly Possible	easily 1,000,000

Classical Mechanics

Classical Mechanics

The study of the motion of bodies (including the special case in which bodies remain at rest) in accordance with the general principles first enunciated by Sir **Isaac Newton** in his *Philosophiae Naturalis Principia Mathematica* (1687), commonly known as the **Principia**.





Newton's Laws of Motion

1st Law: In an inertial frame of reference, an object either remains at rest or continues to move at a constant velocity, unless acted upon by a force

2nd **Law:** In an inertial frame of reference, the vector sum of the forces F on an object is equal to the mass m of that object multiplied by the acceleration a of the object: F = ma.

3rd **Law:** When one body exerts a force on a second body, the second body simultaneously exerts a force equal in magnitude and opposite in direction on the first body.

Classical Mechanics

If the present state of an object is known it is possible to predict by the laws of classical mechanics how it will move in the future (determinism) and how it has moved in the past (reversibility).



Mathematical formalism that can be used to calculate, predict and explain how macroscopic objects behave







Isaac Newton



Gottfried Leibniz



Leonhard Euler

Classical Mechanics *in everyday life*



Classical Mechanics *in everyday life*



Classical Mechanics *in everyday life* Classical mechanics uses <u>common-</u> <u>sense</u> notions of how matter and forces exist and interact. It assumes that matter and energy have definite, knowable attributes such as location in space and speed.

Predicts a precise trajectory for particles (i.e. precise locations and momenta at each instant)



Agrees with everyday experience We use it everyday We learn it from birth Classical Mechanics Classical mechanics provides extremely accurate results when studying large objects that are not extremely massive and speeds not approaching the speed of light



When the objects being examined have about the size of an atom diameter, classical mechanics fails... it becomes necessary to employ **Quantum Mechanics**

Failures of Classical Mechanics

1. Black-body Radiation

Any object with a temperature above absolute zero emits light at all wavelengths. If the object is perfectly black (so it doesn't reflect any light), then the light that comes from it is called blackbody radiation.

Explained by Max Planck (1905)

Quantitization of Energy



Max Planck Nobel Prize Physics (1918) Energy is not shared equally by electrons that vibrate with different frequencies. Planck said that energy comes in clumps. He called a clump of energy a **quantum**.





Failures of Classical Mechanics

2. Photoelectric Effect

When light shines on the surface of a metallic substance, electrons in the metal absorb the energy of the light and they can escape from the metal's surface. Classical physicists expected that when using very dim light, it would take some time for enough light energy to build up to eject an electron from a metallic surface.

Experiments show that if light of a certain frequency can eject electrons from a metal, it makes no difference how dim the light is. There is never a time delay.

Explained by Albert Einstein (1905)



Albert Einstein Nobel Prize Physics (1921)

A beam of light is not a wave propagating through space, but rather a collection of discrete wave packets (photons), each with energy hv.



Failures of Classical Mechanics

3. Hydrogen Atom

When a small tube of hydrogen gas is heated, it begins to glow and emit light. Unlike the blackbody radiation that comes from a hot dense solid or gas, this light consists of just a few colors (wavelengths): a red wavelength, a turquoise, and several violets.

Explained by Niels Bohr (1913)





Niels Bohr Nobel Prize Physics (1922) Energy levels of electrons are discrete and that the electrons revolve in stable orbits around the atomic nucleus but can jump from one energy level (or orbit) to another.





Wave-Particle Duality

Light acts as Particles!

The photoelectric effect proves that light consists of particles called photons.

Light acts as Waves!

When light passes through a double-slit, an interference pattern consisting of bright bands and dark bands is seen on a screen.

This is produced when the wave from one slit combines with the wave from the other slit. If two wave crests meet at the screen, the waves add and you get a bright band. If a wave crest from one slit meets a wave trough from the other slit, the waves cancel and you get a dark band. **This proves that light is a wave**.



Wave-Particle Duality

Wave-Particle Duality

Light acts like a wave if you want to know how it propagates, how it travels from one place to another. To describe how light travels from the double slits to the screen, you have to use the wave characteristics of light.

Light acts like particles (photons) if you want to know how light interacts with matter. To describer how light interacts with the electrons in a metal and how it ejects them from the metal's surface, you have to use the particle characteristics of light.



Louis de Broglie Nobel Prize Physics (1929)

- Electron behave also like waves
- Wave-particle duality applies to all matter in nature.
- Everything propagates like a wave
- Everything interacts like a particle





Quantum Mechanics

Wave-Particle Duality

How to describe?

Waves are spread out

Particles are *localized* they have a definite location not here



not here

Erwin Schrödinger Nobel Prize Physics (1933)

The Schrödinger Equation was proposed in 1926

Changed physics and chemistry forever Plays the role of Newton's Laws in classical mechanics

Rather than considering that a particle travels along a definite path, quantum mechanics acknowledges the wave-particle duality of matter by considering that **a particle is distributed through space like a wave**.

The classical notion of trajectory is hence replaced by a new concept: the wavefunction, Ψ (psi).

 ∇

The **Schrödinger Equation** allows the determination of the wavefunction of any system. In its barest form it states that:

 $\hat{H} \Psi = E \Psi$

H represents the Hamiltonian operator, which operates on a mathematical function Ψ , the wavefunction, to yield the energy (E) of the system. For a given particle of mass m, the Schrödinger equation can take the form:

$$\left\{-\frac{h^2}{8\pi^2 m}\nabla^2 + V\right\}\Psi(\vec{r},t) = \frac{ih}{2\pi}\frac{\partial\Psi(\vec{r},t)}{\partial t}$$

Where h is the Planck constant, is the Laplace operator, and V is the potential energy.

 ∇

The Wavefunction

The **energy and a variety of other properties** of a particle can be determined by solving the Schrödinger equation for Ψ .

Wavefunction contains all the information that is possible to obtain regarding the dynamical properties (including location and momentum) of a given particle.

For a molecular system, Ψ is a function of the positions of the nuclei and of the electrons.

 $\hat{H} \Psi(Q_1, ..., Q_N; q_1, ..., q_n) = E \Psi(Q_1, ..., Q_N; q_1, ..., q_n)$

Quantum Mechanics

The non-relativistic Hamiltonian for a molecular system

Includes operators that account for the kinetic energy of the nuclei and of the electrons, and for the electrostatic interactions between all the charged particles in the system (in atomic units).

$$\hat{H} = -\frac{1}{2} \sum_{X=1}^{N} \frac{1}{m_X} \nabla_X^2 + \sum_{X=1}^{N} \sum_{Y=1}^{X-1} \frac{Z_X Z_Y}{r_{XY}} - \frac{1}{2} \sum_{i=1}^{n} \nabla_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{i-1} \frac{1}{r_{ij}} - \sum_{X=1}^{N} \sum_{i=1}^{n} \frac{Z_X}{r_{iX}} \sum_{i=1}^{n} \frac{Z_X}{r_{iX}} + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{Z_Y}{r_{ij}} - \sum_{i=1}^{N} \sum_{j=1}^{n} \frac{Z_Y}{r_{ij}} + \sum_{i=1}^{n$$

Zx represents the charge of a nucleus X, rij is the distance between electrons i and j rix is the distance between a nucleus X and an electron i.

Energy and Other Properties can be Calculated Mathematically

From Quantum Mechanics to Quantum Chemistry Sadly however...

The Schrödinger equation can only be **solved analytically** for very simple systems.

E.g. include the hydrogen atom and other hydrogenic species (i.e. a one-electron atom or ion of general atomic number Z, as He⁺ and Li^{2+).}

Cannot be solved for many particle systems because of the **correlated motion of particles**. It contains pairwise attraction and repulsion terms, which in practice means that no particle is moving independently of all others.

For more complex systems other approximations are required

From Quantum Mechanics to Quantum Chemistry

Paul Dirac, expressed the problem in 1929 as follows:

"The fundamental laws necessary for the mathematical treatment of large parts of physics and the whole of chemistry are thus fully known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved."



Paul Dirac, 1929 Nobel Prize Physics (1933)

From Quantum Mechanics to Quantum Chemistry

Born-Oppenheimer Approximation (1927)

Assumes that nuclear and electronic motion can be decoupled

Hartree-Fock Theory (1928-1930)

Application of the Variational Principle + Use of a Single Slater Determinant to construct a trial Function

The Roothan Approximation (1951)

Representation of the spin-orbitals as a linear combination of a known set of basis functions

The Basis Set Approximation (1960s)

Molecular orbitals are expressed as a linear combination of atomic orbitals





Max Born Robert Nobel Prize Physics (1954) Oppenheimer





Douglas Hartree Vladimir Fock



CJ Roothan

From Quantum Mechanics to Quantum Chemistry





Max Born Robert Nobel Prize Physics (1954) Oppenheimer





Vladimir

Fock

Douglas Hartree

DO



CJ Roothan



Quantum Chemistry ...a reality

1960s

Nobel Prize in Chemistry (1998)



"to Walter Kohn for his development of the densityfunctional theory (DFT) and to John Pople for his development of computational methods in quantum chemistry."





Walter Kohn

John Pople

W. Kohn's theoretical work has formed the basis for simplifying the mathematics in descriptions of the bonding of atoms, a prerequisite for many of today's calculations (Density Functional Theory)

J. Pople developed the entire quantum-chemical methodology now used in various branches of chemistry

Quantum Chemistry ...a reality Geometry of a molecule Bond lenghts and angles Dipole moment Energy of reaction Reaction barrier height Vibrational frequencies IR spectra NMR spectra Reaction rates Partition function Free energy Atomization energies Binding Energies Ionization Potentials Electron Affinities Heats of Formation Excited States Etc...



Quantitative Information on Molecules and their Interactions Understanding of the molecular processes that cannot be obtained from experiments alone



Quantum Chemistry

Methods



Scaling Behaviour of Diferent Quantum Methods

With the number of basis functions (N)

Scaling Behaviour	Method(s)
N ³	DFT
\mathbf{N}^4	HF
N^5	MP2
\mathbf{N}^{6}	MP3, CISD
\mathbf{N}^7	MP4
N^8	MP5, CISDT
N^9	MP6
\mathbf{N}^{10}	MP7, CISDTQ

DFT = Best Compromise Between CPU Time and Accuraccy

Quantum Chemistry

Methods



Quantum Pharmacology

1976

"It is tempting to add a question mark to the title of this book. Even the generous spirited may feel that it is premature to start applying the methods of molecular quantum mechanics to the problems of medicinal chemistry, while sceptics will certainly feel that pharmacology is far too complex to be clarified by purely theoretical calculations"

"These doubtful views are not in line with the current state of either molecular pharmacology or molecular quantum mechanics. Both subjects have advanced to the point where the problems of one are susceptible of the methods of the other"





Richards, W. G. (1976). Quantum pharmacology (1st ed.). London ; Boston: Butterworths. <u>ISBN 9780408709507</u>.



Quantum Pharmacology





Figure 1 (a) Depth-cued stick representation and (b) a space-filling representation of the molecule thioburimamide. Drawn with the CHEMGRAF suite of programs originating from E. K. Davies, Chemical Crystallography Laboratory, Oxford University; now distributed by Chemical Design Ltd., Oxford.



Figure 7 Computer representation of electrostatic molecular potential on the surface of a nitrosamine (see ref. 10) produced by Dr P. Quarendon, IBM UK Scientific Centre.





Figure 6 Quasi-three-dimensional electron density maps of (a) acetylcholine and (b) nicotine. Contour at 0.01 a.u. of charge. Photographs by Lance Mangold, Department of Molecular Biophysics, Oxford, on an Evans and Sutherland monochromatic display system using multiple exposure and colour filtration.



Figure 8 Methotrexate located in the binding site of dihydrofolate reductase, drawn using an ICL PERQ computer using software developed by R. Hubbard (York University) and C. B. Naylor (OXford University).

Quantum Biochemistry

Nobel Prize in Chemistry (2013)

QM/MM Methods



"for the development multiscale models for complex chemical systems"

of

Martin Karplus Michael l evitt

Arieh Warshel

"This year's Nobel Laureates in chemistry took the best from both worlds and devised methods that use both classical and quantum physics. For instance, in simulations of how a drug couples to its target protein in the body, the computer performs quantum theoretical calculations on those atoms in the target protein that interact with the drug. The rest of the large protein is simulated using less demanding classical physics."

"Today the computer is just as important a tool for chemists as the test tube. Simulations are so realistic that they predict the outcome of traditional experiments."

From Quantum Mechanics to Quantum Chemistry

QM/MM Methods



Computational Enzymology

Study of Enzymatic Mechanisms by Quantum Mechanical/Molecular Mechanics Methods

Computational enzymology is a scientific subdiscipline that applies computational molecular simulation and modeling to enzymes to simulate enzymecatalyzed reactions, in an attempt to explain and rationalize experimental evidence.



Computational Enzymology

OM/MM in the Study of Enzymatic Mechanisms



- X-ray Crystallography
- EPR Data
- EXAFS Results
- UV Spectroscopy
- Mutagenesis
- Kinetic Results
- Biomimetic Compounds









Supercomputer



Typical Procedure

QM/MM in the Study of Enzymatic Mechanisms

1. Model Preparation

- From X-Ray, NMR or Homology
- Model substrates
- Assign protonation states
- Solvate system
- Assign MM parameters

2. Initial Optimization

- Relaxation with Molecular Dynamics
- Optimize structure with hybrid QM/MM
- QM region: Quantum Mechanics (B3LYP/6-31G(d))
- MM region: Biomolecular FF (e.g. AMBER)
- Optimize the structure of the reactants





Typical Procedure

QM/MM in the Study of Enzymatic Mechanisms

3. Linear Transit Scans

- Define possible reaction coordinates
- Exploration of the reaction coordinate from R-> P
- Evaluation of the variation in energy
- Identification of the maxima and optimization of the TS
- Optimization of all minima and intermediate structures

4. Activation and Reaction Energies

- Single-Point Energy Calculation with B3LYP/6-311++G(3df,3pd)
- Frequency calculation for thermal and entropic corrections
- PCM calculation in water
- Determination of the Activation Free Energy
- Determination of the Reaction Free Energy





Typical Procedure

No

QM/MM in the Study of Enzymatic Mechanisms

5. Comparison with Experimental Data

- Does the Activation Free Energy determined is in agreement with the experimental activation free energy determined from the k_{cat}?
- The atoms involved in the interactions formed in minima are in agreement with the experimental information arising from the experimental methods?
- The chemistry involved in the computational system explains all the experimental data?

Valid Mechanism Proposal (until new experimental available)

Yes

Need to Propose New Alternative Mechanism and repeat the calculations! OM/MM in the Study of Enzymatic Mechanisms

Importance in Drug Discovery

Enables an atomic level understanding of the catalytic activity of enzymes associated to specific diseases or important pathways

E.g. explanation of results, critical amino acid residues, proposed mutations, novel strategies to inhibit enzyme function and activity

The Structure of the Transition State provides a model for the development of novel inhibitors

Enzymes stabilize much more the transition state than the reactants

TS = **Pharmacophore**

TS= model for similarity-based VS





To use different methodologies in the treatment of different parts of the system.

QM/MM Hybrid Methods

The **QM region** comprises the part of the system directly involved in a given chemical reaction under study

The **MM region** encompasses the remaining of the system, where events like the formation and breaking of bonds do not take place.



MM as a Perturbation

In QM/MM Methods a chemical reaction is treated as a transformation involving only the QM region but that is influenced by the surrounding environment (the MM region), typically the solvent or the remaining portion of the protein in the case of an enzymatic reaction.

So:

The **MM region acts only indirectly on the electronically important region.** Hence, this interaction can be regarded as a mere **perturbation**, which can take several forms.

This perturbation can assume a **stereochemical** nature if the MM region imposes the adoption of a given geometry by the QM portion, but usually also includes **electrostatic** and **polarization** effects.

Advantages

- More computational time is spent in describing a small number of important atoms involved in the reaction, and which require a quantum mechanical description
- Most of the atoms in the systems participate only indirectly, and are described by molecular mechanics, ensuring its inclusion in the problem, albeit with at a lower level of detail

Examples of Application

- Enzymatic Role Played by Enzymes
- Influence of the Solvent in the Chemical Reactivity
- Study of localized chemical reactions in surfaces

The QM Layer

+

- The most popular choices have been semiempirical methods like AM1, and PM3, Hartree-Fock (HF) theory, Moller-Plesset Perturbation theory (MP), and most of all, the **Density Functional theory (DFT) method**.
- Among density functionals, **B3LYP** is the most common choice, but any alternative can be used.



Other Alternatives: Mo8 suite, MPWB1K, BB1K, etc

The QM Layer

+

- In QM/MM **geometry optimization**: changing the density functional has often small impact on the geometries
- In the determination of **Gibbs free energies**: The specific density functional or method can have a significant impact



The MM Layer

- The most commonly used force-fields in biomolecular studies are AMBER, CHARMM, and GROMOS.
- QM/MM results depend more on the QM level than on the specific force field used to treat the MM Layer.

$$\begin{split} E_{Total} &= \sum_{Bonds} K_r (r - r_{eq})^2 + \sum_{Angles} K_\theta (\theta - \theta_{eq})^2 \\ &+ \sum_{Dihedrals} \frac{V_n}{2} \Big[1 + \cos(n\phi - \gamma) \Big] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\varepsilon R_{ij}} \right] \end{split}$$

Example illustrating the generic form of the potential energy functions used

÷

- Gaussian (ONIOM method)
- ORCA
- QSite
- ChemShell
- LICHEM
- MODQ₃M
- QMS
- AMBER, CHARMM, GROMACS...

Size of the QM and MM Regions

QM Region

÷

< 250 Atoms (if DFT used)

Typical Values 100-160 atoms

MM Region

Typical Values

1,000 to 1,000,000



Choose a Structure





Approach

Choose a Structure

Prepare the Structure

Geometry Optimization







RC Reactant & Product





Sérgio F. Sousa – BioSIM – UCIBIO





bel

Example





Example



Example

