EURO²

Electronic structure calculations by Gaussian 16

Erol Yıldırım
Middle East Technical University

ncc@ulakbim.gov.tr

Key Concepts

- Computational Chemistry
- Common Computational Investigations
- Quantum Mechanics
- Tools of Computational Chemistry

EURO²

"All Theoretical Chemistry is really Physics; and all Theoretical Chemists know it".

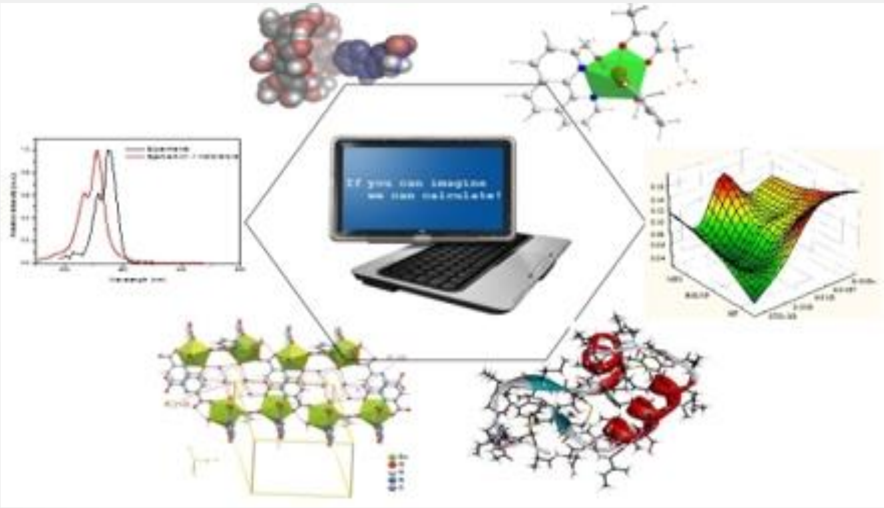
Richard Feynman

However,

"*There is a difference between knowing the rules of chess and being able to play*"

Computational Chemistry

- Computational chemistry (also called molecular modelling; the two terms mean about the same thing) is a set of techniques for investigating chemical problems on a computer rather than using chemicals.
- It uses the results of theoretical chemistry, incorporated into efficient computer programs, to calculate the structures and properties of molecules.



Common Computational Investigations

- Questions commonly investigated computationally are:

✓ Molecular geometry: the shapes of molecules – bond lengths, angles and dihedrals.

✓ Energies of molecules and transition states: this tells us which isomer is favored at equilibrium, and (from transition state and reactant energies) how fast a reaction should go.

✓ Chemical reactivity: for example, knowing where the electrons are concentrated (nucleophilic sites) and where they want to go (electrophilic sites) enables us to predict where various kinds of reagents will attack a molecule.

✓ IR, UV and NMR spectra: these can be calculated, and if the molecule is unknown, someone trying to make it knows what to look for.

✓ Large Scale Classical Simulations: Self-Organization, self-assembly.

✓ Large Scale Classical Simulations: Diffusion, melting, crystallization, folding… etc.

Quantum Mechanics

- The domain of physics that describes how electrons and protons interact is Quantum Mechanics.[1]

$$\hat{H}\Psi(\tau) = \mathcal{E}\Psi(\tau) \quad \hat{H} = \sum_{a=1}^{M}\sum_{b<a}^{M}\frac{Z_a \cdot Z_b}{r_{ab}} - \sum_{i=1}^{N}\sum_{a=1}^{M}\frac{Z_a}{r_{ia}} + \sum_{i=1}^{N}\sum_{j<i}^{N}\frac{1}{r_{ij}} - \sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{a=1}^{M}\frac{1}{2m_a}\nabla_a^2$$

- Models that solve the Schrödinger Equation are called *ab initio* (from the beginning). This is the realm of Quantum Chemistry.

[1] (Note: There is more than one "Quantum Theory" and some problems in Chemistry require to go beyond the Schrödinger Equation. For example, the Dirac Equation for systems with heavy atoms to include the effects of Special Relativity Theory or Quantum Electrodynamics for highly accurate descriptions).

Tools of Computational Chemistry

- Computational chemists have a selection of methods at their disposal. The main tools available belong to five broad classes:

✓ Molecular Mechanics

✓ Molecular Dynamics

✓ *Ab initio* calculations

✓ Semiempirical methods

✓ Density Functional Theory

Molecular Mechanics

- Molecular mechanics is based on a model of a molecule as a collection of balls (atoms) held together by springs (bonds).
- By knowing the spring lengths, their angles, and how much energy it takes to stretch and bend the springs, we can calculate the energy of a given collection of balls and springs, i.e., of a given molecule.
- Geometry is changed until the lowest energy is found enables us to do a geometry optimization.
- Molecular mechanics is fast: a fairly large molecule like a steroid (e.g., cholesterol, $C_{27}H_{46}O$) can be optimized in seconds on a good personal computer.

Molecular Dynamics Simulations

- Molecular dynamics calculations apply the laws of motion to molecules.
- Thus, one can simulate the motion of an enzyme as it changes shape on binding to a substrate, or the motion of a swarm of water molecules around a molecule of solute.
- Quantum mechanical molecular dynamics also allows actual chemical reactions to be simulated.

Ab initio Calculations

- *Ab Initio* calculations (*ab initio*, Latin: “from the start”, i.e., from first principles”) are based on the Schrödinger equation.
- This is one of the fundamental equations of modern physics and describes, among other things, how the electrons in a molecule behave.
- The *ab initio* method solves the Schrödinger equation for a molecule and gives us an energy and *wavefunction*.
- The *wavefunction* is a mathematical function that can be used to calculate the electron distribution (and, in theory at least, anything else about the molecule).

Ab initio Calculations

- *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 is only that that the exact application of these laws leads to equations much too complicated to be solvable.* – **P.A.M. Dirac**
- The challenge in computational chemistry is to simplify the calculation enough to be solvable, but still accurate enough to predict the desired physical quantity.
- There is an enormous toolbox of theoretical methods available, and it will take skill and creativity to solve real-world problems.

Semiempirical Calculations

- Semiempirical calculations are, like *ab initio*, based on the Schrödinger equation.
- Here, more approximations are made in solving it, and the very complicated integrals that must be calculated in the *ab initio* method are not actually evaluated.
- Instead, the program draws on a kind of library of integrals that was compiled by finding the best fit of some calculated entity like geometry or energy (heat of formation) to the experimental values.
- This plugging of experimental values into a mathematical procedure to get the best calculated values is called *parameterization*.

Semiempirical Calculations

- It is the mixing of theory and experiment that makes the method "*semiempirical*".
- It is based on the Schrödinger equation, but parameterized with experimental values (*empirical* means experimental).
- Semiempirical calculations are slower than molecular mechanics but much faster than *ab initio* calculations.
- Semiempirical calculations take roughly 100 times as long as molecular mechanics calculations, and *ab initio* calculations take roughly 100–1,000 times as long as semiempirical.

Density Functional Theory Calculations

- Density functional calculations (DFT calculations) are, like *ab initio* and semiempirical calculations, based on the Schrödinger equation.
- However, unlike the other two methods, DFT does not calculate a conventional wavefunction, but rather derives the electron distribution (*electron density function*) directly.
- A *functional* is a mathematical entity related to a function.
- Density functional calculations are usually faster than *ab initio*, but slower than semiempirical.

Different Density Functionals

Exchange	Pure	Hybrid		Range separated hybrid	Long range correction
HFS	VSXC	B3LYP	X3LYP	HSEH1PE	LC-wPBE
XAlpha	HCTH	B3P86	BMK	OHSE2PBE	CAM-B3LYP
HFB	HCTH93	B3PW91	M06-HF	OHSE1PBE	wB97XD
	HCTH147	B1B95	M06-2X	wB97XD	
	B97D3	mPW1PW91	PBEh1PBE	wB97	
	M06L	PBE1PBE		wB97X	
		mPW1PBE		M11	

Tools of Computational Chemistry

- Very large biological molecules are studied mainly with molecular mechanics.
- Novel molecules, with unusual structures, are best investigated with *ab initio* or possibly DFT calculations.

Computational Chemistry

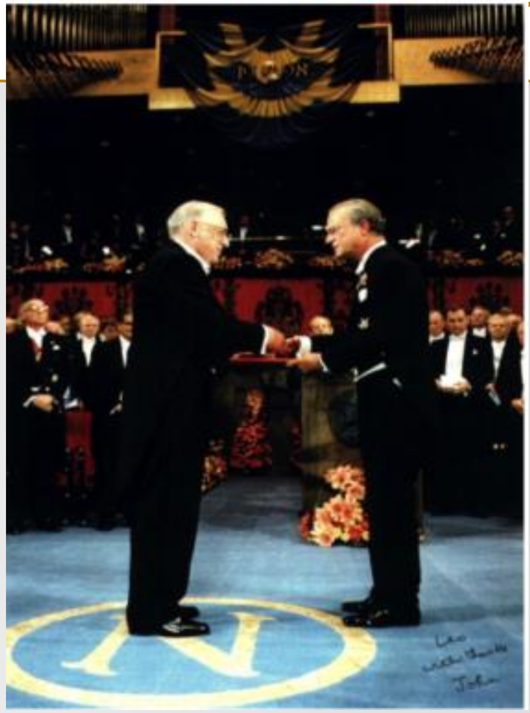
Lecture # 02

Key Concepts

- In this lecture we ‘ll learn:
- Nobel Recognition of Computational Chemistry
- Advantages of Computational Chemistry
- Disadvantages of Computational Chemistry
- Theoretical Model
- Molecular Structure

Nobel Recognition of Computational Chemistry

- The 1998 Nobel Prize in Chemistry was awarded to ***Walter Kohn*** "for his development of the density functional theory" and ***John Pople*** "for his development of computational methods in quantum chemistry".
- In 2013, ***Martin Karplus***, ***Michael Levitt***, and ***Arieh Warshel*** have been awarded the Nobel Prize in Chemistry for the development of computer-based methods to model complex systems.



Advantages of Computational Chemistry

- Calculations are *easy to perform*, whereas experiments are often difficult.
- Calculations are becoming *less costly*, whereas experiments are becoming more expensive.
- Calculations can be *performed on any system*, even those that don't exist, whereas many experiments are limited to relatively stable molecules.
- Calculations are *safe*, whereas many experiments have an intrinsic danger associated with them.

Disadvantages

- Calculations can be *very expensive* in terms of the amount of time required.
- Calculations can be performed on any system, even those that don't exist!

Computational chemistry is not a replacement for experimental studies, but plays an important role in enabling chemists to:

- Explain and rationalize known chemistry
- Explore new or unknown chemistry

Theoretical Model

- The theoretical foundation for computational chemistry is the time-independent Schrodinger wave equation:

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

- Ψ is the wavefunction. It is a function of the positions of all the fundamental particles (electrons and nuclei) in the system.
- $\hat{\mathrm{H}}$ is the Hamiltonian operator. It is the operator associated with the observable energy.
- E is the total energy of the system. It is a scalar (number).
- The wave equation is a postulate of quantum mechanics.

The Hamiltonian

- The Hamiltonian, $\hat{H}$, is an operator. It contains all the terms that contribute to the energy of a system:

$$\hat{H} = \hat{T} + \hat{V}$$

- $\hat{T}$ is the kinetic energy operator:

$$\hat{T} = \hat{T}_e + \hat{T}_n$$

$$\hat{T}_e = -\frac{1}{2}\sum_i \nabla_i^2 \qquad \hat{T}_n = -\frac{1}{2M_A}\sum_A \nabla_A^2$$

- ∇^2 is the Laplacian given by:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

The Hamiltonian

- $\hat{V}$ is the potential energy operator::

$$\hat{V} = \hat{V}_{nn} + \hat{V}_{ne} + \hat{V}_{ee}$$

- $\hat{V}_{nn}$ is the nuclear-nuclear repulsion term:

$$\hat{V}_{nn} = \sum_{A<B} \frac{Z_A Z_B}{R_{AB}}$$

- $\hat{V}_{ne}$ is the nuclear-electron attraction term:

$$\hat{V}_{ne} = -\sum_{iA} \frac{Z_A}{R_{iA}}$$

- $\hat{V}_{ee}$ is the electron-electron repulsion term:

$$\hat{V}_{ee} = \sum_{i<j} \frac{1}{r_{ij}}$$

Atomic Units

- All quantum chemical calculations use a special system of units which, while not part of the SI, are very natural and greatly simplify expressions for various quantities.
  - The length unit is the ***bohr*** ($a_0 = 5.29 \times 10^{-11}$m)
  - The mass unit is the ***electron mass*** ($m_e = 9.11 \times 10^{-31}$kg)
  - The charge unit is the ***electron charge*** ($e = 1.60 \times 10^{-19}$C)
  - The energy unit is the ***hartree*** ($E_h = 4.36 \times 10^{-18}$ J)

- For example, the energy of the H atom is -0.5 hartree. In more familiar units this is −1,313 kJ/mol

The Born-Oppenheimer approximation

EURO²

- The Born-Oppenheimer Approximation is the assumption that the electronic motion and the nuclear motion in molecules can be separated. It leads to a molecular wave function in terms of electron positions and nuclear positions.
- Nuclei are much heavier than electrons (the mass of a proton ≈ 2000 times that of an electron) and therefore travel much more slowly.
- The electronic wavefunction depends upon the nuclear positions but not upon their velocities, i.e., the nuclear motion is so much slower than electron motion that they can be considered to be fixed.

The Chemical Connection

- So far, we have focused mainly on obtaining the ***total energy*** of our system.
- Many chemical properties can be obtained from *derivatives* of the energy with respect to some *external parameter*
- Examples of external parameters include:
    - Geometric parameters (bond lengths, angles etc.)
    - External electric field (for example from a solvent or other molecule in the system)
    - External magnetic field (NMR experiments)

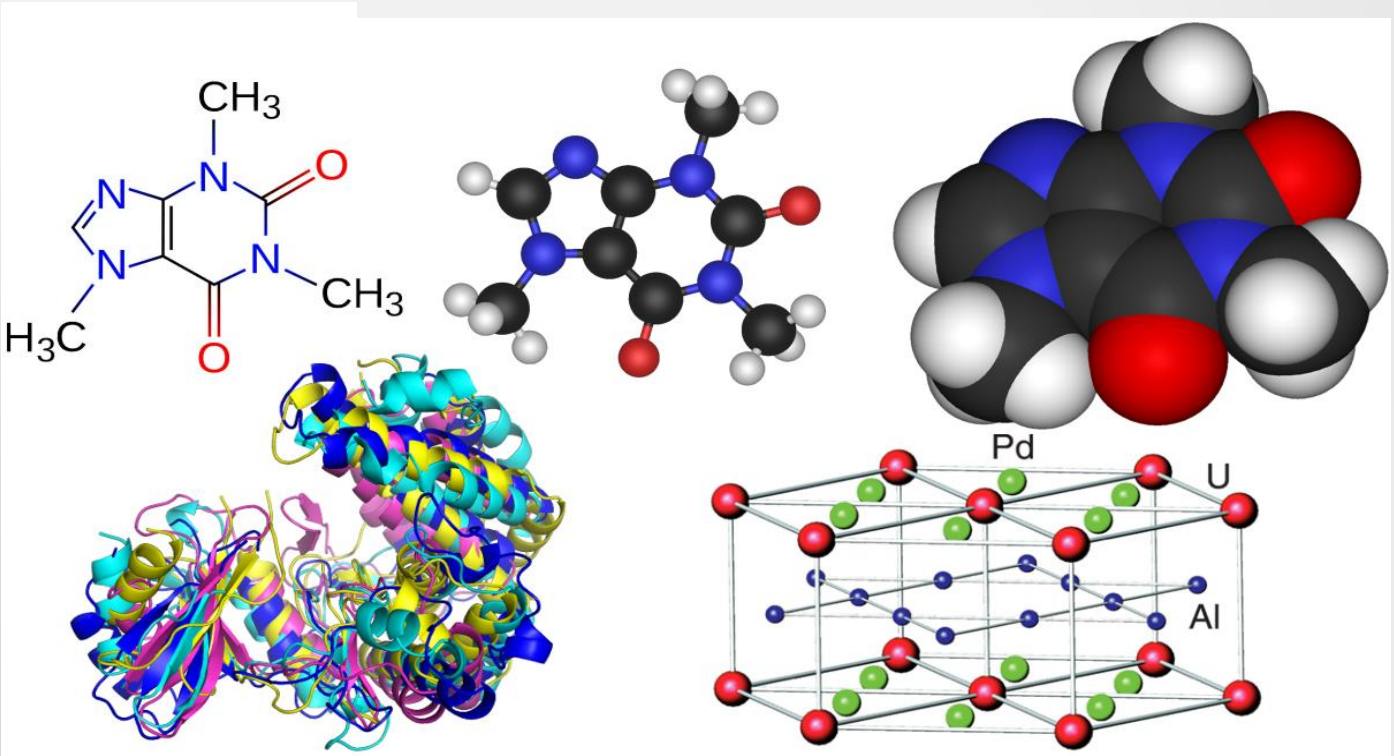
Computable Properties

Many molecular properties can be computed, these include:

- Bond energies and reaction energies
- Structures of ground-, excited- and transition-states
- Atomic charges and electrostatic potentials
- Vibrational frequencies (IR and Raman)
- Transition energies and intensities for UV and IR spectra
- NMR chemical shifts
- Dipole moments, polarizabilities and hyperpolarizabilities
- Reaction pathways and mechanisms

Molecular Structure

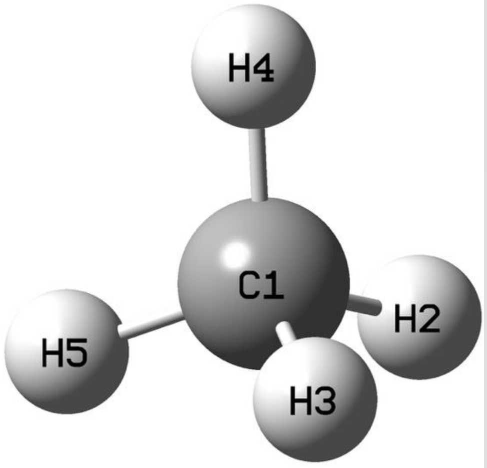
- The first thing most chemists think about when they hear the name of a compound is the structure.



Molecular Structure

Every atom has three cartesian coordinates, methane:

```
C     -1.69999999     1.55000011     0.00000000
H     -1.34334556     0.54119011     0.00000000
H     -1.34332715     2.05439830     0.87365150
H     -1.34332715     2.05439830    -0.87365150
H     -2.76999999     1.55001330     0.00000000
```



Or, we only describe the internal relationships between the atoms:

```
C
H     1  1.07
H     1  1.07    2  109.47
H     1  1.07    2  109.47    3   120.0
H     1  1.07    2  109.47    3  -120.0
```

distance of this atom to atom #1 = 1.07 Å

angle between this atom, the atoms to which the distance is defined (#1) and atom #2 = 109.47°

dihedral angle between this atom, the atoms to which distance and angle are defined (#1 & #2) and atom #3 = 120.0°

Molecular Structure

Or, we only describe the internal relationships between the atoms:

```
C
H     1 1.07
H     1 1.07     2 109.47
H     1 1.07     2 109.47     3  120.0
H     1 1.07     2 109.47     3 -120.0
```

This is called a Z-Matrix, one possible representation of *internal coordinates*. We can see that there are 6 fewer coordinates than in the xyz (*cartesian coordinate*) case (you will remember "3N-6 molecular degrees of freedom" from spectroscopy or thermodynamics).

Further, in this example we can take advantage of the symmetry of the molecule by realizing that all bond lengths are identical (1.07 Å), all angles are identical (109.47°) and both dihedral angles are identical (±120.0°).

Even more, only the bond length is a true variable. Because of molecular symmetry, the angles and the dihedral angle are constants.

Molecular Structure

- Following are different parts of a *Gaussian* calculation:

```
%nprocshared=2

%chk=water.chk

# opt freq b3lyp/sto-3g scrf=(solvent=water) geom=connectivity

Water Molecule Optimization

0 1
 O                 -0.26319297   -0.17756711    0.02240362
 H                  0.69680703   -0.17756711    0.02240362
 H                 -0.58364756    0.72736872    0.02240362

 1 2 1.0 3 1.0
 2
 3
```

- Shared Processors
- Checkpoint File
- Route Card
- Title
- Charge & Multiplicity
- Structure Coordinates
- Connectivity Information

Gaussian Input File

1. Link 0 Commands: -set up memory limits, etc. Line starts with %. (Optional).

2. Route Section: -specifies the details of the calculation
-can be multiple lines with max. 80 characters
-each line in Route Section must start with #

3. Blank Line: -tells program Route Section is done

4. Title

5. Blank Line: -tells program Title is done

6. Charge and Multiplicity

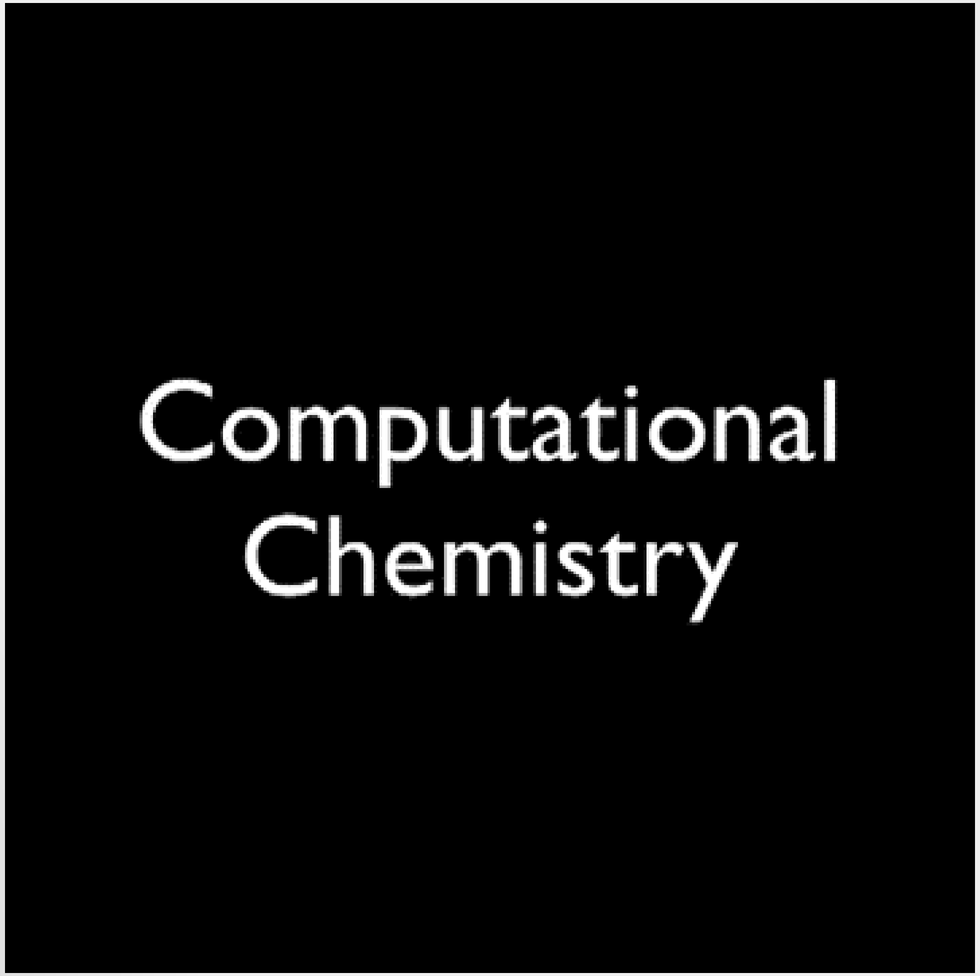
7. Molecular Geometry: -provide the atomic coordinates
-Cartesian or Z-matrix format

8. Blank Line: -tells program the input file is done

Route Card / Route Line

- The route line contains (but not limited to) the following information:
- Property to be calculated (Energy, Optimization, Frequency, UV, NMR, Scan etc.)
- Functional (Theory, like DFT, Hartree Fock etc.)
- Basis Set
- Solvation Information (Solvent to be used)
- Other parameters (SCF tightness, convergence criterion, etc.)

What is a basis set? The molecular spin-orbitals that are used in the Slater determinant usually are expressed as a linear combination of some chosen functions, which are called basis functions. This set of functions is called the basis set.


Computational
Chemistry

Acknowledgement

Introduction to Computational Chemistry

Muhammad Ali Hashmi, University of Education, Lahore, Pakistan

Thanks

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101101903. The JU receives support from the Digital Europe Programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia