Electronic Properties Of The Protein From The Oxidoreductase Family

Matekalo, Jelena

University of Split
Faculty of Science

ELECTRONIC PROPERTIES OF THE PROTEIN FROM THE OXIDOREDUCTASE FAMILY

Master thesis

Jelena Matekalo

Split, September 2023.

I want to extend my heartfelt appreciation to my mentor, Željka Sanader Maršić, and co-mentor, Martina Perić Bakulić, for their outstanding guidance, leadership, and invaluable advice throughout my master's thesis journey.

Additionally, I would like to express my deep gratitude to my dear family—my mother Ana, my sister Željka, and my brother Marko. Their unwavering belief in me and their continuous encouragement have provided me with the motivation and strength to reach this milestone.

Temeljna dokumentacijska kartica

Sveučilište u Splitu
Prirodoslovno – matematički fakultet
Odjel za fiziku
Ruđera Boškovića 33, 21000 Split, Hrvatska

Završni / Diplomski rad

Elektronska svojstva proteina iz obitelji oksidoreduktaza

Jelena Matekalo

Sveučilišni diplomski studij Fizika, smjer Biofizika

Sažetak:
Prijenos elektrona između redoks proteina PQQ-GDH i ugljikove nanocijevi, koja djeluje kao bioelektroda, značajno je područje istraživanja je u posljednjih nekoliko desetljeća. U ovom razdoblju fokus je na ispitivanju prikladnosti ugljikovih nanocijevi kao bioelektroda. Razlog tome je biomedicinske prirode, odnosno prijenos elektrona između proteina i bioelektroda temelj je razvoja malih medicinskih uređaja. U stanicama proteini se nalaze u staničnom mediju i većinom imaju slobodno gibanje. Međutim, da bi bili dio bioelektroda, moraju biti imobilizirani na površini. Zato je bitno rauzmjeti detaljno mehanizam njihove aktivnosti, kako bi se moglo utjecati na tu aktivnost i osigurati da su enzimi aktivni i kad su imobilizirani. Začetak ovog istraživanja je na promatranju elektronske strukture aktivnog mjesta sustava proteina pirolokinolin kinon ovisna glukoza dehidrogenaza (PQQ-GDH) prije te nakon što se dogodi reakcija prijenosa elektrona. Istraživanje je odrađeno metodom teorije funkcionala gustoće (DFT), jednom od metoda kvantne mehanike, kojom je određena molekularna geometrija te elektronska struktura proučavanog molekularnog sustava.

Ključne riječi: PQQ-GDH, enzim, prijenos elektrona, teorija funkcionala gustoće, HOMO-LUMO jaz, molekularna orbitala, Mulliken naboj

Rad sadrži: 26 stranica, 14 slika, 16 literaturnih navoda. Izvornik je na engleskom jeziku.

Mentor: doc. dr. sc. Željka Sanader Maršić

Neposredni voditelj: doc. dr. sc. Martina Perić Bakulić

Ocjenjivači: doc. dr. sc. Željka Sanader Maršić
doc. dr. sc. Martina Perić Bakulić
izv. prof. dr. sc. Bernarda Lovrinčević

Rad prihvaćen: 22. 9. 2023.

Rad je pohranjen u knjižnici Prirodoslovno – matematičkog fakulteta, Sveučilišta u Splitu.

Basic documentation card

University of Split
Faculty of Science
Department of Physics
Ruđera Boškovića 33, 21000 Split, Croatia

Bachelor / Master thesis

Electronic Properties of the Protein form the Oxidoreductase Family

Jelena Matekalo

University graduate study programme Physics, orientation Biophysics

Abstract:
Electron transfer between the redox protein PQQ-GDH and carbon nanotubes, acting as a bioelectrode, has been accomplished in recent decades. In this time, the focus has been on investigating the suitability of carbon nanotubes as bioelectrodes. The reason for this lies in the biomedical applications, specifically the electron transfer between proteins and bioelectrodes, which forms the basis for the development of small medical devices. In cells, proteins are found in the cellular medium and mostly have free movement. However, to become a part of a bioelectrode, they need to be immobilized on the surface. Therefore, it is crucial to thoroughly understand the mechanism of their activity in order to influence this activity and ensure that enzymes remain active even when immobilized. The initiation of this research involves observing the electronic structure of the active site of the PQQ-GDH system before and after the electron transfer reaction occurs. The study was conducted using the Density Functional Theory (DFT) method, a Quantum Mechanical (QM) method, to determine the molecular geometry and electronic structure of the molecular system.

Keywords: PQQ-GDH, enzyme, electron transfer, density functional theory, HOMO-LUMO gap, molecular orbitale, Mulliken charge

Thesis consists of: 26 pages, 14 figures, 16 references. Original language: English

Supervisor: Assist. Prof. Dr. Željka Sanader Maršić

Leader: Assist. Prof. Dr. Martina Perić Bakulić

Reviewers: Assist. Prof. Dr. Željka Sanader Maršić
Assist. Prof. Dr. Martina Perić Bakulić
Asoc. Prof. Dr. Bernarda Lovrinčević

Thesis accepted: September 22, 2023.

Thesis is deposited in the library of the Faculty of Science, University of Split.

Contents

1. Theoretical Introduction

1.1. Electron Transfer (ET)

Electron transfer (ET) reactions are fundamental reactions in many fundamental biological processes such as photosynthesis, respiration, DNA photodamage repair, *etc*. These are reactions in which a single electron is transferred from one molecule to another or from one part of a molecule to another. The understanding of the ET is based on studying and understanding biological transformations on the subatomic scale. The mentioned biological transformations occur due to the execution of the subatomic processes within the molecules that participate in and lead the ET between different substrates. Molecules that catalyse the release of electrons are called oxidases, while molecules that catalyse the acceptance of electrons are called reductases. This chemical reaction is known as redox reaction and includes cofactors that release and accept electrons called donor and acceptor. Most oxidoreductases contain one cofactor, while there are also oxidoreductases with multiple cofactors where one of cofactors is catalytic centre and others only participate in the ET. When two substrates and two products participate in a redox reaction it is usually called a bi-bi reaction, schematically represented in Figure 2. [1, 2, 3]

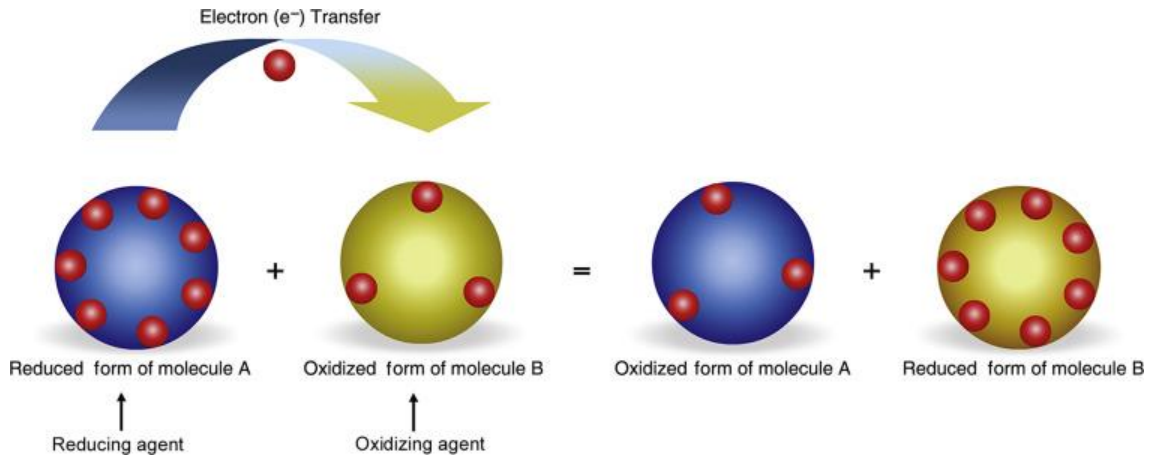

Figure 1: Schematic representation of the ET in a bi-bi reaction. [4]

One of the frequent examples of a bi-bi reaction is glucose oxidation, where co-factor flavin adenine dinucleotide (FAD) participates in the ET: [1]

$$D - glucose + O_2 \rightarrow D - glucono - 1{,}5 - lactone + H_2O_2 \qquad 1.1$$

1.1.1. ET in Photosynthesis

One of the most researched examples of fundamental biological processes in which electron transfer occurs is photosynthesis. Photosynthesis is the source of almost all carbon compounds and the entire amount of oxygen that enables aerobic metabolism. Photosynthesis is the biochemical process in which light energy is converted into chemical energy stored in organic molecules occurred in the chloroplast of the photosynthetic organisms. Protein complexes that participate in photosynthesis are photosystem II, cytochrome b6f complex, photosystem I, ferredoxin-NADP reductase (FNR) and adenosine triphosphate (ATP) synthase. The main photoreceptors in chloroplasts are chlorophylls, which have strong absorption bands in the visible part of the spectrum. The "trapping" of light energy takes place in the chlorophyll, which is the key event of photosynthesis. An electron separated from its original molecule can reduce other molecules and this is how solar energy is stored in chemical form. Electron transfer begins with the oxidation of the water to the oxygen and then absorption of light excites an electron from its ground state to a higher energy state (excited state) which is shown in Figure 3. After excitation, the electron goes through the complex path that includes molecules: pheophytin, plastoquinone, mobile plastoquinones, cytochrome complex, plastocyanin, P700 enzyme, phylloquinone, ferredoxin. The last electron transporter in this series is the flavoprotein ferredoxin-NADP reductase, an enzyme that transfers electrons from reduced ferredoxin to $NADP^+$ according to the equation:

$$2Fd + 2H + + NADP^+ \rightarrow 2Fd + NADPH + H^+ \qquad 1.2$$

and thus completes the sequence of electron transfer that began with the oxidation of water. A molecule that is finally reduced during the action of the photosynthetic system is nicotinamide adenine dinucleotide phosphate (NADP) and water is oxidized.

Figure 3 shows the electron transfer chain in the reaction centre of photosynthetic bacteria, the so-called Z-scheme. The Z-scheme represents the steps in the light reactions, showing the pathway of electron transport from water to $NADP^+$, what leads to the release of oxygen, the "reduction" of $NADP^+$ to NADPH. The described electron transport shows that photosynthesis is the example of truly complex electron transfer that occurs in nature. [5]

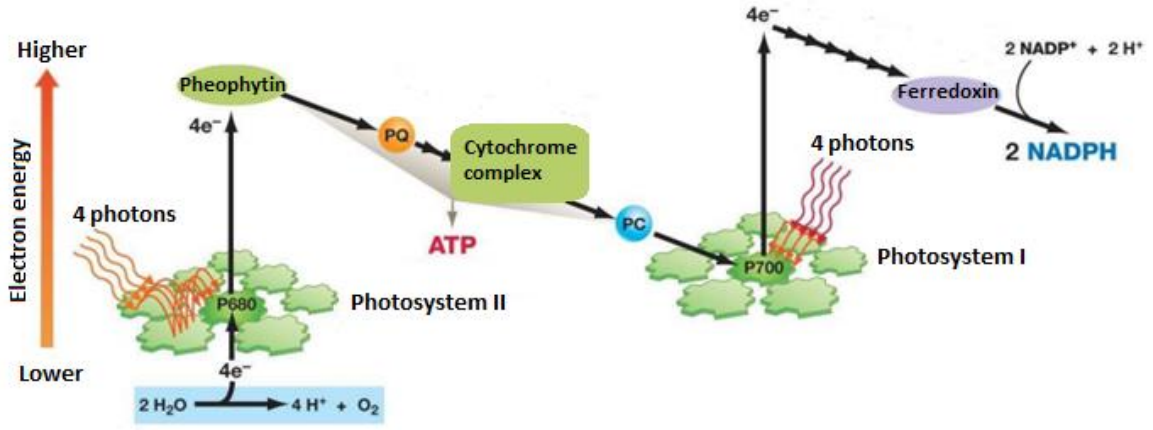

Figure 2: Z-scheme of the ET in photosynthesis (adopted and modified from reference [5]).

1.1.1 DET and MET

Two mechanisms describe the ET process: *direct electron transfer* (DET) and *mediated electron transfer* (MET). DET is a mechanism of electron transfer from the catalytic centre of the enzyme, where co-factor is located, to the receptor where the enzyme directly transfers the electron to the surface of the electrode what is schematically represented in Figure 4. Transfer rate kinetics and high cell voltages are provided by DET and therefore it is necessary to set the distance between the redox centre of the enzyme and the electrode surface to a distance less than 2 nm. Since the catalytic centres of the enzyme are often located deep inside the enzyme, DET can be low efficiency and therefore it is necessary to properly orient the enzyme towards the electrode surface and increase the interaction surface between them. There are some ways to properly orient the enzyme: covalent binding, adsorption, enzyme binding, *etc*. When it is not possible to achieve DET, mediator molecules are used that facilitate the transfer of electrons to the electrode surface and the mechanism is called MET, Figure 4. The intermediary molecule or mediator can be an organic molecule or metal complex that can access the catalytic centre located deep inside the protein. The mediator can accept or donate an electron and transport it to the surface of the electrode where occurs opposite reaction. Due to the large distance between the donor and the acceptor, protein-protein interaction requires the determination of amino acid residues that contribute to the electron tunnelling and that lead to the determination of the right ET pathway. Amino acid residues are carriers of potential intermediate molecular orbitals which reduce the height of the tunnelling barrier and reveal which of the two mechanisms leads the ET process. [3, 6]

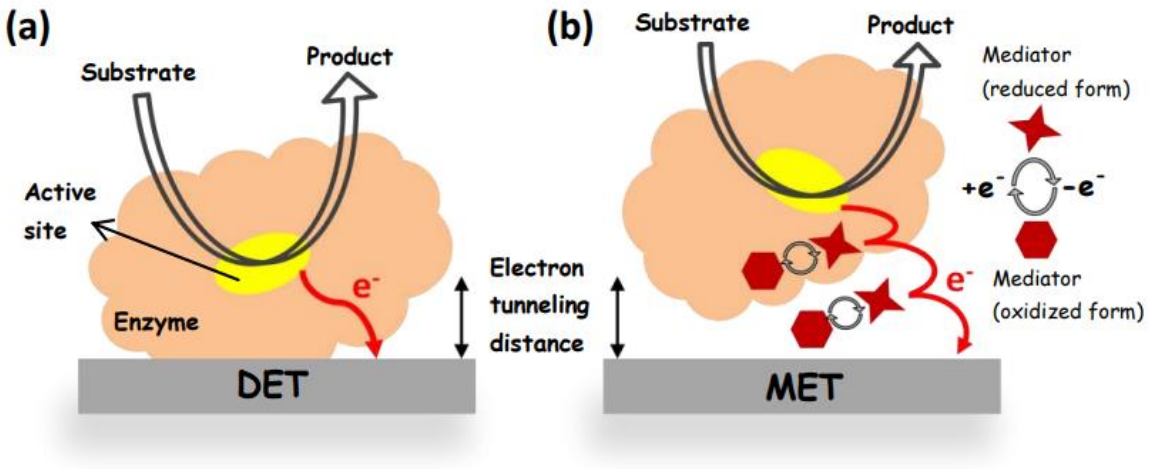

Figure 3: General representation of a simple arrangement for DET (a) and MET (b). [7]

1.1.2. Aromatic Amino Acids

Amino acids that participate in the electron transport process within proteins are aromatic amino acids. Aromatic amino acids are phenylalanine (F), tryptophan (W), histidine (H) and tyrosine (Y) illustrated in Figure 4. Aromatic amino acids are polar, unsaturated cyclic molecules. The residues of these amino acids contain an aromatic ring that provides additional stability to the amino acid due to the Π-electron above and below the plane of the aromatic ring, which is also called the Π-cloud. An aromatic-aromatic interaction occurs between pairs of interacting aromatic amino acids residues based on the following three criteria: I) the centres of the aromatic rings of two interacting residues are 4.5-7 Å apart; II) the dihedral angle is between 30° and 90°; III) the free energies of the formation of such an interaction are between 0.6 and 1.3 kcal/mol. [8]

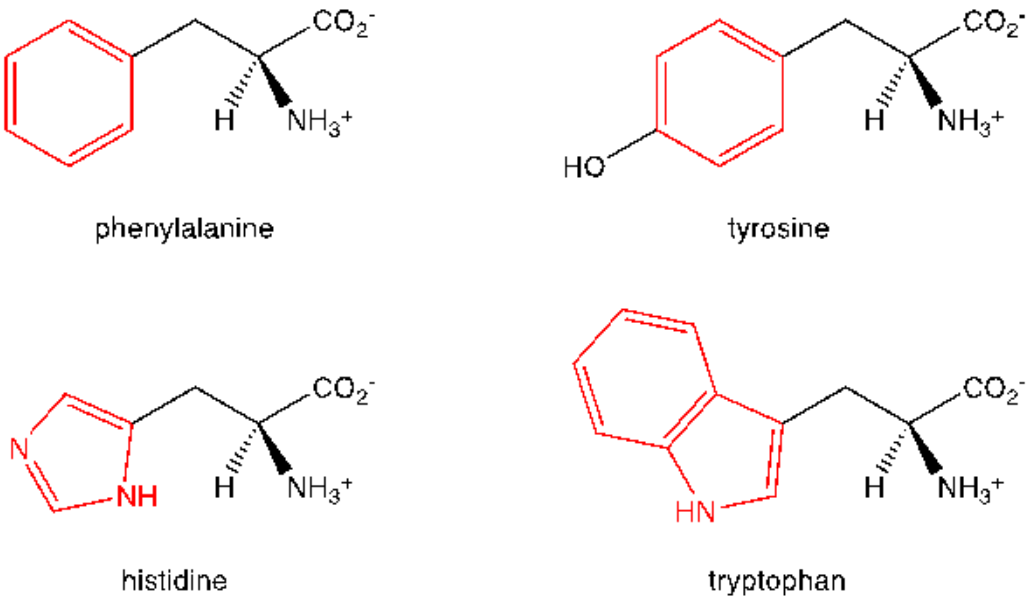

Figure 4: Aromatic amino acids phenylalanine (F), tyrosine (Y), histidine (H) and tryptophan (W) contains red highlighted aromatic rings. [9]

1.2. Enzymatic Biofuel Cell (EBFC)

In this master thesis, focus is on the enzyme that catalyses the transfer of electrons from one molecule (donor of electron) to another (acceptor of electron). These are oxidoreductases enzymes that facilitate the oxidoreduction reaction, meaning that they catalyse the coupled oxidation and reduction of their substrates. These enzymes often include different cofactors, which are electron donors/acceptors during the oxidoreduction reaction. The details of ET reaction can be studied by quantum mechanics computer simulation which had an enormous development and growth in usage in the last few decades. There are two large areas where this development is directed: into biomedical research and into enzymatic biofuel cells.

One significant application of ET in biomedicine are enzyme's biosensors that detect specific targets in the sample based on the interaction between the target and the enzyme immobilized on the electrode. The number of electrons transferred from the target molecule to the electrode is recorded as a signal, which is proportional to the concentration of the target in the sample. The blood glucose biosensor is one of the first and widely known examples whose basic concept is the interaction of glucose and the glucose oxidase enzyme immobilized on the surface of the electrode, which has the task of inducing the transfer of electrons from the glucose to the electrode. The number of recorded electrons transferred to the electrode is proportional to the number of glucose molecules present in the blood.

Enzymatic biofuel cells (EBFC) are an alternative to conventional fuel cells that convert chemical energy into electrical energy using noble gasses or their alloys as catalysts for the

oxidation of pure fuels at an anode and the reduction of oxidants at a cathode. Unlike conventional fuel cells, EBFC uses redox enzymes both on the surface of a bioanode, where the fuel is oxidized, and on the surface of a biocathode, where the oxidants are reduced, illustrating in Figure 6. The advantage of enzymes as catalysts is that enzymatic reactions do not require extreme conditions because they work at the atmospheric pressure, room temperature and neutral pH, and their efficiency is also very high.

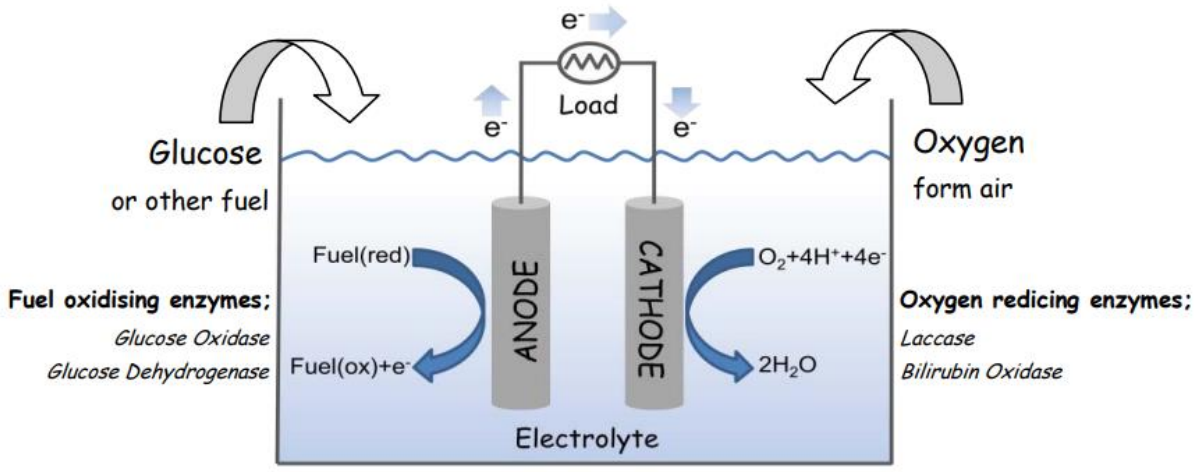

Figure 5: Scheme illustrating of an enzymatic biofuel cell. [7]

The key process monitored in EBFC is transfer of electrons between the oxidoreductase and the electrode. A process that is a demanding task because computer simulations make it difficult to achieve a proper electron treatment of large and complex systems such as enzymes that consist of up to tens of thousands of atoms. Important questions for detailed insight into ET reaction mechanism include: I) the structure and dynamics of the enzyme while immobilized on a surface; II) identification of residues that contribute to electron transfer from donor to acceptor; III) speed of electron transfer between donor and acceptor. The introduction of quantum mechanics is crucial here because the aforementioned characteristics depend on the overlap of the orbitals involved in ET. To know which orbitals to study, it is necessary to know which amino acid residues of the enzyme form the electron transfer pathways. [3, 6]

1.2.1. Pyrroloquinoline Quinone-dependent Glucose Dehydrogenase (PQQ-GDH)

Pyrroloquinoline quinone-dependent glucose dehydrogenase (PQQ-GDH) enzyme from the oxidoreductase family, shown in Figure 6, is used in the design of enzymatic anodes in

EBFCs. X-ray diffraction is used to resolve the crystallographic structure of soluble PQQ-GDH whose resolution with the cofactor is 0.22 nm, and with reduced PQQ and glucose, the resolution is 0.19 nm. PQQ-GDH is a homodimeric enzyme with 452 amino acid residues in each chain and is recognizable by the fact that each monomer has a super barrel structure containing six 4-stranded antiparallel BETA-sheets which are illustrated in Figure 8. The characteristic of the enzyme is having the cofactor pyrroloquinoline quinone (PQQ) as its prosthetic group. PQQ is bound by polar interactions at a broad, positively charged site at the top of the barrel. The function of this enzyme is to oxidize the sugar glucose which binds directly above $PQQH_2$ and makes a large number of hydrophobic interactions with the $PQQH_2$ surface. PQQ-GDH is not sensitive to oxygen, highly catalytically active, capable of direct electron transfer (DET) and has been used for the development of biosensors and biofuel cells. [6]

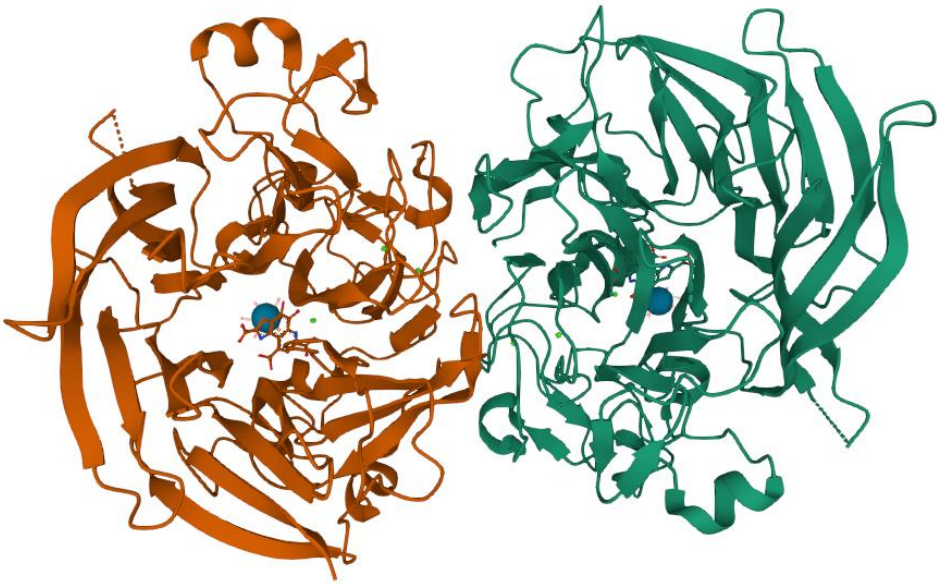

Figure 6: Soluble quinoprotein glucose dehydrogenase from acinetobacter Colcoaceticus in complex with $PQQH_2$ and glucose. [11]

2. Methods

Molecular modelling is a branch of physical chemistry based on the prediction and calculation of properties of molecules or systems of molecules. Two important approaches within molecular modelling are molecular dynamics (MD) and quantum mechanics (QM). MD serves to simulate and study molecular motion and interactions, providing insights into the configuration of molecular systems over time. MD models the motion of atoms and molecules by numerically integrating equations of motion. It employs force fields, which are mathematical representations of molecular interactions, to calculate forces between particles by iteratively calculating forces and updating particle positions. MD simulates the dynamic behaviour of a molecular system. QM provides to understand the electronic structure of molecules, predicting properties like bond lengths and energies with high precision. QM uses mathematical equations to describe the behaviour of electrons and their interactions with atomic nuclei. These equations are solved to obtain electronic properties, typically using computational methods like Hartree-Fock method or density functional theory (DFT). [12] Commonly studied properties include the molecule's electronic properties, molecular structure, and conformational changes. To describe electron transfer, it is essential to have knowledge of the electronic properties of the system, which is why we have turned to the method of quantum mechanics. The task of the master thesis is to verify the electronic properties of PQQ-GDH.

2.2. QM Methods

In quantum chemistry, a critical task is solving the Schrödinger equation with the electronic molecular Hamiltonian. This equation helps to understand the electronic structure of molecules, which determines their chemical properties. While it is possible to solve the equation for simple atoms like hydrogen, it becomes very challenging for complex molecules with many particles like electrons and nuclei. Therefore, it is needed to use approximations to get meaningful results. The theory of electronic structure allows scientists to predict various properties of molecules, limited only by our computational abilities.

One important approach in quantum chemistry is the molecular orbital theory, where electrons in a molecule move around all the atoms instead of being confined to individual bonds. Quantum mechanics helps to describe the spatial and energetic properties of electrons in molecular orbitals. These orbitals surround groups of atoms and contain valence electrons, which play a crucial role in chemical reactions. There are two main methods based on this theory: the Hartree-Fock (HF) method also known as self-consistent field (SCF) method and Density Functional Theory (DFT). [12]

2.2.1. Hartree-Fock Method

The HF method approximates the wave function of a system with a single mathematical expression called a Slater determinant. By solving some equations, it is possible to find the wave function and energy of the system. The main drawback is that there is no direct correlation between electrons; instead, each electron moves within a cloud formed by all the other electrons. [12]

2.2.2. DFT Method

Density Functional Theory (DFT), a quantum mechanics method, is widely recognized for its use in determining the molecular geometry and electronic characteristics of molecular systems. DFT, unlike HF, looks at electrons in a more basic way, focusing on their overall spread across three-dimensional space. It uses special math functions to show how electrons interact. DFT became popular in solid-state physics in the 1970s. Later, in the 1990s, it became very useful in quantum chemistry too, with better approximations for electron interactions. DFT is also easier to calculate with computers compared to HF, which needs more complicated calculations for electron behaviour.[12]

The centre of DFT is the probability electron density function, denoted as $\rho(r)$, which represents the electron density or charge density. The electron density function is determined solely by position, involving just three variables (x, y, z), whereas the wavefunction for an n-electron molecule depends on 4n variables: three spatial coordinates and one spin coordinate for each electron. It can be demonstrated that $\rho(r)$ is connected to the "component" one-electron spatial wavefunctions Ψi, referred to as molecular orbitals, within a single-determinant wavefunction Ψ:

$$\rho = \sum_{i=1}^{n} n_i |\psi_i|^2 \qquad 2.\ 1$$

By utilizing the provided $\rho(r)$, it is possible to compute various ground state properties, such as the energy. Furthermore, the first Hohenberg-Kohn theorem establishes that any property associated with the ground state of a molecule can be expressed as a function of the electron density function in the ground state:

$$E_0 = F[\rho_0] = E[\rho_0] \qquad 2.\ 2$$

Therefore, the precise electronic energy of the ground state in an n-electron system, where paired electrons are represented by identical spatial one-electron orbitals, can be expressed as the summation of several components: the kinetic energies of the electrons, the potential energies associated with the attraction between electrons and nuclei, the Coulomb interaction between the total charge distributions at r_1 and r_2, and the exchange-correlation energy of the system[16]:

$$E[\rho] = -\frac{\hbar^2}{2m_e}\sum_{i=1}^{n}\int \psi_i^*(r_i)\nabla_1^2\psi_i(r_1)dr_1 - j_0\sum_{I=1}^{N}\frac{Z_I}{r_{I1}}\rho(r_1)dr_1 + \frac{1}{2}j_0\int\frac{\rho(r_1)\rho(r_2)}{r_{12}}dr_1dr_2 + E_{XC}[\rho] \qquad 2.3$$

Important components to set the input DFT calculation are basis set and functional. The basis set is a set of mathematical functions used to represent the molecular orbitals of the system. The basis set functions are typically centred on the atomic nuclei and are used to construct the wavefunction as a linear combination of these functions. In molecular quantum mechanics calculations, there are two main types of basis sets: atomic orbital basis sets in which atomic orbitals describe the spatial distribution of electrons around individual atoms and molecular orbital basis sets which are the sets of molecular orbitals that span the entire molecular system. They are obtained by combining the atomic orbitals of the constituent atoms, and they describe the electronic structure of the molecule. This molecular orbital is called LCAO, meaning Linear Combination of Atomic Orbitals.

Some of the commonly used basis sets are minimal basis sets, basis sets with diffuse orbitals, polarized basis sets, valence basis sets, correlated basis sets etc. Minimal basis sets typically include only a few fundamental orbitals for each atom in the molecule. They are fast to compute and are used for quick estimations but are often limited in accuracy as they do not account for all electronic interactions.

Basis sets with diffuse orbitals include orbital functions with diffuse shapes, meaning they are broader and more extended than standard orbitals. This is useful for describing the electronic density around atoms with large atomic radii, such as metals or atoms with a higher number of electrons. Diffuse orbitals aid in better describing electronic interactions between these atoms.

Polarized basis sets include additional sets of orbitals that take into account electronic polarization, making them useful for accurately describing molecular properties.

Valence basis sets are optimized for accurately describing valence electrons in molecules.

Correlated basis sets include orbital functions that consider electronic correlation among electrons, making them suitable for calculations that require precise handling of electronic correlation.

The choice of basis set is critical in quantum mechanics calculations because it affects the accuracy of the results. A small basis set may not accurately describe the electronic structure of the system, leading to less precise calculations. On the other hand, a large basis set requires more computational resources but can provide more accurate results. Some of minimal basis sets are STO-3G and 3-31G to more extensive sets like 6-31G and 6-311G. The basis set size affects the accuracy of the calculations, with larger basis sets providing more precise results. [15]

Another important feature of the DFT is functional. In the context of DFT in quantum chemistry, "functional" refers to the electron density functional or, briefly, the "DFT functional." The DFT functional is a mathematical expression that describes how the electron density is distributed in space and how this density contributes to the energy of the electron system. In other words, the DFT functional is a mathematical expression that connects electron density with the energy state of the system. There are various DFT functionals, each with different levels of precision and applications. Some of the most used functionals are local density approximation (LDA), generalized gradient approximation (GGA), hybrid functionals (for example B3LYP), meta-generalized gradient approximation (meta-GGA) etc.

LDA offers computational simplicity and speed, rendering it particularly suitable for examining structural properties, but its accuracy may be limited when applied to energetic attributes and dynamic behaviour.

GGA surpasses LDA in predicting energy-related properties by incorporating electron density gradients into its calculations. It might exhibit restricted accuracy when gealing with systems characterized by strong electronic correlations.

Hybrid functionals, for example B3LYP, strike an optimal balance between accuracy and computational efficiency by amalgamating GGA with Van der Waals correlations. These functionals may not deliver precise outcomes for systems characterized by pronounced correlations among electrons.

The meta-GGA approach enhances the description of electron correlation and energy properties. It demands more computational resources compared to GGA functionals. [14]

The aim of this master thesis was to analyse, precisely, to optimize only the active site of the enzyme using quantum mechanics (QM) methods. Usually, for the analysis of the entire protein, the quantum mechanics and molecular mechanics (QMMM) method is used, but it was out of scope for the present research. In quantum mechanics calculations, optimization refers to the process of finding the minimum energy configuration of a molecular system by

adjusting the positions of its atoms or nuclear coordinates. The objective of optimization is to determine the most stable and energetically favourable arrangement of the atoms, corresponding to the equilibrium geometry or ground state of the system. The optimization process involves iteratively changing the nuclear coordinates and calculating the total energy of the system at each step using the chosen quantum mechanical method.

2.3. Jmol

Jmol is a powerful and versatile molecular visualization program used in the field of computational chemistry and molecular modelling. Some essential features and characteristics of Jmol which had been used are molecular visualisation and interactive manipulation. Jmol allows users to visualize molecular structures in 3D. It can handle a wide range of molecular file formats, including PDB (Protein Data Bank), XYZ and more. It also provides interactive features that allow users to rotate, zoom, and translate the molecular structure in real-time. This functionality helps users to explore the molecular structure from different perspectives. One of advantages of Jmol is that it can visualize various molecular properties, such as electrostatic potential surfaces, electron density maps and molecular orbitals, helping researchers gain insights into the electronic structure of molecules.

Jmol is also widely used in educational settings as it allows students and researchers to visualize complex molecular structures and understand their spatial arrangements, bonding and interactions. One of the main reasons of widely using this tool is due to it is open-source software. [13]

2.4. Input Setup

In quantum mechanics, optimization refers to finding the minimum energy configuration of a molecular system by adjusting the positions of its atoms or nuclear coordinates. At the beginning of each calculation, user has to submit the input into quantum mechanics software such as Gaussian, Turbomole or Dalton. This input consists of the molecular geometry, quantum mechanical method, basis set, optimization algorithm, convergence criteria and some other specific parameters depending on system.

The first step in performing quantum mechanical optimization is to specify the molecular geometry of the system. This involves providing the coordinates of all atoms in the system, along with their elemental identities. The coordinates can be given in Cartesian coordinates (x, y, z) or other suitable coordinate systems (such as Z-matrix). The coordinates can be extracted from protein structure by Jmol.

It is important to choose an appropriate quantum mechanical method and functional to describe the electronic structure of the system. Commonly used methods include Hartree-Fock (HF) or density functional theory (DFT). The choice of method depends on the system's size, complexity, and the level of accuracy required. In quantum mechanics (QM) calculations, a functional is a mathematical expression used to determine various properties of a system, particularly in relation to the distribution of electrons and their energies and it's a key concept in density functional theory (DFT).

Define convergence criteria to determine when the optimization process has reached a stable configuration. Common criteria include a specific energy threshold or a maximum allowed change in the nuclear coordinates between iterations.

If the system is in a solvent environment, you may consider using solvation models like the Self-Consistent Reaction Field (SCRF) method or explicit solvent models to account for the solvent's influence on the molecular structure and energy.

2.5. Initial Structure of PQQ-GDH Enzyme

In the first calculation, the structure of the active site was examined before the reaction, electron transfer from glucose through PQQ-GDH to carbon nanotube, takes place. Therefore, the hydrogens were on the initial atoms, and the electronic structure was checked, i.e., orbitals and the charge, Mulliken charges (partial atomic charges), of the system were analysed. The focus was on molecular subsystem of PQQ-GDH that contains the amino acid histidine144 located in the immediate vicinity of the protein’s active site, the cofactor PQQ and glucose and along with one calcium atom in their vicinity visualised in Figure 7. Calcium atom is involved in the binding of PQQ to the GDH. In the Jmol program, XYZ coordinates of the atoms were extracted from the protein structure pdb code: 1CQ1, which was downloaded from the RSCB PDB website.

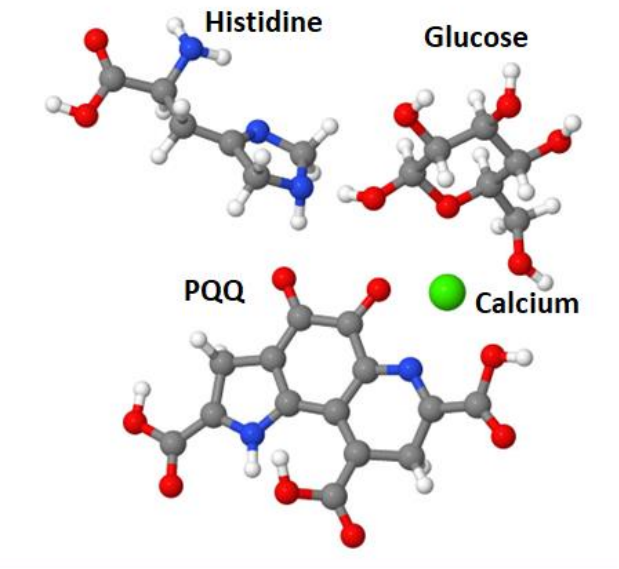

Figure 7: Molecular model representation of the initial structure of PQQ-GDH enzyme.

The complete optimization of the system was performed using the specified PBE functional and def2svp basis set.

The PBE (Perdew-Burke-Ernzerhof) functional is commonly used in a wide range of computational chemistry calculations. It is a type of exchange-correlation functional used in density functional theory (DFT) calculations to approximate the exchange and correlation energies of electrons in a molecular system. PBE is used for performing geometry optimizations, where the positions of atoms in a molecular or solid-state system are adjusted to find the most stable configuration. Some of common types of calculations for which the PBE functional is used are energy minimization, electronic structure calculations, molecular dynamics simulations, vibrational analyses etc. [14]

The def2svp basis set is a standard basis set used in computational chemistry for accurate quantum chemical calculations. The name, "def2svp", denotes a specific level of Gaussian basis set. "def" refers to the "double-ζ (DZ)" basis set, which means that two Gaussian functions are used for each electronic orbital. "2s" and "2p" indicate that two Gaussian functions are used for each s orbital (spherical orbital) and two Gaussian functions are used for each p orbital (d orbital). The "def2svp" basis set typically includes the following types of functions: polarization, diffuse, double-ζ (DZ) and triple-ζ (TZ) functions. It is specifically designed for molecular systems and is suitable for a wide range of elements.

The optimization was conducted with fixed positions of the Cα atoms to prevent significant displacements relative to the protein. However, to consider the protein in some way, we utilize the option provided by Gaussian, called SCRF (Self-Consistent Reaction Field). When using the SCRF option, a calculation is carried out in the presence of a solvent by placing the

solute (the molecule of interest) in a cavity within the solvent's reaction field. This approach creates the solute cavity using a set of overlapping spheres. The optimization is performed in a solvent, for which we specify the dielectric constant. The idea was to use the dielectric constant of the protein itself, i.e., to select a solvent with a dielectric constant of approximately 3.5. This choice allows us to consider the influence of the solvent on the system and accounts for the protein's effect during the optimization process.

2.6. Final Structure of PQQ-GDH Enzyme

The optimization procedure of the initial model took 161 models, and the last one represents the optimal position of all the atoms in the system. So, the final structure of the first step was used for the second step of calculations. By opening the final structure in Jmol, hydrogen atoms were positioned according to the sketch provided in Figure 8. Two hydrogen atoms from glucose were placed on PQQ. The relocated hydrogen atoms would be found at those positions after the reaction occurs, i.e., at their final positions which is visualised in Figure 8. After optimization, the electronic structure is observed once again. The system was again optimized using the same method as in the first calculation.

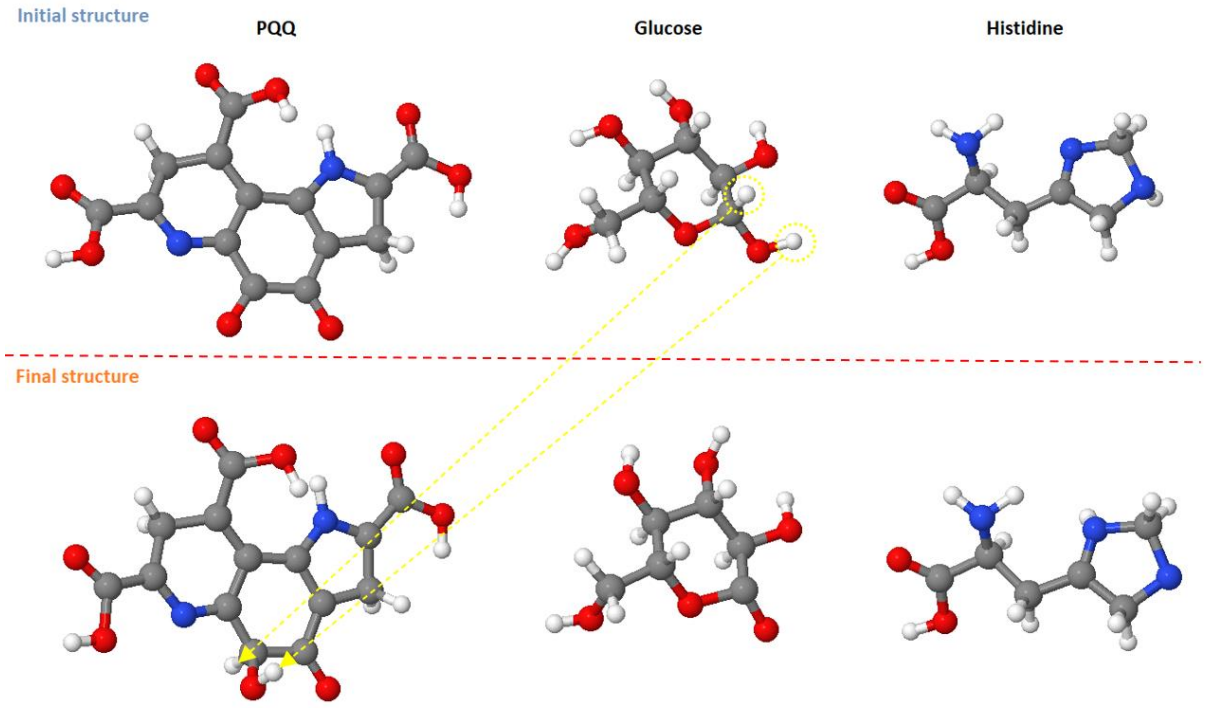

Figure 8: Schematic representation of the positions of hydrogen atoms in the initial and final structure, after the reaction occurred.

The calculation settings were as follows: optimization with the PBE functional and def2svp basis set, to begin with. The objective was to observe the distribution of Mulliken charges in the final structure and visualize some molecular orbitals.

By performing this second step of the calculation, the aim was to gain insights into the charge distribution and molecular orbitals of the optimized system after positioning the hydrogen atoms according to the provided sketch in Figure 8.

3. Results and Discussion

The output of the first calculation consists of optimization steps of the system, 161 of them, which do not significantly differ in the conformation of the molecular system. This is expected as the model is taken from the crystal structure of the protein which means that the atoms are at their proper places. The final model of geometry optimization is the one with the lowest energy for the system. This is the model in which the system relaxed into one of the local minima close to the initial structure on the potential energy surface. Through the optimization of the system's geometry, unfavourable interactions arising from too large or too small distances between atoms, distances of valence angles, covalent bonds, etc., have been eliminated. The last model of geometry optimization is used to prepare the input for the second calculation. Both of these structures, initial and final, have equal energy range of valence molecular orbitals, between -5.500 eV and -0.800 eV.

3.1. HOMO-LUMO gap

Molecular electrical properties are identified by the energy difference between the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). The molecule's capacity to donate and receive electrons can also be described by the HOMO and LUMO energy values, making this gap a measure of its electronic conductivity. The highest energy level of the occupied molecular orbital signifies its ability to donate electrons, while the lowest energy level of the unoccupied molecular orbital signifies its ability to accept electrons.

These molecular orbitals have a significant impact on electronic characteristics and offer valuable insights into biological mechanisms. The energy difference between the frontier molecular orbitals, known as FMO, provides data regarding the structural kinetic stability. Furthermore, FMO offers insights into the system's kinetic energy and chemical reactivity, allowing us to identify the most reactive sites within the molecule under study.

For the initial structure of PQQ-GDH enzyme, energy of HOMO is -3.429 eV and for the LUMO is -3.048 eV. Energies in the final structure are -3.510 eV for the HOMO and -2.748 eV for the LUMO. The computed HOMO-LUMO gap value is 0.381 eV for the initial structure and 0.762 eV for the final structure. The energy level of the HOMO-LUMO gap, along with its corresponding energy value, is shown in the Figure 9. A low HOMO-LUMO gap value indicates the high chemical reactivity and biological activity of the studied system, as well as the presence of intramolecular charge transfer. Furthermore, a molecule with a low-

energy HOMO-LUMO gap is considered to be a soft molecule, suggesting greater reactivity due to its larger size and higher degree of polarization.

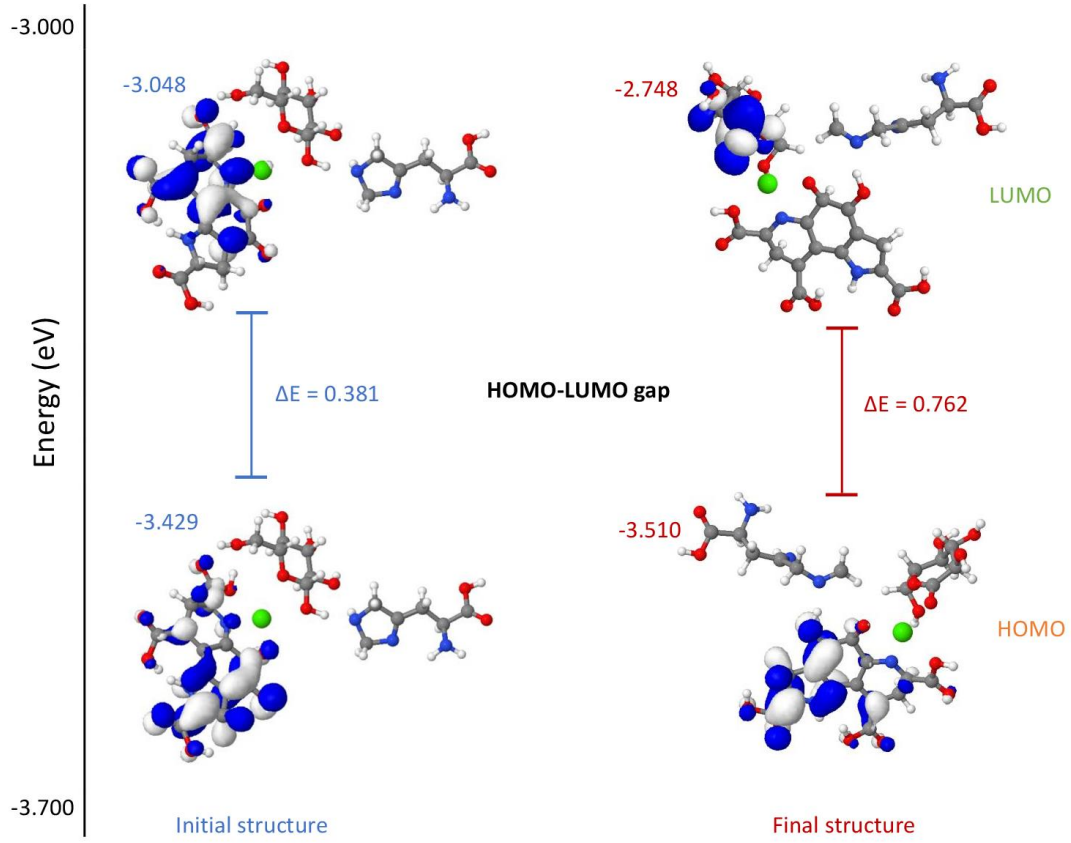

Figure 9: Orbital energy levels diagram of HOMO and LUMO orbitals and HOMO-LUMO gap for the initial and the final structure of PQQ-GDH enzyme.

3.2. Analysis of Orbital Energy Levels

Figures 10 and 12 visualise LUMO+2, LUMO+1, LUMO, HOMO, HOMO-1 and HOMO-2 orbitals for the initial and final structure of PQQ-GDH enzyme. The HOMO orbitals are primarily concentrated on the PQQ molecule. These HOMO orbitals mainly consist of elements like oxygen, nitrogen, carbon which correspond to the aromatic ring structure. In contrast, the LUMO orbitals are distributed throughout the entire system.

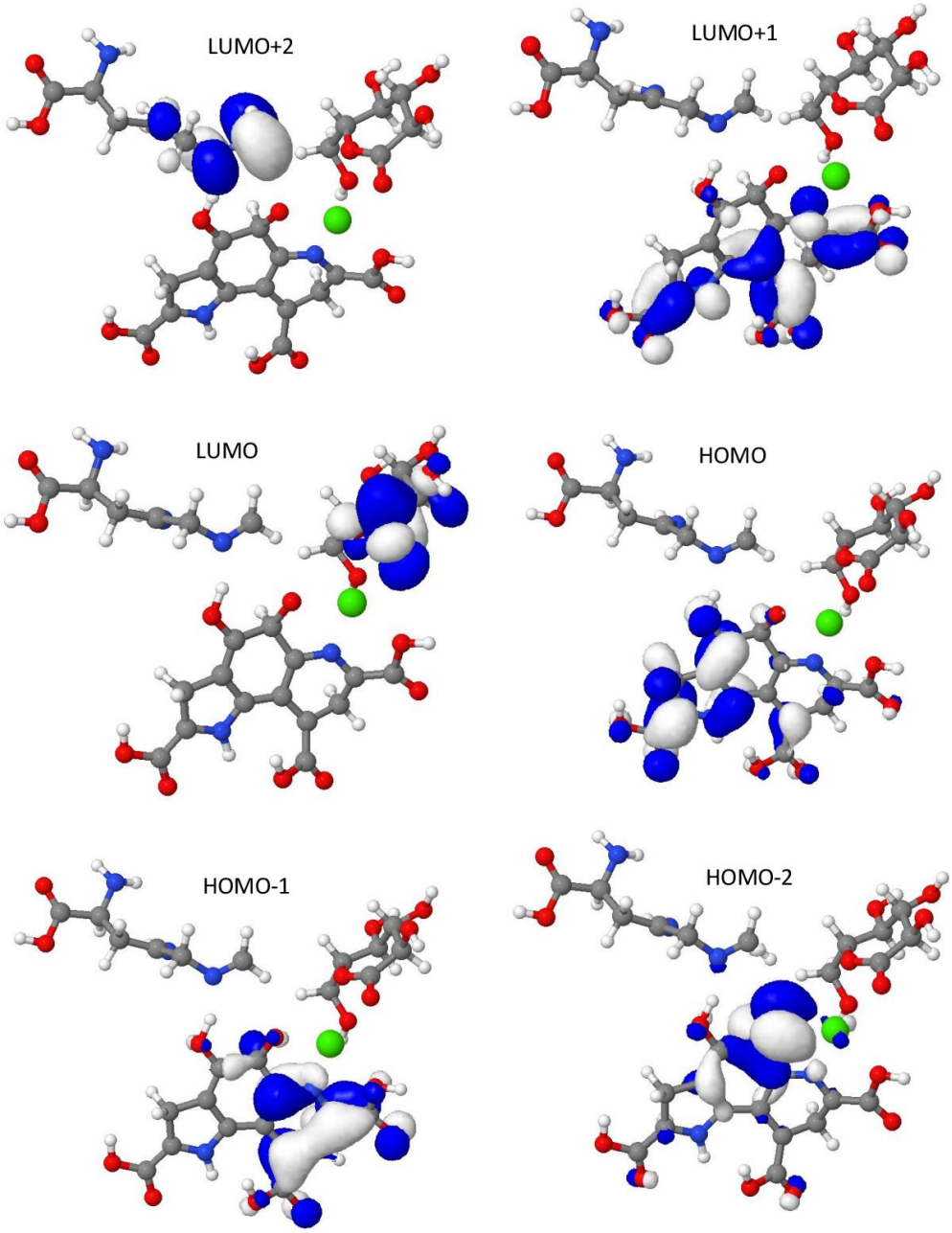

Figure 10: Molecular model representation of the LUMO+2, LUMO+1, LUMO, HOMO, HOMO-1 and HOMO-2 orbitals for the initial structure of PQQ-GDH enzyme.

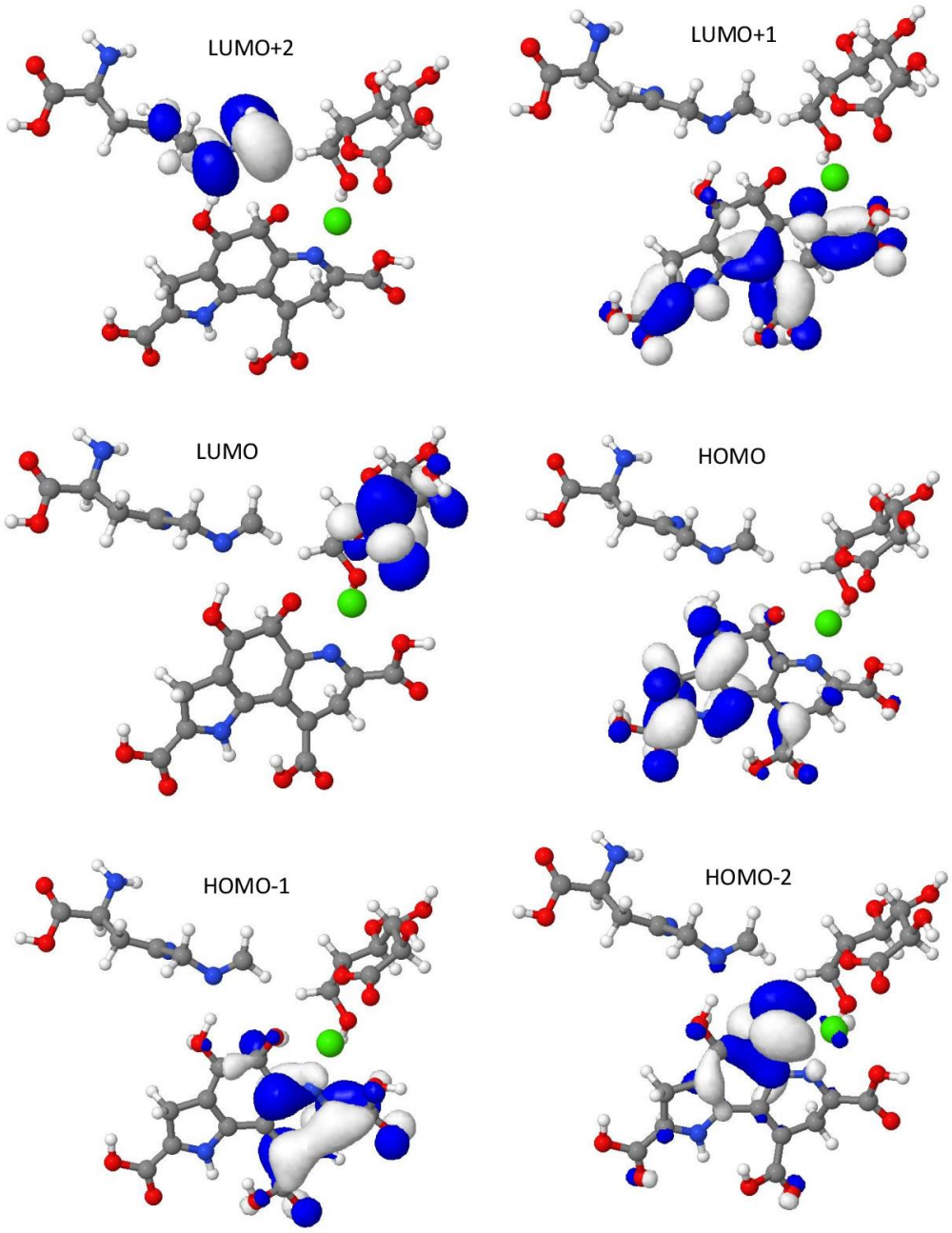

Figure 11: Molecular model representation of the LUMO+2, LUMO+1, LUMO, HOMO, HOMO-1 and HOMO-2 orbitals for the final structure of PQQ-GDH enzyme.

The Figure 12 displays both the energy levels and visual representations of the frontier molecular orbitals, which include HOMO-4, HOMO-3, HOMO-2, HOMO-1, HOMO, LUMO, LUMO+1, LUMO+2, LUMO+3, and LUMO+4. Valence orbital energy levels ranging from -5.500 eV to -0.800 eV for both structures, the initial and the final structure of PQQ-GDH enzyme.

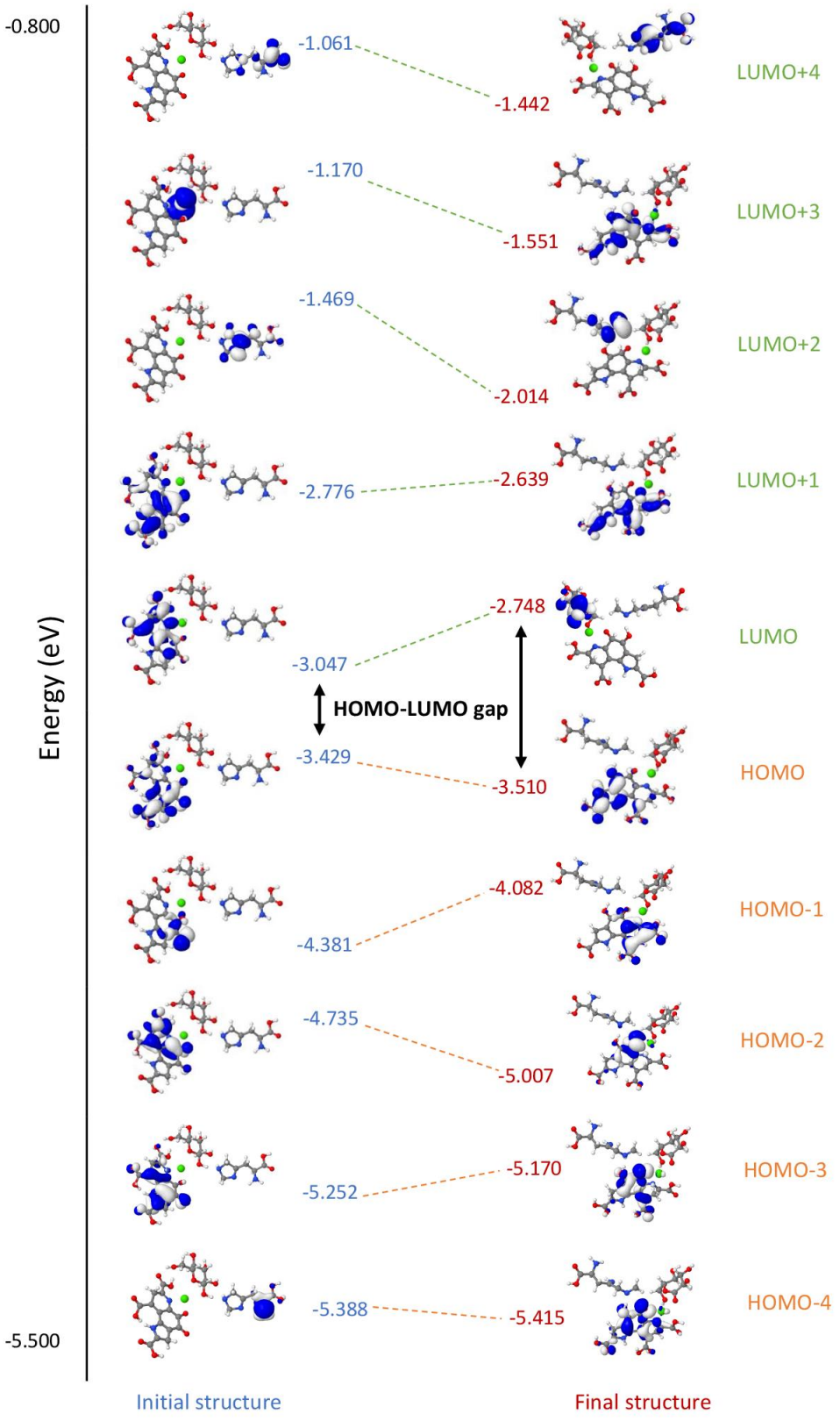

Figure 12: Orbital energy levels diagram of the 5 unoccupied and 5 occupied orbitals for the initial and final structure of PQQ-GDH enzyme.

3.3. Mulliken Population Analysis

Mulliken charge provides a measure of how the electron distribution within a molecule shifts when atoms move, impacting a range of molecular properties including electronic structure, dipole moment, polarizability, and more. The Figure 13 illustrates the distribution of Mulliken atomic charges for the initial structure and Figure 14 illustrates distribution of partial atomic charge for the final structure of PQQ-GDH enzyme. Photos are made in Molden visualization program. Hydrogen atoms tend to carry positive Mulliken charges ranging from 0.0125 e to 0.2128 e for the initial structure and 0.0019 e to 0.2372 e for the final structure of the PQQ-GDH enzyme. Positively charged hydrogen atoms are functioning as acceptor atoms. Oxygen atoms carry negative charges ranging from -0.3934 e to -0.1534 e for the initial structure and -0.5426 e to -0.1868 e for the final structure and act as donor atoms. Carbon atoms have equally positive and negative charge.

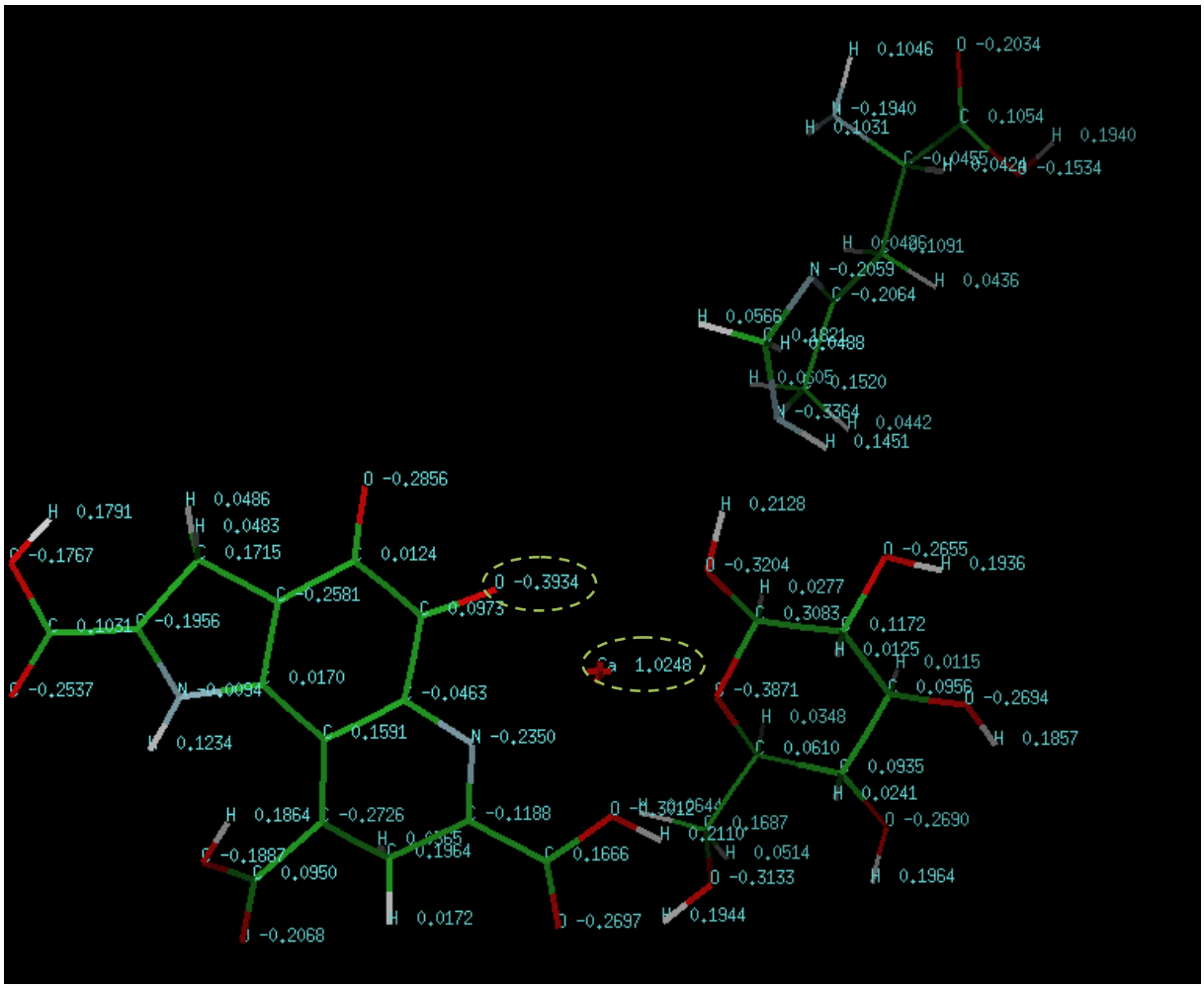

Figure 13: Schematic representation of the distribution of Mulliken atomic charges for the initial structure of PQQ-GDH enzyme.

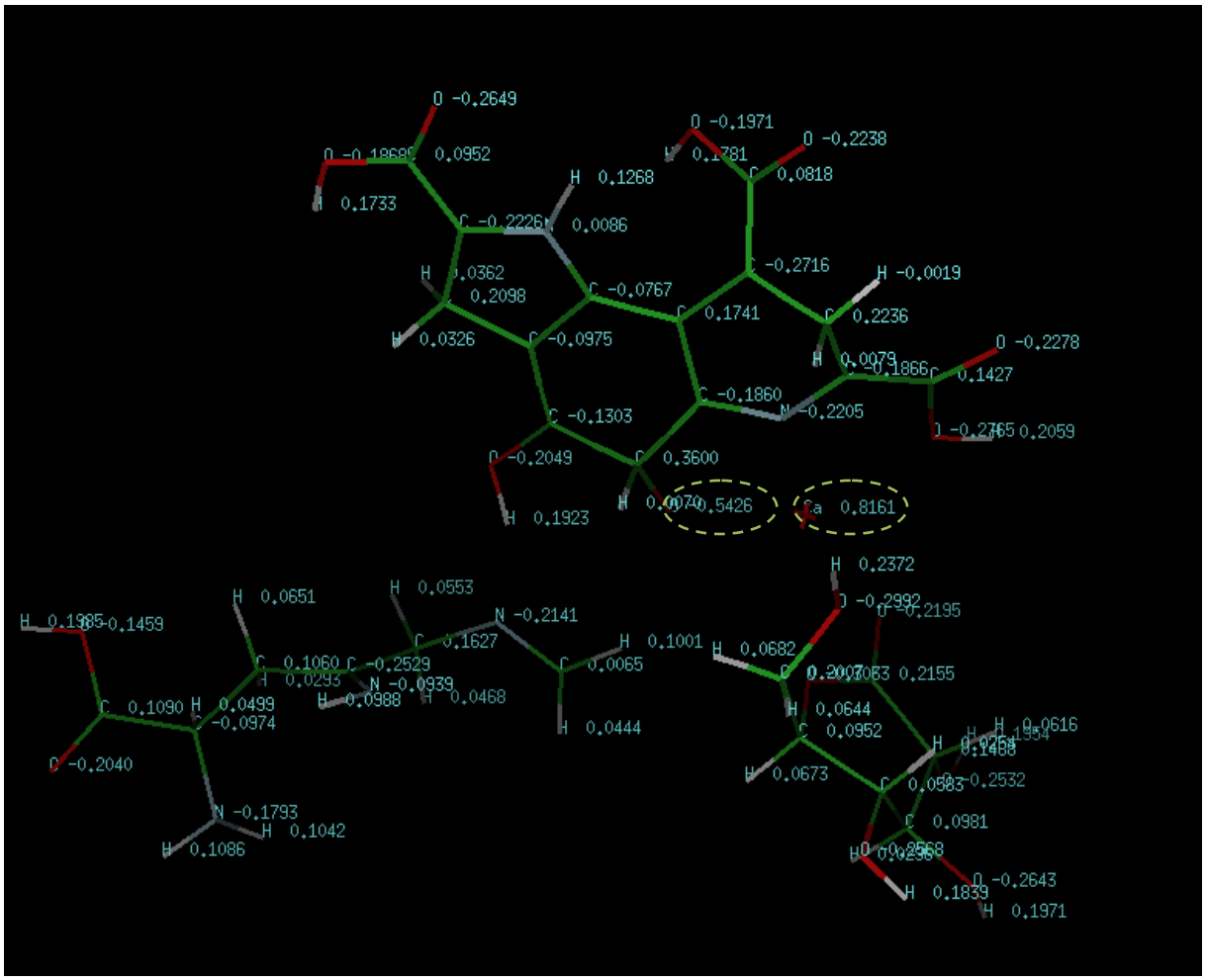

Figure 14: Schematic representation of the distribution of Mulliken atomic charges for the final structure of PQQ-GDH enzyme.

The most electronegative atom, i.e., the atom with the highest number of electrons in the initial structure, is the oxygen atom in PQQ with a charge of -0.3934 e. In the final structure, the same atom also has the highest number of electrons and a charge of -0.5426 e. This atom is located in proximity to the most electropositive atom in the system, which is calcium atom with a charge of 1.0248 e in the initial structure and 0.8161 e in the final structure. These atoms are yellow framed in Figures 13 and 14. Additionally, from Figures 13 and 14, it is evident that other highly electronegative atoms are in the vicinity of positively charged calcium. Therefore, this location should be considered in the design of this enzyme to be placed on the carbon nanotube (CNT) which should be the next step to study within this scientific research.

4. Conclusion

PQQ-GDH is an enzyme protein that plays a crucial role in glucose metabolism and the oxidation of glucose in organisms and that's why it was chosen as a protein of interest suitable for this research. It is a key enzyme protein with applications, particularly in the context of metabolic processes and glucose analysis. Enzymes like PQQ-GDH are frequently used in analytical instruments to measure blood glucose levels, which are essential for diagnosing diabetes and monitoring glycemia, a topic addressed in this master's thesis. It also finds application in biotechnology and in the development of biofuels, where it is used to convert glucose from biomass into biofuels.

Density Functional Theory (DFT), a method that replaces the complex many-body wavefunction with the electron density, thus becoming computationally more efficient, has proven to be a powerful computational technique in quantum chemistry and physics, providing insights into the behaviour of electrons in the initial and final structures of PQQ-GDH enzyme.

Both the initial and final structures exhibit a consistent energy range for valence molecular orbitals, spanning from -5.500 eV to -0.800 eV. Molecular electrical characteristics are determined by the energy difference between the Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO). The calculated HOMO-LUMO gap is 0.381 eV for the initial structure and 0.762 eV for the final structure. A smaller HOMO-LUMO gap indicates heightened chemical reactivity and biological activity within the system, along with the presence of intramolecular charge transfer. Furthermore, a molecule with a low-energy HOMO-LUMO gap is considered "soft," implying greater reactivity due to its larger size and enhanced polarization.

The Mulliken charge offered insight into how the electron distribution within a molecule shifts when atoms relocate, impacting various molecular properties including electronic structure, dipole moment, polarizability, and more. The oxygen atom in PQQ is the most electronegative atom in both the initial and final structures, carrying a charge of -0.3934 e and -0.5426 e, respectively. This atom is located in close proximity to the most electropositive atom in the system, calcium atom, which has a charge of 1.0248 e in the initial structure and 0.8161 e in the final structure. Furthermore, Figures 13 and 14 reveal the presence of other highly electronegative atoms in the vicinity of calcium. Therefore, when designing this enzyme for placement on the CNT, this specific location should be taken into account.

5. Literature

[1] T. Vidakovic-Koch, *Electron Transfer Between Enzymes and Electrodes*, Springer, **167**, 2017.

[2] Ruslan N. Tazhigulov, James R. Gayvert, Mlissa Wei and Ksenia B. Bravaya, *eMap: A Web Application for Identifying and Visualizing Electron or Hole Hopping Pathways in Proteins*, J. Phys. Chem. B, **123** (32), 6946-6951 (2019).

[3] M. Holzinger, Y. Nishina, A. Le Goff, M. Tominaga, S. Cosnier, S. Tsujimura, *Molecular Design of Glucose Biofuel Cell Electrodes*, Wiley-VCH, **1** (11), 287-306 (2018).

[4] M. L. Brusseau, J. Chorover, *Environmental Processes Affecting Contaminant Transport and Fate,* Elsevier, 2019.

[5] L. Bracun, *Photosynthesis,* National and University Library in Zagreb, https://zir.nsk.hr/islandora/object/pmf:2199/preview, (3. 7. 2022.).

[6] Z. Sanader Marsic, *ET in Oxidoreductases (Application Form)* (2022).

[7] M. Gonzalez Guerrero, *Enzymatic Microfluid Fuel Cells: from Active to Passive Power Sources,* disposit digital de documents de la UAB, https://ddd.uab.cat/pub/tesis/2015/hdl_10803_322082/mjgg1de1.pdf, (7. 8. 2022.).

[8] R. Anjana, M. Kirti Vaishnavi, D. Sherlin, S. Pavan Kumar, K. Naveen, Pa. Sandeep Kanth and K. Seker, *Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: a study based on orientation and distance*, PubMed, **8** (24), 1220-1224 (2012).

[9] Chegg, https://www.chegg.com/homework-help/organic-chemistry-8th-edition-chapter-26-solutions-9780840054449, (4. 7. 2022.).

[10] Phys.org, https://phys.org/news/2016-09-team-accurately-permeability-carbon-nanotubes.html, (4. 7. 2022.).

[11] RCSB PDB, https://www.rcsb.org/structure/1cq1, (10. 7. 2022.).

[12] Z. Sanader Marsic, *Lecture materials: Comparison of methods* (2023.)

[13] Jmol, https://jmol.sourceforge.net/, (7.8.2023.)

[14] TCM, University of Cambridge, https://www.tcm.phy.cam.ac.uk/castep/documentation/WebHelp/content/modules/castep/tskcastepsetelecxc.htm#:~:text=PBE%20(Perdew%20et%20al.%2C,fairly%20reliable%20for%20bulk%20calculations., (6.9.2023.)

[15] Q-CHEM, https://manual.q-chem.com/5.3/chap_basis.html, (6.9.2023.)

[16] Lewars, E. G., *Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics*, Springer, Second Edition.