METHODS OF CALCULATIONS

1. Theoretical background

1.1 Møller-Plesset perturbation theory

The alternative way to solve the electron correlation problem was proposed by Møller and Plesset [Møller and Plesset 1934]. Their method is based upon Rayleigh-Schrödinger perturbation theory, in which the true Hamiltonian operator, H is expressed as the sum of a zeroth-order Hamiltonian, H_0 and a perturbation, V:

$$H = H_0 + V \tag{1}$$

The eigenfunctions of the true Hamiltonian operator are Ψ_i with corresponding energies E_i. The eigenfunction of the zeroth-order Hamiltonian are written $\Psi_{i(0)}^{(0)}$ with energies $E_i^{(0)}$. The ground state wavefunction is thus $\Psi_0^{(0)}$ with energy $E_0^{(0)}$. To devise a scheme by which it is possible to gradually improve the eigenfunctions and eigenvalues of H_0 we can write the true Hamiltonian as follows:

$$H = H_0 + \lambda V \tag{2}$$

λ is a parameter that can vary between 0 and 1; when λ is zero then H is equal to the zeroth-order Hamiltonian but when λ is one then H equals its true value. The eigenfunctions Ψ_i and eigenvalues E_i of H are then expressed in powers of λ:

$$\Psi_i = \Psi_i^{(0)} + \lambda\Psi_i^{(1)} + \lambda^2\Psi_i^{(2)} + \ldots = \sum_{n=0} \lambda^n \Psi_i^{(n)} \tag{3}$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \ldots = \sum_{n=0} \lambda^n E_i^{(n)} \tag{4}$$

$E_i^{(1)}$ is the first-order correction to the energy, $E_i^{(2)}$ is the second-order correction and so on. These energies can be calculated from the eigenfunctions as follows:

$$E_i^{(0)} = \int \Psi_i^{(0)} H_0 \Psi_i^{(0)} d\tau \tag{5}$$

$$E_i^{(1)} = \int \Psi_i^{(0)} V \Psi_i^{(0)} d\tau \tag{6}$$

$$E_i^{(2)} = \int \Psi_i^{(0)} V \Psi_i^{(1)} d\tau \tag{7}$$

$$E_i^{(3)} = \int \Psi_i^{(0)} V \Psi_i^{(2)} d\tau \tag{8}$$

To determine the corrections to the energy it is therefore necessary to determine the wavefunctions to a given order. In Møller-Plesset perturbation theory, the unperturbed Hamiltonian H_0 is the sum of the one-electron Fock operators for the N electrons:

$$H_0 = \sum_{i=1}^{N} f_i = \sum_{i=1}^{N} [H^{core} + \sum_{j=1}^{N} (J_j - K_j)] \tag{9}$$

The Hartree-Fock wavefunction, $\Psi_0^{(0)}$ is an eigenfunction of H_0 and the corresponding zeroth-order energy $E_0^{(0)}$ is equal to the sum or orbital energies for the occupied molecular orbitals:

$$E_0^{(0)} = \sum_{i=1}^{occupied} \varepsilon_i \tag{10}$$

In order to calculate higer order wavefunctions we need to establish the form of the perturbation, V. This is the difference between the real Hamiltonian H and the zeroth-order Hamiltonian, H_0. Remember that the Slater determinat description, based on an orbital picture of the molecule, is only an approximation. The true Hamiltonian is equal to the sum of the nuclear attraction terms and electron repulsion terms:

$$H = \sum_{i=1}^{N} (H^{core}) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{r_{ij}} \tag{11}$$

Hence the perturbation V is given by:

$$V = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{r_{ij}} - \sum_{j=1}^{N} (J_j - K_j) \tag{12}$$

The first-order energy $E_0^{(1)}$ is given by:

$$E_0^{(1)} = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} [(ii|jj) - (ij|ij)] \tag{13}$$

The sum of the zeroth-order and first-order energies thus corresponds to the Hartree-Fock energy

$$E_0^{(0)} + E_0^{(1)} = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[(ii|jj) - (ij|ij)\right] \tag{14}$$

To obtain an improvement on the Hartree-Fock energy it is therefore necessary to use Møller-Plesset perturbation theory to at least second order. This level of theory is referred to as MP2 and involves the integral $\int \Psi_0^{(0)} V \Psi_0^{(1)} d\tau$. The higer-order wavefunction $\Psi_0^{(1)}$ is expressed as linear combinations of solutions to the zeroth-order Hamiltonian:

$$\Psi_0^{(1)} = \sum_j c_j^{(1)} \Psi_j^{(0)} \tag{15}$$

The $\Psi_j^{(0)}$ in equation (15) will include single, double, etc., excitations obtained by promoting electrons into the vitual orbitals obtained from a Hartree-Fock calculation. The second-order energy is given by:

$$E_0^{(2)} = \sum_{i}^{occupied} \sum_{j>i}^{occupied} \sum_{a}^{virtual} \sum_{b>a}^{virtual} \frac{\iint d\tau_1 d\tau_2 \chi_i(1)\chi_j(2)\left(\frac{1}{r_{12}}\right)\left[\chi_a(1)\chi_b(2) - \chi_b(1)\chi_a(2)\right]}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \quad (16)$$

These integrals will be non-zero only for double excitations, according to the Brillouin theorem. Third and fourth-order Møller-Plesset calculations (MP3 and MP4) are also available as standard options in many *ab initio* packages.

The advantage of many-body perturbation theory is that it is size independent, unlike configuration interaction-even when a truncated expansion is used. However, Møller-Plesset perturbation theory is not variational and can sometimes give energies that are lower than the true energy. Møller-Plesset calculations are computational intensive and so their use is often restricted to single-point calculations at a geometry obtained using a lower level of theory. They are at present the most popular way to incorporate electron correlation in molecular quantum mechanical calculations, especially at the MP2 level. A Møller-Plesset calculation is specified using the level of theory used (e.g. MP2, MP3) together with the basis set. Thus MP2/6-31G* indicates a second-order Møller-Plesset calculation with the 6-31G* basis set.

1.2 Basis Set Superposition Error (BSSE)

The interaction energy of a complex can be obtain by first calculating the energy of isolated reactant molecules, then calculating the energy of the complex, and finally subtracting the energy of the isolated reactant molecules from the complex.

$$A + B \rightarrow AB$$

$$E_{int} = E_{AB} - (E_A + E_B) \tag{17}$$

Where E_{int} is the interaction energy between A and B, while E_{AB} is the energy of complex AB and E_A, E_B are the energies of isolated reactant molecule A, B in their own basis functions, respectively.

However, the energy difference always overestimates of the true value. The discrepancy arises from a phenomenon known as basis set superposition error (BSSE). BSSE occurs because the basis functions on each molecule in the complex provide a better description of the electronic structure around the other molecule. It is clear that the BSSE would be expected to be particularly significant when small, inadequate are used which do not provide for an adequate representation of the electron distribution far from the nuclei, particularly in the region where non-covalent interactions are strongest. One way to estimate the BSSE is via the counterpoise correction method of Boys and Bernardi (Boys and Bernardi, 1970) using the following equation (18)

$$\delta_{BSSE} = E_A^* + E_B^* - E_A^+ - E_B^+ \tag{18}$$

where δ_{BSSE} is the BSSE term, while E_A^*, E_B^* are the energy of A, B in the basis function of AB and the geometry of complex AB. E_A^+, E_B^+ are the energy of A, B in their own basis functions at geometry of the complex AB. So the corrected interaction energy (E_{INT}^{corr}) is given by equation (19)

$$E_{INT}^{corr} = E_{INT} - \delta_{BSSE} = E_{AB} - (E_A^* + E_B^*) - (E_A - E_A^+) - (E_B - E_B^+) \tag{19}$$

2. System studied

2.1 Neutral system

The present study investigates the electronic and structural properties of phosphoenolpyruvate carboxykinase (PEPCK) enzyme complexed with oxaloacetate (OAA). This structure was obtained from protein data bank (code 2QF1). Surrounding 13 amino acids in the radius of about 7 Å centered at oxaloactate were selected in this study. Amino acids in the studied system included Arg87, Gly236, Gly237, Lys243, Lys244, His264, Ser286, Ala287, Asp311, Phe333, Gly334, Arg405 and Phe485. The schematic representation of the investigated system is given in Figure 8. All amino acids, assumed to be in their neutral form, were terminated. The N and C termini of the residues were capped with an acetyl group (CH_3CO-) and methyl amino group (-$NHCH_3$), respectively (Figure 7). Hydrogen atoms were added to the X-ray structure to generate the complete structure of the model by using Sybyl7.2 program. However, all heavy atoms of the ligand-binding site complex were fixed at the positions taken from the X-ray structure. The added hydrogen atoms were adjusted to the initial complex structure by the geometry optimization using the PM3 method in the Gaussian 03 program package (Frisch *et al.*, 2003).

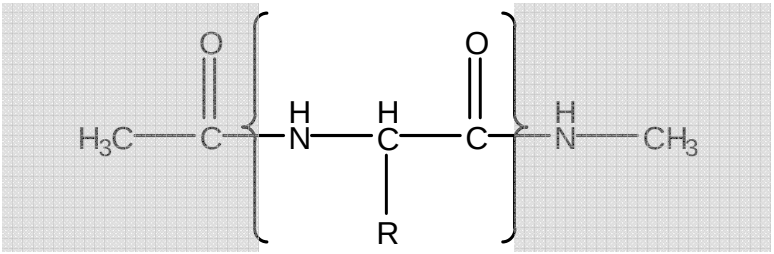

Figure 7 Capped groups of the terminal ends for amino acid chain.

The residues surrounding OAA included in this study are listed as following:

Ala86-Arg87-**Ile88**
Tyr235-Gly236-Gly237-**Asn238**
Gly242-Lys243-Lys244-**Cys245**
Glu263-His264-**Met265**
Pro285-Ser286-Ala287-**Cys288**
Asp310-Asp311-**Ile312**
Phe332-Phe333-Gly334-**Val335**
Ser404-Arg405-**Phe406**
Pro484-Phe485-**Phe486**

Where the capped groups of the terminal ends for amino acid chain showed in the bold character. The model composed of nine chains of the model study around the binding pocket of OAA-PEPCK complex.

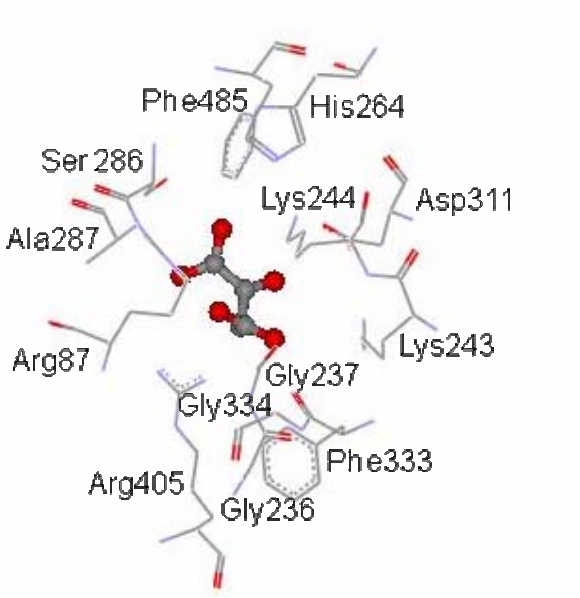

Figure 8 Schematic representation of oxaloacetate bound to the PEPCK. binding site, consisting of 13 residues.

2.2 Charge system

Another system that I studied was the charge system. In this system, it composed of oxaloacetate and individual amino acids surrounding the binding site. There are 5 amino acids that have charge Arg87 (+1), Lys243 (+1), Lys244 (+1), Asp311 (-1) and Arg405 (+1). The N and C termini of the amino acids were capped with an acetyl group (CH_3CO-) and methyl amino group (-$NHCH_3$), respectively. Hydrogen atoms were added to the X-ray structure to generate the complete structure of the model by using Sybyl7.2 program. However, all heavy atoms of the ligand-binding site complex were fixed at the positions taken from the X-ray structure. The added hydrogen atoms were adjusted to the initial complex structure by the geometry optimization using the PM3 method in the Gaussian 03 program package (Frisch *et al.*, 2003).

3. Interaction energies calculations

The interaction energies between oxaloacetate and individual amino acids (defined as X_i) around the binding site were calculated from the optimized complex structure obtained from PM3 semiempirical of calculation. The interaction energy (INT) is defined as following:

$$INT_{(OAA\text{-}Xi)} = E_{(OAA\text{-}Xi)} - [E_{(OAA)} + E_{(Xi)}]$$

Where $E_{(OAA\text{-}Xi)}$ is the energy of the complex structure of OAA bound to the pocket of PEPCK, $E_{(OAA)}$ and $E_{(Xi)}$ are the energies of OAA and individual amino acid, respectively.

The single point calculation at the MP2/6-31G(d, p) level with the basis set superposition error (BSSE) correction on which configuration, taken from the optimized geometry of OAA-PEPCK complex was performed for interaction energy. All calculations were performed by Gaussian 03, running on linux 3.4 GHz personal computer.

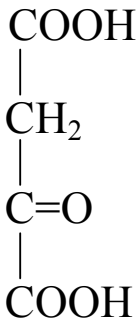

Figure 9 Chemical structure of oxaloacetic.