THE MOLECULAR DYNAMICS SIMULATION OF 2,6-DIHYDROXYBENZOIC ACID CRYSTAL IN ETHANOL, METHANOL AND P-XYLENE AT 20°C USING COMPASS FORCE FIELD

NURUL KALYISAH BNTI ABD KHAR

Thesis submitted in partial fulfillment of the requirements
for the award of the degree of
Bachelor of Chemical Engineering

Faculty of Chemical and Natural Resources Engineering
UNIVERSITI MALAYSIA PAHANG

JANUARY 2012

ABSTRACT

2,6-dihydroxybenzoic acid is an active pharmaceutical ingredient which will exhibits polymorphism when it is in crystalline form. Different polymorphs exhibit different physical properties. From previous study found, the selection of appropriate solvent plays a significant role in the formation of the desirable FI and FII polymorph. In this research, molecular dynamics (MD) were applied to simulate the behavior of 2,6-dihydroxybenzoic (DHB) acid crystal in ethanol, methanol and p-xylene at 20^{o}C using COMPASS force field. The objective of this research is to study the effect of various solvents on 2,6-dihydroxybenzoic acid crystal. MD simulation is performed using Material Studio for ethanol, methanol, p-xylene and 2,6-dihydroxybenzoic acid by applying COMPASS Force Field together with Forcite and Amourphous Cell Module. The dynamics run for pure solvent is performed initially in NVE ensemble at 200 ps and followed by NPT ensemble at 800 ps. From the trajectory files, the density, radial distribution function (rdf) and diffusion coefficient were analyzed and calculated. It is expected that MD simulation able to correlate rdf with the specific functional group in solvents structure. Based on the rdf, the probabilities of finding specific intermolecular interactions of hydrogen bond which is a relatively strong form of intermolecular attraction in specific solvents can be assessed. It is a type of attractive intermolecular force that exists between two partial electric charges of opposite polarity. The difference atomic interaction from the rdf trends show that the type of solvent use in 2,6-DHB crystallization process may lead to the polymorphism.

Keywords : radial distribution function, diffusion coefficient, hydrogen bonding

ABSTRAK

Asid 2,6-dihidroxibenzoik merupakan bahan farmaseutikal aktif yang akan menpamerkan polymorphism apabila berada dalam bentuk kristal. Polimorf yang berbeza menunjukkan sifat-sifat fizikal yang berbeza. Kajian sebelum ini mendapati pemilihan pelarut yang sesuai memainkan peranan penting dalam pembentukan polimorf FI dan FII yang dikehendaki. Dalam kajian ini, molekul dinamik (MD) telah digunakan untuk mensimulasikan tingkah laku kristal asid 2,6-dihidroxibenzoik (DHB) di dalam etanol, metanol dan p-xylene pada suhu 20^{o}C menggunakan medan daya COMPASS. Objektif kajian ini adalah untuk mengkaji kesan pelbagai pelarut terhadap polimorf kristal asid 2,6-dihidroxibenzoik. Simulasi MD dilakukan dengan menggunakan Material Studio untuk etanol, metanol, p-xylene dan asid 2,6-dihidroxibenzoik dengan menggunakan medan daya COMPASS bersama-sama dengan Modul Forcite dan Amourphous Cell. Dinamik yang dijalankan untuk pelarut asli dilakukan pada mulanya dalam ensembel NVE pada 200 ps dan diikuti oleh ensembel NPT pada 800 ps. Dari fail-fail trajektori, ketumpatan, fungsi taburan jejarian (rdf) dan pekali resapan dianalisis dan dikira. Ia dijangka bahawa simulasi MD dapat mengaitkan rdf dengan kumpulan berfungsi yang tertentu dalam struktur pelarut. Berdasarkan kepada rdf, kebarangkalian untuk mencari interaksi antara molekul khususnya ikatan hidrogen yang merupakan satu bentuk tarikan yang agak kuat antara molekul dalam pelarut tertentu boleh dinilai. Ia adalah sejenis daya tarikan antara molekul yang wujud di antara dua cas separa elektrik yang bertentangan kekutuban. Perbezaan interaksi atom dari graf rdf menunjukkan bahawa jenis penggunaan pelarut dalam proses penghabluran 2,6-DHB boleh membawa kepada polymorphism.

Katakunci : fungsi taburan jejarian, pekali resapan, ikatan hidrogen

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

t	Time
D	Diffusion coefficient
C	Concentration
g (r)	Radial distribution function
r	Radius
ρ	Density
N	Number of molecules

LIST OF ABBREVIATIONS

DHB	Dihydroxybenzoic Acid
COMPASS	Condensed-phase optimized molecular potential for atomistic simulation studies
NVE	Microcanonical Ensemble
NPT	Isobaric-Isothermal Ensemble
rdf	Radial distribution function
MSD	Mean square displacement
OPLS	Optimized Potentials for Liquid Simulations

CHAPTER 1

INTRODUCTION

1.1 Introduction

2,6-dihydroxybenzoic acid is an active pharmaceutical ingredients (APIs) which are frequently delivered to the patient in the solid state as part of an approved dosage form such as in the form of tablets or capsules. Solids provide a convenient, compact and generally stable format to store an API or drug product. When it is in crystalline form, it exhibits polymorphism. According to Alfred et al., polymorphism in molecular crystals is a common phenomenon and is of great interest to the pharmaceutical field. The solid state form is a key quality attribute of a crystalline product. Severe consequences of drug products might happen due to the inconsistencies in the solid phase produced during the manufacturing and storage of drug substance. Therefore, it is essential to understand the solid state behavior of the drug and select the optimal solid form for development. Generally, the solid phase of a material whether formed of organic molecules, inorganic ions or extended covalent network, can exhibit different structures which possess the same chemical composition. However, different polymorphs will exhibit different physical properties. In this case, selection of appropriate solvent plays significant role in the formation of desirable polymorph.

1.2 Problem Statement

Polymorphism is a common phenomenon in the chemical and pharmaceutical industries. Polymorphs generally have different solubility which can also assist in identifying their stability. Currently, the experimental work cannot describe or give the insight of the process at the molecular level of solvent effect on 2,6-dihydroxybenzoic

acid polymorphism. Therefore, another alternative has been applied to solve this problem which is by using molecular dynamics (MD) simulation technique. The MD simulation is an effective way to study the relationship between the structures of the molecules. Besides, MD simulations can describe the motion of individual atoms in a way which is not possible in laboratory experiments. Subsequently, the choice of an appropriate solvent is critical to crystallization of chemical product. In the pharmaceutical industry, it is necessity to control the crystal size and shape distribution for the products crystal (Ma et al., 2002). Crystal size and morphology affect the ease of separating, washing, drying, packaging, handling and storage. In the case of pharmaceutical products, crystal morphology affects dissolution characteristics, bioavailability and the ease with which the crystals are compressed into tablets. Factors such as temperature, mixing intensity, solvent, supersaturation and addition of seed crystals have been routinely used to control the size and morphology of product crystals (Jones et al., 2005). There are instances where some solvents belonging to a particular class give completely different polymorph from other solvents in the same class. In view of the importance of product crystal polymorph, in this research several solvents are selected to be studied for the crystallization of 2,6-dihydroxybenzoic acid crystal in order to obtain the desired polymorph.

1.3 Objective of Study

This research is conducted in order to achieve the following objective which is to study the effect of various solvents on 2,6-dihydroxybenzoic acid crystal polymorph.

1.4 Scope of Study

For this research, three solvents having widely varying polarity were selected such as ethanol, methanol and p-xylene for the crystallization of 2,6-dihydroxybenzoic acid. The selection of solvents were based on the ability to form the hydrogen bond acceptor or donor between the 2,6-DHB solute molecule. These solvents possess different properties such as polarity and aromaticity. This study is done at 20^{o}C. This temperature is fixed the same for all the solvents used so that it gives an isothermal

condition. Besides that, the pressure is constant at 1bar or 0.0001GPa. The third generation force field, which is COMPASS Force Field, will be used to simulate the liquid system. According to Acclerys, COMPASS is a force field which enables accurate prediction of structural, conformational, vibrational, and thermophysical properties for a broad range of molecules in isolation and in condensed phases, and under a wide range of conditions of temperature and pressure.

1.5 Rationale and Significance

The type of solvent is a major factor in polymorphic selectivity and resulting crystal morphology. Thus, the rationale for this study is to understand the polymorphism of 2,6-dihydroxybenzoic acid crystal by the effect of solvent type at temperature 20^{o}C. After that, from this research we can compare the result obtained with the previous researched. This is in order to gain more understanding about the effect of solvent towards the polymorphism of 2,6-dihydroxybenzoic acid crystal whether it gives Form I (metastable), Form II (stable) or monohydrate.

CHAPTER 2

LITERATURE REVIEW

2.1 Molecular Dynamic Simulation

2.1.1 Definition

MD simulation is an effective way to study the relationship between the structures of molecules and the properties of the bulk system (C.G. Hanke et al., 2001). It provides a molecular level picture of structure. The method allows the prediction of the static and dynamic properties of molecules directly from the interactions between the molecules. Besides, MD simulation also provides an atomistic picture of how solute-solvent interactions can significantly affect the initial association of solute molecules to the point that they determine the polymorphic outcome of the crystallization. (Hamad et al., 2006)

2.1.2 Concept

In molecular dynamics, successive configurations of the system are generated by integrating Newton's laws of motion. The result is a trajectory that specifies how the positions and velocities of the particles in the system vary with time. Newton's laws of motion can be stated as follow where first, a body continues to move in a straight line at constant velocity unless a force acts upon it. Second, force is equals to the rate of change of momentum and third is for every action there is an equal and opposite reaction. The trajectory is obtained by solving the differential equations embodied in Newton's second law, $F = ma$:

$$\frac{d^2y}{dt^2} = \frac{F_{x_i}}{m_i}$$

This equation describes the motion of a particle of mass m_i along one coordinate (x_i) with F_{xi} being the force on the particle in that direction (Andrew, 1996).

2.1.3 Van der Waals Forces

According to Encyclopedia Britannica, van der Waals forces are relatively weak electric forces that attract neutral molecules to one another in gases, in liquefied and solidified gases and in almost all organic liquids and solids. The forces may arise from three sources. First, although the molecules in some materials are electrically neutral, it may be permanent electric dipoles. One side of a molecule is always positive and the opposite side is negative. It is because of fixed distortion in the distribution of electric charge in the structure of some molecules. The tendency of such permanent dipoles to align with each other will results in a net attractive force. Second, inducing polarization which is due to the presence of molecules that are permanent dipoles temporarily distorts the electron charge in other nearby polar or nonpolar molecules. The interaction of a permanent dipole with a neighbouring induced dipole results to an additional attractive force. Third, even though no molecules of a material are permanent dipoles, a force of attraction exists between the molecules accounting for condensing to the liquid state at sufficiently low temperatures.

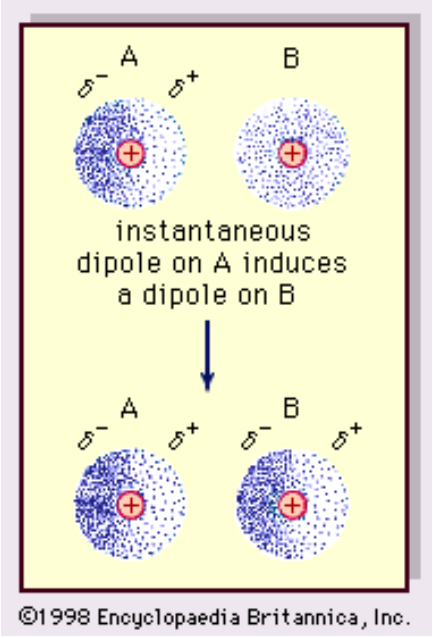


Figure 2.1 : The weak dipole attraction of the van der Waals bond.

2.1.4 Intermolecular and Intramolecular Interaction

Intramolecular forces operate within the molecules while intermolecular forces operate between molecules. Significantly, intramolecular forces are much stronger than intermolecular forces. For intramolecular forces, these are the covalent bonds within a molecule of a molecular substance, the forces between ions in ionic compound and the forces between atoms in metal. Intermolecular forces operate between the molecules of a covalent substance, the atoms of a monoatomic element or the ions of one substance and the molecules of another. Typically, intermolecular interactions are weak forces such as van der Waals and stronger short range ones including hydrogen bonding that are believed to determine the packing of organic molecules during the crystal growth process. A different packing of the same molecules leads to the formation of a new crystal structure. Li et al studied 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile, so called ROY which has the largest number of solved polymorphs by developing an electronic concept or function within the framework of conceptual density functional theory, namely crystallization force and able to reveal intermolecular interactions in the crystals.

2.1.5 Radial Distribution Function

Radial distribution function is defined as the probability or occurrence of atoms with type y at a distance r from atoms of type x in the liquid. Furthermore, according to Everyscience, radial distribution function, $P_{nl}(r)$ is the probability that an electron in the orbital with quantum numbers n and l will be found at a distance r from the nucleus. It is related to the radial wave function by the following relationship:

$$P_{n,l}(r) = 4\pi r^2 R_{n,l}(r) \;;$$

Normalized by,

$$\int_0^\infty P_{nl}(r)dr = 1$$

The factor $4\pi r^2$ arises because the radial distribution function refers to the probability of finding an electron not at a specific point in space but on a spherical shell of area $4\pi r^2$, at a distance r from the nucleus. The integral results from the fact that the total

probability of finding the electron is one, as it must be found somewhere around the nucleus. Besides, radial distribution function also refers to molecular structure of liquids. According to Encyclopedia Britannica within the molecular structure of liquid, for a complete understanding of the liquid state of matter, an understanding of behaviour on the molecular level is necessary. Such behaviour is characterized by two quantities called the intermolecular pair potential function, *u*, and the radial distribution function, *g*. The pair potential gives information about the energy due to the interaction of a pair of molecules and is a function of the distance, *r* between their centers. Information about the structure or the distances between pairs of molecules is contained in the radial distribution function. In this case, radial distribution function measures the probability of finding the centre of one molecule at a distance *r* from the centre of a second molecule. For values of *r* less than those of the molecular diameter, *d*, *g* goes to zero. This is consistent with the fact that two molecules cannot occupy the same space. The most likely location for a second molecule with respect to a central molecule is slightly more than one molecular diameter away, which reflects the fact that in liquids the molecules are packed almost against one another. The second most likely location is a little more than two molecular diameters away, but beyond the third layer preferred locations damp out, and the chance of finding the centre of a molecule becomes independent of position. From Figure 2.2 below, considering a homogeneous distribution of the atoms or molecules in space, the radial distribution function represents the probability to find an atom in a shell *dr* at the distance *r* of another atom chosen as reference point.

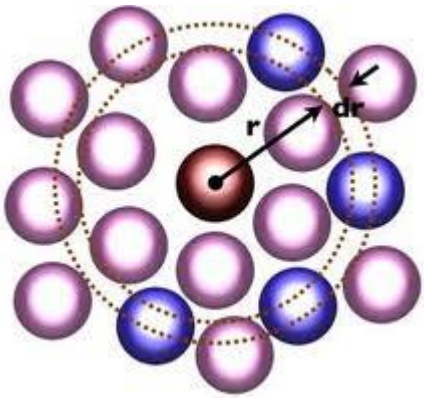


Figure 2.2 : Space for the evaluation of radial distribution function.

2.1.6 Diffusion

Molecular diffusion can be defined as the transfer or movement of individual molecules through a fluid by means of the random, individual movements of the molecules. Diffusion of molecules is due to a concentration gradient. Other driving forces for diffusion also occur because of temperature, pressure electrical potential and other gradient. In many industrial processes such as separation operations, diffusion of solute in liquids is very important especially in such separation operations as liquid-liquid extraction or solvent extraction, gas absorption and distillation. In nature, diffusion occurs in many situation such as diffusion of salt ni blood. (Geankoplis, 2003).

2.2 Crystallization

2.2.1 Definition

Crystallization is a solid-liquid separation process, in which mass transfer of a solute from the liquid to a pure solid crystalline phase occurs. It is a process where solids particles are formed from a homogeneous phase for example in the formation of solid crystals from a liquid solution. It is used extensively as a purification and separation process in industry due to its ability to provide high purity separations. In crystallization the solution is concentrated and usually cooled until the solute concentration becomes greater than its solubility at that temperature. Then the solute comes out of the solution, forming crystals of approximately pure solute. The yield, purity, sizes and shapes of the crystals are important in commercial crystallization (Geankoplis, 2003). In the manufacture of pharmaceutical compounds, crystallization is a critical unit operation which performing a purification process in the synthesis of an API. Crystallization of an API in particular is essential for product qualities such as chemical purity and correct polymorphic form which need to be controlled to meet set specifications (Kim S. et al., 2005).

2.2.2 Crystal Growth

Crystal growth can be defined as the increase in size of crystals as solute is deposited from solution that will determine the final crystal size distribution which is an important product attribute. Crystal can grow faster than they nucleate at low supersaturation which resulting in a larger crystal size distribution. However, crystal nucleation dominates crystal growth at high supersaturation which resulting in smaller crystals. It is also known as the series of processes by which an atom or molecule is incorporated into the surface of a crystal which causing an increase in size. The processes are either transport process or surface process. Transport process include transport of atoms through solution while surface process include attachment of atoms to the surface, movement of atoms on the surface and attachment of atoms to edges and kinks. Growth can be transport or surface controlled depending on the slowest process that control overall crystal growth (Pablo and Michael).

2.2.3 Crystal Polymorphism

Polymorphism is a common phenomenon in the chemical and pharmaceutical industries. Polymorphism can be defined as the ability of the same chemical substances which are in solid state to exist in different crystalline structures (Findlay et al. 1951) which is a regular, repeating arrangement of atoms or molecules in the solid state. It means that, a compound may form crystalline solids which have different three dimensional structures resulting in polymorph (Britain, 1999). Different polymorphs of a given substance are chemically identical but will exhibit different physical properties and behavior, such as different solubility and melting points. Besides, crystal polymorphism is an important factor related to the physicochemical and biological properties of drug substances and formulations. Gautam and Arijit from the Indian Institute of Science has found that when 4,4’-bipyridine and 4-hydroxbenzoic acid were dissolved together in a solvent such as methanol, they would co-crystallize to form two different polymorphs. Polymorphism has profound implications for a variety of phenomena including pigment properties, solid state reactivity and also pharmaceutical performance as it affects every phase of drug development, from initial drug discovery to final clinical evaluation, including patent protection and competition in the market.

One of the implication examples in pharmaceuticals is the dissolution rate of a drug which can be dependent on the polymorphic form. Therefore, a critical challenge is the early identification of possible polymorphs (H.H. Tung, 2009).

2.2.4 Driving Force for Crystallization

Supersaturation is a solution that contains more of the dissolved material than could be dissolved by the solvent under the solubility amount. It is also defined as the difference in chemical potential between a molecule in solution and that in the bulk of the crystal phase (Pablo and Michael). It is the driving force for solution crystallization process within crystal nucleation and growth. Crystallization process and product quality can be controlled by controlling the level of supersaturation during the process.

2.2.5 Solvent Role in Crystallization

A solvent is a liquid chemical that dissolves solute, resulting in a solution that is soluble in a certain volume of solvent at a specified temperature. It can react with other chemicals to form weak bonds that suspend both chemicals in a stable form. For a chemical to be a good solvent, it must be able to surround the dissolving chemical and prevent separation of the two substances while temperature and pressure are maintained. Not all solvents work with all chemicals, so the correct solvent must be chosen. Solvents can be classified into two categories which are polar and non-polar. As demonstrated in previous studies (Khoshkhoo et al.,1993), on the basis of knowledge of the temperature dependence of the solubility, recrystallization experiments, in which the supersaturation was systematically varied, were carried out in an attempt to isolate each of the polymorphic forms from each solvent system. These recrystallization experiments reveal that not all of the known polymorphic forms can be crystallized from any given solvent by varying the supersaturation. Indeed some solvents selectively favour the crystallization of a particular form or forms. Besides, by varying the solvent, it will affect crystal morphology, crystallization kinetics, formulation and drug stability (H.H. Tung, 2009).

2.3 Force Field

2.3.1 Definition

In the context of molecular modeling, a force field refers to the form and parameters of mathematical functions used to describe the potential energy of a system of particles typically molecules and atoms. Force field functions and parameter sets are derived from both experimental work and high-level quantum mechanical calculations. "All-atom" force fields provide parameters for every type of atom in a system, including hydrogen, while "united-atom" force fields treat the hydrogen and carbon atoms in methyl and methylene groups as a single interaction center. "Coarse-grained" force fields, which are frequently used in long-time simulations of proteins, provide even more crude representations for increased computational efficiency (Molecular, 1998). The goal of a force field is to describe entire classes of molecules with reasonable accuracy. In a sense, the force field interpolates and extrapolates from the empirical data of the small set of models used to parameterize the force field to a larger set of related models. Some force fields aim for high accuracy for a limited set of element types, thus enabling good prediction of many molecular properties. Other force fields aim for the broadest possible coverage of the periodic table, with necessarily lower accuracy.

2.3.2 Parameterization of the force field

In addition, a force field can also be defines as a set of parameters for each type of atom. For example, a force field would include distinct parameters for an oxygen atom in a carbonyl functional group and in a hydroxyl group. The parameter set includes polarizability, atomic mass, and partial charge for individual atoms, and equilibrium values of bond lengths and angles for pairs, triplets, and quadruplets of bonded atoms. Although many molecular simulations involve biological macromolecules such as proteins, the parameters for given atom types are generally derived from observations on small organic molecules that are more tractable for experimental studies and quantum calculations. Parameter sets and functional forms are defined by force field developers to be self-consistent. Because the functional forms of the potential terms vary extensively between even closely related force fields, the

parameters from one force field should never be used in conjunction with the potential from another (Stefano Marino, 2006).

2.3.3 COMPASS force field

COMPASS stands for Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies. It is a powerful force field that supports atomistic simulations of condensed phase materials. Since most parameters are initially derived based on *ab initio*, then it is called an *ab initio* force field. It is the first *ab initio* force field that has been parameterized and validated using condensed-phase properties in addition to various *ab initio* and empirical data for molecules in isolation. Consequently, it enables accurate and simultaneous prediction of structural, conformational, vibrational, and thermophysical properties, that exist for a broad range of molecules in isolation and in condensed phases, and under a wide range of conditions of temperature and pressure (Accelrys, 2011).

CHAPTER 3

METHODOLOGY

3.1 Simulation Work

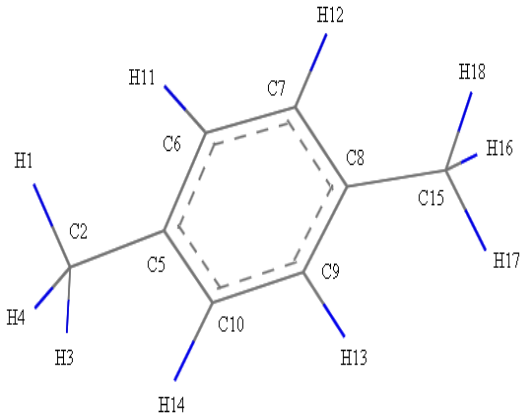
The MD simulations technique applied in this work was the same employed in earlier MD studies of 2,6-DHB in toluene and chloroform solvent (F. Adam et al). Prior to simulation, the single molecules of methanol, ethanol and p-xylene were optimized and minimized using the generic COMPASS force field in Material Studio. The cubical boxes containing number of molecules by 216 for methanol, ethanol and p-xylene pure solvents were created using Amorphous Cell in Material Studio. Using the same way, the cubical boxes of 2,6-DHB in methanol, ethanol and p-xylene solvents respectively were also created containing 216:108 of solvent:solute ratios. Then, energy minimization of simulation boxes was applied. The molecular structures of the solvents and solute molecules are shown in Figure 3.1. Initially, the boxes were equilibrated using Forcite Module by performing the molecular dynamics simulation in NVE ensemble for 200ps and followed by NPT ensemble for 800ps with the total of 1ns simulation time at the set point 293K and 0.0001GPa. For the dynamics run in the NVE ensemble, the temperature and pressure of the system is scaled down until the values achieved are consistent with the setting conditions in order to get a reasonable total energy for the system. Besides, dynamics run in the NPT ensemble will maintain the temperature and pressure of the system and control the simulation box size to achieve the density of the real system. Once the temperature, pressure and energy of the system were in equilibrium at the desired values, the radial distribution function of the systems were calculated and analyzed from the trajectory files. For analyzing, cutoff for half of A length with 0.5 interval were chosen to perform the trajectory files. The radial distribution function depends on the density and temperature. Therefore, it is serves as

an indicator of the nature of phase assumed by the simulated system. The radial distribution function, g_{xy} *(r)* gives the probability or occurrence of atoms with types y at a distance r from atoms of type x in the liquid as the equation below

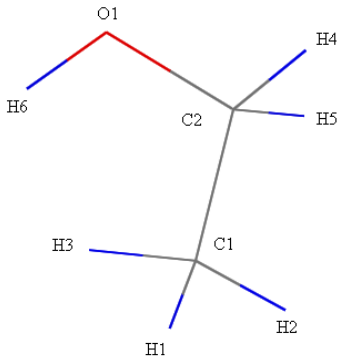
$$g_{xy}(r) = \frac{\langle N_y(r, r+dr)\rangle}{\rho_y 4\pi r^2 dr}$$

3.2 Molecular Labeling

From the Figure 3.1 below, the molecular structures were labeled by naming every single atom in the molecules. It is because all the atoms play a significant role in the polymorphism of 2,6-DHB based on their ability to form hydrogen bonding with particular neighbouring atoms in the solution system. The label for 2,6-DHB structure were adopted from the previous MD study in order to facilitate in performing the comparison between this research with the previous one.



(a) p-xylene



(b) ethanol