CARDIFF UNIVERSITY

APPLICATION OF ATOMISTIC MODELLING TO MOLECULAR SOLIDS CONTAINING HYDROGEN BONDS

by

HASMERYA MAAROF

A thesis submitted to

CARDIFF UNIVERSITY

For the degree of

DOCTOR OF PHILOSOPHY

School of Chemistry

Aug 2009

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date 18/11/09

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD (insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed ..... (candidate) Date 18/11/09

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed ..... (candidate) Date 18/11/09

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date 18/11/09

Application of Atomistic Modelling to Molecular Solids Containing Hydrogen Bonds

Abstract

The work presented in this thesis is mainly concerned with crystal structures containing hydrogen bonds.

Chapter 1 and 2 mainly discuss the background and basic concepts used in this study such as the importance of hydrogen bond in crystal engineering, co-crystals and polymorphism, and recent studies of urea co-crystals.

Chapter 3 is a study about urea/α,ω-dihydroxyalkanes co-crystal structures. It begins with parameterising DMAREL to obtain lattice energy from a set of homologous co-crystals where DMA multipoles were generated from different method, GDMA and MOLPRO. The simulated lattice energy, structures and interaction energy were discussed and compared whether there is possibilities for the co-crystals to appear in different urea ribbon structures (parallel and anti-parallel) as these could not be crystallised experimentally. The energy data shows that urea/α,ω-dihydroxyalkanes co-crystal of anti-parallel ribbon type structure are more thermodynamically favoured compared to the parallel structure. In latter part of this chapter, attempts to construct and simulate the anti-parallel urea ribbon co-crystal type structures from initial experimental structures were discussed.

In Chapter 4, sulfur pair potential was modelled to fit for use in DMAREL for TTCA structure simulation. The original potentials were taken from Lennard Jones potential of α-S_8 crystal structure. Initially, the potentials were modelled against α-S_8 and thiourea crystal. Improved potentials were applied to a set of S-contained structure, specifically with similar environment to TTCA to validate the reliability of this potential against other molecules. Potentials works fairly well for 5 out of 10 molecules simulated, where TTCA shows poorest performance against the potential even though it has improved from the original sulfur potential. Contrasting crystal structure between TTCA and CA when substituted between each other is discussed.

The final chapter, Chapter 5 is the continuation from work in Chapter 3. In urea co-crystal, it was found that urea structure was not exactly planar. We then continue on the search of different conformation of urea molecule in as and solid structure. First, the conformation of urea monomer were discussed and followed by calculation of larger planar urea clusters where one of the urea is substituted with either C_2 or C_s conformers. Urea clusters were build systematically mimicking dense urea crystal structure. Planar structure was finally obtained by using 5 urea molecules involving four hydrogen bonds.

Dedicated to my dearest parents

Acknowledgements

First and foremost, I would like to express my praise and gratitude to God Almighty for giving me the chance to successfully completing this piece of research work. Surely this knowledge that I have is only a tiny piece of His many other undiscovered knowledge.

I would like to acknowledge the following people who have participated in one way or another during the time spent working on this PhD., and the writing of this thesis.

Dr. Dave Willock and Professor Kenneth D.M. Harris, my supervisors, for their guidance, continuous support and enthusiasm during the time spent working in their laboratory. Also, not forget to mention, their encouragements and suggestions throughout my writing up duration.

James Landon, Paula, and Kara Howard who have helped me in understanding some computer simulation techniques used in this study.

Dr Jamie Platts, Dr. Philip Davies and Dr. Liling Ooi for their useful discussion and suggestions of my research backgrounds and findings.

I would like to thank all members in KDMH group, both path and present; Dr. Eugene Cheung, Dr. Fang Guo, Dr. Colan Hughes, Dr. Zhongfu Zhou, Dr. Javier Marti Rujas, Dr. Ming Can Xu, Dr, Zhigang Pan, Anabel Morte-Rodenas, Lim Gin Keat, Kemal Testamur, Nathan Dedman, Ben and Dr. Kotaro Fujii. It was a pleasure and great fun working with them.

Friends; Sis Masni, Sis Sawiah, Hasliza Bahruji, Sharifah Norashikin, Fang Jia, Ashriya, and many others. Thank you for your great encouragement and countless prayers. I really acknowledge all your supports.

My thanks and gratitude also goes to my dearest Mum and Dad, both my little brothers (Kheiry & Reza) and sister (Saleha) for their love, supports and countless prayers throughout these years.

Finally, I would like to thank Malaysian Government for the financial support and Cardiff University for the opportunity to precede my postgraduate study. It has been a great experience for me!

Thank you!

Table of contents

List of Figures

List of Tables

Glossary of Terms and Abbreviations

(In Alphabetical Case and Order)

DFT	Density Functional Theory
AO	Atomic Orbital
HF	Hartree Fock
STO	Slater Type Orbital
LCAO	Linear combination of atomic orbital
MO	Molecular orbital
AIM	Atom in Molecule
GDMA	Distributed Multipole Analysis of wavefunctions calculated by the Gaussian system of programs
DMAREL	Energy minimisation packages to simulate rigid molecules with multipoles
MOLPRO	System of *ab initio* programs for molecular electronic structure calculations (Quantum chemistry package)
CONDOR	Cardiff University's High Throughput Computing (HTC)
MP2	Møller–Plesset second order perturbation theory
DMA	Distributed Multipoles Analysis
rms	Root mean square
LJ	Lennard-Jones
1,4-dhb	1,4-dihydroxybutane
1,6-dhh	1,6-dihydroxyhexane
1,8-dho	1,8-dihydroxyoctane
1,10-dhd	1,10-dihydroxydecane
1,12-dhdd	1,12-dihydroxydodecane
1,14-dhtd	1,14-dihydroxytetradecane
TTCA	Trithiocyanuric acid
CA	Cyanuric acid

Chapter 1

Introduction

1.1 Aims

This thesis is part of a larger scheme to understand about crystal structure containing hydrogen bonds. Hydrogen bonds can contain contributions from either:

i) Electrostatics;
ii) Polarisation;
iii) Charge transfer

This thesis will concentrate on the first contribution, which is expected to be the major contributor in the molecule studies and which determines the directionality of hydrogen bonds in crystal. Crystal structures that have been achieved experimentally often needs detailed understanding especially about the structure of molecular process during the crystallisation, where sometimes the same molecule may crystallised into different type of crystal arrangements (polymorphs). The possibility to predict or engineer similar type of molecule crystals or crystals with the same molecular arrangements may also be achieved based on understanding on its hydrogen bond network and electronic/energy properties.

In our study, we will first try to understands the lattice energy properties of urea/α,ω-dihydroxyalkanes co-crystals by comparing between two urea ribbons structure type. These will determine whether the mentioned co-crystal may or may not exist in both structures (probability to form a polymorph). Experimental urea co-crystal structure obtained from previous work will then be constructed in search of its missing structure in the odd number C_n analog series in urea anti-parallel dimers.

A study of how crystal structures known to have similar structure, but with different symmetry and space group are also studied to look at the relationship of both different space groups, and the possibility of both molecules exist in each other's space groups. Finally, we will have a look at the formation of urea structures and how the number of hydrogen bonds and neighbouring urea affect the planarity of pyramidal structure.

1.2 Role of Computational Studies in Solid State Chemistry

Computational studies are now a central technique in solid state science [1-6]. The field is widely developed ranging from reproduction of experimental data to calculations in predicting detailed description of highly complex systems. Solid state which relates to the computational studies are usually took place in the crystal molecular forms.

Crystal packing and lattice energy simulation has been of one of the interest area in solid state computational study [7,8]. General aspects that should be considered in simulating a crystal structure are:

a) Evaluation of the energy function such as the force field, charges, polarization
b) Space group symmetry

Crystal packing simulation usually involves electrostatic energies and intermolecular forces between molecules in a crystal structure. An example of this is shown by Eckhardt which involves simulation studies in predicting structural and energetic feature of crystal compounds [9]. Method development of the new atom-atom Coulombic energy (PIXEL) were improved and compared from the advanced point-charge method for use in crystal energy determination. Another approach is using *ab initio* calculation in studying solid phase urea with a force field optimised in condensed phase [10]. Crystalline structure of urea were simulated in the temperature range from 12 to 293 K from this study gave densities as well as cell parameters within 2% difference from the experimental data.

Crystal prediction studies which involve computational simulation are quite popular with the large demand in crystal engineering field. Every year [11,12], new and simplified methods in crystal prediction were developed rapidly. The molecules studies evolved from small and rigid molecules [7,13] to larger and flexible molecules [14]. This also improves as the capability of recent computer technology increased.

1.3 Hydrogen Bonds in Crystals

Hydrogen bonding is one of the concepts that influence a lot to the physical properties of molecule compounds, especially molecules with high polarity. There are two main contributors to hydrogen bonding; the acceptor, A and donor, X (Figure 1.1). In the formation of hydrogen bond, the heavy atom of the proton donor group, X-H withdraw electron density and leaves the proton partially unshielded. This will then interact with the acceptor, A which have lone pair electrons or polarizable π electrons [15].

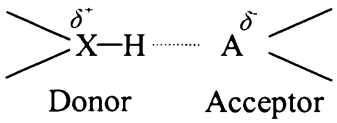


Figure 1.1: Schematic picture of donor and acceptor in hydrogen bond

Since a hydrogen bond can interact in a long range distance, a donor can interact with two or three acceptors simultaneously [16]. It is possible to have more than three hydrogen bond acceptors in principle, but they are rarely found in practice because it requires very high spatial densities of acceptors. A notable example is urea in which each oxygen acceptor interacts with four N-H donors.

These hydrogen bonds can be classified into three types, which are strong, moderate and weak hydrogen bonds (Table 1.1).

Table 1.1: Properties of hydrogen bonds [15,16]

	Strong	**Moderate**	**Weak**
A-H···X interaction	Mostly covalent	Mostly electrostatic	Electrostatic/ dispersion
Bond lengths	X-H ≈ H···A	X-H < H···A	A-H << H···X
H···X (Å)	~1.2 – 1.5	~ 1.5-2.2	2.2-3.2
A···X (Å)	2.2 – 2.5	2.5-3.2	3.2-4.0
Bond Angles (°)	175-180	130-180	90-150
Bond Energy (kcal mol^{-1})	14-40	4-15	<4
Relative IR υ_s vibration shift (cm^{-1})[a]	25%	10-25%	10%
Examples	• Gas-phase dimers with strong acids or strong bases • Acid salts • Proton sponges • Pseudohydrates • HF complexes	• Acids • Alcohols • Phenols • Hydrates • All biological molecules	• Gas phase dimers with weak acids or weak bases • Minor components of 3-centre bonds • C-H···O/N bonds • O/N-H···π bonds

[a]Observed υ_s relative to υ_s for a nonhydrogen bonded X-H.

Theoretical studies have played an important role in understanding the properties of hydrogen-bonded systems. One of the most widely used methods is density functional theory (DFT). DFT methods have been used a lot in solid-state physics methodology that over the last three decades has been transplanted and transformed, very successfully into chemistry. In DFT, calculation of electron density was used in describing electronic structure of a molecule. According to recent reports [17-20],

only hybrid functional can provide an accurate description for the systems with hydrogen bond.

Over the years, this method has been improved to produce more accurate approximations of exchange correlation functional.

1.3.1 Graph Set Theory

Hydrogen bonded patterns can be considered as an interest in studies of intermolecular interaction. First introduced by Etter, graph-set theory is used for categorizing hydrogen-bond motifs in a way that complex hydrogen bond patterns can be described more systematically and consistently [21,22]. The method is based on viewing hydrogen-bond pattern topologically as if they were intertwined nets with molecules as the nodes and hydrogen bonds as the lines rather than detailed geometrical parameters of its hydrogen bond networks.

The set of hydrogen bonds in the molecule structure are called an array, which specifically assigned first to motifs and these will then connect together to produce a network. Motifs are assigned to determine the preferred pattern using a set of notation, $G_d^a(r)$. Where G describes the type of pattern; C (chain), R (ring), D (dimer or non cyclic finite pattern), and S for intramolecular hydrogen bond. The subscript d and superscript a are refers to donors and acceptors, while r indicate the number atoms involved in the motif. If the donor and acceptor are one and the same, the subscript and superscript are ignored.

For example, in urea dense crystal structure have two types of motifs, ring and chain as shown in Figure 1.2

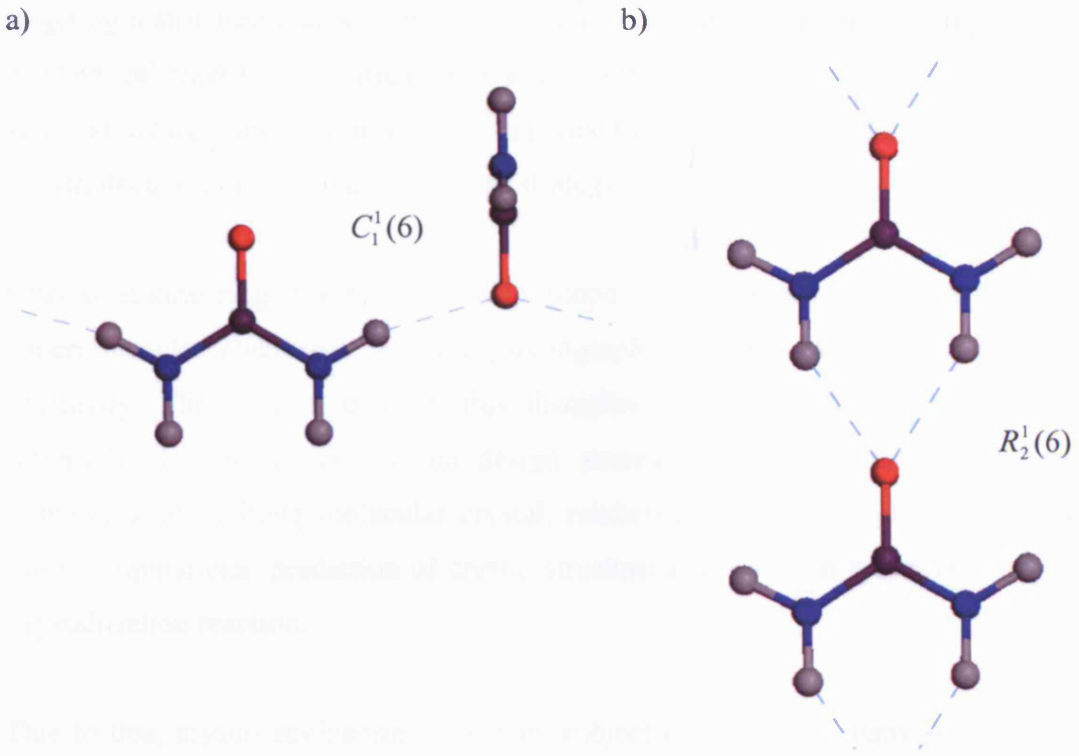


Figure 1.2 a) Example of urea dimer with molecular plane perpendicular, b) Example of urea dimer with molecule coplanar [Atom: Blue – Nitrogen, Red – Oxygen, Maroon – Carbon, and Grey – Hydrogen]

which then combined to make a network of $C(6)\ R_2^1(6)$. The topological descriptors are also useful in decoding packing networks to differentiate polymorphic crystal systems [23].

Investigating the effect of hydrogen bonds, and molecular structures could provide an understanding and designing the intermolecular interactions that dictate crystal packing. This will eventually play important role in crystal engineering.

1.4 Crystal engineering

Crystal engineering is a branch of supramolecular chemistry which is concerned with the design and synthesis of extended structures with predictable form and function [24]. It involves designing different types of molecular solids and often aimed at

targeting a structure that will give rise to a specific desired solid-state property, such as chemical reactivity, electrical, optical or magnetic properties [25]. Functional solid state structures are synthesized from neutral of ionic building blocks using intermolecular interaction in the design strategy.

Crystal engineering covers very wide scope of disciplines. It can overlap with supermolecular chemistry, X-ray crystallography, material science and solid state chemistry. The concept used in this discipline is the manipulation of nature of intermolecular interaction as the design strategy. This has great connection in explaining or defining molecular crystal, relationship to polymorphism, knowledge base computational prediction of crystal structure and efforts at mapping pathway of crystallisation reaction.

Due to this, crystal engineering has been subject of interest to many solid state and structural chemist, where it is important to design other kind of organized phases and assemblies. Work on variations in three dimensional arrangements of a complex and combinatorial-type synthesis of supramolecular assemblies of each molecular entity will unravel many possible structural features and molecular recognition features. Besides, it is important for computational studies concerning intermolecular interactions such as molecular simulations for polymorph predictions and crystal structure prediction [26]

One part of crystal engineering work involves crystal design and function. In pure organic crystal design, molecule with three dimensional with comparable (and strong) interactions at three directions is predicted to have three dimensional crystal structure.

Aromatic molecules, on the other hand will have two dimensional structural controls in the plane of the aromatic rings. This is due to the natural behaviour of this type of molecule which have $\pi...\pi$ C-H...O molecular interaction [27]. Therefore, this type of molecular interaction between molecules will be much weaker than others (i.e: O-H...O and N-H...O hydrogen bonding) in different direction. The three dimensional structure will also be difficult to describe in designing this type of molecules.

Braga presented a paper on the recent trends in crystal engineering. There, he mentioned that there are four aspects on the development and future prospect in crystal engineering [28]. First, it is in intermolecular interaction on the evaluation and application to crystal design, second, in the network design and application, third, on the approach of crystal engineering to crystal synthesis and lastly, on the polymorphism, solvate and chiral crystal resolution.

As a conclusion, crystal engineering involves not only the study and understanding of intermolecular interactions and the application in the bottom up construction of sophisticated solid state superstructures starting from suitably chosen ionic and/or molecular building blocks, but has also moving forward towards an all purpose mature discipline, a science without borders, where the motivation can be utilitarian and economical but also aesthetical and/or fuelled up by pure, quintessential, scientific curiosity [29].

1.5 Co-crystal

Crystal engineering brings us to the topic of organic solid structure where classification of organic solids structure can easily be explained in Figure 1.3. From the chart, organic solids can exist in either crystalline or non-crystalline structure. All crystalline structures can be polymorphic, which means it can exist in multiple structure arrangements. Single component of crystalline structure is called a dense structure where depending on the size; it can be either a single crystal or a nano-crystal material.

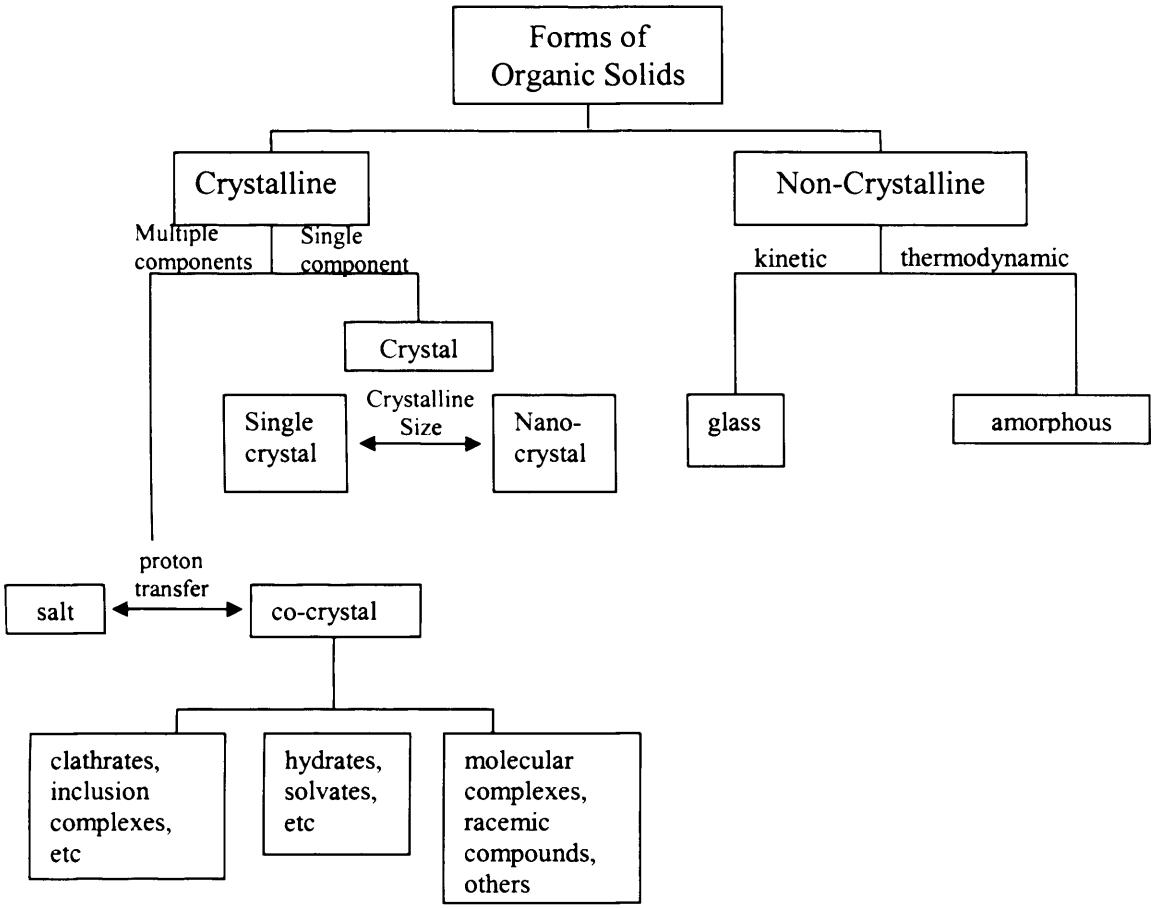


Figure 1.3: Classification chart of organic solids. Reproduced from Reference 30

Co-crystals can be defined as a crystal which contains two or more components together. It encompasses molecular compound, molecular complexes, hydrates, solvates, inclusion compounds, channel compounds, clathrates, and possibly other types of multi-components crystals [31]. Stahly [30] defines a salt as a multi-component system because it contains a single ionic compounds, although it also consists of multiple ions. Formation of salts of organic compounds involves proton transfer from an acid to base.

Co-crystals structures can also be a host-guest structure where host is the main component of the crystal and guest is another molecule which fills the space of the host. For example, urea molecules can exist in dense crystal structure [32,33], co-crystal [34] and urea inclusion compound [35] type of co-crystal with other molecule.

1.6 Polymorphism

As mentioned previously, crystalline solid may lead to polymorphism. Polymorphism is the ability of a solid material to exist in more than one form of crystal structure. It can be found in any type of crystalline material including polymers, minerals, and metals and is related to allotropy, which refers to elemental solids. For example, out of 245 crystals structure screened, 90% have multiple crystalline structure and non-crystalline form [30]. The difference in energies between one polymorph to another are quite small, typically < 3kcal mol^{-1} [36]. This explains the coexistence of different polymorphs under the same thermodynamic conditions. However, even a hypothetically 'perfect' method would yield a single crystal structure as the most stable one under given thermodynamic conditions, this does not necessarily mean that this structure can be formed in a real crystallisation experiment [37]. Due to the large probability of a crystal to form polymorph, computational studies of polymorphism has taken quite an interest in solving the different polymorph structure.

Polymorphism is especially important in pharmaceutical industry. This is because it may cause different solubility quality or performance of a drug bioavailability or stability. Polymorphism can also be sorted by its mechanical properties. This has been discussed by Reddy *et al* [38] on polymorphism within layered structure by doing mechanical shearing and separating from harder crystals of a visual indistinguishable polymorph that resist such shearing.

One of many challenges of polymorphism in the future is the polymorph selection where the probability of searching crystal polymorphs of a compound either by experimental or crystal prediction techniques.

1.7 References

(1) Ochsenfeld, C. *Computational Methods in Science and Engineering, Vol 1* **2009**, *1108*, 151-157.

(2) Brittain, H. G. *Journal of Pharmaceutical Sciences* **2008**, *97*, 3611-3636.

(3) Winkler, B. *Zeitschrift Fur Kristallographie* **1999**, *214*, 506-527.

(4) Catlow, C. R. A.; Gale, J. D.; Grimes, R. W. *Journal of Solid State Chemistry* **1993**, *106*, 13-26.

(5) Sher, A.; Vanschilfgaarde, M.; Berding, M. A. *Journal of Vacuum Science & Technology B* **1991**, *9*, 1738-1739.

(6) Hopfinger, A. J.; Tripathy, S. K. *Crc Critical Reviews in Solid State and Materials Sciences* **1980**, *9*, 285-333.

(7) Day, G. M.; Price, S. L.; Leslie, M. *J. Phys. Chem. B* **2003**, *107*, 10919-10933.

(8) Jack , D. D.; Angelo, G., 2002; Vol. 85; pp 3949-3964.

(9) Eckhardt, C. J.; Gavezzotti, A. *The Journal of Physical Chemistry B* **2007**, *111*, 3430-3437.

(10) H. Sun, P. W. C. K., 2005; Vol. 26; pp 169-174.

(11) Hulme, A. T.; Price, S. L. *Journal of Chemical Theory and Computation* **2007**, *3*, 1597-1608.

(12) Nowella, H.; Price, S. L. *Acta Crystallogr.,Sect.B: Structural Science* **2005**, *61*, 558-568.

(13) Karamertzanis, P. G.; Pantelides, C. C. *Journal of computational chemistry* **2005**, *26*, 304-324.

(14) Karamertzanis, P. G.; Pantelides, C. C. *Molecular Physics* **2007**, *105*, 273-291.

(15) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press, 1997.

(16) Steiner, T. *Angewandte Chem. Int. Ed. Engl.* **2002**, *41*, 48-76.

(17) Novoa, J. J.; Sosa, C., 1995; Vol. 99; pp 15837-15845.

(18) Maerker, C.; Schleyer, P. v. R.; Liedl, K. R.; Ha, T.-K.; Quack, M.; Suhm, M. A. *Journal of Computational Chemistry* **1997**, *18*, 1695.

(19) Lozynski, M.; Rusinska-Roszak, D.; Mack, H.-G., 1998; Vol. 102; pp 2899-2903.

(20) Georgieva, I.; Binev, D.; Trendafilova, N.; Bauer, G. *Chem. Phys.* **2003**, *286*, 205-217

(21) Etter, M. C.; MacDonald, J. C. *Acta Cryst.* **1990**, *B46*, 256-262.

(22) Etter, M. C. *Acc. Chem. Research* **1990**, *23*, 120-126.

(23) Bernstein, J.; Davis, R. E.; Shimoni, L.; Chang, N.-L. *Angewandte Chem. Int. Ed. Engl.* **1995**, *34*, 1555-1573.

(24) Burrows, A. D. In *Supramolecular Assembly Via Hydrogen Bonds I*; Springer-Verlag Berlin: Berlin, 2004; Vol. 108; pp 55-95.

(25) Ahn, S.; PrakashaReddy, J.; Kariuki, B. M.; Chatterjee, S.; Ranganathan, A.; Pedireddi, V. R.; Rao, C. N. R.; Harris, K. D. M. *Chemistry - A European Journal* **2005**, *11*, 2433-2439.

(26) Marivel, S.; Suresh, E.; Pedireddi, V. R. *Tetrahedron Lett.* **2008**, *49*, 3666-3671.

(27) Desiraju, G. R. *Angewandte Chem. Int. Ed. Engl.* **2007**, *46*, 8342-8356.

(28) Braga, D.; Brammer, L.; Champness, N. R. *Crystengcomm* **2005**, *7*, 1 - 19.

(29) Braga, D.; Desiraju, G. R.; Miller, J. S.; Orpen, A. G.; Price, S. S. L. *Crystengcomm* **2002**, *83*, 500-509.

(30) Stahly, G. P. *Crystal Growth & Design* **2007**, *7*, 1007-1026.

(31) Dunitz, J. D. *Crystengcomm* **2003**, *5*, 506.

(32) J.E. Worsham, J. *Acta Cryst.* **1957**, *10*, 319-323.

(33) Docherty, R.; Roberts, K. J.; Saunders, V.; Black, S.; Davey, R. J. *Faraday Discussion* **1993**, *95*, 11-25.

(34) Custelcean, R. *Chem. Commun.* **2008**, 295-307.

(35) Lee, S.-O.; Kariuki, B. M.; Harris, K. D. M. *New Journal of Chemistry* **2005**, *29*, 1266 - 1271.

(36) Nangia, A. In *Models, Mysteries and Magic of Molecules*; Boeyens, J. C. A., Ogilvie, J. F., Eds.; Springer Netherlands, 2008; pp 63-86.

(37) Gdanitz, R. J. In *Theoretical Aspects and Computer Modelling* pp 185-199.

(38) Chickos, J. S. In *NIST Chemistry Web Book*; Linstrom, W. G. M. a. P. J., Ed.; National Institute of Std and Tech, Gaithersburg, MD, http://webbook.nist.gov, 1998; Vol. 69; pp 20899.

Chapter 2

Theoretical Background

2.1 Introduction

The purpose of this chapter is to introduce the fundamental concepts and software applications used in most of the research discussed in this thesis. Here, discussions about quantum mechanics, density functional theory (DFT), theory of intermolecular forces, lattice energy and distributed multipoles is emphasized. A basic introduction to the software applications (DMAREL and GDMA) and parameters used are also explained briefly.

2.2 Quantum Mechanics

Quantum mechanics (QM) as usually applied to molecular problems is the mathematical description of the behaviour of electrons, and thus of molecular chemistry characteristics. In this section, we will discuss how we determine energy of a molecular system.

2.2.1 Schrödinger Time-Independent Equation

Electrons are known as quantized particles. They can be described by wave function. As wave, it can be interpret as:

$$H\Psi = E\Psi \qquad (2.1)$$

Where H is the Hamiltonian operator, a mathematical operation which consist of kinetic and potential properties. E is a scalar function which represents the total energy. In the Schrödinger equation, E is a constant and also referred to as an eigenvalue. Ψ is a function, referred to as an eigenfunction, wave function or state function. Equation 2.1 is also called as the Schrödinger time-independent wave equation which describes stationary states of electron wave function.

$$\left[-\left(\frac{h^2}{8\pi^2 m}\right)\nabla^2 + V(x,y,z)\right]\Psi(x,y,z) = E\Psi(x,y,z) \qquad (2.2)$$

Equation (2.1) can be expand as above (2.2) describing a single particle of mass m of nuclei and electron moving in three dimensional (x, y, z) potential field, V. The terms in square bracket are the Hamiltonian operator where h is the Plank's constant and ∇^2 is the Laplacian operator in three dimensional Cartesian coordinates as in equation (2.3)

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial y^2}\right) \qquad (2.3)$$

In solving equation 2.2, there are a few conditions on Ψ that has to be fulfilled. First, Ψ is required to be a single-valued function because we want $|\Psi|^2$ to give an unambiguous probability for finding a particle in a given region. Also, first derivatives of Ψ should be piecewise continuous and that Ψ itself is continuous. A general restriction on Ψ is that it is a normalisable function, where integral of $|\Psi|^2$ over all space must not be equal zero or infinity and is equal to the total number of electrons in the system. For single electron states, normalisation means

$$\int_{allspace} |\Psi|^2 \, d\tau = 1 \qquad (2.4)$$

where the probability in finding it is one.

The Schrödinger equation is complicated, yet a very powerful equation where it can be a tool in determining the molecular energy of a system. It can be simplified by using Born-Oppenheimer approximation where, motion of nuclei is separated from motion of the electrons. This is because electrons move faster than the nuclei, which makes the nuclei looks stationary. Thus, the electronic wave function can be solved by ignoring the nuclear motion for the system.

2.2.2 Density Functional Theory

Density functional theory (DFT) has become very popular in recent years due to it being a method which is less computationally intensive than other methods with similar accuracy. It is used to calculate quantum energy based on electron density instead of a wave function as in the *ab initio* method, where the energy functional is written as a sum of two terms:

$$E[\rho(\mathbf{r})] = \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + F[\rho(\mathbf{r})] \qquad (2.5)$$

The first term arises from the interaction of the electrons with an external potential $V_{ext}(\mathbf{r})$ typically due to the Coulomb interaction with the nuclei. $F[\rho(\mathbf{r})]$ is the sum of the kinetic energy of the electrons and the contribution from the electronic interactions. The minimum value in the energy corresponds to the exact ground state electron density which enabled us to use variational approach [1].

Originated from Hohenberg and Kohn [2], it is originally applied to finding the ground state electronic energy of a molecule. It was then developed by Kohn and Sham [3] who formulated a method similar in structure to the Hartree-Fock (HF) method.

Electron density is expressed as a linear combination of basis functions similar in mathematical form to HF orbitals. The linear combinations give one electron wave function which then gives density. A determinant is then formed from these functions called Kohn-Sham orbitals. It is the electron density from this determinant of orbitals that is used to compute the energy. The orbitals describe the behaviour of electrons in a molecule, just as the other quantum methods do.

A density functional is then used to obtain the energy for the electron density. A functional is a function of a function, in this case, the electron density. The exact density functional is not known. Therefore, there are a whole list of different functionals that may have advantages or disadvantages. Some of these functionals were formed from fundamental quantum mechanics and some by parameterizing

functions to best reproduce experimental results. Therefore, there are essences of *ab initio* and semiempirical methods in DFT.

By using electron density, the integrals for Coulomb repulsion need to be done only over the electron density, which is a three dimensional function, thus scaling N^3, where N is the number of electrons involve in the calculation. Additionally, at least some electron correlation can be included in the calculation. This produced faster calculation than HF calculation and the computation are more accurate as well. Better DFT functionals results gives accuracy similar to the MP2 calculation.

Complex set of DFT functionals utilizes the electron density and its gradient. These are called gradient-corrected methods. There are also hybrid methods that combine functionals from other methods with the exchange energy of a Hartree-Fock calculation, usually the exchange integrals. In general, gradient-corrected or hybrid calculations gives the most accurate results that other earlier DFT methods.

One of the most popular DFT methods is the B3LYP [4] hybrid functional method, due to its accuracy for a large range of compounds, particularly organic molecules. Besides methods used, the accuracy of DFT methods also depends on basis sets chose during the calculation.

2.2.3 Basis Sets

Equation 2.2 can only be solved directly if it is a one electron, one nuclei system. For more than one nuclei system, this can be solve by adding a guess wave function ϕ as a linear combination of atomic wave function φ;

$$\phi = \sum_{i=1}^{N} a_i \varphi_i \qquad (2.6)$$

where the set of N functions φ_i is called the 'basis set' and each associated with coefficient a_i. This combination is known as the linear combination of atomic orbital (LCAO) approach.

Basis sets describe the atomic orbital used to calculate the required molecule orbitals. It is a mathematical function to construct Hartree Fock (HF) wave function. So, choosing the right basis sets is important to simulate the best molecule energy or interaction involved. A few things have to be put into consideration in choosing the best basis sets; which is the choice of core orbitals, valence, polarisation and diffuse function.

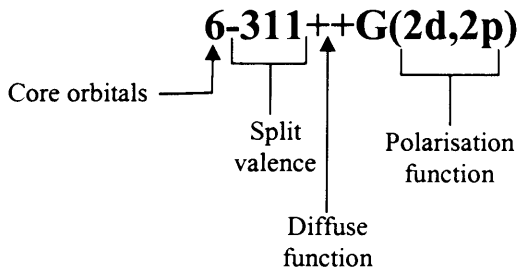


Figure 2.1: Definition of nomenclature in basis set

In primitive (single-ζ (zeta)) basis sets, STO-3G (Slater-Type Orbital approximated by 3 Gaussians) is also known as the 'minimal' basis set where one basis function defined each type of orbital core through valence. For example, for Li to Ne, there are five core functions, 1s, 2s, $2p_x$, $2p_y$, and $2p_z$. This number is the absolute minimum required, though it is nowhere near the infinite basis set limit. To increase the flexibility of the molecule, two or more basis function can be used for each AO, called double-ζ, triple-ζ basis or higher (multiple-ζ) basis. The advantage of using larger basis sets is that the calculation will produce closer result to HF limit.

Split-valence describes the density distribution of valence orbital. This is important especially for understanding chemical bonding of atoms. It improves the description in core orbitals by describing it into arbitrary many functions. The number of the hyphen (Figure 2.1) indicate the number of primitives used in the valence function.

Core and valence functions do not provide mathematical flexibility to adequately describe wave function for pyramidal geometry because s and p function are centered on atoms. For first row atom, flexibility is described in the form of basis function. However for second row and above, higher function (d) are important to predict geometry of a molecule. It usually uses * (star) as the nomenclature where it indicate polarisation of d function on heavy atoms and two star for p function polarisation on hydrogen atom. However, in modern set of calculation with double-ζ and triple-ζ, more than one set of polarisation function need to be added. For example, in Figure 2.1, 2 set of d function is added to heavy atom polarisation and 2 set of p functions for hydrogen atom. However, one should take notice in keeping a balance basis function in applying basis sets.

Diffuse functions (+, plus) are needed to keep the possibility of weakly bound electron to localize far from the remaining density. This function improves on the energies and other molecular properties. Again, the first plus indicate presence of diffuse function in heavy atoms, while the second indicate diffuse function on hydrogen atom.

2.2.4 Basis Set Superposition Error

In a supermolecular approach, the interaction energy ΔE of a complex AB is obtained by performing a calculation of the total energy for the complex and subtracting from it the sum of the total energies of the monomer A and B:

$$\Delta E = E^{AB} - E^{A} - E^{B} \qquad (2.7)$$

The applied method was assumed to be size consistent, which means that calculation on the super system at infinite distance between the monomers should yield the sum of the free monomer energies.

Basis sets superposition error (BSSE) effect arises in calculating interaction energy in a supermolecular system where the monomer energies are calculated using its own monomer basis set. Each monomer automatically takes advantage of the basis set

from the whole dimer or supermolecular energy calculated. Based on the gradient techniques usually used in optimising geometry of a complex, it ignores the improvements of the monomer energies which may be taking place and the outcome of the optimisation will be affected by BSSE. This will result for the intersystem distances to be too short and the corresponding ΔE to be overestimated.

Size of BSSE depends on the details of calculation. Typical values of BSSE are within the range of 0.5-3.0 kcal mol^{-1} for each hydrogen bond in a system.

BSSE is usually largest when using small basis set. This can be overcome by using large basis sets in the calculation, or counterpoise correction method. To calculate using counterpoise correction, one need to recalculates the fragment energies at the same geometry in the same basis, i.e. using the same two-electron repulsion integral, but using a set of one electron integrals, in which the nuclear charges of 'absent' (or 'ghost') fragments has been put to zero. The counterpoise-corrected interaction energy for a given geometry, ΔE^{CP} then becomes

$$\Delta E^{CP}(R) = E^{AB}(R) - E^{A}(\text{full basis}) - E^{B}(\text{full basis}) \qquad (2.8)$$

Where E^A and E^B are the fragment energies.

2.2.5 Frequency calculation

Energy calculation and geometry optimisation provides an idealized picture of the system, assuming that there is no nuclear motion. In equilibrium states, molecular vibrations lead to characteristic of its infra red and Raman spectra.

The frequency calculations can indicate the nature of the stationary point. These frequencies show the minimum point in the potential energy surface (PES) of a structure. Imaginary frequencies, on the other hand, indicate the saddle point which is the transition state (TS) on the potential energy surface (PES).

Minimum PES (no imaginary frequency):

$$\frac{\partial E}{\partial x_i} = 0 \qquad \frac{\partial^2 E}{\partial x_i^2} > 0 \tag{2.9}$$

Transition state on PES (imaginary frequency):

$$\frac{\partial E}{\partial x_i} = 0 \tag{2.10}$$

$$\frac{\partial^2 E}{\partial x_i^2} > 0 \quad x_i \neq \text{reaction coordinate} \tag{2.11}$$

$$\frac{\partial^2 E}{\partial x_i^2} < 0 \quad x_i = \text{reaction coordinate} \tag{2.12}$$

Frequency calculation also yields thermochemical information about the molecule which can relate to reaction kinetics, heats of reaction, and heats of formation.

2.2.6 Zero-Point Energies and Thermodynamics Corrections

Thermodynamic properties are use to observe single molecule, from microscopic regime to macroscopic environment. Behaviour of molecules is governed by empirically determined laws of thermodynamics. Most chemical reactions and many chemical properties are defined in terms of some of the fundamental variables of thermodynamics, such as enthalpy, entropy, free energy and others.

Internal energy at 0K for a molecule can be defined as

$$U_0 = E_{elec} + \sum_i^{modes} \frac{1}{2} h\omega_i \tag{2.13}$$

Where E_{elec} is the energy for the stationary point on Born-Oppenheimer PES, h is the Planck's constant and ω is the vibrational frequency. It is also known as the energy of the ground state of a molecule in reference to the quantum harmonic oscillator and its zero point oscillation. By calculating the vibration modes along the vibration frequency we will obtain its ZPE energy. This value will then be added to the total energy of the optimised geometry, providing a more accurate energy calculation.

Energy can be described as ensemble of properties known as partition function. The contributions to the partition function, entropy, internal energy and constant volume heat capacity from vibrational motions are composed of a sum of the contributions from each vibrational mode. Vibrational mode is alternately symmetrical and antisymmetrical and only real modes are considered. Vibrational modes which have imaginary frequency indicate transition structure of a molecule. The harmonic oscillator is not allowed to have zero energy. The smallest allowed value of vibrational energy is called the zero point energy. If the oscillating particle have zero energy it would also have zero momentum, and would therefore be located exactly at the position of minimum potential energy.

2.2.7 Mulliken charge

Mulliken charge is one of the partial atomic charge population analysis models which involve a direct partitioning of the molecular wave function into atomic contributions from arbitrary orbital-based scheme. In this model, the electrons from the orbitals are devided among the atoms of the molecule with different in which different atomic orbitals (AO) basis functions contribute to the overall wave function.

In a wave function, total number of electron, N can be written as

$$N = \sum_{j}^{electrons} \int \psi_j(\mathbf{r}_j)\psi_j(\mathbf{r}_j)d\mathbf{r}_j \qquad (2.14)$$

where **r** is the integration variable which contains x, y, and z of all the electrons j.

This wave function can be expanded in its AO basis sets

$$N = \sum_{j}^{electrons} \left(\sum_{r} c_{jr}^{\ 2} + \sum_{r \neq s} c_{jr} c_{js} S_{rs} \right) \tag{2.15}$$

where r and s index AO basis function, r in molecular orbital (MO) j, and S is the overlap matrix element.

The first term in equation 2.15 represents electron at nuclei (AO) basis function. The second term is the product of two different AO basis functions which represents the electron 'shared' between basis function. This will then be divided evenly between the two atoms.

Mulliken partial charge over atom k can be divided as

$$q_k = Z_k - N_k \tag{2.16}$$

where Z is the nuclear charge and N_k is computed according to equation 2.15.

2.3 Theory of Intermolecular Forces

2.3.1 Definition of pair potentials in solids

Pair potentials can be use to describe the energetics of solids. Molecular mechanics is used to correlate energetics within structure which to attempt to understand the reason for a particular crystal based on intermolecular forces. Within the molecule, the potential is written as if the atom/molecule were held together by balls and springs. Atom/molecule behaves differently at different length of atom separation (R), (Figure 2.2). In a pair potential, U(R), changes with the change of distances, R. At

U(R) = 0, refers to infinite separation of the atoms. As the atom-atom distance reduces the potential will appoach minimum energy (ε_{min}) at R_{min}. As the distance becomes even shorter, potential will increase infinitely. The pair potential comprises two main contributions. A repulsive part, which cater for small R distance, and attractive part, which cater for large R distance.

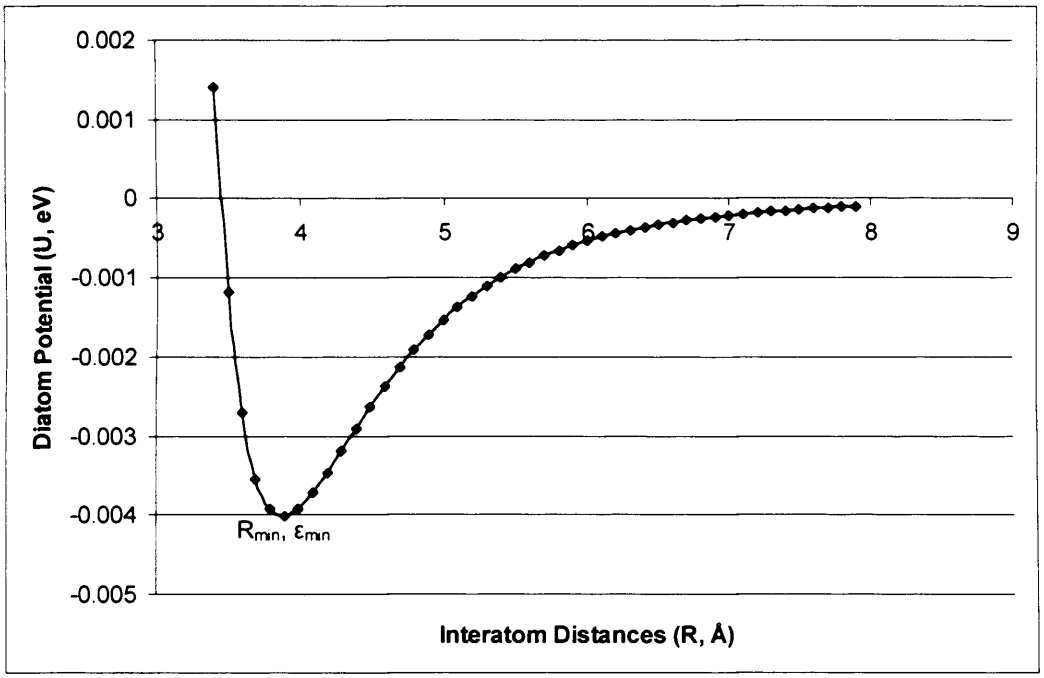


Figure 2.2: C···C atom potential interaction

In the solid state, this approach has been used to calculate *a priori* mechanical and structural properties of solids for silica for use in the Car-Parrinello molecular dynamics simulation (CPMD) [5]. Other than that, *ab intio* calculation was used in evaluating the potentials of *n*-pentane and *n*-pentane/silicate-1 [6] for describing the competition between potential forces exerted by the zeolite and molecular movement. This is an alternative approach for developing the function that represents interaction between pairs of molecules or the interaction of a molecule with a lattice by obtaining the potential function from the *ab initio* energy points. They show that the attractive potential fields dominate the molecular movement at low temperature and the dominance decreases as the temperature increases.

The assumption made, is that the dominant electronic interactions are of sufficiently short range and that a few local parameters will well describe the structure. Based on

the exact nature of chemical bonding in a molecule/solid, the actual structure are adopted from the balance between attractive and repulsive forces. The most basic potential is the Lennard-Jones potential, which is often used to describe interactomic forces between atoms within organic molecules [7].

$$U(R_{ij}) = \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \tag{2.17}$$

where A and B > 0 depend on the identity of the atoms (*i,j*). The potential consists of a van der Waals attractive part of longer range than the harder repulsive potential that sets in at shorter distances.

Another alternative expression of Lennard-Jones potential can be written as:

$$U(R_{ij}) = 4\varepsilon\left(\left(\frac{\sigma}{R_{ij}}\right)^{12} - \left(\frac{\sigma}{R_{ij}}\right)^{6}\right) \tag{2.18}$$

where A from equation 2.12 is $4\varepsilon^{12}$, B = $4\varepsilon\sigma^{6}$, $\sigma = \left(\frac{A}{B}\right)^{\frac{1}{6}}$ and $\varepsilon = \frac{B^2}{4A}$

The $\left(\frac{1}{R}\right)^{12}$ term describes repulsion and the $\left(\frac{1}{R}\right)^{6}$ term describes the attraction forces or dispersion. The repulsion term depend exponentially on the distance, and its physical origin is related to the Pauli Principle, where when the electronic clouds surrounding the atoms start to overlap, the energy of the system increases abruptly [8].

Another sets of model potential that will be used mainly in this thesis is the Buckingham potential.

Buckingham potential can be written as

$$U_{ij} = A\exp\left(-\frac{R_{ij}}{\rho}\right) - \frac{C}{R_{ij}^{6}} \qquad (2.19)$$

where R is the separation distance of *i* and *j* atoms.

Buckingham potential is an improvement of Lennard Jones equation with more realistic exponential expression. Atom pair potential [9] that will be use in this thesis is listed in Appendix 1

2.3.2 Distributed Multipole Analysis (DMA)

Molecules can respond to electrical fields and electric charges. This is described as molecular electrical properties, where it explains the longest range interaction between molecules, electrostatic attraction or repulsion of the molecular charge distributions. This long-range attraction is a fundamental part of what makes van der Waals or a weak hydrogen bond. In classical electrostatics, the starting point for discussing molecular electrical properties is an understanding of point charges. In DMA, distributed multipole expansion of a molecule were described around a number of centres where it provides a starting point for the construction of such model [10]. This gave detailed information about the charge distribution and an accurate way of the spatial distribution of electric charge within a molecule.

A DMA multipoles series consist of a charge, dipole, quadrupole and higher terms and is located at each atom centre. The series can be derived from *ab inito* or density functional theory (DFT) calculation using Gaussian basis sets. Detailed description includes at each sites are:

a) Charges, where it describes the electronegativity effects in a chemically intuitive way;
b) Dipoles, arising from overlap of s and p orbitals and describing lone pairs and other atomic distortions;

c) Quadrupoles, arising from the overlap of p orbitals and associated with pi bonds;
d) Octapoles and hexadecapoles can be included if very high accuracy is required

The DMA method provides a powerful [11] method for representing molecular wave function for long range electrostatic calculation. It has also widely used in molecular simulation such as crystal structure prediction for small organic molecules, where significant progress can often be made while using rigid molecular structures [12-15].

2.3.3 Lattice Energy Calculation

Lattice energy is the energy lowering when one mole of a substance is formed from its constituent gas-phase molecule. This can be estimated from electrostatic consideration of its crystal structure. Lattice energy is related to crystal structure from its covalent character and electron-electron interactions in atomic solids.

Within this study, lattice energy calculation was mainly done using DMAREL; A program which minimizes the energy of a lattice cell mainly using electrostatic model and van der Waals represented by Buckingham potential. DMAREL calculates and minimizes lattice energy of a crystal structure to determine the adequacy of intermolecular potential. It performs a symmetry-constrained (rigid molecule) relaxation of crystal structure which then searches for the minimum relaxation of lattice energy. This includes position and rotation of the molecules.

The Ewald sum is usually used for calculating electrostatics interaction while doing the minimization. This method was well known for calculating structure in periodic boundary condition. The model used for neutral periodic system where charged point atoms mutually interacting using Coulomb potential. Two types of calculation are made in this method. First, each atom is neutralised by superposition of a spherical gaussian cloud of opposite charge centred on the atom. This is also known as real space. The second part is superimposing a second set of Gaussian charges using the same charges as the original point atoms and again centred it on the point atoms. This nullified the effect of the first set of gaussian. The potential due to these Gaussians is

obtained from Poisson's equation and is solved as a Fourier series in reciprocal space sum. The complete Ewald sum requires an additional correction, known as the self energy correction, which arises from a gaussian acting on its own site, and is constant. Ewald's method therefore replaces a potentially infinite sum in real space by two finite sums: one in real space and one in reciprocal space; and the self energy corrections (Figure 2.3).

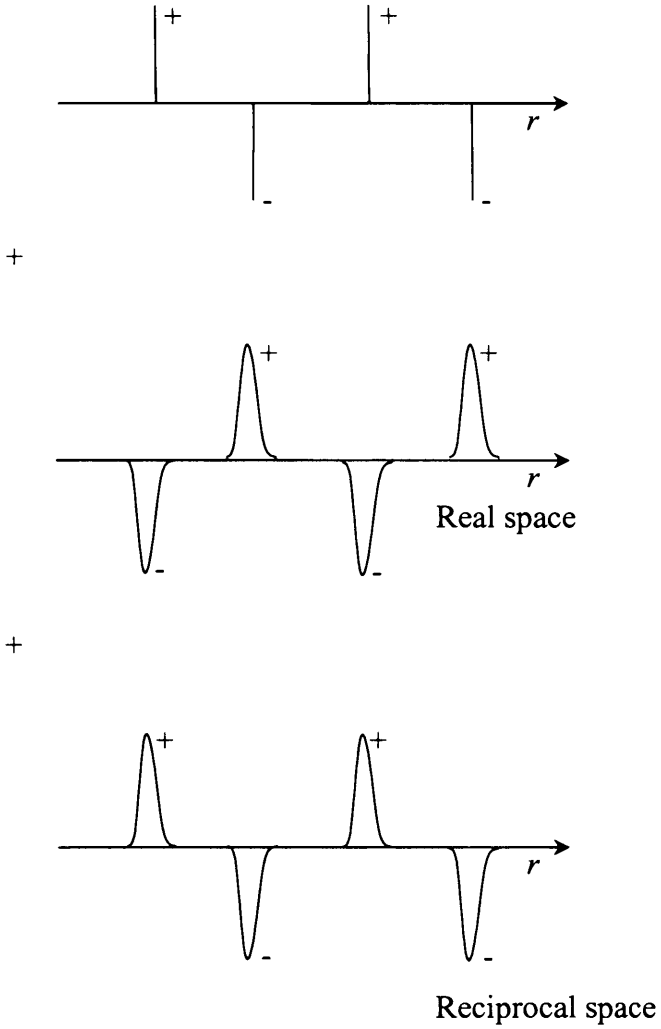


Figure 2.3: In the Ewald summation method, the initial set of charges are surrounded by a Gaussian distribution (calculated in real space) and thereafter, added a cancelling change distribution which calculated in reciprocal spaces. Reproduce from reference 16.

However, the lattice energy minimisation calculation neglects the effects of zero-point motion on lattice parameters. This affects and causes the change of lattice parameters

when minimising. It is unlikely that any potential will give perfect agreement with experiment on a range of structures. However, for hydrogen bonded structures, errors of less than ~3% in lattice parameters should be achievable [9]. Neumann *et al* described of having a discrepancy with experiment at 1% cell difference change is possible when DFT is combined with empirical van der Waals correction method [17].

In situations where the sublimation energy could not be accessible, discrepancy index was used to test the accuracy of lattice structure calculation. The discrepancy index calculated the deviation of structure from the experimental. It is defined [18] as:

$$F = \left(\frac{\Delta\theta}{2}\right)^2 + (10\Delta x)^2 + \left(100\frac{\Delta a}{a}\right)^2 + \left(100\frac{\Delta b}{b}\right)^2 + \left(100\frac{\Delta c}{c}\right)^2 + \Delta\alpha^2 + \Delta\beta^2 + \Delta\gamma^2 \quad (2.20)$$

Where the first term is the value for the net rotation of the molecule (deg, $\Delta\theta$), the second term is the net translation of the centre of mass of the molecule (Å, Δx) in the system, the next three term are the root mean square (rms) change in lattice cell and last three terms are the changes in cell angle. During calculation, rms error values were used if there is more than one independent molecule in the asymmetric unit cell.

The agreement between optimised and experimental structure can also be used in determining the accuracy of the calculation. It was considered that there are possibilities to have lattice energy differences between 2 to 3 kcal mol^{-1} (8.4 to 17 kJ mol^{-1}) due to the thermodynamic effect of the crystal [19]. This type of measurement is used by DMAREL application where it involved anisotropic potential model for electrostatic interactions, which in reality contributes 90% of the lattice energy of polar molecules.

2.4 Computational Methods and Parameters

2.4.1 DMAREL

As mentioned in the previous topic, the DMAREL [14] application was used to simulate static crystals of molecules whose electrostatic interactions are represented by a distributed multipole model. The computer program uses a modified Newton Rapshon procedure, which make occasional use of some analytical second derivatives, to find a stable crystal structure by reducing to zero the forces and torques of the molecules, and the lattice strain.

The starting point in simulating the molecule is usually taken from Cambridge Structural Database (CSD), and then converted in the form of P1 Cartesian representation. Hydrogen atoms of the molecular structures are positioned using a standard bond length of C-H (1.08 Å) [20] and N-H (1.01 Å) along the experimental bond direction. Lattice vectors, position and orientation of each molecule in the unit cell are independently optimised, not constraining its crystal symmetry during the simulation in finding the local minima. Symmetry of the force field in the starting structure however, could prevent symmetry breaking, but there is also a possibility that global minima could occur in a less symmetrical structure.

The DMA electrostatic model was obtained from the DFT wave function of isolated molecule calculated using 6-31G (d,p) basis set, using either MOLPRO or Gaussian03-GDMA program suite. One important thing that should be noted while generating the electrostatics for isolated structures is to make sure that the same local axis definition is use in both the setup for the preparation of DMA and for DMAREL input file. Failing to do so, might produced high discrepancy index even though the structure is successfully relaxed.

Ewald sum was used to cancel out the accumulation of charge-charge interaction from the lower order moments of long range electrostatic calculation and for higher poles electrostatics, a user-defined cutoff based on molecular centre-of mass separation is used. Total lattice energy was then obtained from the direct summation of inter-molecular charge-charge, charge-dipole, dipole-dipole and higher multipole

interaction energy. The main error in evaluating of electrostatic contribution to lattice energy, excluding the effect of penetration into the molecular charge distribution is the accuracy of the wave function. This known deficiency of QM method tends to overestimate multipoles. Therefore, a scaling factor of 0.9 was included during the calculation [12].

The short range intermolecular interactions were calculated using gradient based fitting of empirical atom-atom pair potential. The intermolecular potential was assumed to have isotropic atom-atom form represented by

$$U_{ik} = \frac{\varepsilon_{\iota\kappa}}{\lambda_{\iota\kappa} - 6}\left[6\exp(\lambda_{\iota\kappa})\exp\left(-\lambda_{\iota\kappa}\frac{R_{ik}}{R_{0,\iota\kappa}}\right) - \lambda_{\iota\kappa}\left(\frac{R_{0,\iota\kappa}}{R_{ik}}\right)^6\right] \qquad (2.21)$$

where atoms i and k are of types ι and κ respectively and resembles the well depth. R_0 is the minimum energy separation and λ is a 'steepness' parameter (assumed to have value of 13.5 for all interactions [9]). This potential depends only on the separation R_{ik} of the atoms in different molecules and the types of atoms concerned. The polar heavy-atom...H_p potential fitting was used where it previously showed the best results of urea structure relaxation [9]. Combining rules used for heteroatoms interactions parameters are given by

$$\varepsilon_{\iota\kappa} = (\varepsilon_{\iota\iota}\varepsilon_{\kappa\kappa})^{\frac{1}{2}}, \quad R_{0,\iota\kappa} = \frac{1}{2}(R_{0\iota\iota} + R_{0,\kappa\kappa}) \qquad (2.22)$$

The accuracy of the crystal structure relaxations was describe using the discrepancy index as mention earlier (Equation 2.11). From here, we can compare the quality of DMA and parameters used by comparing the changes occurs before and after the relaxation. The smaller the number, the better the relaxation is. It is then compared to the sublimation energy, obtained from the experimental.

The computational requirement of a DMAREL minimization depends on the number and size of the molecules in the unit cell, the range used for the lattice summations, and also the number of maximum iterations to be used in finding the minimum. Example of DMAREL input file is shown in Appendix 2.

Recently, DMAREL has been use to calculate and analyse atomistic properties of phonons frequencies and thermodynamic quantities for crystals of rigid organic molecules by Day *et al* [13]. Here, they used parameterized simple exp-6 model to describe vibrations in simple van der Waals crystals and some hydrogen bonded rigid crystals. Besides that DMAREL was also frequently used in studies regarding crystal structure prediction and polymorphism since it is a very useful program in calculating lattice energy crystals [15,21-23].

2.4.2 GDMA calculation

As mentioned earlier in section 2.4.1, GDMA is use to generate a DMA, a very important electrostatic component in lattice energy calculation. In order to produce the urea multipole description of a molecule, it uses the formatted checkpoint file produced by Gaussian program. The DMA multipoles used in this study is generated up to hexadecapole.

Normal DMA calculations depends on basis sets diffuse function. The charge distribution may differ minimally at large basis set where then affected the distributed multipole description. As the size of basis sets increase, the probability of this kind of ambiguity increases and this will make the problem worse by diverging the multipole values.

In modern calculation, large basis sets is usually use in simulating molecules. GDMA improves on the multipole allocation algorithm so the partitioning of the density between atoms is carried out in real space, as in atom in molecules (AIM) procedure. It uses numerical quadrature for the diffuse functions while retaining the original

method for the more compact ones [24]. Example of input file is shown in Appendix 3.

2.5 References

(1) Leach, A. R. In *Molecular Modelling: Principles and Applications*; Pearson Education Limited, 2001; pp 127.

(2) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(3) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(4) Lozynski, M.; Rusinska-Roszak, D.; Mack, H.-G., 1998; Vol. 102; pp 2899-2903.

(5) Carre, A.; Horbach, J.; Ispas, S.; Kob, W. *Europhys. Lett.* **2008**, *82*, 17001 (6 pp.)-17001 (6 pp.).

(6) Loisruangsin, A.; Fritzsche, S.; Hannongbua, S. *Chem. Phys.* **2008**, *344*, 1-12.

(7) Burdett, J. K. *Chemical Bonding in Solids*; Oxford University Press, 1995.

(8) In *http://en.wikipedia.org/wiki/Lennard-Jones_potential*, 2009; pp 1-3.

(9) Willock, D. J. *Philosophical Magazine B* **1996**, *73*, 127-138.

(10) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233-239.

(11) Dykstra, C. E. *Chem. Rev.* **1993**, *93*, 2339-2353.

(12) Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *J. Phys. Chem.* **1996**, *100*, 7352-7360.

(13) Day, G. M.; Price, S. L.; Leslie, M. *J. Phys. Chem. B* **2003**, *107*, 10919-10933.

(14) Willock, D. J.; Price, S. L.; Leslie, M. *Journal of Computational Chemistry* **1995**, *16*, 628-647.

(15) Coombes, D. S.; Nagi, G. K.; Price, S. L. *Chem. Phys. Lett.* **1997**, *265*, 532-537.

(16) Leach, A. R. In *Molecular Modelling: Principles and Applications*; Pearson Education Limited, 2001; pp 336.

(17) Neumann, M. A.; Perrin, M.-A. *J. Phys. Chem. B* **2005**, *109*, 15531-15541.

(18) Filippi, G.; Gavezzotti, A. *Acta Crystallographica Section B* **1993**, *49*, 868-880.

(19) Gavezzotti, A.; Filippini, G. In *Theoretical Aspects and Computer Modelling of the Molecular Solid State*; Gavezzotti, A., Ed.; John Wiley & Sons, 1997; Vol. 1; pp 61-97.

(20) Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. *J. Chem. Soc., Perkin Trans. 2* **1987**, S1 - S19.

(21) Ouvrard, C.; Price, S. L. *Crystal Growth & Design* **2004**, *4*, 1119-1127.

(22) Tremayne, M.; Grice, L.; Pyatt, J. C.; Seaton, C. C.; Kariuki, B. M.; Tsui, H. H. Y.; Price, S. L.; Cherryman, J. C. *J. Am. Chem. Soc.* **2004**, *126*, 7071-7081.

(23) Lewis, T. C.; Tocher, D. A.; Price, S. L. *Crystal Growth & Design* **2005**, *5*, 983-993.

(24) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128-1132.

Chapter 3

Urea/α,ω-Dihydroxyalkanes (Diol) Co-Crystal Stucture Studies

3.1 Introduction

In this chapter, we will study the structure of urea/α,ω-dihydroxyalkanes (diol) co-crystals and their lattice energies. Formation of urea/α,ω-diol co-crystal was first observed while studying the urea inclusion crystallization of urea, α,ω-diaminoalkanes (1,7-diaminoheptane and 1,8-diaminooctane) and 1,4-dihydroxybutane [1]. The attempt only crystallized co-crystal of urea/1,4-dihydroxybutane as the binary co-crystal.

Since then, several attempts have been made to crystallize sets of urea/α,ω-diol [2] with longer alkane chain (C5 to C14). As a results, a homologous urea/α,ω-diol co-crystal series was successfully crystallized out of only even carbon number diol chains from urea/1,4-dihydroxybutane (urea/1,4-dhb) to urea/1,14-dihydroxydodecane (urea/1,14-dhdd). However, from these homologs there are two different crystal structures which contain the same ratio of urea : diol (2:1) molecules in the crystal packing. The space groups within this series are the same ($P\bar{1}$) except for urea/1,4-dihydroxybutane (urea/1,4-dhb) and urea/1,8-dihydroxyoctane (urea/1,8-dho) which have $P2_1/c$ space group (Table 3.1).

Table 3.1: Properties of co-crystal structure obtained from experimental powder X-ray data at 150K

Diol Co-crystals	**Lattice Parameter**						**Symmetry**	**Space Group**
	a (Å)	b (Å)	c (Å)	α(°)	β (°)	γ (°)		
1,4[a]	5.205	7.364	14.606	90.000	98.604	90.000	Monoclinic	$P2_1/c$
1,6[b]	5.145	7.399	8.3448	95.851	96.187	101.25	Triclinic	$P\bar{1}$
1,8[c]	10.43	5.170	14.132	90.000	111.660	90.000	Monoclinic	$P2_1/c$
1,10[c]	5.163	7.367	11.101	98.618	96.302	102.007	Triclinic	$P\bar{1}$
1,12[b]	5.165	12.262	14.561	88.518	81.900	79.267	Triclinic	$P\bar{1}$

[a] Structure from reference [1]

[b] From redetermination of structure by Javier Marti-Rujas

[c] From original work on co-crystals [2]

For urea/1,4-dihydroxybutane structure (Figure 3.1), urea molecules are hydrogen bonded to each other to form double-stranded ribbons that run parallel to the *a*-axis. The two strands of ribbon are linked by two networks of N–H···O hydrogen bonds between urea molecules, involving the C=O group and one NH_2 of each urea molecule, designated $C(4)R_4^2(8)$ and $R_2^2(8)$ in graph set system. The third network involves NH_2 group of each urea molecule extends from the edge of the ribbon. For these NH_2 groups, each N-H bond is engaged in a N–H···O hydrogen bond with the OH group of a 1,4-dihydroxybutane molecule, and each OH group is involved in two N–H···O hydrogen bonds with different urea molecules, designated $C(4)R_3^5(10)$ (Figure 3.1). The fourth type of network is between the sheet of urea/1,4-dihihydroxybutane co-crystal, where the OH group of the diol is hydrogen bonding with the C=O of a urea that is perpendicular to the ribbon plane, designated $R_4^4(12)$. The ring of network only involves 5 molecules, 3 being urea and two diols. This type of network creates a zigzag type of urea/diol arrangement.

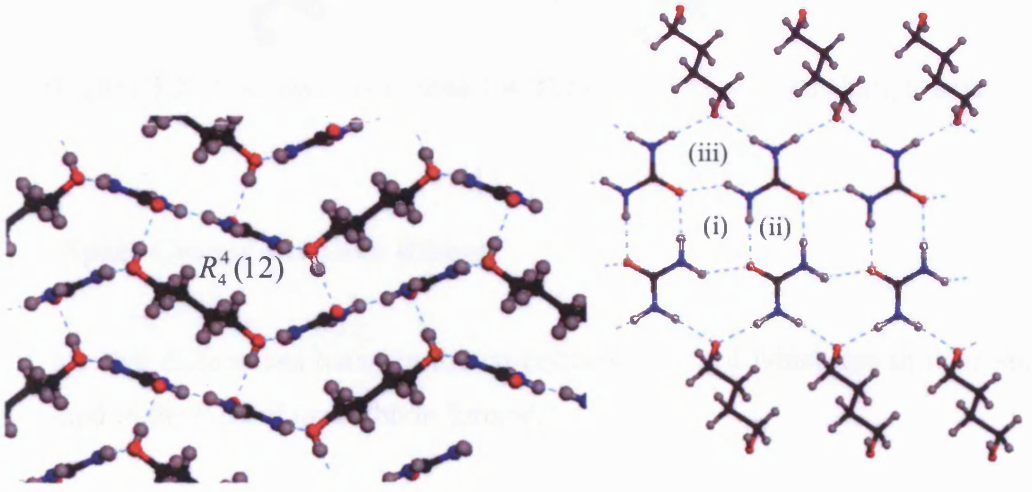


Figure 3.1 : Hydrogen bond network in urea/1,4-dihydroxyalkane co-crystal. (a) viewed along the direction of the propagation of the ribbons (*a*-axis), and (b) perpendicular to the average plane of a corrugated sheet (*ac*-plane) with the ribbons running in the horizontal direction. Graph set networks are (i) $C(4)R_4^2(8)$, (ii) $R_2^2(8)$ and (iii) $C(4)R_3^5(10)$

From the structures reported, only urea/1,4-dhb has the combination of anti-parallel type urea ribbon, but also, has the same space group as the parallel type of urea ribbon, $P2_1/c$. It is the only co-crystal within these homologous co-crystal series to have anti-parallel urea ribbons, with the mentioned space group. This space group indicates that a screw axis exists in the co-crystal unit cell (Figure 3.2). In this case, the screw axis is between the urea ribbon and 1,4-dhb diol sheet.

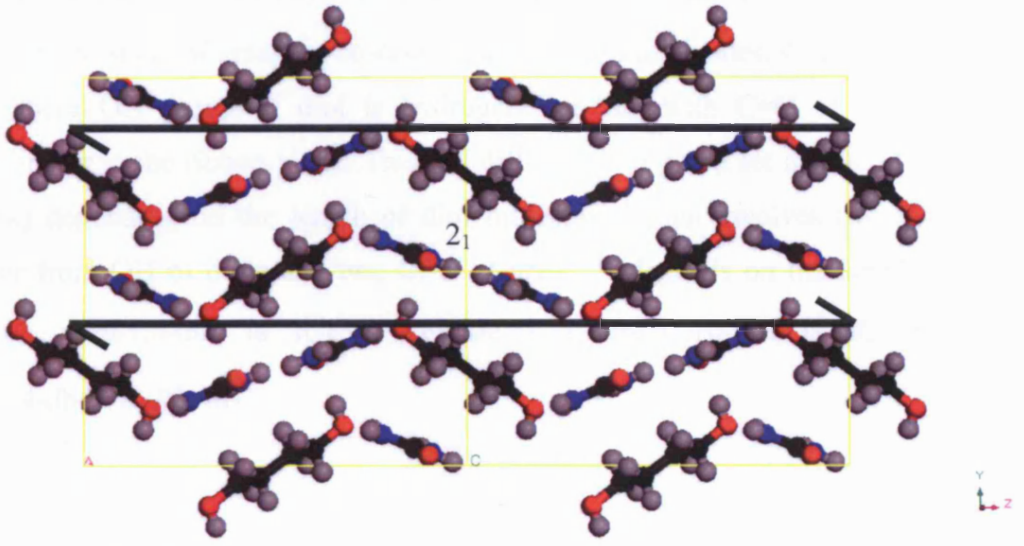


Figure 3.2: 2_1 screw axis in urea/1,4-dhb co-crystal, viewed along *a*-axis

3.1.1 Space Groups and Urea Ribbon

There are two differences between the co-crystals obtained which are in their space groups and in the type of urea ribbon formed.

The co-crystals with anti-parallel ribbon structure other than urea/1,4-dhb have $P\bar{1}$ space group. They have a similar urea motif structure based on hydrogen-bond network that forms double-stranded ribbons running nearly parallel to the *a*-axis, as the urea/1,4-dhb co-crystal mentioned earlier. They feature three types of hydrogen networks for the co-crystal sheet and one network connecting the sheets. In the first network, the two strands of the ribbon are linked together by a network of N-H···O hydrogen bonds between urea molecules involving the C=O group and one NH_2 group of each urea molecule. The ribbons contain two different types of cyclic

hydrogen bonded arrays alternating along the centre of the ribbon. The first is designated $R_2^2(8)$ in the graph set system for the anti-parallel hydrogen bond ring of two urea molecules. The second is $C(4)R_4^2(8)$ which involves hydrogen bonds between anti-parallel pairs to form the ribbon chain. This type of network often occurs for primary amides [3,4]. The third type of network is between sheets of urea ribbons, involving NH_2 groups of two urea molecules, each hydrogen bonding with O of a diol molecule. This is designated by $C(4)R_3^5(10)$ in graph sets. The fourth type of network is between the sheet of urea/diol co-crystal, involving 6 molecules, 4 being urea and 2 diols, where OH group of diol is hydrogen bonding with C=O of urea that is perpendicular to the ribbon plane. This has different *r* in graph set notation (different network) depending on the length of diol molecule, which involves two hydrogen acceptor from OH of diols and two C=O of urea. It depends on the length of diol molecule, urea/1,6-dhh is $R_6^4(30)$ (Figure 3.3), urea/1,10-dhd is $R_6^4(38)$, and urea/1,14-dhdd is $R_6^4(46)$.

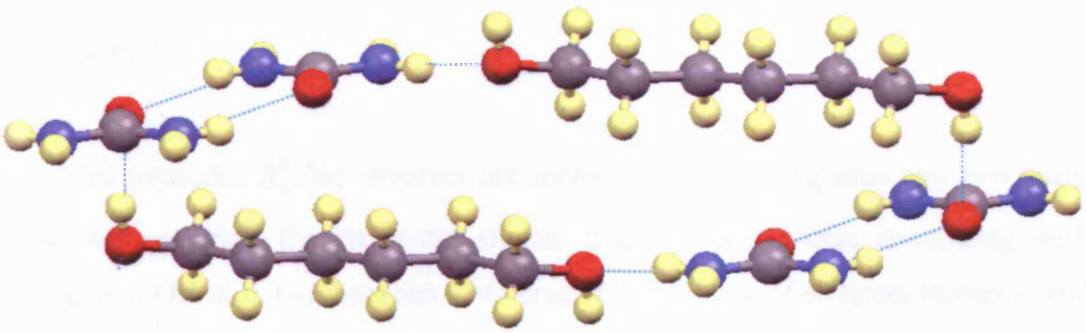


Figure 3.3: Hydrogen bond network connecting the co-crystal sheets. In this figure, urea/1,6-dhh network, $R_6^4(30)$.

On the other hand, co-crystals with $P2_1/c$ space group have parallel type of urea structure that form double-stranded ribbons running nearly parallel to the *b*-axis which can be designated with graph set notation as $C(4)R_3^2(8)$ (Figure 3.4). This is the only type of parallel urea ribbon in urea/diol co-crystal successfully crystallized. The hydrogen bonds that formed the urea/diol co-crystal sheet have different network

from the anti-parallel ribbon type. The second hydrogen network which form the co-crystal sheet, $C_2^1(4)R_3^2(10)$ connects the urea ribbon and diol (Figure 3.4).

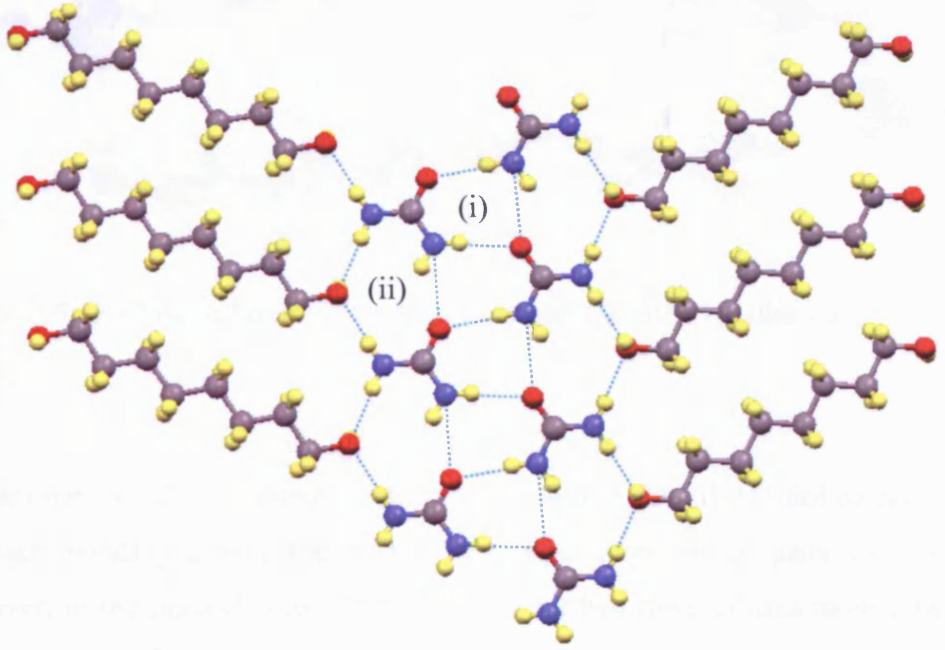


Figure 3.4: Hydrogen network of urea/1,8-dho co-crystal viewed along *a*-axis. Two types of hydrogen networks are shown on the plane, (i) $C(4)R_3^2(8)$ and (ii) $C_2^1(4)R_3^2(10)$

The third network, $R_6^5(36)$ involves six molecules, four being urea and two diols. These are engaging the co-crystal sheets, with two C=O urea connecting with hydrogen in OH diol, two between C=O urea with NH urea of different molecule and another two between O in OH diol with NH of urea molecule (Figure 3.5). Generally, only three types of network exist form this type of co-crystal.

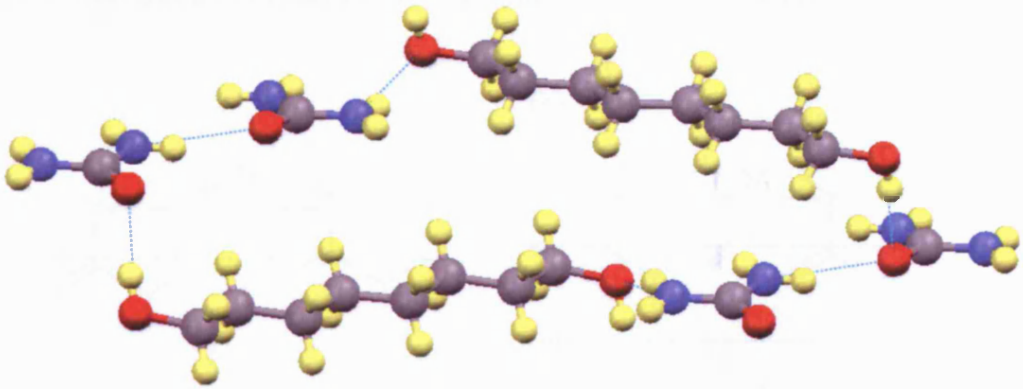


Figure 3.5: Hydrogen bond network in between the urea/1,8-dho co-crystal sheets, $R_6^5(36)$.

The anti-parallel ribbon system found in urea/1,4-dhb and all $P\bar{1}$ diol co-crystals has hydrogen bonding across the ribbon which involves single pairs of molecules. However, in the parallel urea ribbon system, the two rows of urea molecules within the ribbon are offset so that each urea is hydrogen bonded to two molecules on the other side of the ribbon. For both types of ribbon, the NH_2 pattern outside ribbon is very similar, with distances of around 5.2 Å between the NH_2 groups. So, interactions between ribbons and diols could be expected to be similar.

Graphic description of the urea ribbon is clearly described in Figure 3.6.

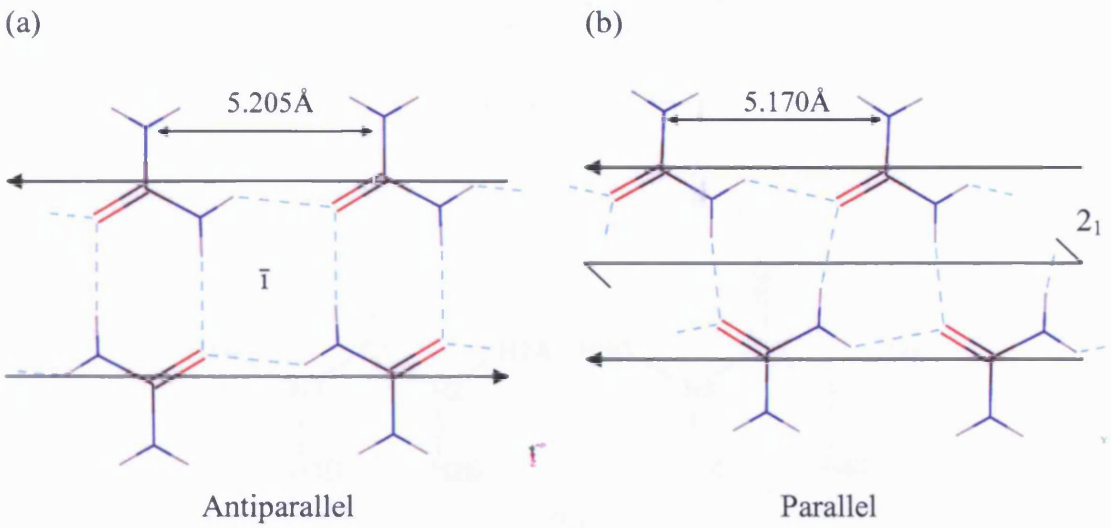
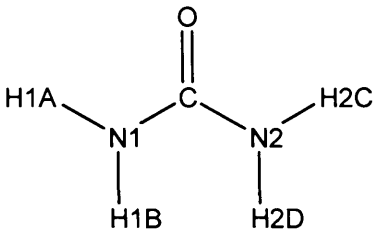
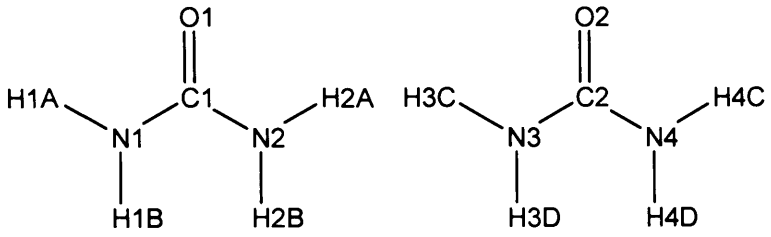


Figure 3.6: Schematic representation of the two urea motifs observed in the crystal structures between urea and α,ω-dihydroxyalkanes: (a) anti-parallel orientation of urea strands with inversion centre, and (b) parallel orientation of urea strands with 2_1 screw axis.

Detailed information about the hydrogen bond length and angle of experimental urea/diol co-crystal structure is given in Table 3.2 below. Urea structure from 1,4-dhb to 1,10-dhd of urea/diol co-crystal are based on urea molecule in Figure 3.7 (a), where N(1) is bonding with H(1A) and H(1B), while N(2) is bonding with H(2C) and H(2D). In urea/1,12-dhdd co-crystal, there are two different types of urea, labeled and shown in Figure 3.7 (b).



(a)



(b)

Figure 3.7 (a) Urea structure for urea/diol co-crystal of 1,4-dhb to 1,10-dhd (b) Urea structure for urea/1,12-dhdd co-crystal.

Table 3.2: Selected D-H...A hydrogen bond distances (*d*) and angles (<) in co-crystal structure

Co-Crystal	**D-H**	**A**	$d_{D...A}$ (Å)	$<_{D-H...A}$ (°)
Urea/ 1,4-dihydroxybutane	N(1)–H(1A)	O_{diol}	3.002	175.04
	N(1)–H(1B)	O_{diol}	3.008	168.22
	N(2)–H(2C)	O_{urea}	2.983	170.87
	N(2)–H(2D)	O_{urea}	2.958	140.96
	O–H_{diol}	O_{urea}	2.791	163.37
Urea/ 1,6-dihydroxyhexane	N(1)-H(1C)	O_{diol}	2.978	168.51
	N(1)-H(1D)	O_{diol}	3.038	158.32
	N(2)-H(2C)	O_{urea}	3.004	165.90
	N(2)-H(2D)	O_{urea}	2.944	145.93
	O-H_{diol}	O_{urea}	2.815	172.14
Urea/ 1,8-dihydroxyoctane	N(1)-H(1A)	O_{diol}	2.916	136.52
	N(1)-H(1B)	O_{diol}	2.952	165.31

	N(2)-H(2C)	O_{urea}	2.952	164.91
	N(2)-H(2D)	O_{urea}	2.986	157.79
	O-H_{diol}	O_{urea}	2.770	173.36
Urea/ 1,10-dihydroxydecane	N(1)-H(1A)	O_{diol}	2.995	165.92
	N(1)-H(1B)	O_{diol}	2.985	159.45
	N(2)-H(2C)	O_{urea}	2.966	166.81
	N(2)-H(2D)	O_{urea}	2.923	146.01
	O-H_{diol}	O_{urea}	2.780	168.91
Urea/ 1,12-dihydrocydodecane	N(1)-H(1A)	$O(1)_{urea}$	2.990	168.12
	N(1)-H(1B)	$O(1)_{urea}$	2.931	147.08
	N(2)-H(2A)	$O(4)_{diol}$	3.022	165.49
	N(2)-H(2B)	$O(4)_{diol}$	2.992	157.06
	N(3)-H(3C)	$O(3)_{diol}$	2.972	163.25
	N(3)-H(3D)	$O(3)_{diol}$	2.990	164.54
	N(4)-H(4C)	$O(2)_{urea}$	2.992	166.88
	N(4)-H(4D)	$O(2)_{urea}$	2.916	140.43
	$O(3)-H(3)_{diol}$	$O(1)_{urea}$	2.751	173.15

All hydrogen bonds exist in the co-crystals satisfy the criteria of moderate hydrogen bonds. This shows that the co-crystals are mostly governed by electrostatic interaction. However, as the type of diol grows longer for each type of co-crystals, the strength of moderate hydrogen bond change towards a weaker type of hydrogen bond. This is shown especially in urea/1,12-dhd type of co-crystal, where the angle becomes smaller, and hydrogen bond become less symmetrical with big differences in hydrogen bond distances for different urea oxygen acceptors.

These two types of co-crystal arrangements motivated us towards using calculations to investigate their energy properties in relation to the different urea ribbons (parallel and antiparallel). This will further lead us to develop methods to predict the crystal structure for other structures that could not be obtained from the previous experiment.

For example, suggesting a 1,4-diol co-crystal structure with parallel urea ribbon rather than the observed anti-parallel motif.

3.1.2 Diol Structures

Table 3.3: Properties of diol crystal structures from CDS database

Diol	**Lattice Parameter**						**Symmetry**	**Space Group**
	A (Å)	b Å)	c Å)	α (°)	β (°)	γ (°)		
1,4[5]	5.01	13.79	7.46	90.00	107.27	90.00	monoclinic	$P2_1/n$
1,6[6]	8.03	5.10	18.30	90.00	111.06	90.00	monoclinic	$P2_1/c$
1,8[5]	4.80	5.09	17.54	90.00	90.32	90.00	monoclinic	$P2_1/n$
1,10[7]	21.15	5.16	4.94	90.00	90.69	90.00	monoclinic	$P2_1/a$
1,12[8]	4.96	5.19	24.50	90.00	90.75	90.00	monoclinic	$P2_1/c$

Table 3.3 shows the crystallographic properties of diols that are used for calculating interaction energy of urea/diol co-crystals. As shown, all of the dense diols are in monoclinic symmetry with $P2_1$ space groups. Each diol are hydrogen bonded with two other neighbouring diols, one acting as the hydrogen bond donor and the other as hydrogen bond acceptor, generating diol sheets. However, there is no hydrogen bond connecting between the sheets. Graph set designation for each dens diol are; 1,4-dhb $R_4^4(18)$, 1,6-dhh $R_4^4(22)$, 1,8-dho $R_4^4(26)$, 1,10-dhd $R_4^4(30)$, 1,12-dhdd $R_4^4(34)$ and 1,14-dhtd $R_4^4(38)$. The hydrogen bond detailed description is shown in Table 3.4 below.

Table 3.4: Selected D-H...A hydrogen bond distances (*d*) and angles (<) in diol dense structure

Diol Crystal	D-H	A	$d_{D...A}$ (Å)	$<_{D-H...A}$ (°)
1,4	O(1)-H(1)	O(2)	2.762	175.05
	O(2)-H(2)	O(1)	2.772	170.55
1,6	O(1)-H(1)	O(2)	2.807	168.20
	O(2)-H(2)	O(1)	2.792	170.66
1,8	O(1)-H(1)	O(1B*)	2.788	173.70
	O(1B*)-H(1B*)	O(1)	2.788	173.70
1,10	O(1)-H(11)	O(1)	2.804	167.82
	O(1B*)-H(11B*)	O(1B*)	2.804	167.82
1,12	O(1)-H(1)	O(1B*)	2.832	171.04
	O(1B*)-H(1B*)	O(1)	2.832	171.04
1,14	O(1)-H(1)	O(1B*)	2.831	171.92
	O(1B*)-H(1B*)	O(1)	2.831	171.92

Based on distance and angle, all display characteristics of moderate type of hydrogen bond.

3.2 Computational Method

All lattice energy calculation is simulated using DMAREL and distributed multipoles (DMA) generated from either MOLPRO or Gaussian03-GDMA.

Lattice energy calculation of diol crystals are done using two types of multipoles. First using MOLPRO. However, for the longer carbon chain diols, MOLPRO failed to

produce multipoles correctly as the whole molecule was found to be charged. Due to this problem, the DMA calculation was changed to GDMA method.

This can be explained because MOLPRO created DMA in the punch files. Its calculation are based on original Stone method [9] which is an extension of Mulliken analysis using only basis function overlap to define atomic density. For DMA calculated using GDMA [10], diffuse basis sets which derive from numerical integration. This improves the stability of DMAs, with respect to basis set size.

3.2.1 Interaction Energy

Interaction energy (IE) is needed to compare co-crystals. The interaction energy is obtain from the difference of total lattice energy of the co-crystal to the total lattice energy of its dense crystals molecule components (Equation 4.1)

$$\text{IE} = E_{co-crystal} - \left(nE_{urea} + mE_{diol}\right) \qquad (4.1)$$

Where $E_{co-crystal}$ is the calculated energy for a co-crystal cell containing n urea and m diol molecules. E_{urea} and E_{diol} are the lattice energy per molecule for urea and diol respectively. To calculate the interaction energy, lattice energy is calculated out of co-crystal, dense urea and different analogous of dense diol.

3.2.2 Roles of Multipoles in DMAREL Calculation

Having accurate multipole moments is very important in simulating the lattice energy of a crystal. It is the easiest intermolecular forces to quantify from *ab initio* wave functions from its isolated molecules as the charge distribution can be well represented by sets of multipoles at each atomic site. It has the advantage over both the central multipole expansion, which is generally used for small polyatomics and the atomic charge model, which is the usual model for organic molecules and has the ability to readily represent the electrostatic effect of non-spherical features such as

lone pair and π electron density [11]. When the electrostatic interactions is accurately calculated, it often dominates the orientation dependence of the total intermolecular potential, particularly for hydrogen bonding and π-π interactions between uncharged molecules.

In order to have the whole series of lattice energy for urea/diol co-crystals, we produced multipoles using a new program by Stone, whereby it generates Distributed Multipole Analysis of wave function calculated by the Gaussian program system. The program, called GDMA [10] uses the formatted checkpoint file produced by Gaussian to create multipole moments at sites with different order limit (charges, dipoles, quadrupoles, etc) defined by users. With a given wave function, it provides an accurate description of the electrostatic field of the molecule. It also handles diffuse functions satisfactorily giving results which converge rapidly to steady values as the basis set is improved and is much cheaper compared to direct wave function calculation.

3.3 Parameter Test for DMAREL using MOLPRO and GDMA Multipole

To obtain the reference energies for calculating interaction energies, calculation was first applied to dense urea and diols. This also allowed testing of parameters for DMAREL calculations on co-crystal systems. One of the important components in lattice energy calculation using DMAREL is using the right DMAs. This is calculated using MOLPRO and Gaussian by single point calculation of single molecule with MP2 method and 6-31G (d,p) basis sets. Lattice energy of both DMAs for diols are shown in Table 3.5:

Table 3.5 : Total lattice energy calculated using MOLPRO and Gaussian-GDMA DMAs

Crystal	Total Lattice Energy per molecule (kJ mol^{-1})		Literature
	MOLPRO	GDMA	
Urea	-89.67	-87.92	87.7 – 98.6 [12,13]
1,4-dhb	-83.83	-86.10	95.7**
1,6-dhh	-94.83	-96.24	112.0 [14]
1,8-dho	-110.75	-112.97	139.3 [14]
1,10-dhd	-3321.25	-125.59	155.8 [14]
1,12-dhdd	-17650.96	-126.54	n/a
1,14-dhtd	-38370.14	-139.89	n/a

* All calculations are using scale factor of 0.9 in DMAREL

** Total of enthalpy of vaporization and fusion [15] of 1,4-dhb

Table 3.5, shows that for 1,4-dhb to 1,8-dho co-crystals, the lattice energies calculated using MOLPRO or GDMA are reasonable compared to the literature value. However, for the case of 1,10-dhd diols and their higher analogs, calculation using MOLPRO DMAs are far too high compared to the one from GDMA. The calculation of DMA in MOLPRO started to give unreasonable lattice energy results starting with 1,10-diol. Looking at the total charge (Table 3.6) produced by MOLPRO for these diols, it is consistent with the high lattice energy calculated. This high molecule charge is unphysical and results in high lattice energy crystal. The total charge calculated for each molecule are within 6 figures of decimal points due to the accuracy of the multipoles used.

Table 3.6: Total charged produced in MOLPRO and GDMA

Crystal	Total Charge for each molecule (e)	
	MOLPRO	**GDMA**
Urea	0.18×10^{-5}	0.18×10^{-5}
1,4-DHB	-0.18×10^{-5}	-0.14×10^{-4}
1,6-DHH	-0.18×10^{-5}	-0.54×10^{-5}
1,8-DHO	0.18×10^{-5}	-0.50×10^{-4}
1,10-DHD	-3.60	-0.61×10^{-4}
1,12-DHDD	-9.00	0.90×10^{-5}
1,14-DHTD	-14.40	0.25×10^{-4}

The energy was then compared with multipoles modelled using MOLPRO and to determine how much it differs from the experimental related to sublimation energy. Table 3.7 shows the energy profile of urea dense crystal and the error produced from the methods mentioned above.

Table 3.7 Comparison of lattice energy of urea dense crystal relaxed structure and its relative error produced between GDMA and MOLPRO multipoles.

Energy per molecule (kJ/mol)	**MOLPRO multipoles**	**GDMA multipoles**
	-89.67	-87.92
Range of parameter (a,b,c) change (%)	± 0-1	± 0-2
rms % error in lattice parameters[a]	0.85	1.33

[a]The rms % error is calculated as $3\delta^2 = (100\Delta a/a)^2 + (100\Delta b/b)^2 + (100\Delta c/c)^2$

Both energy produced using MOLPRO and GDMA multipoles are still within the range of sublimation energy (-87.65 to -98.6 kJ mol^{-1}) [16-19] , which is also considered as the experimental lattice energy. Contrary to that obtained from the MOLPRO multipoles, GDMA multipoles also produced similar lattice energy to the one observed experimentally. In terms of quality, GDMA produced slightly higher errors. However, it is still within 3% of rms, which is quite acceptable, as stated by

Coombes [13] *et al* in their paper on the role of electrostatic interaction to crystal structure study.

From these results, we try to look back to find out why these multipoles gave different results of energy. We compared the charge distributions produced by each method for urea. GDMA gives smaller charges and has higher value for its distributed moments than MOLPRO. However, MOLPRO have larger sets of charge, while all other distributes moments have smaller values. This is shown in Table 3.8 for oxygen atom charge and dipole values. Originally, in the distributed multipole procedure, each product of primitives is represented by multipoles on the site nearest to its overlap centre. However, GDMA assign the charged sites using numerical quadrature of Becke [20] with some minor difference for the diffuse function, to overcome the strong basis-set dependence of the original version but is simpler to implement than other real-space partitioning schemes.

Table 3.8 : Charge and dipole moments comparison between MOLPRO and GDMA for oxygen atom in urea.

Multipole	**MOLPRO**	**GDMA**
O (q)	-1.025243	-0.513323
O (μ_z)	0.000000	0.000000
O (μ_x)	0.359273	-0.520542
O (μ_y)	0.000000	0.000000

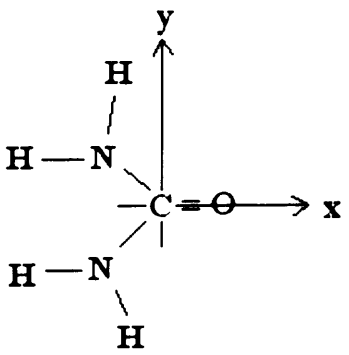


Figure 3.8: Urea isolated molecule defined on *xy*-plane

In the case of larger diols (1,10-dhd and above), charges and dipole moment are shown below (Table 3.9). MOLPRO assigned quite a large value for the charges and multipoles, particularly, for the carbon atoms. This produces high charge (-3.60) to the accumulated charge in the crystal, as shown in Table 3.6. The deviation of error in the total charge increases as the size of diol molecules become larger.

Table 3.9 : Charge and dipole moments comparison between MOLPRO and GDMA for carbon (CAR6) atom in 1,10-dhb.

Multipole	MOLPRO	GDMA
C (q)	0.362574	0.046091
C (μ_z)	0.013772	0.003841
C (μ_x)	0.174893	0.061472
C (μ_y)	0.431397	0.114007

From here onwards, Gaussian-GDMA multipoles will be used for DMAREL calculation for lattice energy.

3.4 Parameter Optimization for GDMA Multipoles for Dense Crystals

Since GDMA multipoles will be implemented in DMAREL, new parameters may be needed. Basic parameters such as the scale factor are fundamentally very important since it is directly involved in correcting the long range calculation.

3.4.1 Scale Factor Correction to Urea Crystal

Once the choice of GDMA multipoles was made, we tried to enhance the lattice energy by looking at the scale factor effect to match relatively to the newest sublimation energy results, -97.6 kJ mol^{-1} [18]. Earlier dipole moments calculated from 6-31G(d,p) wave function for eight small molecules, by Cox and Williams [21], led to the scaling factor 0.9 which is widely used as a 'fudge factor' in molecular modelling. This is also currently used as a default scale factor in DMAREL, which

allows the effect of electron correlation on the wave function and so that the multipoles given to the program are always scaled by this factor prior to use in lattice simulation. However, it is only a crude function and this ratio varied from 0.78 for NH_3 to 0.96 for CH_3CN for molecules considered by Cox and Williams [21]. Coombes [13] *et al* demonstrate that the strength of the electrostatic forces is relative to the other components of the potential. Therefore the quality of the wave functions used to derive the electrostatic model, does affect the calculated structure as well as the lattice energy.

Here we first try to test the multipole scale factor for urea dense crystal between 0.7 to 1.5 in 0.1 intervals. However, it seems that lattice energy change drastically with 0.1 scale factor difference. Scale factor 0.9 gave lattice energy of -87.92 kJ mol^{-1} while 1.0 resulted to be -105.17 kJ mol^{-1}. We narrowed down the search to the range within 0.9 to 1.0 in 0.01 intervals. The energy was slowly decreasing (becoming more negative) as we increase the scale factor (Table 3.10). 0.96 scale factor produced appropriate lattice energy which replicates the recent [18] published sublimation energy. It was also reported that the correlation coefficient for molecules containing double bonded oxygens may vary between 0.94 to 0.98 [22]. Unfortunately, with the increase of scale factor, it also increases the deviation of structure from its experimental geometry. However, from Table 3.10, it shows that the scale factor did not change the F factor much from its starting point while at the same time improving its lattice energy. Moreover, the result is still better than obtained by Coombes *et al* [13], which produced estimated rms % error in lattice parameter to be 1.7 with EST+0.9DMA method.

Table 3.10: Effect of Scale Factor of urea crystal within the range of 0.9 to 1.0 in 0.01 intervals.

Scale Factor	rms % error in the lattice parameter (average of Δa%, Δb%, Δc%)[a]	Δx (molecular movements including strains) (Å)	F[b] (Discrepancy indices)	Lattice Energy (kJ mol^{-1})
0.90	1.33	0.079225	16.62	-87.92
0.91	1.35	0.079114	16.92	-89.53
0.92	1.36	0.079086	17.28	-91.17
0.93	1.38	0.079140	17.70	-92.83
0.94	1.40	0.079277	18.18	-94.52
0.95	1.42	0.079518	18.73	-96.24
0.96	**1.44**	**0.079815**	**19.33**	**-97.97**
0.97	1.47	0.080191	19.98	-99.74
0.98	1.49	0.080644	20.69	-101.52
0.99	1.52	0.081172	21.46	-103.33
1.00	1.55	0.081774	22.29	-105.17

[a] The rms % error is calculated as $3\delta^2 = (100\Delta a/a)^2 + (100\Delta b/b)^2 + (100\Delta c/c)^2$

[b] F, the structural discrepancy indices is defined [23] as:

$$F = \left(\frac{\Delta\theta}{2}\right)^2 + (10\Delta x)^2 + \left(100\frac{\Delta a}{a}\right)^2 + \left(100\frac{\Delta b}{b}\right)^2 + \left(100\frac{\Delta c}{c}\right)^2 + \Delta\alpha^2 + \Delta\beta^2 + \Delta\gamma^2$$

where rms values for Δθ (the net rotation of the molecule (deg) and Δx (the net translation of the centre of mass of the molecule (Å)) are used when there is more than one independent molecule in the asymmetric unit cell.

3.4.2 GDMA multipoles for dense diol crystals

The above method was then repeated using GDMA multipole moments to dense diol crystals. The effect of cut-off was also tested for each diol structure to reconfirm at which cut-off does the lattice energy come into agreement with the experimental values. We want to ensure whether the cut-off interdependency to the length of diols.

Lattice energy was finally converged and the best cut-off calculated are 25.0 Å for 1,4-dhb, 30.0 Å for 1,6-dhh and 35.0 Å for both 1,8-dho, 1,10-dhd, 1,12-dhdd. This shows that cut-off also depends on the crystal cell length. The difference of energies with different cut off is about 1 kJ mol^{-1} to the converged lattice energy. Based on the discrepancy index shown, the mentioned cut-off are used for lattice energy calculation. Finally, 0.96 scale factor was applied to these best cut-off structures.

The lattice energy of GDMA and its comparison to the sublimation energies obtained from the literature is shown in Table 3.11. GDMA does show good sensitivity to the lattice parameters and improves the energy efficiently especially to the longer diols where this fails when using MOLPRO. However, it overestimates the values of sublimation energy. Lattice energy, as mentioned by Kitaigorodsky [24] in analysing differences of calculated lattice energy with the experimental sublimation energy, concludes that discrepancies up to 3-4 kcal mol^{-1} (12–17 kJ mol^{-1}) should not cause any concern when judging the quality of a potential model. This is because of the neglect of the zero point energy and the high experimental error in determining the reference values.

Table 3.11: Energy profiles of diols crystals GDMA multipoles method using 0.96 scale factor.

Diols	$\Delta_{sub} H^{\circ}$ (kJ mol^{-1})	GDMA	
		Relaxed Energy (kJ mol^{-1})	rms % error in the lattice parameter
1,4-dhb	-95.7	-95.10	2.01
1,6-dhh	-112.0	-103.95	0.86
1,8-dho	-139.3	-121.33	1.59
1,10-dhd	-155.8	-135.24	0.97
1,12-dhdd	n/a	-150.66	0.18
1,14-dhtd	n/a	-165.47	0.21

As expected, the longer the diol chain, the higher the sublimation energy.

3.5 Non-planarity of urea in urea/diol co-crystal

The conformation of urea in each co-crystal is theoretically planar. However, due to the X-ray structure determination N-(C=O)-N urea plane in all co-crystal structures were found to be slightly non-planar as shown in Table 3.12.

Table 3.12: Torsion angle of heavy atoms (N-(C=O)-N) in urea of urea/diol co-crystal.

Type of urea/diol co-crystal	Urea N-(C=O)-N Torsion Angle (°)
1,4-dhb	-179.81
1,6-dhh	-179.74
1,8-dho	-179.35
1,10-dhd	-179.32
1,12-dhdd	-179.32

This non-planarity is corrected while calculating the energy minimization by turning the urea molecules into a planar structure.

3.6 Energy relaxation for co-crystals using GDMA multipoles

The same routine was repeated for urea/diol co-crystals to determine the lattice energy and most importantly to observe the interaction energy. A Cut-off of 35.0 Å is used for this calculation since it's the best cut-off value tested for all co-crystals. Table 3.13 shows that urea/diol co-crystal produce good relaxed structure with the similarity of discrepancy indices produced. Good sensitivity was also shown in its rms % error which is about or less than 3%.

Table 3.13: Comparison of energy and physical properties in urea/diol co-crystal. All co-crystals are in anti-parallel urea ribbon conformation except urea/1,8-dho co-crystal

Urea/diol co-crystal	Total Energy (kJ mol^{-1})	Total Energy (kJ mol^{-1}) 2:1 ratio	rms error[a] (%)	F (Discrepancy Index)
1,4-dhb	-671.45	-335.73	2.10	30.60
1,6-dhh	-343.00	-343.00	1.97	28.54
1,8-dho	-630.57	-315.28	2.55	32.73
1,10-dhd	-373.00	-373.00	3.26	38.93
1,12-dhdd	-386.24	-386.24	3.27	41.54

[a] The rms % error is calculated as $3\delta^2 = (100\Delta a/a)^2 + (100\Delta b/b)^2 + (100\Delta c/c)^2$

Since each co-crystal have different number of molecule crystals in each lattice cell, it is easier to compare each energy terms in 2:1 ratio, which represents 2 urea and 1 diol in a co-crystal. The total energy of co-crystal 2:1 ratio are plotted in a graph to observe how lattice energy depend on diol length (Figure 3.9).

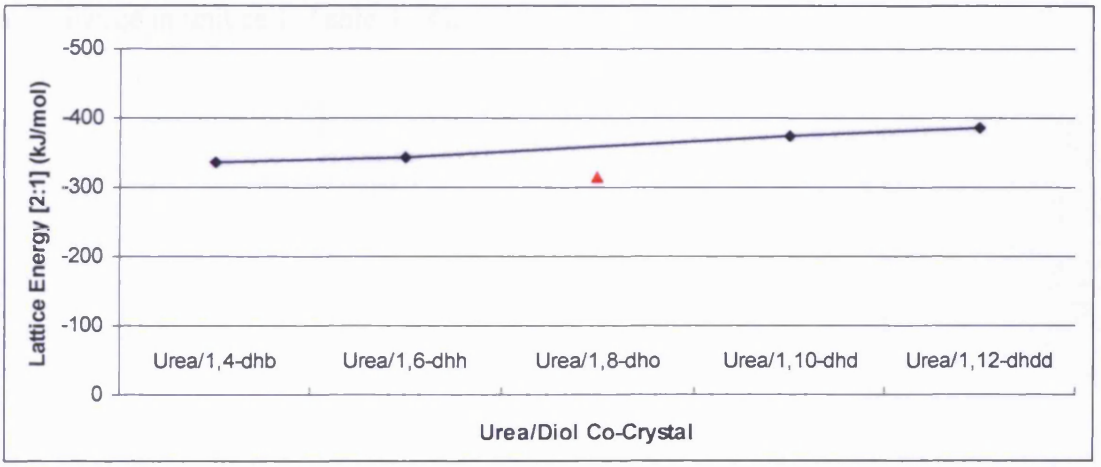


Figure 3.9: Total lattice energy plot of urea/diol co-crystal in 2:1 (urea:diol) ratio. All co-crystal structure are in anti-parallel urea ribbon (blue) except urea/1,8-dho co-crystal (red).

Based on the graph in Figure 3.9, lattice energies of all four urea/diol co-crystal with anti-parallel ribbon structures can be plotted as a straight line. However, the urea/1,8-dho shows that this structure has a higher lattice energy at -315.28 kJ mol^{-1} compared to urea/diol co-crystal with anti-parallel urea ribbon. This may explains why the crystallization of this type of urea/diol co-crystal structure is difficult. The difference between one diol homolog co-crystal to another between 7 to 13 kJ mol^{-1}. If an interpolation were drawn to determine the possibility of lattice energy of urea/1,8-dho in with anti-parallel urea ribbon structure, it is predicted to be about -357.58 kJ mol^{-1} (based on graph reading calculated in excel).

3.6.1 Discrepancy index: An overall comparison

The use of the optimised parameters (cut-off and scale factor) could now produced reliable relaxed structure with reasonable lattice energy. This is proven with small discrepancy indexes produced by each crystal (Figure 3.10). Discrepancy index is the total of several factors in relaxed geometry compared to its initial experimentally determined structure. The factors contribute to the discrepancy index are rigid-body rotational displacement, rigid-body translational displacement, change in cell length and change in unit cell (Table 3.14).

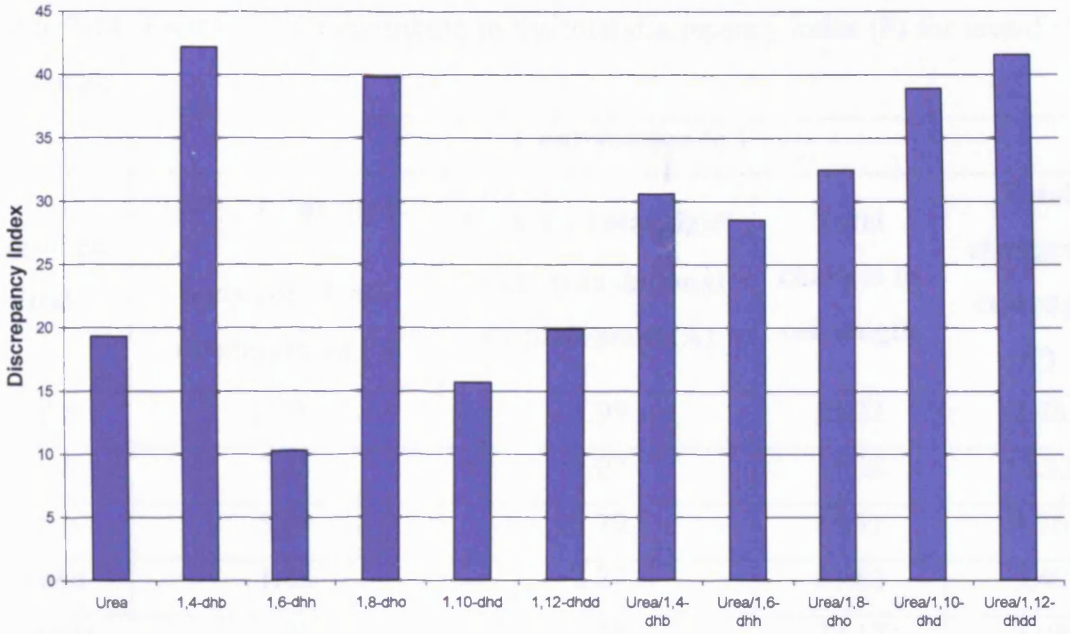


Figure 3.10: Comparison of discrepancy indices for all dense crystals and urea/diol co-crystals

The main factor which contributed to the discrepancy indices for each co-crystals is due to the change in cell length during relaxation. The most affected by this changes is urea/1,12-dhdd co-crystal with the discrepancy indices 41.54. In the case of urea/1,4-dhb co-crystal, it is also shown that the rigid-body rotational displacement also contributed significantly to its discrepancy index.

Table 3.14: Factors which contribute to the total discrepancy index (F) for urea/diol co-crystal.

	Contribution to F			
Urea/diol co-crystal	$\left(\frac{\Delta\theta}{2}\right)^2$, **Total rigid-body rotational displacement (°)**	$(10\Delta x)^2$, **Total rigid-body translational displacement (Å)**	**Total changes in cell length**	**Total changes in cell angle (°)**
1,4-dhb	12.91	1.99	13.22	2.48
1,6-dhh	2.25	1.07	11.68	13.53
1,8-dho	5.64	5.79	19.31	1.76
1,10-dhd	1.45	2.27	31.82	3.40
1,12-dhdd	1.93	4.35	32.17	3.08

3.7 Interaction energy of experimental urea/diol co-crystal

The interaction energy of a co-crystal shows how stable the formation of the co-crystal is compared to the alternative pure dense phases of a crystal. This can be obtained by calculating the lattice energy difference of the co-crystal to its dense crystals. Table 3.15 shows that GDMA multipoles produced negative interaction energy for all anti-parallel urea ribbon type co-crystals. However, in urea/1,8-dho with parallel urea ribbon type of co-crystal, positive interaction energy is calculated. This implies that the crystals needs to form in a special condition, depends on its solvent and kinetic factors. This may also explain the difficulty in obtaining this type of co-crystal from other homolog diols to form experimentally.

Table 3.15: Interaction energy obtained using the new scale factor in 2:1 ratio. All are in parallel urea ribbon conformation except for urea/1,8-dho co-crystal.

Urea/diol Co-crystal	Interaction Energy (kJ/mol)
	GDMA Multipole
1,4-dhb	-44.64
1,6-dhh	-43.11
1,8-dho	1.99
1,10-dhdd	-41.82
1,12-dhdd	-39.64

Another possible explanation on why urea/1,8-dho of anti-parallel type of structure was not obtained experimentally is because this might be due to the nucleation energy for anti-parallel structure of urea/1,8-dho is higher than the parallel type urea ribbon (Figure 3.11).

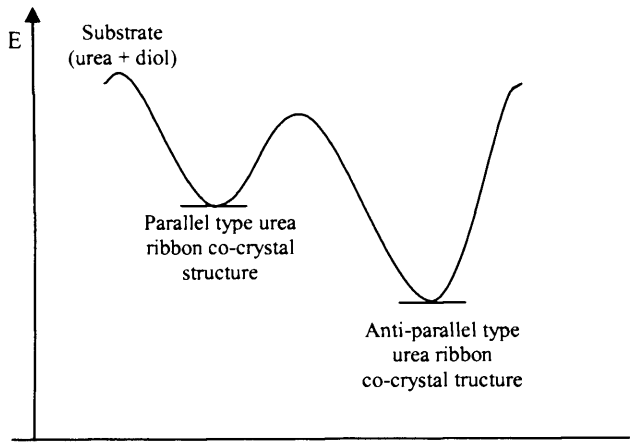


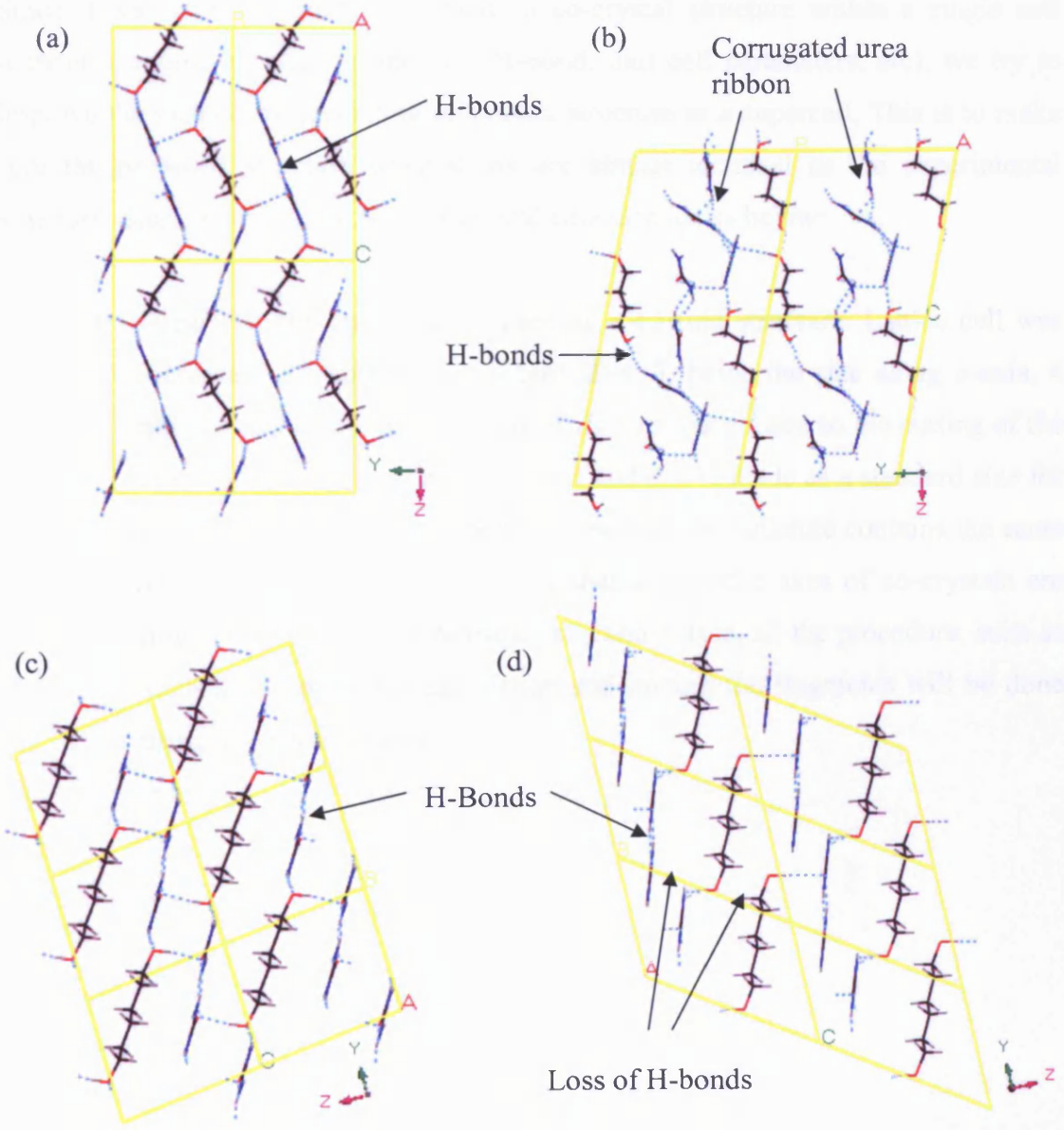
Figure 3.11: Schematic illustration of nucleation energy of urea/diol co-crystal formation.

3.8 Constructing of Urea/Diol Structure from Experimental

It is therefore of interest to simulate urea/diol co-crystal structure of parallel-ribbon type and vice versa. In predicting structure and energy of the urea/α,ω-dihydroxyalkane polymorph series, we decided to start with co-crystal that

has anti-parallel urea ribbon structure (space group $P\bar{1}$) since there is only 1,8-dihydroxyalkane missing in this series. This was done by shortening a longer diol of co-crystal into the required length in a single cell. During relaxation, the multipoles from optimised diol structure was used. We also tried to predict the energy of shorter diols obtained in experiment (C4 and C6) as a control method, to compare the result with the experimental structure and its energy value. However, other than shortening the diols, other parameters are not changed, including the unit cell parameter. It was expected that upon relaxing the energy, the cell parameter should also shrink.

Figure 3.12 : Comparison of co-crystal orientation and hydrogen network depicted in larger cell. Axis in predicted cell is in different orientation from the experimental due to its origin co-crystal structure cell. (a) Experimental structure of urea/1,4-dhb co-crystal viewed in 1×2×2 lattice cell size , (b) Simulated structure of urea/1,4-dhb co-crystal (viewed in 1×2×2 lattice cell size) established from urea/1,12-dhdd structure, (c) Experimental structure of urea/1,6-dhh co-crystal, viewed in 1×3×2 lattice cell size, and (d) Simulated structure of urea/1,6-dhh co-crystal (viewed in 1×3×2 lattice cell size) established from urea/1,10-dhd structure.



Suprisingly, the results gave unexpected implication to the size of the unit cell. The unit cell becomes more oblique and most of the hydrogen bonds are not in order. Less

H-bonds occur between diols and urea ribbons. Density and volume of the cells changes crucially for 1,4-dihydroxybutane at -11.91% (density) and 13.51% (volume) compared to experimental. However, 1,6-dihydroxybutane differs only by -1.80% (density) and 1.83% (volume). This shows that, modification from long (1,12-dhdd) to shorter diol, 1,4-dhb (differ by 8 carbons) will make the structure more inaccurate, and we conclude that, the unit cell parameter should be changed before performing structure relaxation. (Figure 3.12)

Since it was quite difficult to replicate a co-crystal structure within a single cell without changing its other properties (H-bond, unit cell parameters, etc), we try to improve the method by predicting co-crystal structure as a supercell. This is to make sure the proposed structure orientations are almost identical as the experimental structure. Steps in producing new co-crystal structure are as below:

a) Urea/1,10-dhh co-crystal is used as a starting structure. Lattice cell was increased to a supercell structure, 2×4×3; twice the size along *a*-axis, 4 along *b*-axis and 3 along *c*-axis. This size was chosen so the cutting of the diols in supercell is straightforward and this is made as a standard size for all proposed co-crystal. Therefore, each of the structure contains the same amount of molecules. Assuming that a, b and c-axes of co-crystals are nearly aligned with its cartesian x, y and z-axes, all the procedure, such as atom deletion, vector calculation and joining the fragments will be done using x,y,z coordinates.

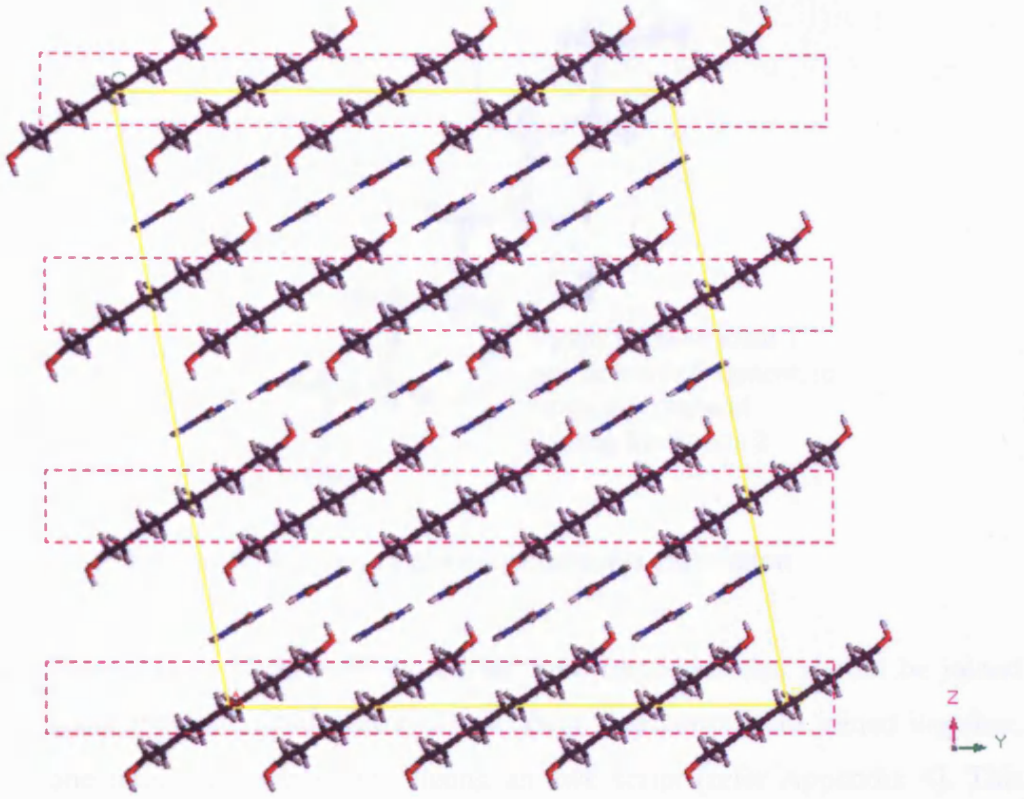


Figure 3.13 : Schematic diagram of urea/1,10-dhd in supercell (2×4×3 lattice size). Fragments in pink dotted box will be deleted, while the remaining will be joined together.

b) A vector between the two diol fragments is then computed from the C atom in one end of the diol to the atom it would replace before deleting the appropriate carbon atoms (Figure 3.14).

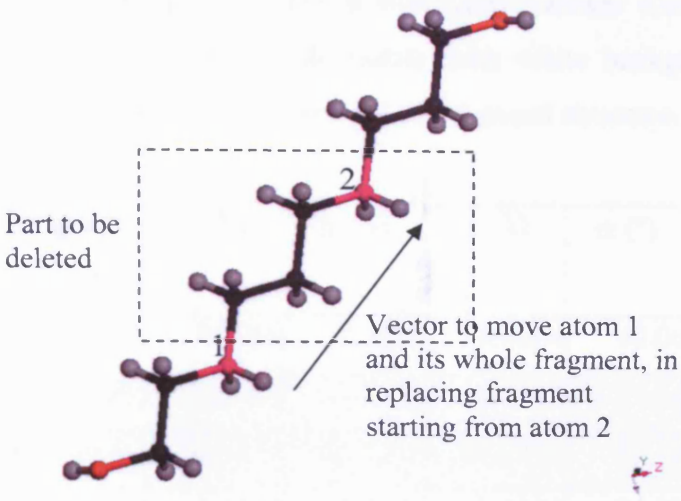


Figure 3.14 : Atom deletion and vector calculation

c) From Figure 3.13, there would be three fragments that should be joined using the calculated vector. These three fragments were joined together, one fragment after another using an awk script [refer Appendix 4]. This consistent relocation will keep the C-C bond and all the H-bonds similar to the experimental value.

d) Finally, the length of the unit cell is shortened twice (as there are two diol layers in the supercell) its z-component of the vector translation where most of the translation has taken place. Lattice angles are also adjusted so the H-bonds of the urea ribbons and H-bonds between urea and diols are keep intact to each other. During the energy calculation, multipoles from optimized structures are used.

After applying this method, it seems that all predicted structure show an improvement to its proposed relaxed structure. Table 3.16 shows the structure packing properties of experimental urea/α,ω-dihydroxyalkane co-crystal and its proposed structure constructed from different original structure. Proposed supercell [2×4×3] structure packing properties shows proportional value to the experimental structure rather than the proposed structure in single unit cell, whereby in 1,6-dhh, α angle of proposed structure is way too large compared to the experimental.

Table 3.16 : Structure Packing Of Urea/ α,ω-dihydroxyalkane Co-crystal (anti-parallel urea ribbon/dipole). Urea/diols co-crystals with white background are the experimental structure and darker background are of proposed structure.

Urea/diols co-crystal	Original Structure	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
1,4-dhb		5.2046	7.3638	14.6064	90.00	98.60	90.00
1,4-dhb*	1,12-dhdd	4.6957	11.3440	12.2047	79.34	96.82	90.70
1,4-dhb* [2×4×3]	1,10-dhd	10.4110	28.7015	22.9711	92.38	99.22	100.87
1,6-dhh		5.1642	7.4945	8.4337	94.82	96.02	101.49
1,6-dhh*	1,10-dhd	5.1820	7.3439	10.6282	124.45	87.55	101.42
1,6-dhh* [2×4×3]	1,10-dhd	10.3845	29.5603	26.3500	97.83	92.60	101.24
1,8-dho* [2×4×3]	1,10-dhd	10.3580	29.7929	30.5110	99.31	94.14	100.56
1,10-dhd		5.1630	7.3670	11.1010	98.62	96.30	102.00
1,12-dhdd		5.1650	12.2620	14.5610	88.52	81.90	79.27

* Proposed structure properties after structure minimisation

The H-bond arrangement of urea ribbons and H-bond between urea and diols survived after structure relaxation. However, different length of diols involved in the co-crystal shows different result. Some of the H-bonds are too stretched (Figure 3.15 (f)), which then produced high interaction energy (Table 3.17). This occurs to the proposed 1,8-dihydroxyoctane structure. In 1,6-dihydroxyhexane structure H-bonds are not crucially stretched (Figure 3.15 (d)), nevertheless the urea ribbons arrangements are similar to its experimental structure. For 1,4-dihydroxybutane proposed structure, even though the arrangement is not the same as in experimental (Figure 3.1 (a)), we manage to minimised its interaction energy in relaxed structure compare to proposed structure in single cell.

Figure 3.15 : Initial and relaxed structure of proposed co-crystal. The blue lines are the H-bonds in the co-crystal structure system.

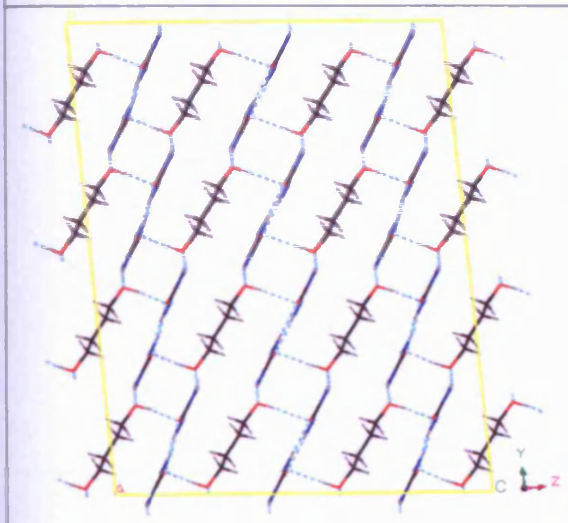
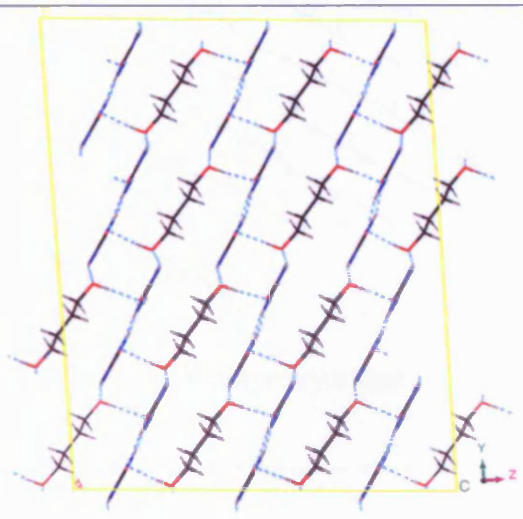
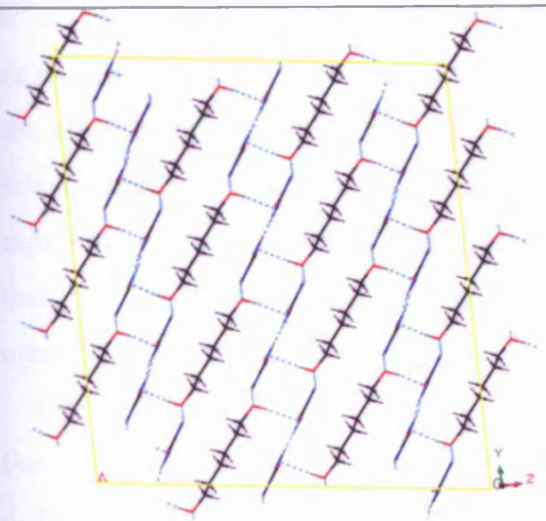
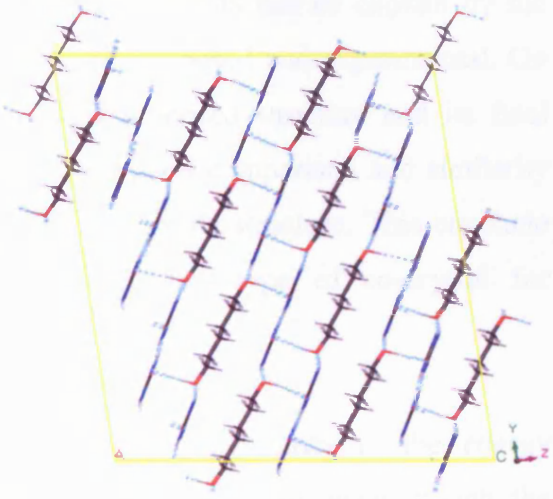
Initial Structure	**Relaxed Structure**
 a) 1,4-dihydroxybutane	 b) 1,4-dihydroxybutane
 c) 1,6-dihydroxyhexane	 d) 1,6-dihydroxyhexane

Table 3.17 : Lattice energy (kJ mol^{-1}) of Co-crystal Structure (anti-parallel urea ribbon/dipole). Urea/diols co-crystals with white background are the experimental structure and darker background are of proposed structure. Final energy of supercell has been scaled back to one unit cell for comparison.

Co-crystal	Urea	Diols	Final Energy (kJ mol^{-1})	Lattice Energy (kJ mol^{-1})	
				per unit cell	2:1 ratio
1,4-DHB	4	2	-573.73	-671.46	-335.73
1,4-DHB*	4	2	-399.94	-563.24	-281.62
1,4-DHB* [2×4×3]	48	24	-	-714.49	-357.25
1,6-DHH	2	1	-300.24	-343.00	-343.00
1,6-DHH*	2	1	-294.39	-365.03	-365.03
1,6-DHH* [2×4×3]	48	24	-	-342.90	-342.90
1,8-DHO*	4	2	-677.05	-788.08	-394.04
1,8-DHO* [2×4×3]	48	24	-	-668.86	-334.43

* Proposed structure energy properties after energy minimization. Final lattice energy which is the energy of the whole [2×4×3] lattice structure is not shown due to its large value

Lattice energy of the systematically generated proposed structure shows an improvement compared to the one generated manually from its initial experimental lattice cell. It also shows similar result to the experimental lattice energy. However, for the new generated anti-parallel urea ribbon structure type of urea/1,8-dho, it shows higher lattice energy compared to all other bulk constructed structures with lattice energy at -334.43 kJ mol^{-1}. This might be due to the loss of hydrogen bonds that occurs during relaxation. On the other hand, this may also explained why it is difficult to crystallize experimentally.

We suggested that the molecule becomes more flexible at longer diol and performed a molecular dynamic calculation with the bulk proposed structure. However, there were not enough data calculated that can be presented in this thesis.

3.9 Conclusion

In an attempt to determine the lattice energy of urea/α,ω-dihydroxyalkanes (urea/diol) co-crystal, we introduced a new GDMA multipole for use in DMAREL. This multipole solved the problem that occurs in this homologous set of co-crystal structure compared to MOLPRO multipoles.

The DMA simulation from the new multipoles successfully and accurately describes urea, all pure diol co-crystals with discrepancy indices below 45. The lattice energy for all co-crystal structure shows that urea/diol with anti-parallel urea ribbon type of co-crystal has lower lattice energy compared to the parallel type urea ribbon structure by about 42.30 kJ mol^{-1} when compared to the anti-parallel type. Interaction energy also shows that parallel urea ribbon type of co-crystal structure is higher (1.99 kJ mol^{-1}) than the anti-parallel type which is about -39 to -44 kJ mol^{-1}.

Co-crystals constructed by deleting atoms, eg 1,6-dhh generated from 1,10-dhd give lattice energies and structures in good agreement with experimental observations. However, the remaining discrepancy between the two suggests that model including molecular flexibility will be required to give genuinely predictive results. Some initial molecular dynamics calculations were carried out to this end but the available project time was insufficient for a full investigation.

3.10 References

(1) Lee, S.-O.; Kariuki, B. M.; Harris, K. D. M. *New Journal of Chemistry* **2005**, *29*, 1266 - 1271.

(2) Marti-Rujas, J., Cardiff University, 2005.

(3) Etter, M. C.; MacDonald, J. C. *Acta Cryst.* **1990**, *B46*, 256-262.

(4) Etter, M. C. *J. Phys. Chem.* **1991**, *95*, 4601-4610.

(5) Thalladi, V. R.; Boese, R.; Weiss, H.-C. *Angewandte Chem. Int. Ed.* **2000**, *39*, 918-922.

(6) M.Lindgren; T.Gustafsson; J.Westerling; A.Lund *Chemical Physics* **1986**, *106*, 441.

(7) Nakamura, N.; Sato, T. *Acta Cryst., Sect. C: Cryst. Struct. Commun.* **1999**, *55*, 1685-1687.

(8) Nakamura, N.; Setodoi, S. *Acta Cryst., Sect. C: Cryst. Struct. Commun.* **1997**, *53*, 1883.

(9) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233-239.

(10) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128-1132.

(11) Scheiner, S. In *Wiley Tutorial Series in Theoritical Chemistry*, 1994; pp 297.

(12) Chickos, J. S. In *NIST Chemistry Web Book*; Linstrom, W. G. M. a. P. J., Ed.; National Institute of Std and Tech, Gaithersburg, MD, http://webbook.nist.gov, 1998; Vol. 69; pp 20899.

(13) Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *J. Phys. Chem.* **1996**, *100*, 7352-7360.

(14) Knauth, P.; Sabbah, R. *Can. J. Chem* **1990**, *68*, 731–734.

(15) In *NIST Standard Reference Database Number 69*; U.S. Secretary of Commerce on behalf of the United States of America, 2008.

(16) Suzuki, K.; Onishi, S.; Koide, T.; Seki, S. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 127.

(17) DeWit, H. G. M.; Miltemburg, J. C. V.; DeKruif, C. G. *J. Chem. Thermodyn.* **1983**, *15*, 651.

(18) Kabo, G. Y.; Miroshnichenko, E. A.; Frenkel, M. L.; Kozyro, A. A.; Simirskii, V. V.; Krasulin, A. P.; Vorob'eva, V. P.; Lebedev, Y. A. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1990**, 662-667.

(19) Chickos, J. S. In *Molecular Structure Energetics*; Liebman, J. F., Greenberg, A., Eds.; VCH: New York, 1987; Vol. 2; pp 67.

(20) Becke, A. D. *The Journal of Chemical Physics* **1988**, *88*, 2547-2553.

(21) Cox, S. R.; Williams, D. E., 1981; Vol. 2; pp 304-323.

(22) Naray-Szabo, G.; Ferenczy, G. G. *Chem. Rev.* **2002**, *95*, 829-847.

(23) Filippini, G.; Gavezzotti, A. In *Acta Crystallogr.,Sect.B*, 1993; Vol. 49; pp 868-880.

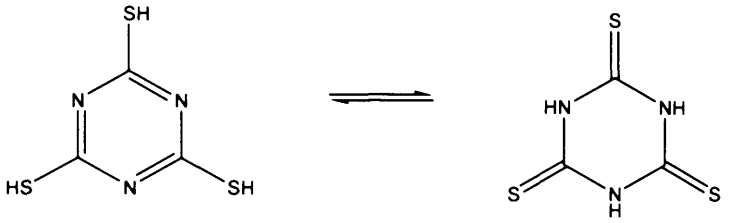
(24) Gavezzotti, A.; Filippini, G. In *Theoretical Aspects and Computer Modelling of the Molecular Solid State*; Gavezzotti, A., Ed.; John Wiley & Sons, 1997; Vol. 1; pp 61-97.

Chapter 4

Sulfur Potential & Lattice Cell Substitution

4.1 Introduction

1,3,5-triazine-2,4,6-trithiol can also be referred to as 2,4,6-trimercaptotriazine or commonly known as trithiocyanuric acid (TTCA). As illustrated in Figure 4.1, it can exist in either thiol (a) or thione (b) forms in aqueous solution.



(a) (b)

Figure 4.1 : Trithiocyanuric acid (TTCA) isomeric structures.

TTCA is a highly soluble water compound $C_3N_3S_3Na_3.9H_2O$ and can form precipitates with heavy metal ions like Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, Hg^{2+}, Pb^{2+}, Hg^{+}, Ag^{+}, and Tl^{+}. It has been used as a removal reagent of heavy metal ions from industrial waste water. It is also very efficient in removing residual palladium and its compounds from reaction mixtures, in which they are used as catalysts where the residues are undesirable, especially for pharmaceutically active ingredients. Recently, it has been used as a synthesizing agent for Zn^{2+}, Fe^{2+} and Mn^{2+} complexes for study of antitumor activity [1]. The presence of nitrogen and sulfur atoms and the analogy to the pyramide nucleobase makes TTCA a very interesting compound from a biological point of view [1].

Co-crystallisation of TTCA was observed from the crystallisation of TTCA with 4,4'-bypyridyl (BP) [2]. It reveals the presence of intermolecular N-H···N hydrogen bonds between TTCA and BP, creating cavites of 10Å diameter that can be occupied with aromatic compounds such as benzene, toluene, *p*-xylene and anthracene. Further discussion about this work was done by Ahn *et al* [3], about the crystal properties of TTCA co-crystals with acetone, 2-butaone, dimethylformamide, dimethyl sulfurdioxide, methanol and acetonitrile, who analysed the hydrogen bonding

functionality and capability of these co-crystals. The co-crystals are classified into different group depending on the topological arrangements of the TTCA molecules in the crystal structure governs by hydrogen bonding.

Pure TTCA is quite hard to crystallise. Previously, this has been achieved with some water molecules as reported by Mahon *et al* [4]. Pure TTCA is difficult to crystallise because it has a strong propensity to form co-crystal (solvate) structure in crystallization experiment from the type of solvents in which is readily soluble. In 2006, TTCA pure crystalline form without any water of crystallisation (Figure 4.1 (b)) was successfully obtained by Guo and Harris *et al* [5], by comparing its powder XRD pattern to a pure commercial TTCA. However, the single crystal pattern can't be use for structure determination. Thus, they successfully solved the pure TTCA structure from commercial sample using the direct-space genetic algorithm technique for the structure solution, followed by Rietvield refinement. It comprises of two molecules of TTCA in a monoclinic unit cell. Each molecule was connected with 6 hydrogen bonds, 3 donors and 2 acceptors forming TTCA sheets parallel to $(1\,\bar{2}\,0)$ plane [5]. Of all 3 sulfur atoms, one of them, (S(2)) do not involve in hydrogen bonding (Figure 4.2). The other two, S(1) and S(3) each formed hydrogen bonding with a neighbouring TTCA molecule involving a pair of N-H···S in a cyclic $R_2^2(8)$ type of arrangement.

These networks of hydrogen bonding create two types of connectivity (Figure 4.2 (b)) in the crystal packing based on the connectivity type suggested by Ahn *et al* [3]. In type 1 (represented by black line), the free N-H bond and C=S bond are at opposite corners of the molecule with $R_2^2(8)$ type of hydrogen bond connectivity (infinite single tape structure). This zigzag tape is crossed (intersected) to type 2 ribbon (or type 3 as suggested by Ahn *et al* [3]), represented by pink line in Figure 4.2 (b), where the $R_2^2(8)$ hydrogen bonding arrangement between adjacent molecules involves two N-H bonds and with only one C=S bond of each TTCA molecule.

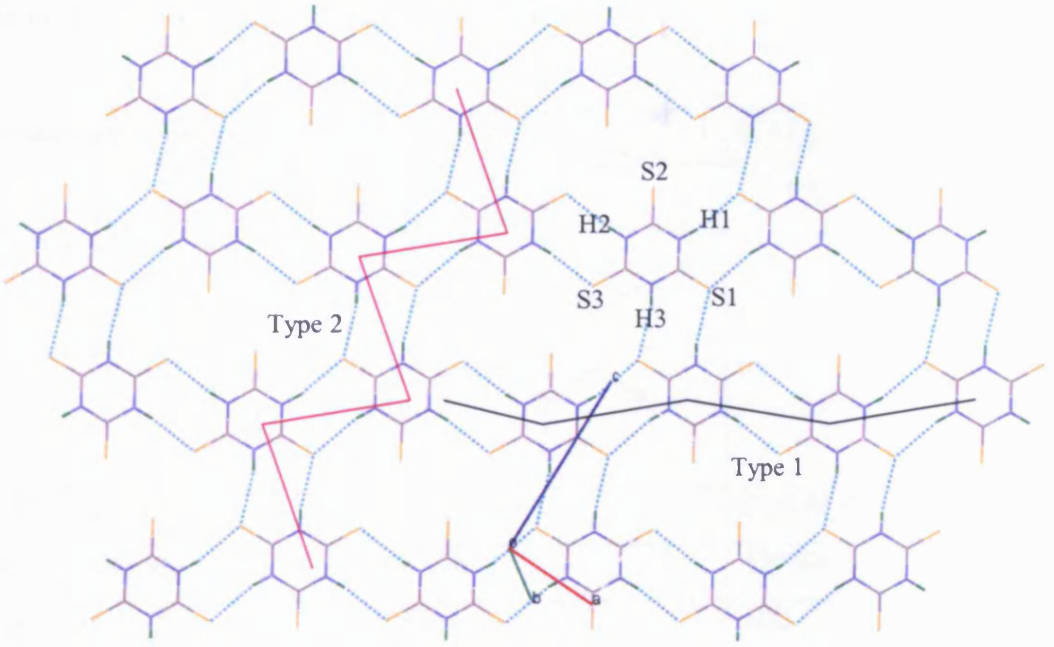
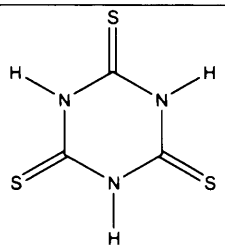
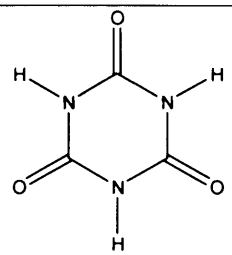


Figure 4.2 : TTCA structure packing and ribbons. (a) Sulfur is yellow, Hydrogen is green, Nitrogen is blue and Carbon is grey (b) Types of ribbons in TTCA; Type 1 – black and Type 2 – pink.

Cyanuric acid (CA) molecule structure is similar to a TTCA structure. The difference is just between oxygen and sulfur atom in relative molecule. As mentioned by Guo and Harris *et al* [5], the TTCA structure shows an important contrast to the crystal structure of CA. Detailed crystallographic data comparison is shown in Table 4.1.

Table 4.1: Crystallographic data of CA and TTCA

Crystallographic Data	**TTCA[5]**	**CA[6]**
CSD	CEHQEM	CYURAC05
Structure		
Empirical Formula	$C_3H_3N_3S_3$	$C_3H_3N_3O_3$
Formula Weight	177.26	129.07
Temperature (K)	150	100
Crystal System	Triclinic	Monoclinic
Space Group	P1	C2/n
Z	2	4
Unit Cell Dimension		
a (Å)	5.857	7.749
b (Å)	7.048	6.736
c (Å)	8.800	11.912
α (°)	102.99	90.00
β (°)	92.87	130.69
γ (°)	110.47	90.00
Volume	328.282	471.46
Density Calculated (g cm^{-3})	1.806	1.818

Cyanuric acid and its supramolecular crystal structure have been discussed experimentally in many studies [7-9]. CA as a crystal molecule is very prone to form hydrogen bond to other molecules. Studies of the CA interaction and its hydrogen bonding functionality to other molecules in co-crystals and to itself have been quite an interest to many people [3,10-14]. Pedireddi *et al* reported from his investigations that CA forms strong cyclic N-H···O hydrogen bond 'tapes' between the adjacent molecules which arranged in two dimension forming layered structures [13]. He

reported in complex crystalisation of CA with dimethylsulfoxide (DMSO), dimethylamine (DMA) and dimethylformamide (DMF), the arrangement of CA tapes were structured so it can accommodate the guest molecules mentioned earlier. CA structure also has similarity as TTCA structure (Figure 4.1(b)).

Cyanuric acid (CYURA05 in CSD code) crystal has 4 molecules in a unit cell. It exists in monoclinic crystal system. CA crystal structures are build of hydrogen bonded layers of molecules. Hydrogen bond sheets are located on *bc* plane. On each molecule, all oxygen and hydrogen atoms are involved in hydrogen bonding, one at each atom. Two types of hydrogen bonding can be seen in this crystal structure, (Figure 4.3) where first type is a zigzag $R_2^2(8)$ tape similar to type 1 in TTCA structure as mentioned previously. Type 2 of CA hydrogen bonding is a chain type which intersected the type 1 ribbon. These are of $C_1^1(6)$ arrangement with one hydrogen donor and a hydrogen acceptor of oxygen atom.

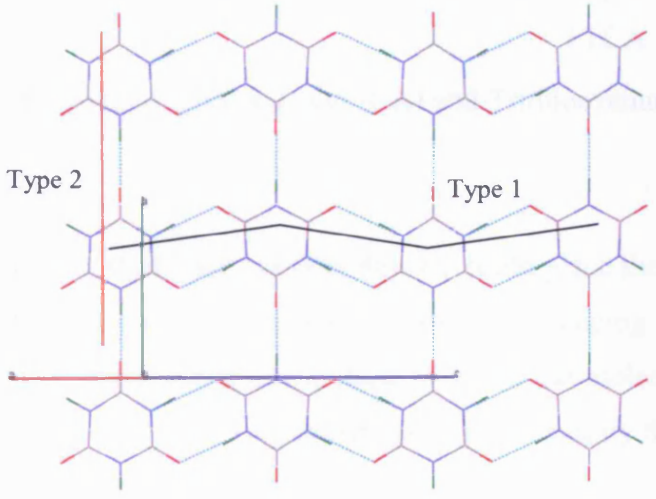


Figure 4.3: Cyanuric acid hydrogen bond sheet perpendicular to *bc* axis. Type of hydrogen bonds: (i) Type 1 – black, (ii) Type 2 – orange.

A study by Martsinovich and Kantorovich [15] reported that this planar hydrogen bond geometry structures of both CA and TTCA are influenced by the electronegative properties of the atom involved (oxygen and sulfur) and thus, affects the order of

stabilities of the two dimensional networks for both structures. In this 2D network structure study, they shown that N-H...O bonds are stronger and have the energy of ~ -0.27 eV for a CA dimer, while for N-H...S bonds, it is ~ -0.18eV.

4.2 Objective and Computational Method

My next area of study is to calculate the crystal lattice energy of trithiocyanuric acid (TTCA) and cyanuric acid (CA). The intention would then be to predict the possibility of TTCA adopting the crystal structure of CA and vice versa (Figure 4.4).

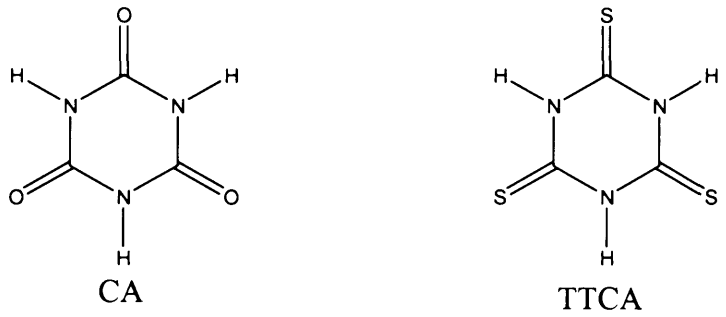


Figure 4.4: Structure of Cyanuric acid (CA) and Trithiocyanuric acid (TTCA)

The motivation for this study was the idea that we could place the TTCA molecule in the CA crystal structure (and vice versa) by simply exchanging S or O atoms. This will allow us to understand why the two, apparently similar molecules adopt different hydrogen bonding networks in the solid state and the possibility that both compounds might be capable of exhibiting polymorphism.

The S-S pair potential required for TTCA is not listed in the current pair potential set available with DMAREL. Due to this matter, the first step in this study is to find the best sulfur potential fitting that can fit well with the DMAREL program. The sulfur potential found in the literature is in the Lennard-Jones (LJ) form and this will be recalculated to fit the parameters to a Buckingham potentials and then refine to fit α-S_8, thiourea and several molecules with thiourea functional groups. Several

parameters were taken into account such as the scale factor (for adjusting the multipoles used from GDMA), min E (minimum energy) and min R (minimum R distance) in the Buckingham parameters of S···S potential and its reliability from the value of the discrepancy index of the relaxed structure. The assessment of the quality of a potential during refinement of sulfur potentials will focus on the discrepancy index produced after the relaxation.

Finally, the S in TTCA crystal structure will be changed to O atom and vice versa for O in CA, to see the effect of interchanging these crystal structures packing.

4.3 Initial Study of Sulfur Potential

In our previous study using DMAREL, the pair potential for sulfur is not available due to the molecules studied in the parameterisation reference papers are not related to sulfur atom interaction. Sulfur potential from Lennard Jones (LJ) equation was then taken from the literature [16] of isotropic sulfur intermolecular potential model in α-S_8 crystals. The repulsive contribution of LJ equation usually is written in terms of the well depth ε and the distance of the closest approach σ as follows:

$$U_{ij} = 4\varepsilon\left(\left(\frac{\sigma}{R_{ij}}\right)^{12} - \left(\frac{\sigma}{R_{ij}}\right)^{6}\right) \qquad (4.1)$$

where U_{ij} is the interaction energy for atoms i and j at separation R_{ij}.

Initial value of isotropic intermolecular potential of S···S atoms are σ = 3.39Å and ε = 1.70 kJ mol^{-1}. The set of potentials used in DMAREL are in the Buckingham form. We then produced it as a new set of sulfur parameters by using *Mathematica* to fit the Buckingham potential from LJ potential of sulfur. Buckingham potential can be written in equation 4.2:

$$U_{ij} = A\exp\left(-\frac{R_{ij}}{\rho}\right) - \frac{C}{R_{ij}^{6}} \qquad (4.2)$$

The optimised value of the fitting gave A = 216310 kJ mol^{-1}, ρ = 0.32993 Å and C = 11327.7 kJ mol^{-1} $Å^6$. Plots in Figure 4.5 shows the fit of the Buckingham model to the LJ potential.

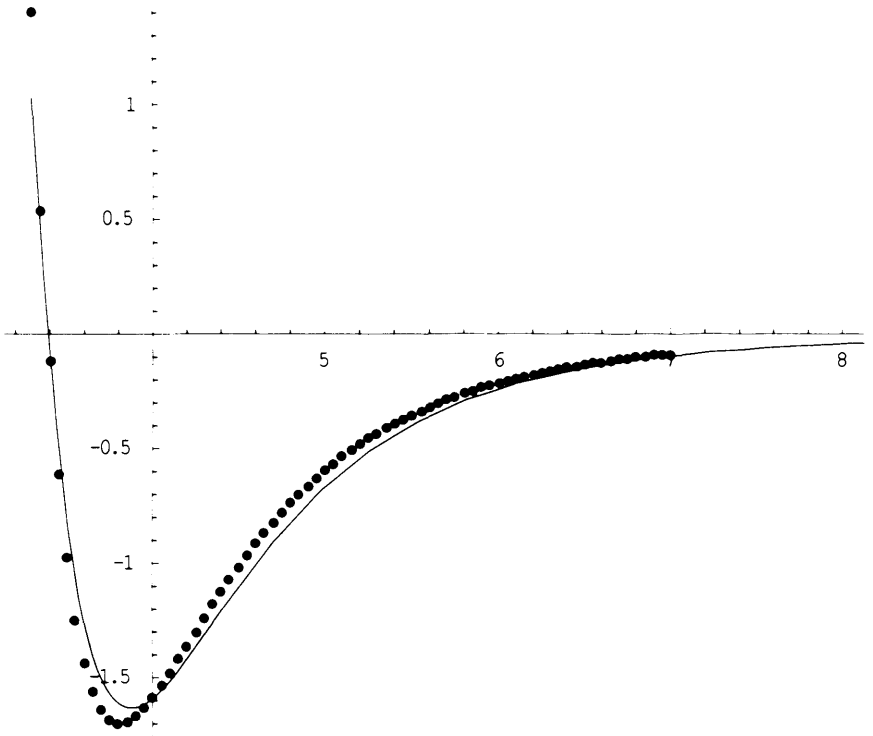


Figure 4.5: Comparison of initial LJ (····) fitting and optimised fit of Buckingham (—) potential.

However, in DMAREL, the Buckingham equation used are as below (Equation 4.3), which consist of R (Å), ε_{lk} (kJ mol^{-1}) and λ (no unit) for combination rules with other atoms [17]. The fitted A, B and C parameters give $R_{\circ,lk}$ = 4.4541 Å, ε_{lk} = 0.3707 kJ mol^{-1}, and a constant value, λ_{lk} = 13.5.

$$U_{ij} = \frac{\varepsilon_{lk}}{\lambda_{lk} - 6}\left[6\exp(\lambda_{lk})\exp\left(-\lambda_{lk}\frac{R_{ij}}{R_{\circ,lk}}\right) - \lambda_{lk}\left(\frac{R_{\circ lk}}{R_{ij}}\right)^6\right] \qquad (4.3)$$

In which atoms *i* and *j* are of type *l* and *k*.

Table 4.2: List of pair-pair potential (min R, min E and λ) based on equation 4.3.

Atom	Min R (Å)	Min E (kJ mol $^{-1}$)	λ (no unit)
O - O	3.6096	0.3347	13.5
C – C	3.8905	0.3875	13.5
H – H (non-polar)	3.3666	0.0414	13.5
H – H (polar)	3.3666	0.0414	13.5
N – N	3.6986	0.6260	13.5
S – S	4.4541	0.3707	13.5

The above potential (Table 4.2) then applied to α-S_8, thiourea and TTCA giving the relaxation energy as shown in Table 4.3.

Table 4.3: Relaxation energy of α-S_8, thiourea and TTCA based on the original potential from literature.

	α-S_8 (orthorhombic)	**Thiourea**	**TTCA**
Sublimation E (kJ mol^{-1})	106.13[18]	112.0[19]	n/a
Initial E (kJ mol^{-1})	-17.36	-96.09	-53.73
Relaxed E (kJ mol^{-1})	-32.09	-103.68	-78.53
Discrepancy Index	192.79	103.48	758.62
Parameter change (%)	a: 8.31 b: 3.39 c: 8.18 α: 0.00 β: 0.02 γ: 0.00	a: -9.64 b: 0.16 c: -0.85 no change on angle during relaxation	a: 20.22 b: -3.61 c: 2.72 α: -13.72 β: 3.45 γ: 8.17
Volume change (%)	21.14	-10.26	13.91
RMS	7.01	5.6	11.96

In α-S_8, the lattice energy calculated was less then the experimental sublimation energy. Difference in initial and relaxed energies calculated in DMAREL show quite a big difference. This is not preferable since the best result is to have initial energy and relaxed energy as less difference as possible. The relaxation also produced 8% of change in *a* and *c* cell vectors and 3% for *b* axis. However, there are no changes in angle parameter. Due to the parameter change, the volume of α-S_8 was also changed by 21%. Figure 4.6 shows the difference of α-S_8 structure before and after relaxation in relaxed lattice cell. As we can see, there is no hydrogen bonding involved in the α-S_8 structure packing.

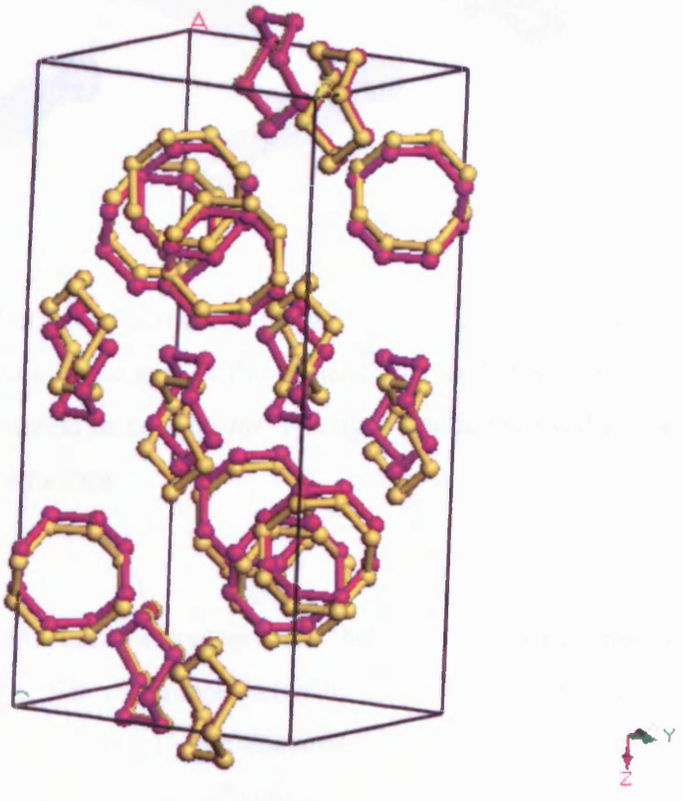


Figure 4.6: View of α-S_8 structure before (yellow) and after (pink) relaxation along *b* axis.

For thiourea structure (Figure 4.7), the relaxed energy obtained was nearly similar to the sublimation energy, which is -103.68 kJ mol^{-1}. The relaxation impacted change in *a* axis at -9.6%, while *b* and *c* axis changed in a good percentage, which is less then 1%. Due to the change in the parameter length, the volume of the relaxed structure

also contracted 10% from its initial volume. This also resulted the discrepancy index at 103.5, which is significantly large.

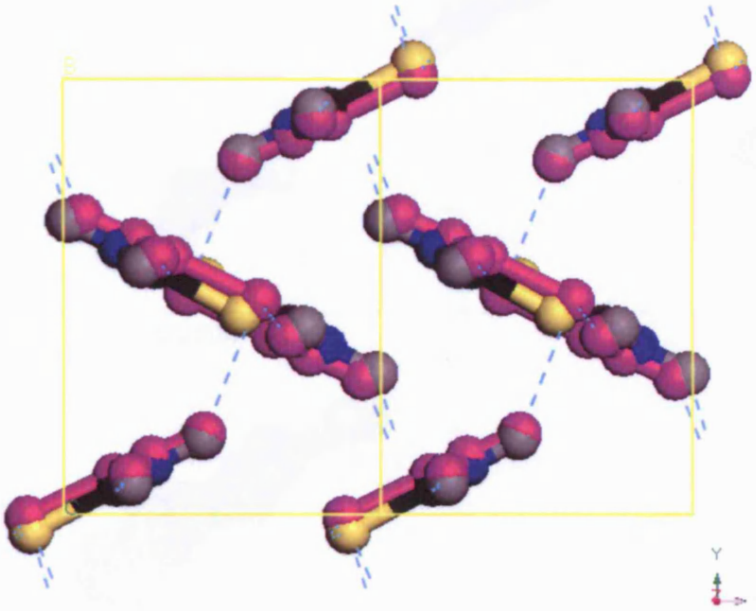


Figure 4.7: View of thiourea structure in two unit cell before and after (pink) relaxation along *c* axis in the initial lattice cell. Hydrogen bonds were also illustrated to note on the hydrogen bonds involved in the thiourea packing structure.

TTCA on the other hand, suffers the most from the use of this sulfur potential fit as illustrated in Figure 4.8. The difference between the initial energy and its relaxed structure energy is about 25 kJ mol^{-1} and discrepancy index at 758.62. This is due to the large volume change at *a* axis (20.22%) followed by angle change of α angle with -13% change.

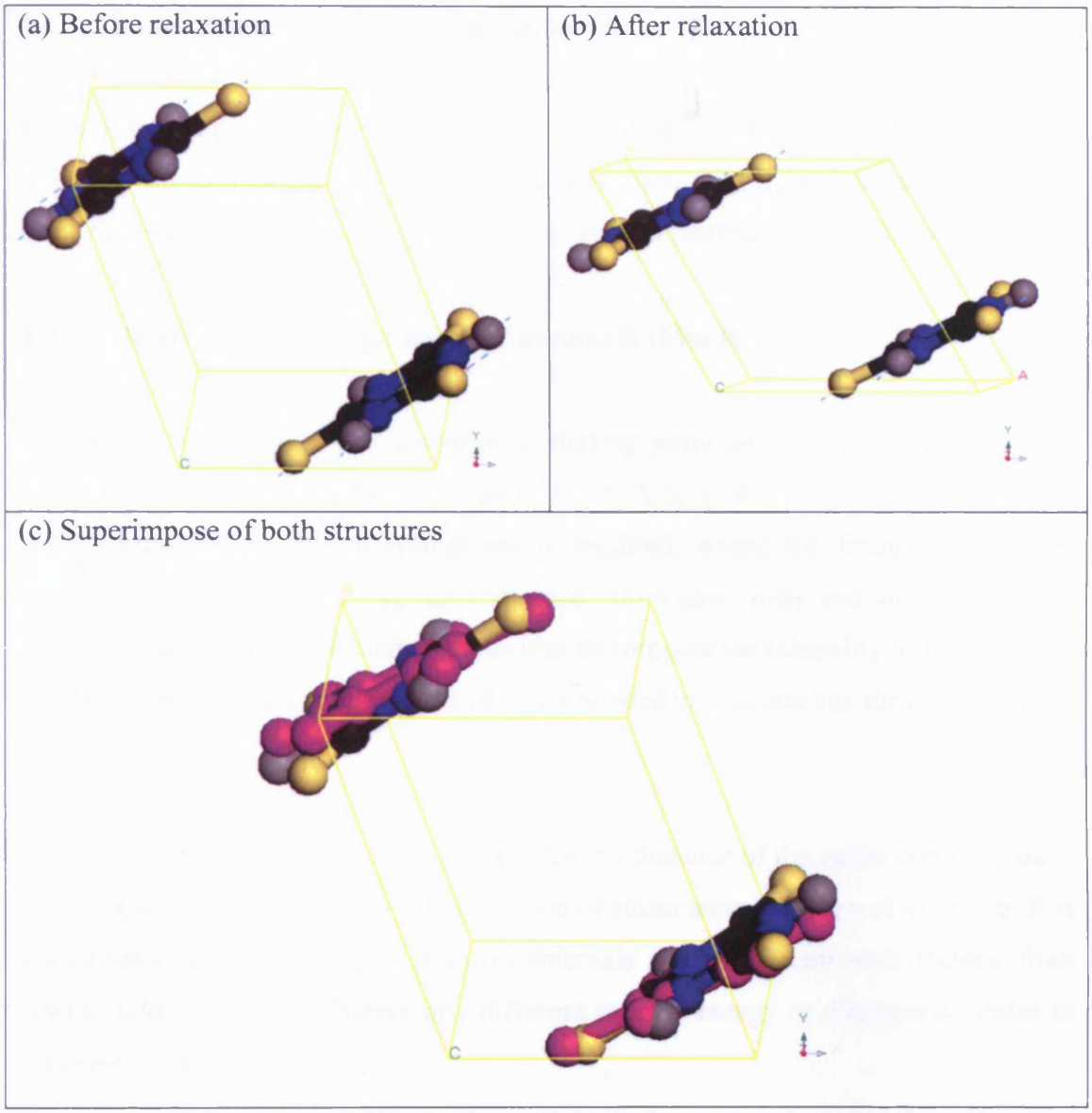


Figure 4.8: Structure of TTCA lattice cell (a) before relaxation, (b) after relaxation and (c) superimpose of both structure with pink structure being the relaxed structure.

The results using the original potential based on the paper by Pastorino [16] clearly shows that further modification should be done to refine the potential to be fitted for use in DMAREL for crystal relaxation.

4.4 Sulfur Potential Test: α-S_8 & Thiourea

From this subtopic, we will begin training the initial S···S potential to fit the potential for use with DMAREL application. This was done to α-S_8 and thiourea molecule since the experimental sublimation energy of both crystal structures is known.

4.4.1 Determination of Optimum Minimum R (Min R)

The potential in Figure 4.5 provides a starting point of S···S potential based on literature calculation of α-S_8 [16]. However, the S species in TTCA are involved in polar bonds and so further refinement is required, where the fitting needs to be optimised and tested with sulfur contained molecules. α-S_8 and thiourea crystal structures are chosen as the initial molecules to compare the reliability of the potential and because thiourea structure is a hydrogen-bonded crystal and has similar functional groups to TTCA.

We tested the minimum R potential for optimum distance of the sulfur potential and ε which describes the van der Waals repulsion of sulfur atom. A range of minimum R is varied between 4.09 Å to 5.09 Å in 0.1 intervals and at different scale factors; from 0.90 to 0.96. This is to observe any different relaxed energy or discrepancy index at different electrostatic levels.

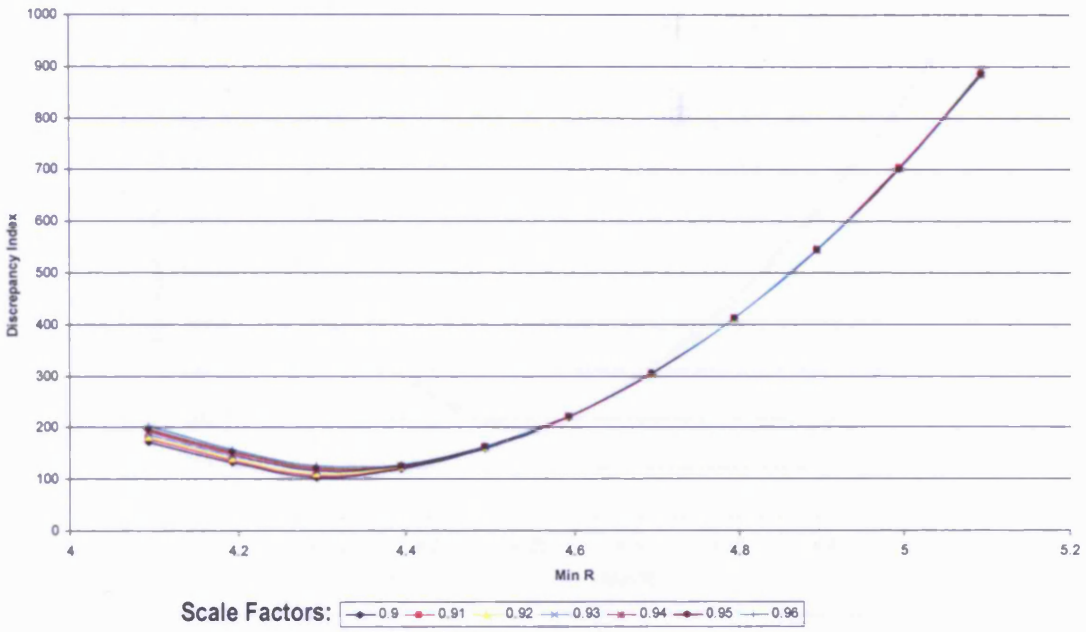


Figure 4.9: Discrepancy index of α-S_8 at different scale factors and min R values

One of the important parameters in DMAREL is the scale factor. This is required as a 'fudge' factor to account for the neglect of electron correlation inherent in an SCF wavefunction [20] in atomic multipoles. Based on Figure 4.9, it was found that scale factor did not affect much of the calculation for both α-S_8 and thiourea. We then continued to calculate at smaller range of min R (0.02 intervals) and at scale factor 0.96 to find the average optimum value of minimum R for both structures. In α-S_8 crystal structures, it was shown a sharp value of discrepancy index for min R at scale factor 0.96, as shown in Figure 4.10.

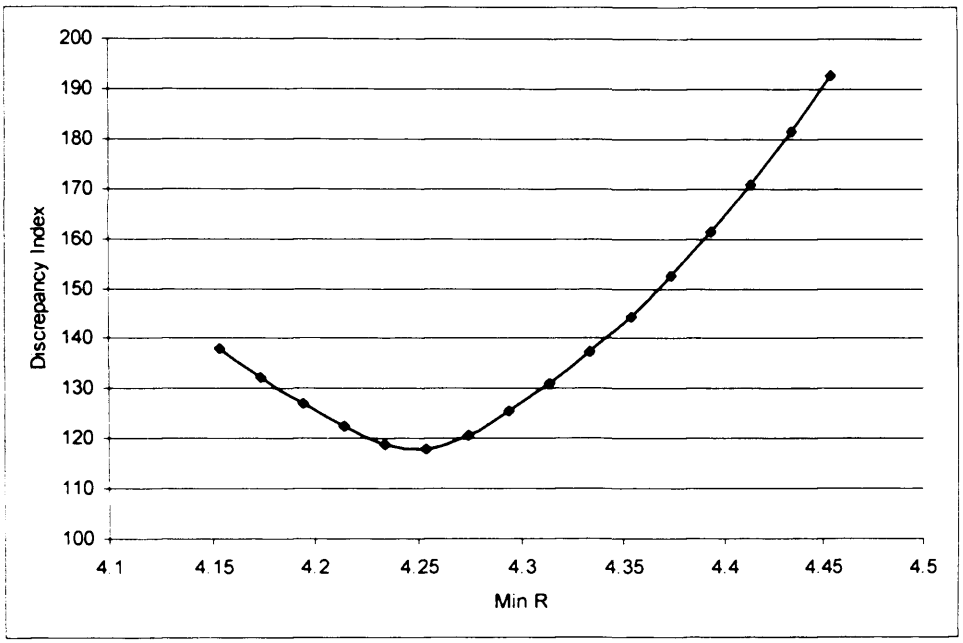


Figure 4.10: Discrepancy index of α-S_8 in smaller range of Min R (Å)

Optimum minimum R for both α-S_8 and thiourea structures are 4.2541Å. This value will be use to calculate min E of S···S atom potential.

4.4.2 Determination of Optimum Minimum E (Min E)

S-S atom potentials were then tested for minimum E with the range of 0.3275 to 1.2075 kJ mol^{-1} in 0.01 intervals. This was a very long range of calculation compared to min R search previously because we seem to have difficulty in finding the minimum. The discrepancy index tends to decrease gradually for α-S_8 and increasing gradually for thiourea. The first attempt was carried out based on the initial min R (4.4541Å) which is obtained from the literature [16] and followed by using the new min R (4.2541Å) obtained previously.

With this new min R and the same range of min E mentioned previously, the discrepancy index for α-S_8 decrease gradually, but no significant minimum was obtained (Figure 4.11). However, this has made about 50% of change for α-S_8 discrepancy index.

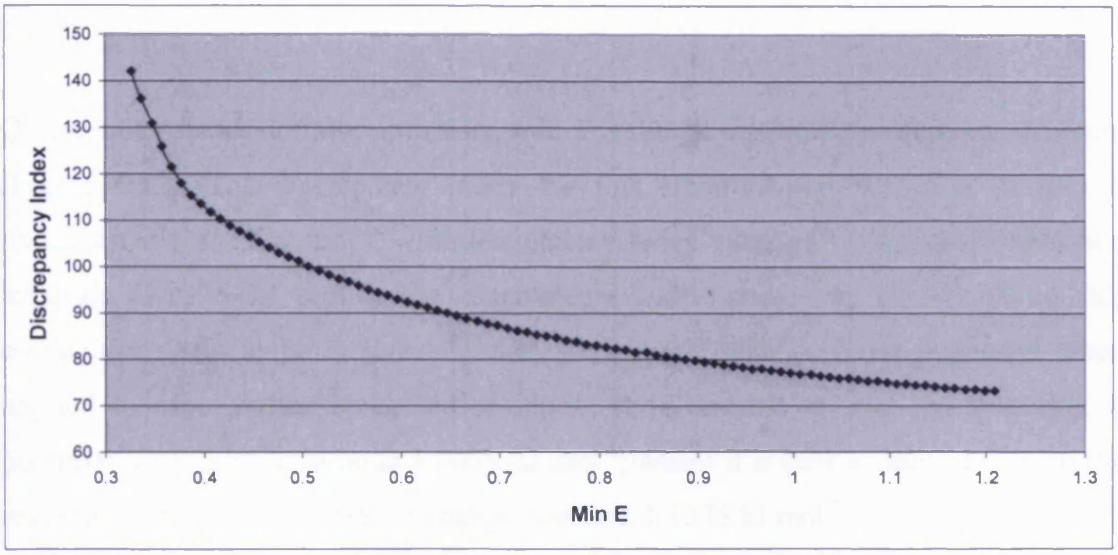


Figure 4.11: Discrepancy index for α-S_8 with different min E ranges and min R at 4.2541Å

Besides the change in discrepancy index, difference in initial and final lattice energy was also observed. Figure 4.12 shows the difference of both lattice energies and within the range of min E calculated, it seems that the lattice energy was gradually decreasing; converging nearly to the sublimation energy found in the literature (106.13 kJ mol^{-1}), where when min E is 1.2075 kJ mol^{-1}, its initial energy is -82.20 kJ mol^{-1} and final lattice energy is -95.86 kJ mol^{-1}.

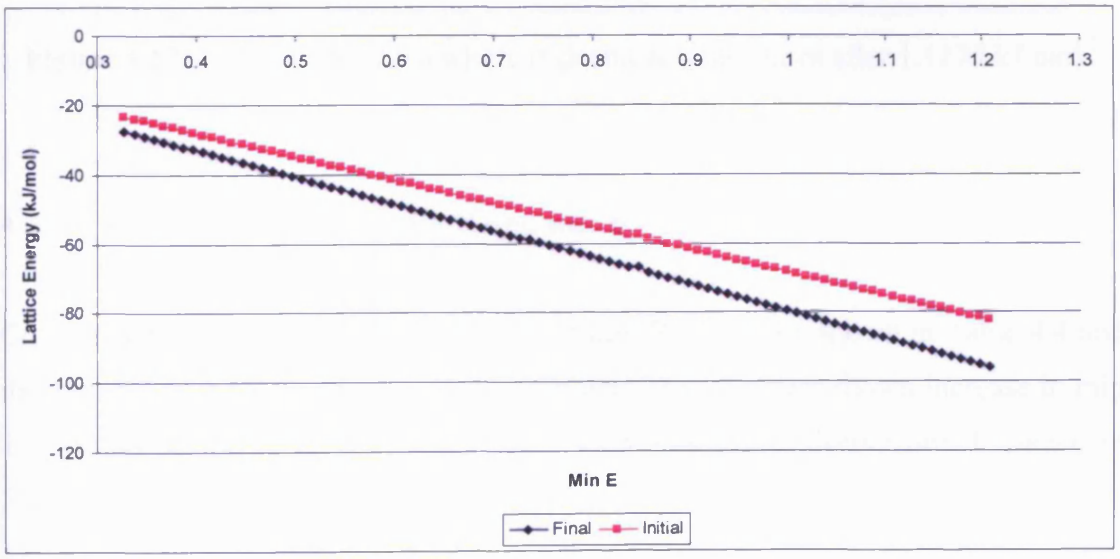


Figure 4.12: Lattice Energy Minimisation of α-S_8 with different min E ranges and min R at 4.2541Å

On the other hand, definite optimised min E value is observed for thiourea structure (Figure 4.13). The discrepancy index for this structure was found at its min E minimum at 1.1175 kJ mol^{-1}, with discrepancy index value of 75. However, after two intervals (1.1375 kJ mol^{-1}), the discrepancy index shoots up to 247. With this consequence and to be at the safe side, avoiding similar situation happened when applied to other sulfur contained structure, it is decided to take the best min E potential for S-S atom to be at 1.1075 kJ mol^{-1}, where it is 0.01 kJ mol^{-1} less than the real minimum due to the drastic change found at 1.1375 kJ mol^{-1}.

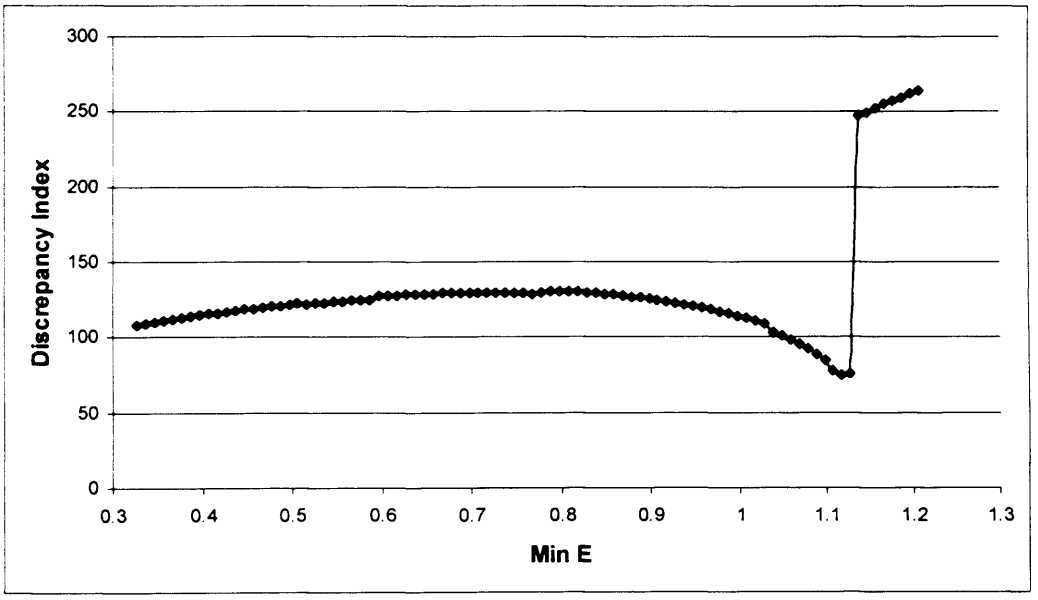


Figure 4.13: Min E of thiourea where it drastically decreases after 1.1175 kJ mol^{-1}

4.4.3 Potential Fitting for min R and min E

Comparison of the initial and final value of min R and E are shown in Table 4.4 and its potential plot in Figure 4.14. The new potential parameters shown increase in min E of triple compared to the initial value. This produced a deeper min E shown in Figure 4.14.

Table 4.4: New value of S···S potential

	Initial	**Final**
Min E (kJ mol^{-1})	0.3707	1.1075
Min R (Å)	4.4541	4.2541
λ	13.5	13.5

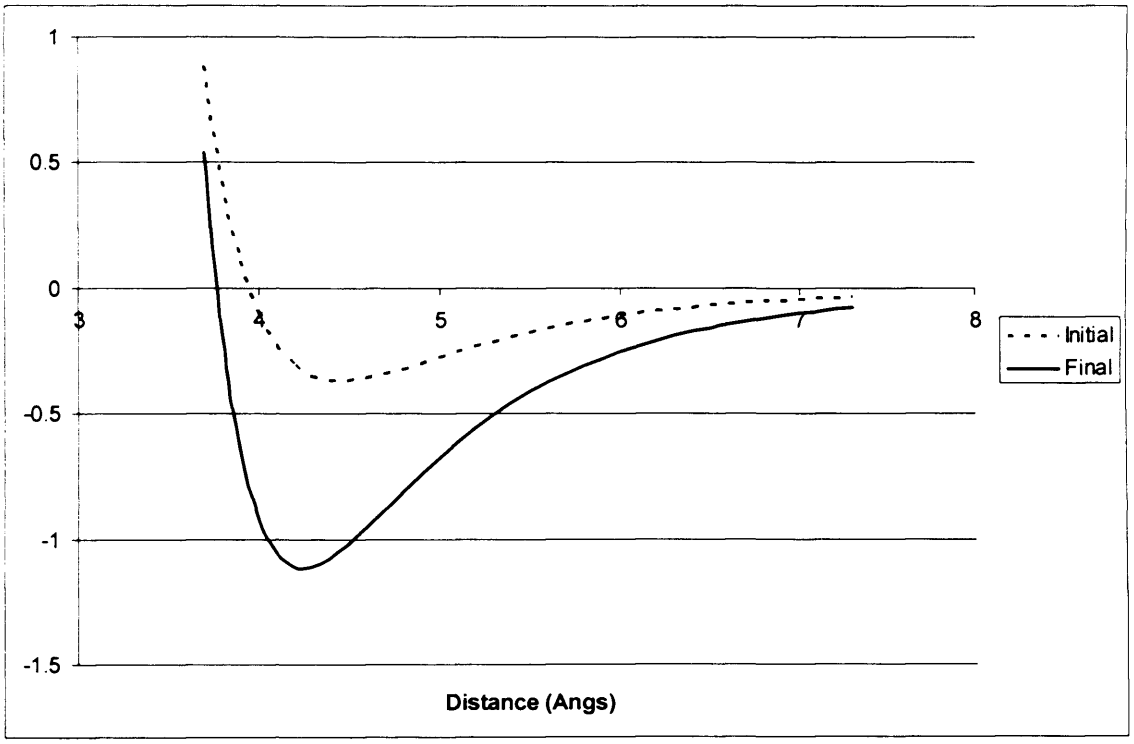


Figure 4.14 : New fitting of optimised Buckingham potential of sulfur atom.

4.4.4 Final min R Fitting

Based on the fitting results previously (Table 4.4), thiourea shown a significant different of discrepancy index as min E increases. However, for α-S_8, discrepancy index and lattice energy keep on decreasing and shows no notable difference. This causes us to do another potential fitting with the new min E (1.1075 kJ mol^{-1}) varying in different min R at 4.2541 ± 1 Å at 0.1 intervals.

Within the range of min R from 3.2541 to 5.2541 Å, the results for min R test for both molecules are discussed. For α-S_8, the initial and final lattice energy calculated started

to converge and have less difference to each other as min R moves towards 4.0541 Å (Figure 4.15). Initial lattice energy calculated from min R = 4.0541 Å is 82.30 kJ mol^{-1} and final at -84.61 kJ mol^{-1}. However, at min R higher than 4.4541 Å, the difference started to deviate again. This shows that the best value of min R is between 3.4541 and 4.4541 Å.

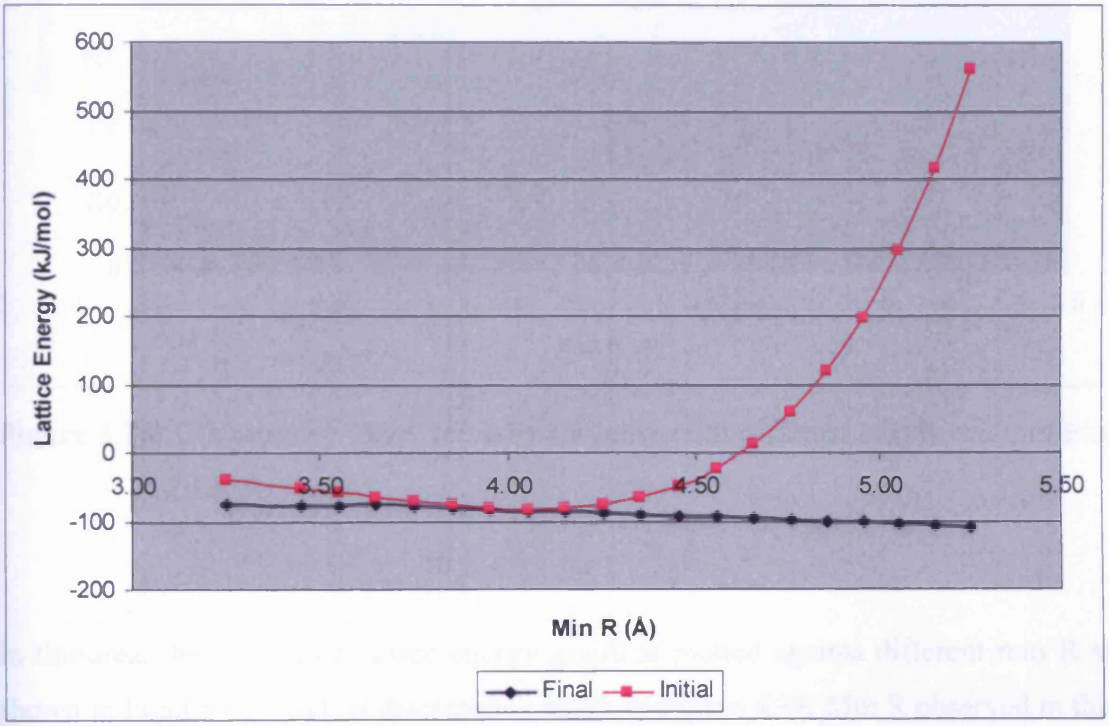


Figure 4.15: Lattice energy minimisation of initial and final α-S_8 structure with different min R and min E at 1.1075 kJ mol^{-1}

The discrepancy index of the above relaxed structure was analysed and it shows that at the minimum difference of initial and final lattice energy, the discrepancy index is also at its lowest value, 47 (Figure 4.16). This is a significantly improved value where the discrepancy index decreased significantly from its initial value (193) and now is less than 100. Discrepancy index for min R = 3.5541 Å is excluded from the graph because it produce a very large discrepancy index (3268). At this point, the discrepancy index was contributed by high total rotational (138.4) and translational (425.1) displacement of α-S_8 and also mostly by change in total cell length (2704.9).

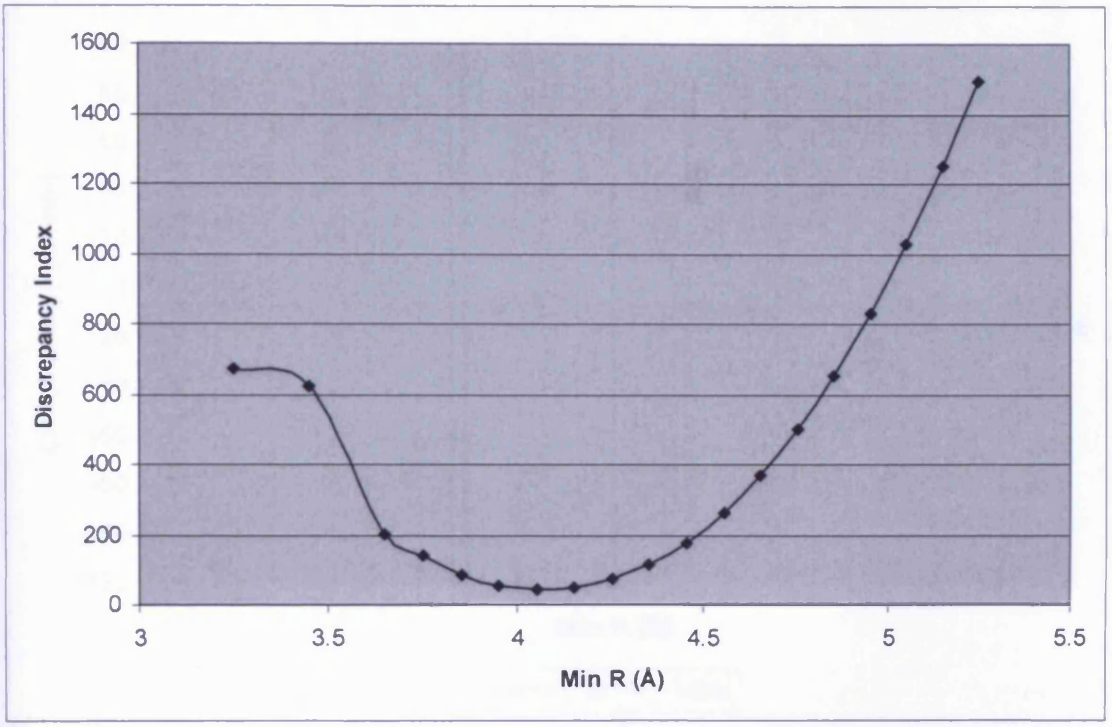


Figure 4.16: Discrepancy index for α-S_8 structure with different min R and min E at 1.1075 kJ mol^{-1}

In thiourea, the calculated lattice energy graph is plotted against different min R as shown in Figure 4.17 and its discrepancy index in Figure 4.18. Min R observed in this test is still at its best discrepancy index, which is the previous min R (4.2541Å). However, the difference between initial and final lattice energy is higher than the potential at atom-atom distance less than 4.1541Å. Apart from that, the discrepancy index after this point (4.2541Å) increases at a significant value, which is 200 in difference. To avoid this potential difference that might happened in relaxing other sulfur contained structure, it is better for us to choose a stable discrepancy index and small difference of initial and final lattice energy. This can be observed at point between 3.7541 to 4.0541 Å min R where the discrepancy indexes are stable, between 97 to 99 and a considerably acceptable lattice energy difference.

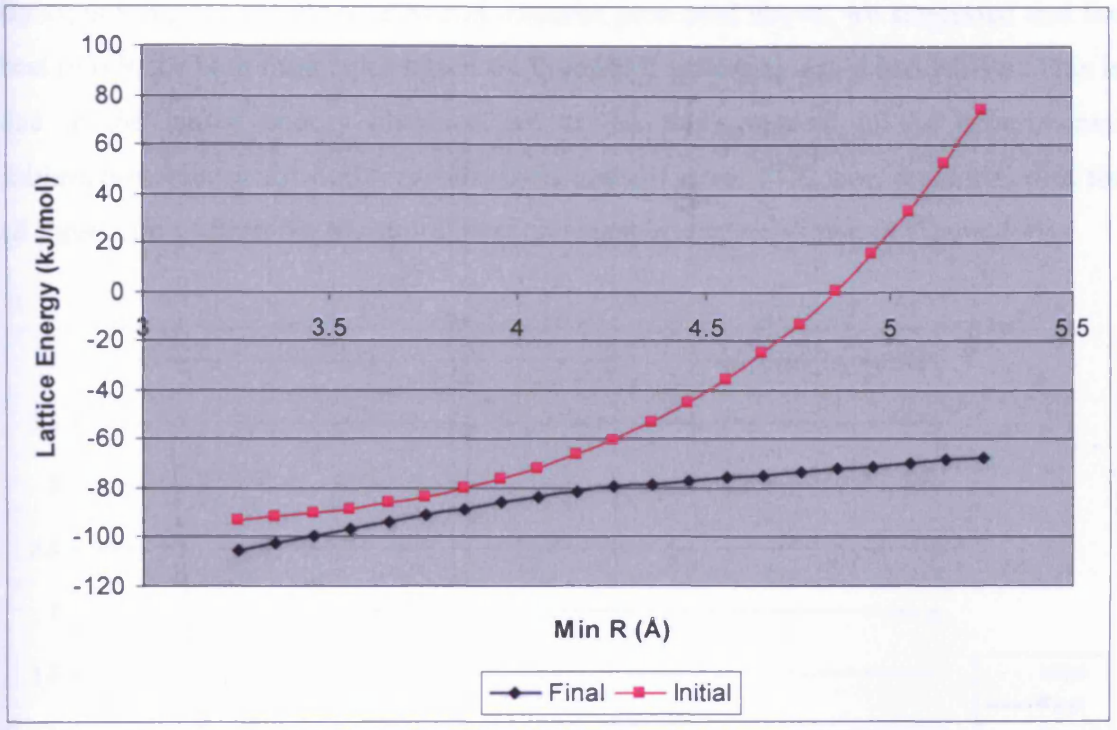


Figure 4.17: Lattice energy minimisation of initial and final thiourea structure with different min R and min E at 1.1075

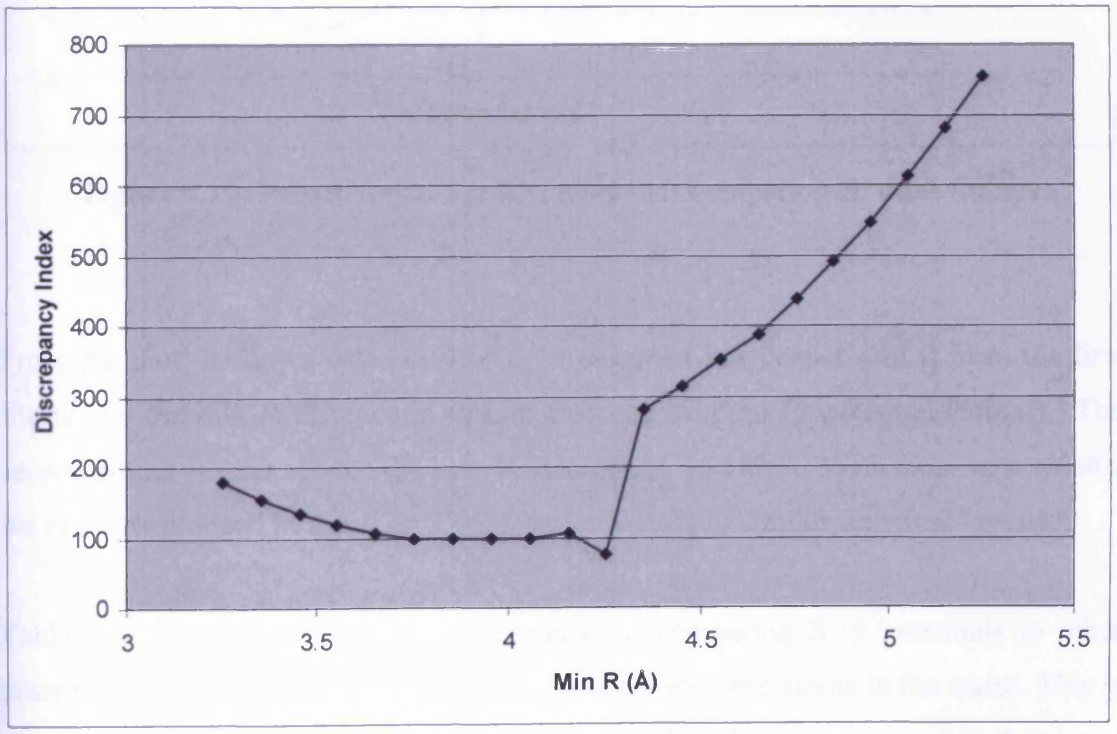


Figure 4.18: Discrepancy index for thiourea structure with different min R and min E at 1.1075 kJ mol^{-1}

By combining the results of α-S_8 and thiourea presented above, we suggested that the best min R for both molecules based on the min R potential test is at 3.9541Å. This is due to the lattice energy obtained are at the point nearest to the experimental sublimation energy for both crystals; α-S_8 and thiourea. S···S atom potential plot for all three fitting, from the literature, first and second test are shown in Figure 4.19.

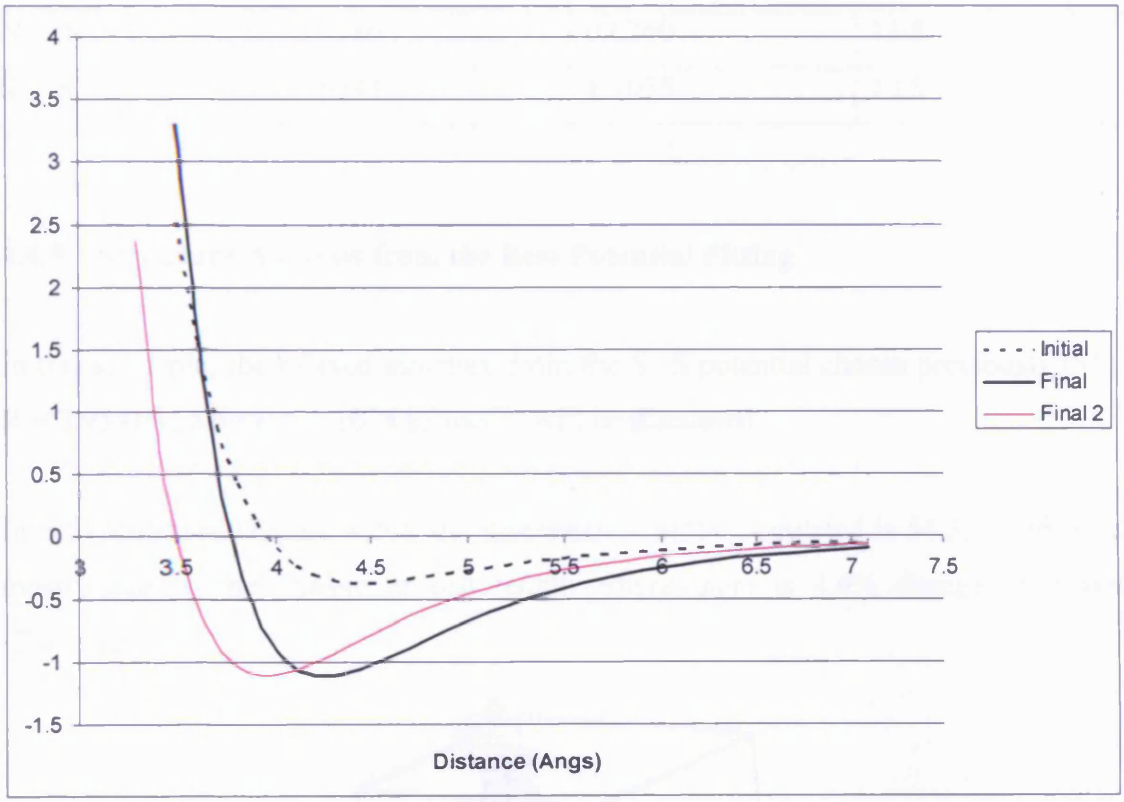


Figure 4.19: Potential plot for S···S atom pair comparing all three fittings

From the plot, it shows that our new S···S potential has deeper min E from the first fitting and the min R distance is shorter than the original LJ potential (Initial). The second fitting further reduce the min R distance to 3.9541 Å. With these new results, we can now proceed to test it on TTCA and molecule to similar functional group.

Table 4.5 shows list of the new potentials. In comparing S···S potentials to other potentials, min E of sulfur atom is rather bigger then other atoms in the series. This is due to the large size of sulfur atom and its high polarisibilities compared to the rest of the atoms. Thus, this value of S···S potential is quite reasonable.

Table 4.5: New pair-pair potential fitting for Min R and Min E

Atom	Min R (Å)	Min E (kJ mol^{-1})	λ (no unit)
O ··· O	3.6096	0.3347	13.5
C ··· C	3.8905	0.3875	13.5
H ··· H (non-polar)	3.3666	0.0414	13.5
H ··· H (polar)	3.3666	0.0414	13.5
N ··· N	3.6986	0.6260	13.5
S ··· S	3.9541	1.1075	13.5

4.4.5 Structure Analysis from the Best Potential Fitting

In this subtopic, the relaxed structure from the S···S potential chosen previously (Min R = 3.9541Å, Min E = 1.1075 kJ mol^{-1}) will be discussed.

In α-S_8 structure (Figure 4.20), the discrepancy index calculated is 54.5, contributed mostly due to the change in cell length where there is 4.6% change in *c* axis (Table 4.6).

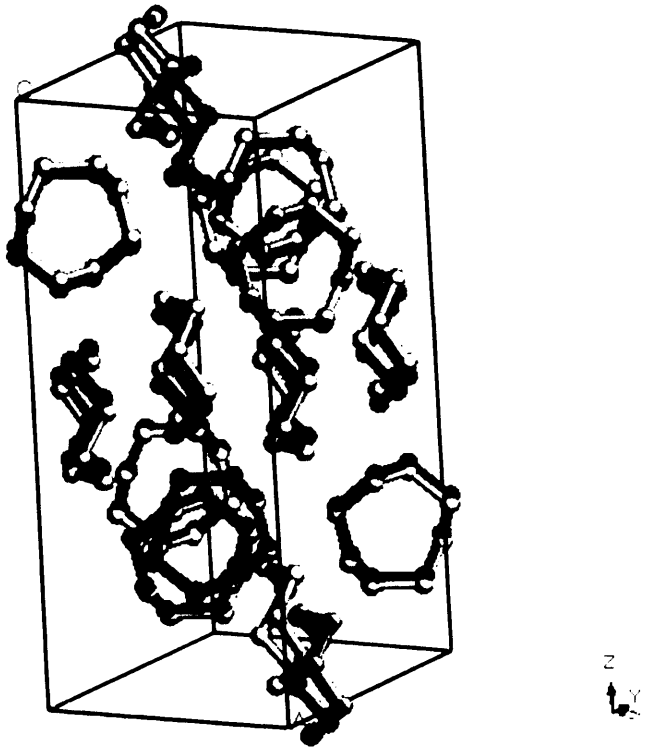


Figure 4.20: View of α-S_8 structure before (yellow) and after (pink) relaxation

Lattice vectors, volume and density information of best S···S potential for α-S_8 are shown in Table 4.6. From the table, it is clearly seen that the minimisation does not give any significant effect towards cell angles which keep its orthorhombic structure. However, due to the change in cell length, the cell volume is increased with 2.6% and decreases the density by 2.5%. Discrepancy index from this minimisation is 54 (Table 4.7). This is a good value compared to its higher initial result, 193. Lattice energy obtained for α-S_8 from this calculation is α-S_8 -82.55 kJ mol^{-1}, differs from experimental by about 23.58 kJ mol^{-1} (~5.6 kcal mol^{-1}) [Experimental value at 106.13 kJ mol^{-1}]. This lattice energy value has significantly improved. However, it is still far from the experimental value.

Table 4.6: Detailed lattice vectors, volume and density information of α-S_8 structure at min R minimum.

Lattice vectors (Angstroms and Degrees)

	a	b	c	alpha	beta	gamma
Initial	10.46	12.87	24.49	90.00	90.00	90.00
Final	10.58	12.48	25.62	90.00	90.01	90.00
diff.(abs)	0.11	-0.38	1.13	0.00	0.01	0.00
diff.(%)	1.08	-2.98	4.62	0.00	0.02	0.00

Cell volume and cell density

	Cell Volume ($Å^3$)	**Cell density (g cm^{-3})**
Initial	3296.74	2.08
Final	3382.39	2.03
diff.(abs)	85.65	-0.05
diff.(%)	2.60	-2.53

For thiourea, improvement on the discrepancy index is also observed (Table 4.7). The difference between initial and final relaxed structure is shown in Figure 4.21. In the relaxed structure, the hydrogen bond of thiourea sheet (C=S···H) stretched apart from 2.42 Å to 2.67 Å. This is also happened to the hydrogen bond between the thiourea sheet. However, the changes is relatively small as opposed to the previous mentioned,

which is from 2.75 Å to 2.78 Å. This shows that the current potential produced stronger repulsion within the hydrogen bond from the experimental structure, though it has improved quite a lot from the initial potential.

Table 4.7: Comparison of α-S_8 and thiourea properties between its initial and final S-S potentials.

	α-S_8		**Thiourea**	
S-S Potential	**Initial**	**Final**	**Initial**	**Final**
Sublimation E (kJ mol^{-1})	106.13[18]		112.0[19]	
Initial E (kJ mol^{-1})	-17.36	-80.49	-96.09	-76.28
Relaxed E (kJ mol^{-1})	-32.09	-82.55	-103.68	-85.88
Discrepancy Index	192.79	54.47	103.48	98.60
Volume change (%)	21.14	-1.75	-10.26	6.77
RMS	7.0	3.6	5.6	5.3

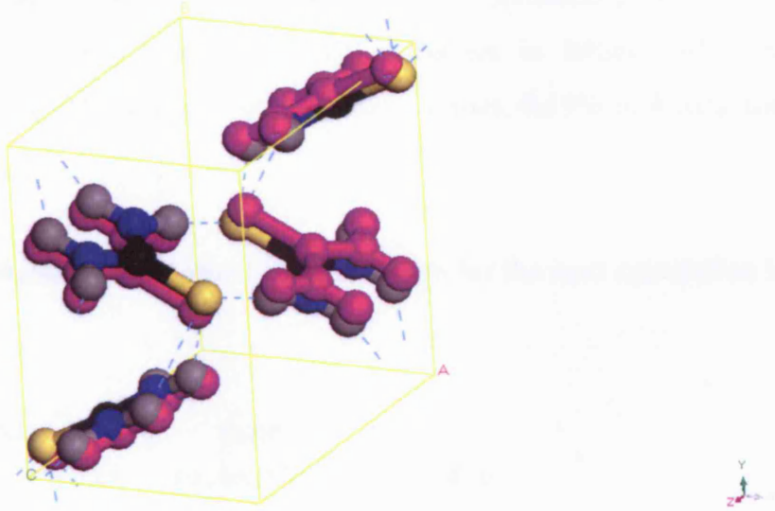


Figure 4.21: Thiourea structure of before and after (pink) relaxation. Hydrogen bond connectivity of the initial structure is also illustrated.

The S···S potentials were then applied to TTCA to observe the effect of this new pair potential in its crystal system. From the calculation, the data obtained do not give any significant changes compared to the earlier potential parameters. The discrepancy index obtained is 749. It is suggested that this might be due to S···HP pair potential, which is the potential between sulfur and hydrogen atoms involved in the hydrogen

bonding within the crystal. In the next subtopic, a calculation varying S···HP pair potential was done to the thiourea molecule to confirm this hypothesis.

4.5 CONDOR Fitting S···HP Potential in Thiourea

Calculations was done by varying the A and C parameters in Buckingham potential of S···HP interaction which is important in the thiourea hydrogen bonding interaction. The A and C parameters of the Buckingham parameters where altered by addition of a percentage of the current value of S···S fit with combining rules. For the S···HP potential, combining rules were not used. Full relaxation of structures was performed using DMAREL on CONDOR carried out by D.J. Willock.

During the relaxation, A parameters affects on the repulsion of the potential and C gave effects on the dispersion. From the calculation, we managed to decrease the discrepancy index as low as 37 and with the lattice energy approaching near to the experimental sublimation energy, -104.2 kJ mol^{-1}. Changes in lattice cell after relaxation also appears to be small, which are -5.24% in *a* axis, 0.19% in *b* axis and 1.97% in *c* axis.

Therefore, the S···HP Buckingham potential that will be use for the next calculation is shown as below:

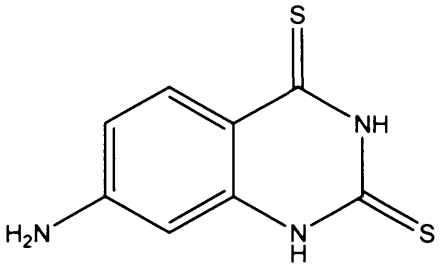
```
BUCK    HP    CODA    SU    CODA
906.518465    0.271137    10.568516  0.0  30.0
ENDS
```

The first line of the input file defines the atom-atom in the set of potential with set of parameters (second line) from equation 4.2. The numerical value in first three rows represents (in sequence) the A, ρ and C parameter, with the distance range from 0.0Å to 30.0Å. the 'END' label defines the end of the set of the defined potential.

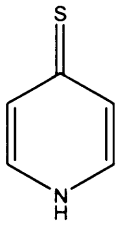
4.6 Application of Potential Test to Thiourea Similar Structure

Optimised potential with the new S···HP fit potential will be tested to other significant molecules which similar to thiourea and TTCA which are:

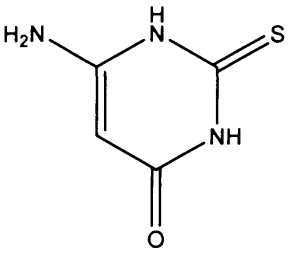
a) ADAHET



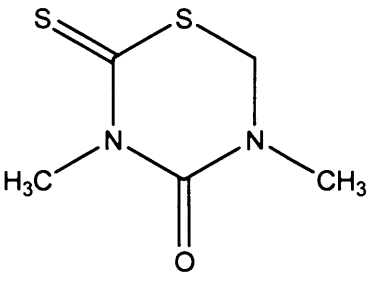
b) AKOVOL



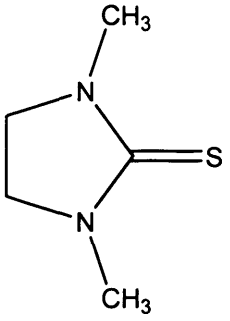
c) AMTURM



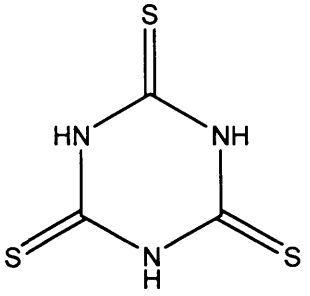
d) BIFHIH



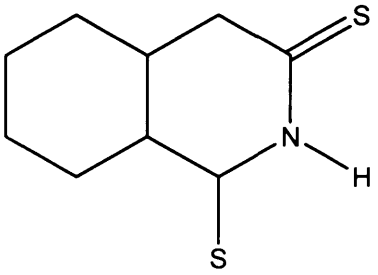
e) BUHVAB



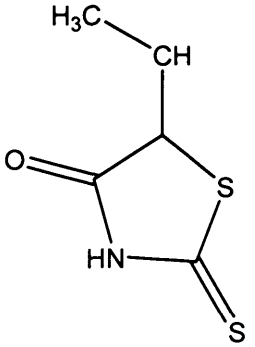
f) CEHQEM (TTCA)



g) BADMUO



h) CEBDES



i) ETTHUR

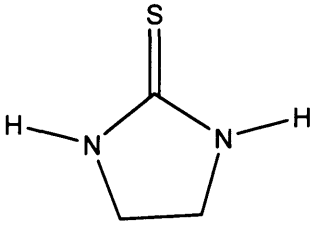


Figure 4.22 : Molecule structures which have similar TTCA environment taken from CDS structures.

These molecules have the same molecular environment for S as in thiourea and TTCA, which N-C=S bond type. All structures has hydrogen bonded network, except for BIFHIH and BUHVAB. Starting from the potentials from previous optimised potential, Table 4.5, the calculation will be fitted to the structures in Figure 4.22. The optimised potential later on will be applied to TTCA and CA molecules for energy study and crystal prediction.

4.6.1 Min E SU-SU Potential Test

For a start, a calculation was done with the above potential ranging Min E from 0.9575 to 1.1575 kJ mol^{-1} in 0.01 interval steps. However, it seems that the total discrepancy for the ten molecules (includes thiourea) data set are decreasing with increasing min E. This range was then further increase/widened from 0.7075 to 1.1575 kJ mol^{-1} in order to find suitable minima in total discrepancy index.

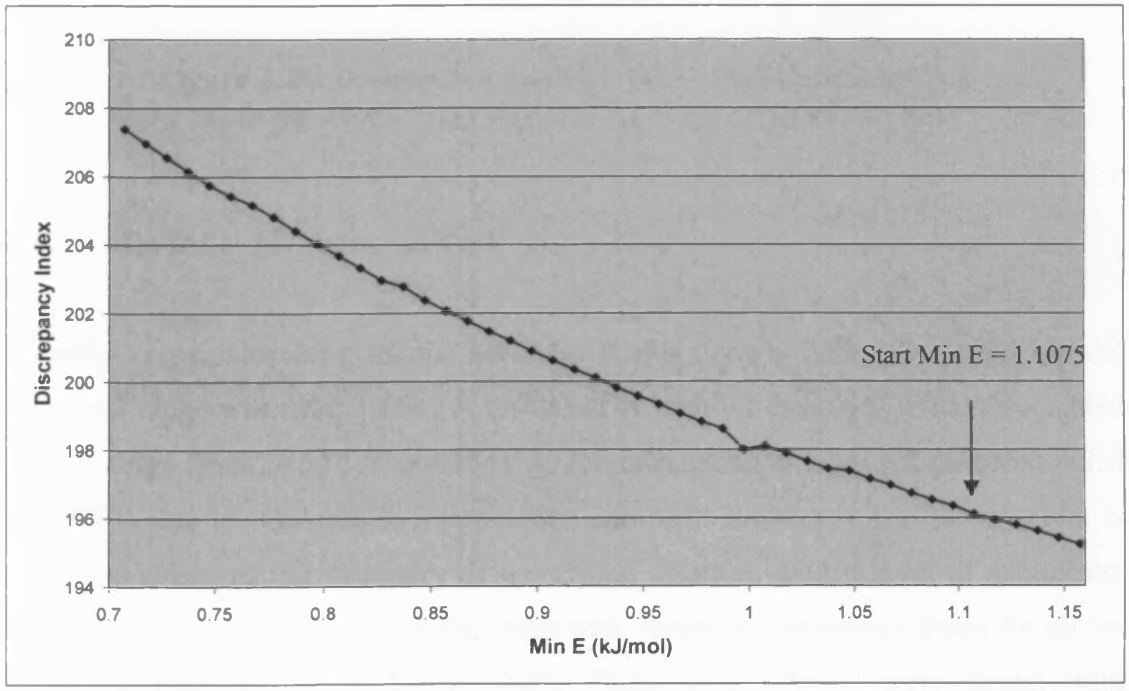


Figure 4.23: Average discrepancy index from different min E

Figure 4.23 shows average discrepancy index of all ten molecules. The discrepancy index decrease gradually with smaller Min E. When the molecular structures were examined individually, they show resemblance pattern of discrepancy index

(gradually decreasing with no absolute minima), but with different values, in exception of TTCA, where there is a range of minima from 0.83 kJ mol^{-1} to 1.05 kJ mol^{-1} (Figure 4.24). However, TTCA still appeared to be the highest discrepancy index of all other molecules. From here, we decided to use 0.8975 kJ mol^{-1} as the new minimum base on the performance shown from all molecules and TTCA.

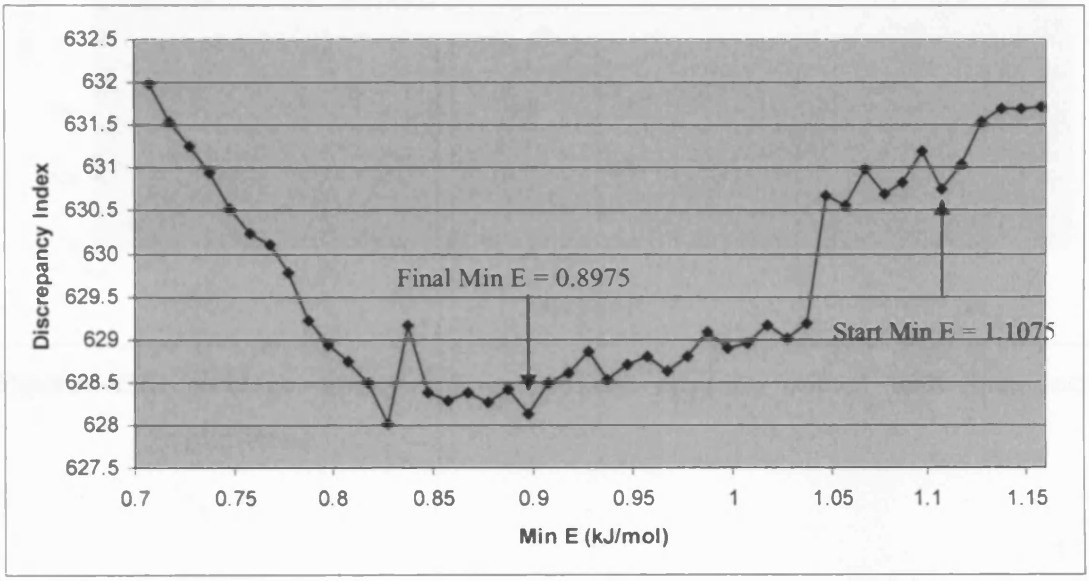


Figure 4.24: Discrepancy index of TTCA with different min E

4.6.2 Min R SU-SU Potential Test

Potential test in searching for the best Min R was done at Min E = 1.1075 kJ/mol. First trial range was from 3.4541 Å to 4.4541 Å with 0.1 intervals. It was then added further range from 2.4541 Å to 4.4541 Å. Job calculation at 3.5541 Å potential failed for TTCA due to 'Quardratic interpolation' problem suggesting that the program is unable to converge the geometry of the crystal structure with this set of parameters and thus, has been omitted from the graph plot. Average discrepancy index for all ten structures were shown in Figure 4.25. Three local minima were found, with discrepancy index (F) 275, 206 and 138, each at min R 3.2541 Å, 3.5541 Å and 4.1541 Å. Generally, this shows that change of min R significantly affect the discrepancy index compared to min E discussed previously. This also suggested that discrepancy index is significantly decreased and improve at larger min R.

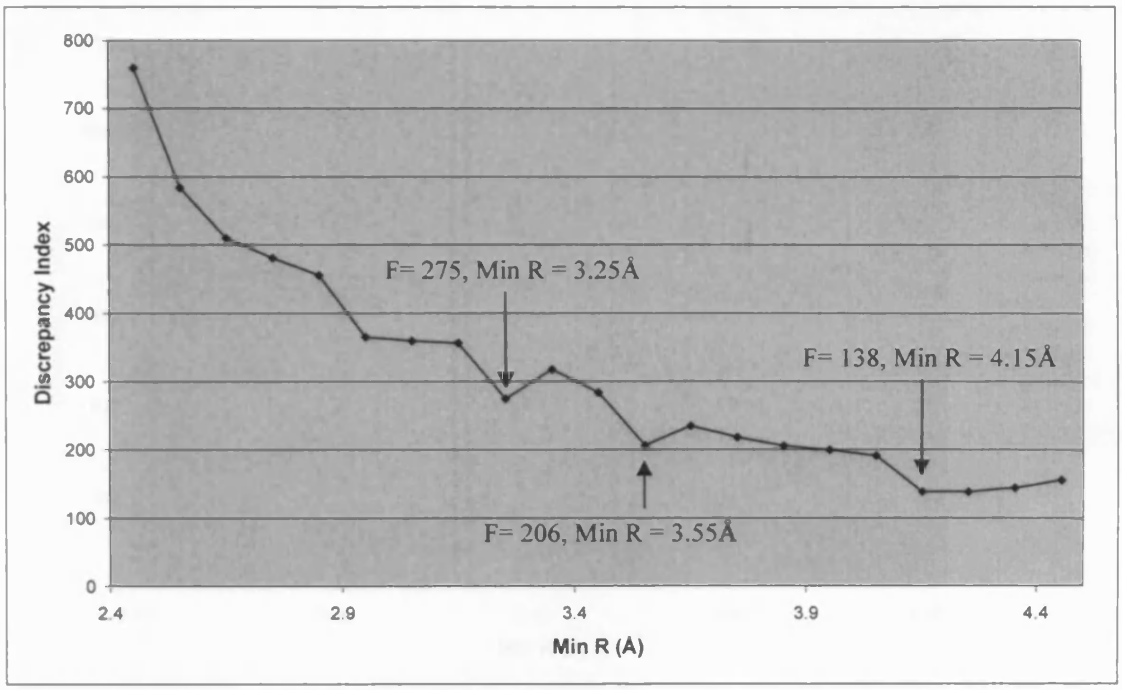


Figure 4.25: Average discrepancy index from different min R with three local minimas.

Figure 4.26 below shows how each molecule contributed to the average discrepancy index discussed above. BIFHIH has quite an unstable pattern of discrepancy index with the change of Min R. AMTURM, however, contributed the lowest discrepancy index, and shows good response with the potential applied. AKOVOL gave a drastic discrepancy change, from 672 (at min R = 3.15Å) to 97 (at min R = 3.25Å) discrepancy index and continue to a stable decrease from 3.2541 Å to a higher value of min R, where it goes as low as 63 discrepancy index at 3.9541 Å. AMTURM, THIOUREA and ADAHET have stable decrease of discrepancy index as the min R grows higher. However, TTCA on the other hand, contributed the largest discrepancy index of all molecules. The lowest discrepancy index value for TTCA is 631 at min R 3.95 Å.

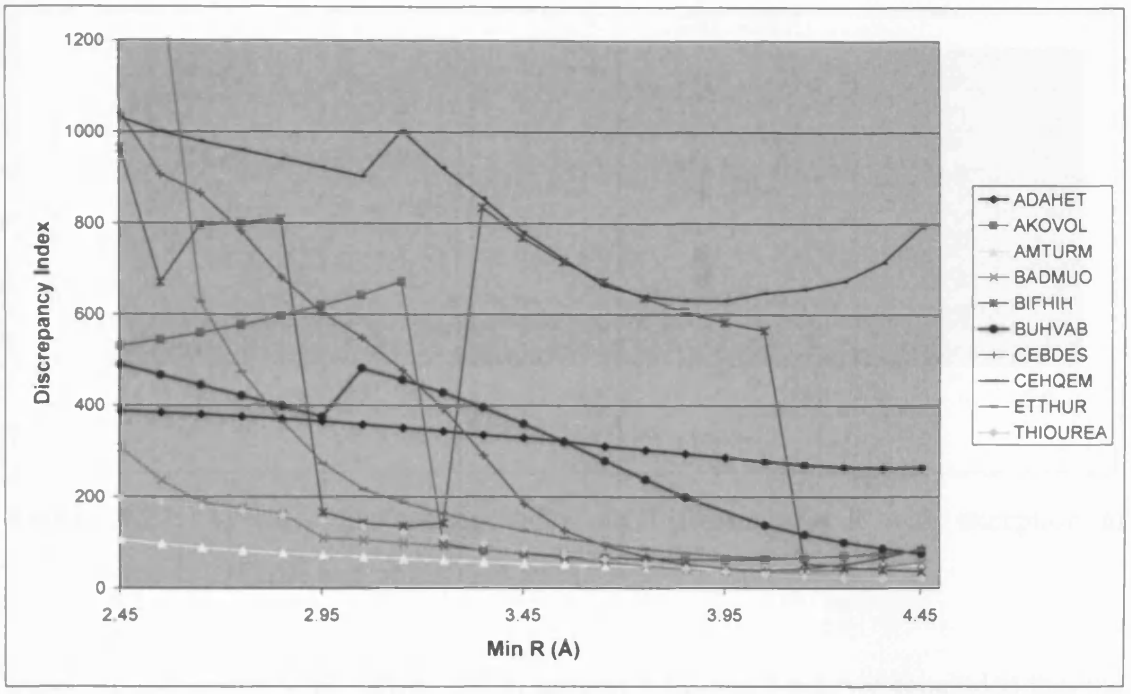


Figure 4.26: Discrepancy index contributed by each molecule from different min R

Leaving out BIFHIH as the unstable molecule and min R at 3.55 (due to 'Quardratic interpolation problem in TTCA), the results are re-plot as Figure 4.27 and a very clear view of minimisation of Min R was observed at 4.1541Å. This value is a little bit higher by 0.2 Å than the previous optimised value. Generally, the total discrepancy index shows better value as the size of Min R goes larger. There is a large downwards jump of total discrepancy index from 3.1541 Å to 3.2541 Å, even though it is only a 0.1 Å difference. This value is caused by sudden fall in discrepancy index of AKOVOL; from 672 to 97. The 575 difference affects total index discrepancy significantly. Other molecules that have significant decrease of discrepancy index from the original Min R range are AMTURM, THIOUREA and TTCA.

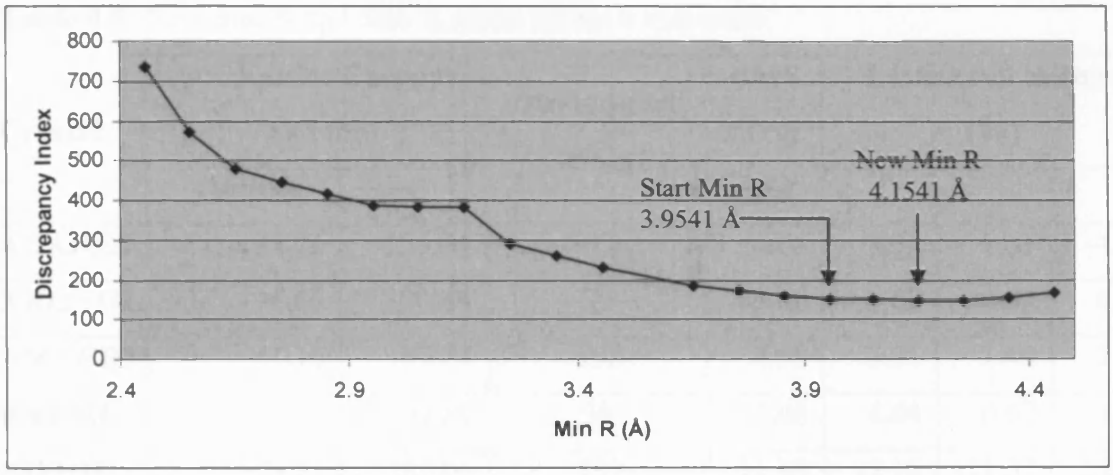


Figure 4.27: Average discrepancy index for different min R with exception of BIFHIH and min R at 3.55 Å.

Based on the calculations carried out in section 4.6.2 and 4.6.1, we concluded that the best S···S parameter potential for use in DMAREL is min E at 0.8975 kJ mol^{-1} and min R at 4.1541Å.

4.6.3 Structure Analysis of New S···S Pair Potential to Crystal Structures

The optimised potential obtained by scanning within a range of min R and min E was applied to all crystal structures in Figure 4.22. The energy properties of the relaxed structure are shown in Table 4.8.

Table 4.8: New min R and Min E result for each molecule

Crystal	Lattice Energy (kJ mol^{-1})		Discrepancy Index	RMS Error (%)	Lattice cell change (%)		
	Initial	Final			a	b	c
ADAHET	-131.71	-155.01	216	6.04	9.31	1.61	-4.50
AKOVOL	-95.69	-106.80	71	3.80	0.02	0.62	6.92
AMTURM	-41.16	-89.94	152	4.59	3.21	1.40	7.13
BADMUO	-94.04	-99.89	38	2.48	4.04	-0.62	1.32
BIFHIH	-73.27	-82.66	552	11.03	-5.12	-11.37	14.46
BUHVAB	-75.23	-81.52	132	5.75	-4.30	0.53	8.97
CEBDES	-56.25	-83.39	57	3.48	-1.63	3.13	4.88
CEHQEM (TTCA)	-105.70	-116.37	653	10.13	16.29	-6.50	0.47
ETTHUR	-104.01	-106.97	67	4.00	3.85	3.71	-4.43
THIOUREA	-101.75	-105.43	26	2.68	-3.84	0.82	2.49

After SU···SU potentials were fitted the with combining rules and using a fixed SU···HP potential, it is found that structures such as thiourea, ETTHUR, CEBDES, BADMUO and AKOVOL do perform well. ADAHET and AMTURM structure produced moderate errors since they are constructed of weak hydrogen bonds. However other structures, particularly TTCA which is the target compound for the comparison of O and S forms, gives large discrepency index. TTCA is a layered structure with two types of hydrogen ribbon tape. In each layer, there is a gap between tape 2 hydrogen ribbon type (Figure 4.28 (a)). After optimisation these gaps became stretched and distorted which resulted in large changes in cell parameters (Figure 4.28 (b)). This happened at the intersection of Type 1 and Type 2 hydrogen ribbon (Figure 4.29).

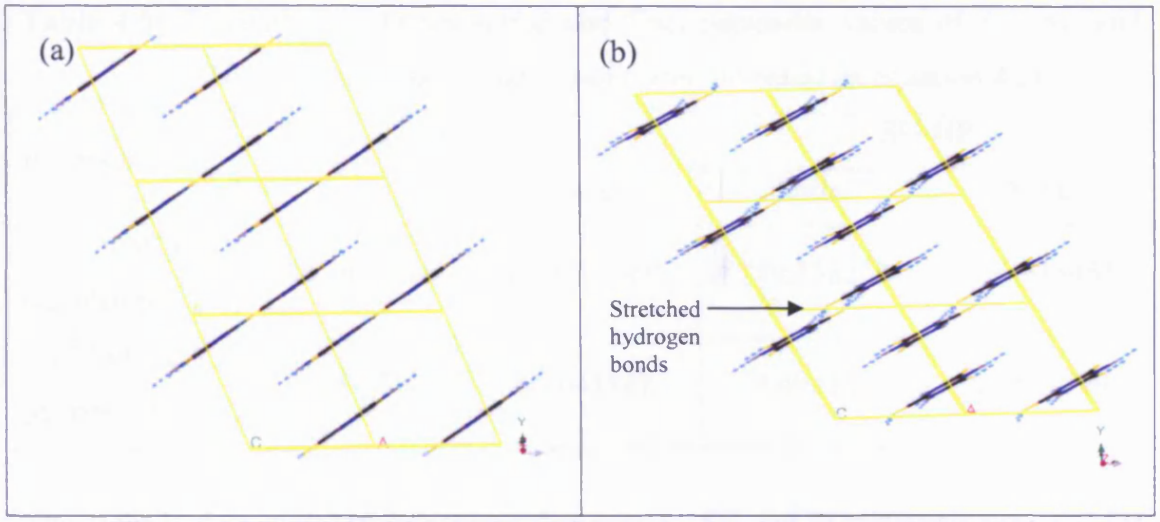


Figure 4.28: TTCA (a) initial and (b) final structures views along c-axis

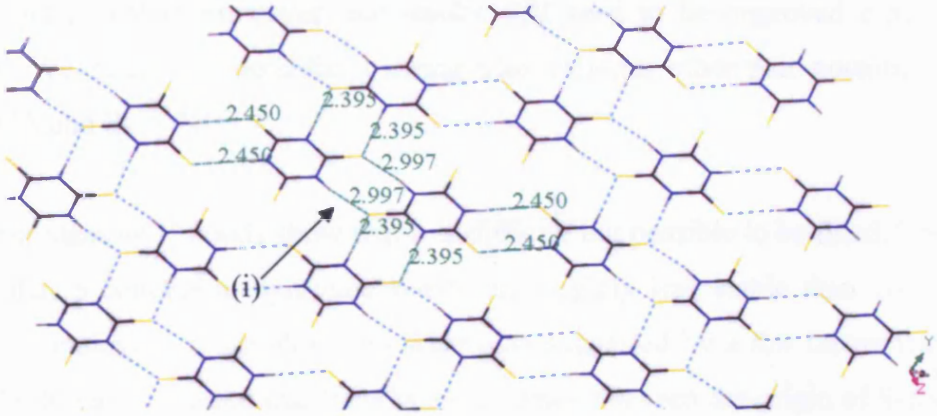


Figure 4.29: Relaxed structure of TTCA with hydrogen bond involved. Arrow shows the intersection between Type 1 and Type 2 hydrogen ribbon (i)

Compared to CA, which is the 'O version' at the experimental structure, the layers in CA maintain in planarity. So it is likely that the SU···C and SU···N parameters are at fault as these will control the S···ring interactions.

We know that A parameter in the Buckingham potential contributes to repulsion and C represents the dispersion between atom-atom potential. Both initial (sulfur potential from literature) and final potential is compared and shown in Table 4.9.

Table 4.9: Comparison between initial and final parameter values of SU···SU and SU···HP potentials (A and C parameter are based on equation 4.2)

Parameter	SU-SU		SU-HP	
	Initial	Final	Initial	Final
A (kJ mol^{-1}) (repulsion)	2241.966726	5428.014935	749.338218	906.518465
C ($Å^6$ kJ mol^{-1}) (dispersion)	54.000273	86.041848	4.492123	10.568516

The optimised potential (final) shows that in SU...SU and SU···HP pair potential for use in DMAREL application has increased in its repulsion and dispersion compared to the initial potential. However, the results still need to be improved especially for TTCA. We suggested potential training also includes other pair potential such as SU...CA and SU...NI.

Sulfur containing H bonds show that it is difficult but possible to be fitted. The reason being that S-containing hydrogen bonds are slightly less stable than O-containing bonds in analogous molecules. This were also suggested by a few theoretical studies [21-23]. It was suggested that there is a difference between the origin of S-containing compared to O-containing hydrogen bonds, where O-containing are suggested to be mostly electrostatics while S-containing involves dispersive charge-dipole and multipole interactions [21].

4.7 TTCA and CA lattice cell substitution

Since TTCA and CA have similar structure, but exist in different crystal packing, an attempt to swap the molecules between the crystal forms was performed. C=S and C=O bonds of each molecule were averaged from experimental structure. Substitution was done by mapping the structures on each other based on the centroid of the aromatic ring. TTCA and CA of the new crystal system all molecules constructed seems to have the same electrostatic multipoles as the one from the experimental structure.

Table 4.10: Energy profile and error data of both experimental and constructed TTCA and CA

Crystal*	Space Group	Lattice Energy (kJ mol^{-1})		Discrepancy Index	RMS error in cell change (%)
		Initial	Final		
TTCA (exp)	P1	-105.70	-116.37	653	10.13
TTCA (const)	C2/n	1019.51	-113.39	4025	32
CA (exp)	C2/n	-154.38	-156.32	22	2.45
CA (const)	P1	-72.37	-133.54	1969	16

* exp is experimental structure and const is constructed structure

Since TTCA in C2/n space group gave very disappointing result due to poor sulfur potential description, we will ignore the discussion on this structure and will concentrate on comparing structures of the simulated CA in P1 space group compared with CA in C2/n space group (experimental).

As mentioned earlier, experimental CA relaxed structure produced very good results with rms lattice change at 2.45% and discrepancy index 22. CA ribbon of Type 1 and 2 are maintained after relaxation.

For constructed CA crystal in P1 space group, large discrepancy index were produced which is at 4025. This is expected since hydrogen bonds length in the initial structure are far larger than normal hydrogen bonds in CA (1.8 Å in experimental, 2.6 Å in constructed). However, it does produce reasonable final lattice energy, where this indicates that this type of structure may exist.

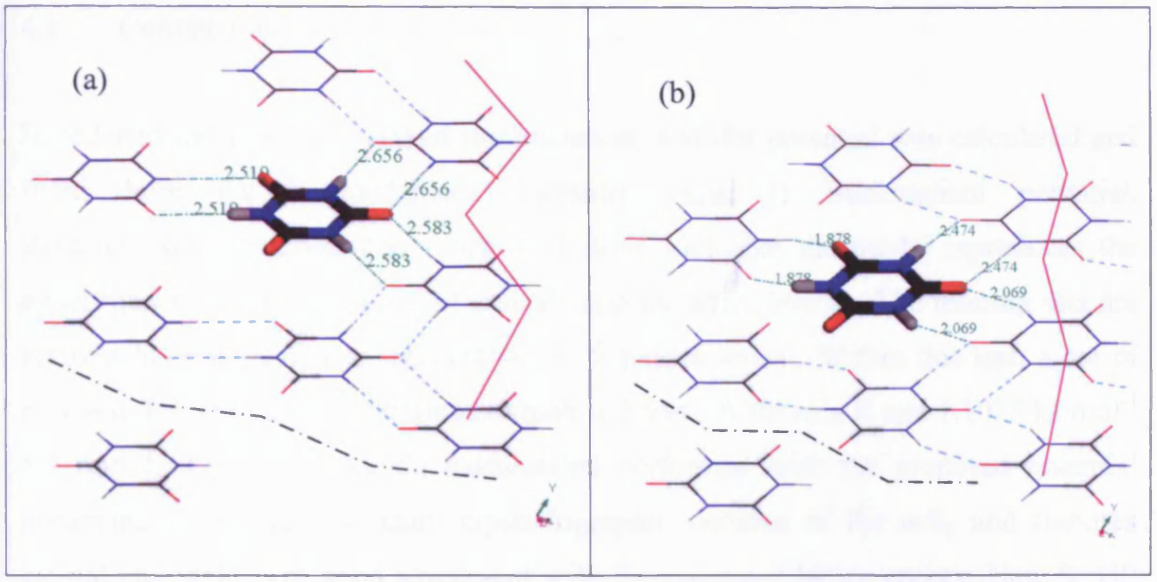


Figure 4.30 : Hydrogen bond network of CA structure in P1 space group (a) Initial and (b) Final. The bold CA is the reference molecule to compare both structures. Network type represented by: Type 1 – Black (-·-·-) and Type 2 - Pink

In the constructed CA, both types of hydrogen bond network still exist, but in a shorter and compact version due to change from sulfur to oxygen atom type. Type 1 O…H length of N-H···O hydrogen bond network shortened from 2.656 Å to 2.474 Å and 2.583 Å to 2.069 Å. In Type 2 network, O…H length of N-H···O hydrogen bond shortened from 2.519 Å to 1.878 Å and 2.583 Å to 2.069 Å. This occurs because oxygen atom has higher electronegativity and smaller size than sulfur [24]. Moreover, due to the shortened hydrogen bond length, cell density was increased by 41% from the initial structure.

To conclude, CA in TTCA structure framework may exist, but further study needed such as varying the lattice cell size of the new predicted structure. This is also suggested by a DFT study of both structures recently carried out by Martsinovich [15].

4.8 Conclusion

In order to calculate the relaxed minimisation, a sulfur potential was calculated and fitted from the Lennard-Jones literature value to Buckingham potential. Parameterisation of potential function is done such that the model reproduces the experimental structure of a set of crystals and its lattice energy. The training sets are set first between α-S_8 and thiourea for S···S pair potential. Within this test, a set of new min R and min E are obtained, which is 3.9541 Å for min R and 1.1075 kJ mol^{-1} for min E. Crystal relaxation calculations performed with the proposed potential acceptably reproduce the main crystallographic features of the α-S_8 and thiourea crystal and yield very good agreement with the estimated lattice energy. New S···HP was also found to be of significant value in producing a good pair potential based on the discrepancy index values. This was then applied to TTCA crystal structure. However, it produced a poor discrepancy index value, 631.

We advanced our study by testing min R and min E against a set of molecules which have similar environment to TTCA. From the training set, lattice energy of min R and min E obtained gave several local minima from the average discrepancy index calculated. Best min R and min E obtained for S···S pair potential of all sulfur contained structures were 4.1541Å and 0.8975 kJ mol^{-1}. AKOVOL, BADMUO, CEBDES, ETTHUR and thiourea produced very good discrepancy index which are all below 100. However, from this new potential, TTCA still contributes the poorest discrepancy index value out of all molecules, which is 653.

Substitution of TTCA and CA was done with the fitted S···S and S···HP pair potential. Poor results were shown in both new substituted CA and TTCA crystal packing. This is also true for a constrained volume relaxation. Discussion is focused on new CA crystal structure where we suggested that it is possible for CA to exist in monoclinic, P1 space group crystal structure as in TTCA crystal packing by direct substitution of molecule. However, we suggest that this should also be followed by flexible change in lattice cell during the crystal relaxation since DMAREL does not cater for large change in lattice cell.

Though we have been able to produce S-S potential fitting for most of the sulfur contained molecule, for future work, fitting test can be done to other atom with S potential to find the contribution of other heteroatom (S···X) pair potential to the lattice energy calculation such as N and C. Besides that, this initial TTCA and CA lattice cell substitution can be applied to MD simulation to find other possible polymorphism structures available for these molecules.

4.9 References

(1) Kopel, P.; Dolezal, K.; Machala, L.; Langer, V. *Polyhedron* **2007**, *26*, 1583-1589.

(2) Ranganathan, A.; Pedireddi, V. R.; Chatterjee, S.; Rao, C. N. R. *J. Mater. Chem.* **1999**, *9* 2407.

(3) Ahn, S.; PrakashaReddy, J.; Kariuki, B. M.; Chatterjee, S.; Ranganathan, A.; Pedireddi, V. R.; Rao, C. N. R.; Harris, K. D. M. *Chemistry - A European Journal* **2005**, *11*, 2433-2439.

(4) Mahon, M. F.; Molloy, K. C.; Venter, M. M.; Haiduc, I. *Inorganica Chimica Acta* **2003**, *348*, 75-81.

(5) Guo, F.; Cheung, E. Y.; Harris, K. D. M.; Pedireddi, V. R. *Crystal Growth & Design* **2006**, *6*, 846-848.

(6) H.Dietrich; C.Scheringer; H.Meyer; K.-W.Schulte; A.Schweig *Acta Crystallogr.,Sect.B: Struct.Crystallogr.Cryst.Chem.* **1979**, *35*, 1191-1197.

(7) Wang, M. X. *Chem. Commun.* **2008**, 4541-4551.

(8) El-Gamel, N. E. A.; Brand, J.; Kroke, E. *J. Coord. Chem.* **2009**, *62*, 1278-1284.

(9) Marivel, S.; Suresh, E.; Pedireddi, V. R. *Tetrahedron Lett.* **2008**, *49*, 3666-3671.

(10) Wu, A. Q.; Guang-Hua, G.; Fa-Kun, Z.; Xi, L.; Guo-Cong, G.; Jin-Shun, H. *Inorg. Chem. Commun.* **2005**, *8*, 182-185.

(11) Burrows, A. D. In *Supramolecular Assembly Via Hydrogen Bonds I*; Springer-Verlag Berlin: Berlin, 2004; Vol. 108; pp 55-95.

(12) Barnett, S. A.; Blake, A. J.; Champness, N. R. *Crystengcomm* **2003**, 134-136.

(13) Pedireddi, V. R.; Belhekar, D. *Tetrahedron* **2002**, *58*, 2937-2941.

(14) Ranganathan, A.; Pedireddi, V. R.; Sanjayan, G.; Ganesh, K. N.; Rao, C. N. R. *Journal of Molecular Structure* **2000**, *522*, 87-94.

(15) Martsinovich, N.; Kantorovich, L. *J. Phys. Chem. C* **2008**, *112*, 17340-17350.

(16) Pastorino, C.; Gamba, Z. *Chem. Phys.* **2000**, *261*, 317-321.

(17) Willock, D. J. *Philosophical Magazine B* **1996**, *73*, 127-138.

(18) Chrisey, D. B.; Boring, J. W.; Johnson, R. E.; Phipps, J. A. *Surface Science* **1988**, *195*, 594-618.

(19) Gatta, G. D.; Jozwiak, M.; Brunetti, B.; Abate, L. *J. Chem. Thermodyn.* **2000**, *32*, 979-997.

(20) Coombes, D. S.; Nagi, G. K.; Price, S. L. *Chem. Phys. Lett.* **1997**, *265*, 532-537.

(21) Platts, J. A.; Howard, S. T.; Bracke, B. R. F. *J. Am. Chem. Soc.* **1996**, *118*, 2726-2733.

(22) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *The Journal of Physical Chemistry A* **1998**, *102*, 3782-3797.

(23) Wierzejewska, M.; Saldykaa, M. *Chem. Phys. Lett.* **2004**, *391*, 143-147.

(24) In *Electronegativity: http://en.wikipedia.org/wiki/Electronegativity*, 2009.

Chapter 5

Urea Molecule Clusters

5.1 Introduction

5.1.1 Urea structure in gas and aqueous

The molecular structure of urea in the pure crystalline state is well known to be planar. However, recently, it was found that in the gas phase, the urea structure can also have pyramidal conformations at the NH_2 functional groups, resulting in structures with C_s and C_2 symmetry (See Figure 5.1).

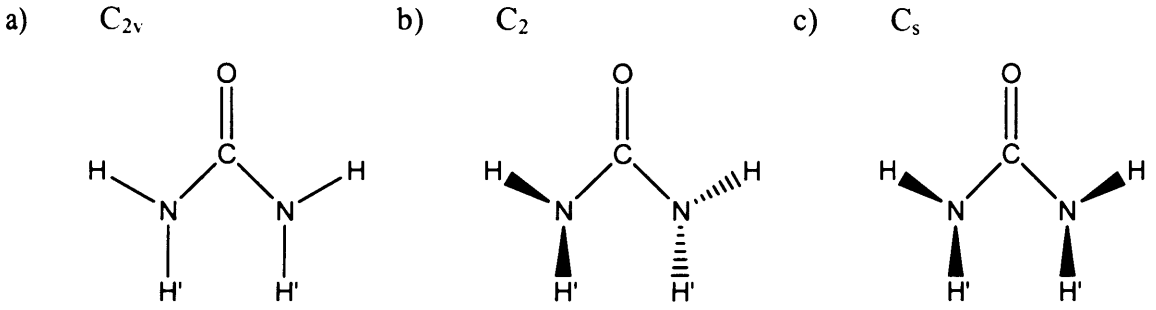


Figure 5.1: Conformations of isolated urea molecule (a) C_{2v}, (b) C_2, and (c) C_s

This was observed in both experimental microwave spectroscopy and theoretical calculations [1-3]. There are also other possible urea structure conformations reported by Masunov and Dannenberg in their study of urea monomers and dimers [4]. These conformers are transition states of the urea molecule that are high in energy and some lower energy structures (Figure 5.2).

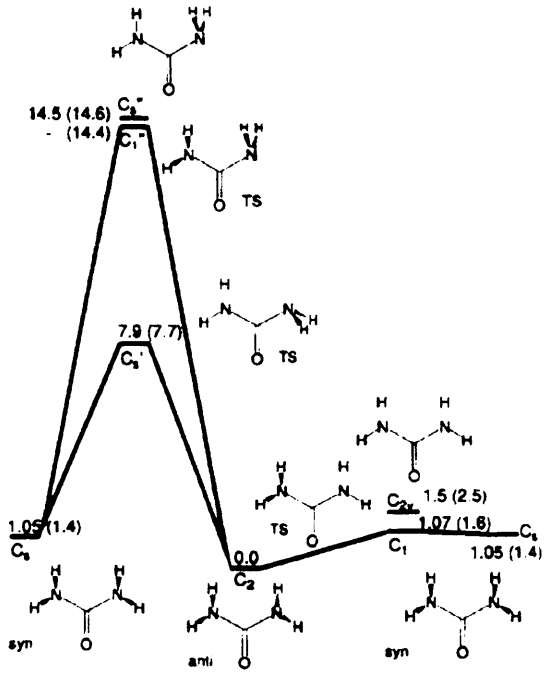
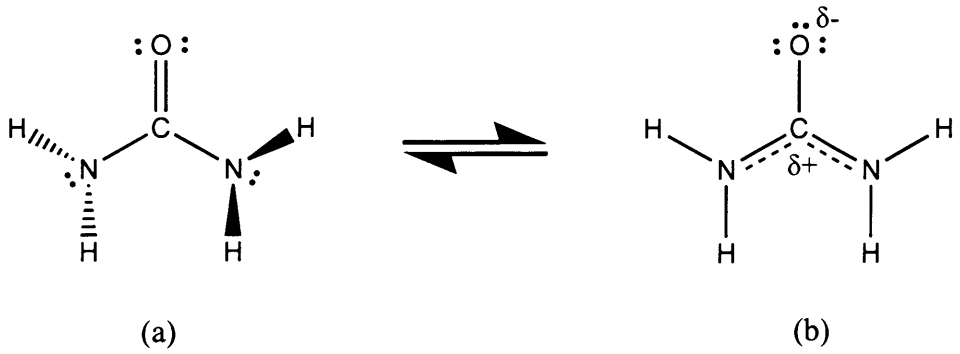


Figure 5.2: Urea conformers previously studied by Masunov and Dannenberg [4]. The numbers show relative energy (kcal mol^{-1}) of each conformer relative to C_2 *anti* conformation calculated using B3PW91/D95(d,p) and MP2/D95(d,p) in parentheses.

Note that these high energy structures resulted from rotation around C-N bonds. The fact that they are high in energy confirms some double bond character in the C-N bond. This is consistent with resonance shown in Figure 5.3. Where in gas phase, NH_2 urea is pyramidal, closer to Figure 5.3 (a) and in solid, urea becomes planar, closer to Figure 5.3 (b) structure with the resonance of N-C-O, forming single bond in C-O and double bond in C=N. C-N bonds would more easily rotate in the (a) form rather than in (b). This means that although the NH_2 is pyramidal, the electronic structure is still a mixture of these two forms.



(a) (b)

Figure 5.3: (a) Lewis structure with lone pairs at N. (b) Lewis structure of N lone pairs delocalised over N-C-N.

In a hydrogen bonding situation, the bonding to other molecules will be enhanced by redistribution of charge, giving a negative oxygen centre that would be expected to be a better hydrogen bond acceptor than in the formal keto group and a more positive NH_2 group, a better hydrogen bond donator.

These three main conformers have received much attention [5] (Figure 5.1). The resonance structures have different symmetry. Pyramidal geometries at N can be C_2 (Figure 5.3 (a)) or C_s with all hydrogens at the same side of the molecular plane. The planar form, Figure 5.3 (b) has C_{2v} symmetry.

There are also studies of urea in aqueous solution, where it was found that urea also form pyramidal conformers to form dimers and were hydrogen bonded with water molecules to form clusters [6,7]. It has strong dipolar and hydrogen bond interactions with both water and itself. In our study of urea molecule, we will concentrate on studying gas phase (C_2 and C_s) and planar forms as they are the most stable conformers that exist in gas phase.

5.1.2 Previously Studied Urea Oligomers Structure

Urea dimer structure were studied by Masunov in comparing the dimers and the monomer structure. Urea dimers studies of chains and urea ribbon structures (Figure 5.4) showed to be in planar conformations after applying counterpoise and thermal correction in an *ab initio* study [4,8]. He reported that the ribbons conformation is more favoured by dimers than small aggregates.

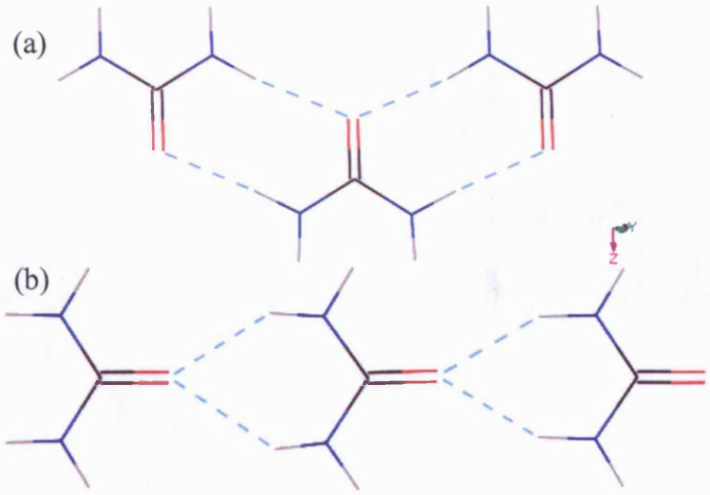


Figure 5.4: Urea (a) ribbon structure is more favoured than urea (b) chain structure

A series of urea microclusters containing up to eight C_2 symmetry of urea molecule were studied by Spoliti *et al* [9] to determine the vibrational spectroscopy data of the compound. Different conformers of structure were obtained including in the clusters of C_s geometry. They also stated that the exact determination of relative stability of the two conformers (C_s and C_2) is strongly dependent on the computational level of theory and the inclusion of ZPE correction.

5.1.3 Urea structure in solid phase

In the solid form, urea crystal structure has space group of $P\bar{4}2_1m$ [10] with a molecular planar conformation. Each urea molecule has eight hydrogen bonds; four from oxygen and one from each hydrogen. Along the urea plane, there are two hydrogen bond networks which involve two hydrogen bonds of NH···O of the same oxygen atom that can be described as $C_2^1(4)[R_2^1(6)]$, and $C_2^1(6)$, and also a hydrogen bond network perpendicular to the urea plane designated as $C_2^1(6)$, as illustrated in Figure 5.5.

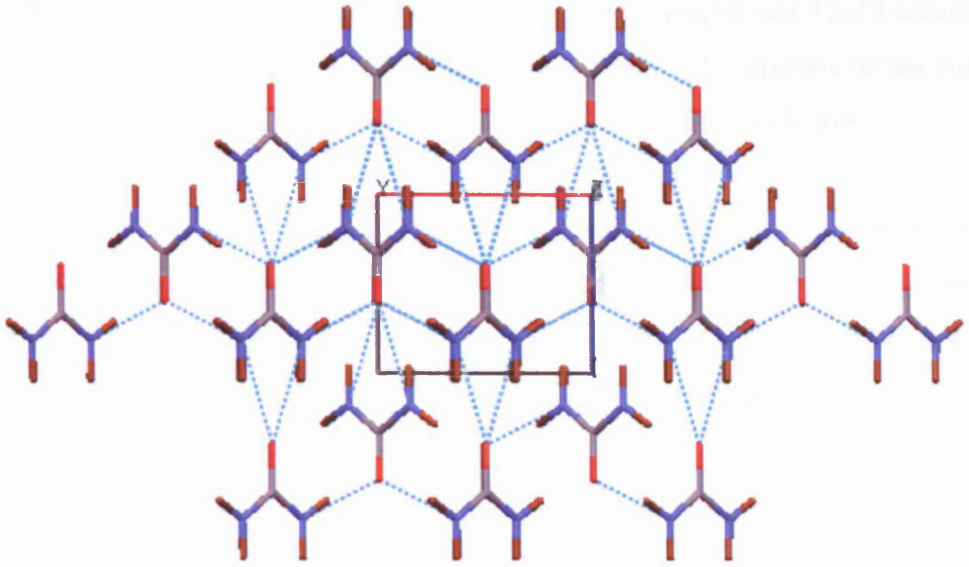


Figure 5.5 : View of hydrogen network in urea crystal along b-axis

Rousseau *et al* compare and explain experimental intermolecular vibration splitting in the IR and Raman spectra of the urea crystal using molecular cluster model [10]. They use RHF/6-31++G and urea with C_{2v} conformation.

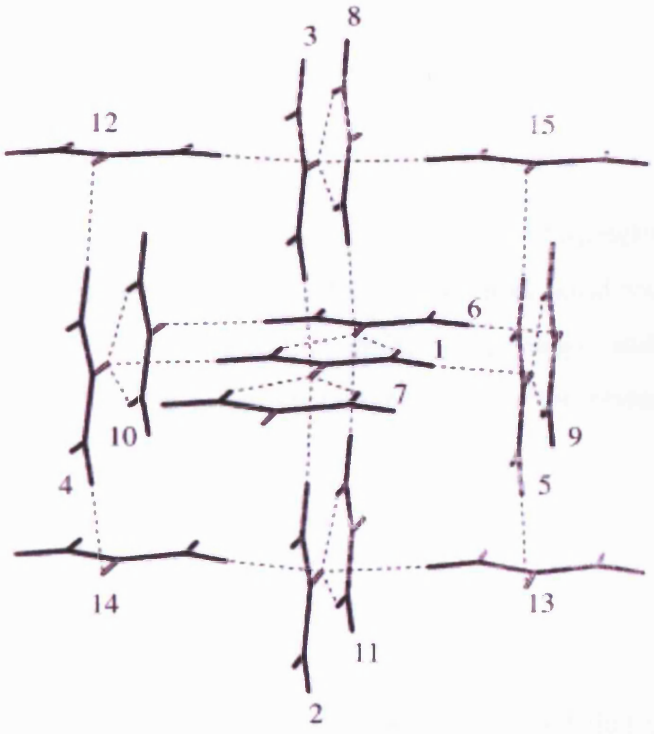


Figure 5.6: Urea molecule labels for studying molecular vibration splitting by Rousseau *et al.* [Reproduced from reference 10]

It was reported (Figure 5.6) that interaction of urea 1 with neighbours 12-15 contributed the smallest in magnitude, based on the force constant value. 2-5 contributes the most in the interaction and then 6-7, closely followed by 8-11. As shown in table below:

Table 5.1: Maximum (k_{max}, in mdyne $Å^{-1}$) and rms average (k_{rms}, in mdyne $Å^{-1}$) values of the calculated Cartesian interaction force constants in the intermolecular blocks between the central molecule (1) and its different neighbours (2-15). Reproduced from reference 10.

k_{max}	k_{rms}	**Neighbour molecules**			
0.303	0.033	2	3	4	5
0.163	0.024	6	7		
0.134	0.018	8	9	10	11
0.066	0.009	12	13	14	15

Vibrational studies of lattice urea crystal were studied theoretically for Raman lattice active vibration [11]. Another vibrational studies is the microwave spectrum of urea where it is reported to be of 5 and 50 GHz. The urea molecule studied was in the form of planar urea, that when it is in the solid formation. Its planarity contraints is in agreement up to at least J=20. It was also reported that the dipole moment of urea molecule is $\mu = \mu_b = 3.83$ D (12.8×10^{-30} Cm) [12].

Engkvist [13] studied the morphology of urea crystal computationally using DMA for the intermolecular potential. Another theoretical study of the urea solid morphology was carried out by Docherty [14] where the changes in polar morphology and absolute polarity of crystalline urea were found to be in good agreement with the observed morphologies of crystals prepared from the vapour phase.

5.1.4 Aims of the study

This chapter is the continuation of Chapter 3 where we noticed that the urea structure in a co-crystal is not exactly planar as we expect. However, in the gas and liquid phase, urea

exists in a non-planar form. To answer these questions, we continue our research where our main aim is to understand how the solid state structure causes urea to change between pyramidal to planar conformation during crystallisation. How many neighbouring molecule and hydrogen bonds are involved or required in this process and to understand the process under laying it.

5.2 Computational Method

Calculations of urea monomer in Gaussian03 were carried out with each of the low energy C_{2v}, C_2 and C_s conformers to observe their electronic energy and other properties, such as optimised structure, zero point energy and energy difference between each method of calculation. Different basis sets were used to satisfy the accuracy of the calculation using MP2 and density functional theory (B3PW91, B3LYP) method. Calculation was tested with 'Tight' and without 'Tight' option, and it was shown that there is no difference in the electronic energy optimisations. Correction for BSSE was performed with the counterpoise method.

From the optimised structure of each conformer, oxygen atom in urea was substituted with CH_2 functional group. This removes the possibility of the N lone pair delocalisation and so should give a reference for the pyramidal structure with no bond delocalisation (Figure 5.3(a)) using the same type of optimisation; MP2 and density functional method (B3LYP). This knowledge is important to be used later on for classifying the degree of pyramidalization of larger formations of urea molecules.

Calculation on the urea dimer in C_{2v} conformer was made to determine which method is best should be used for building the urea clusters. B3LYP and MP2 method were used with different basis sets starting from 6-31G valence, increasing up the valence to 6-311G, and slowly adding polarisation function and diffusion function.

Different arrangements of hydrogen bonded dimer were optimised with the best basis set method with symmetry off calculation and this was also applied for the rest of clusters with different type of possible hydrogen bond interaction. Clusters were constructively built up by increasing the number of urea molecules one by one, so that the hydrogen bonding for the

first molecule in the cluster is gradually increased. Upon optimisation of the clusters, only the pyramidal (C_2 and C_s) urea molecule was set free from any restriction while all other neighbouring ureas were frozen. Degree of planarity and Mulliken charge of urea pyramidal structure after each optimisation was compared. This will give us an idea on the influence of neighbour urea in producing planar urea in the crystal.

5.3 Structure conformation of Urea Monomer

An isolated urea molecule may exist in one of 3 main isomers [3]. The first conformer is the well known planar conformer, C_{2v}. Other two conformations have C_2 and C_s symmetry. C_s symmetry has pyramidal amino groups (NH_2) with the two pairs of H atoms at the same side of N-(C=O)-N plane while in the C_2 conformer the H atoms in the amino group are on opposite sides (Figure 5.1).

Electronic energy of each isomer were calculated and compared at the second order of Møller Plesset (MP2) and hybrid density functional theory (B3PW91 and B3LYP) method using several different basis sets and multiple polarisation functions as in Table 5.2.

Table 5.2: Calculated energy of urea isomers (energy in Hartree, H)

Conformer	**C_2**	**C_{2v}**	**C_s**
MP2			
6-311G(d,p)	-224.741687	-224.736794	-224.739122
6-311G(d,p) + ZPE	-224.676700	-224.674329	-224.674831
6-311G(d,p) + Gibbs	-224.702899	-224.699810	-224.701273
6-311++G(d,p)	-224.756496	-224.752121	-224.754579
6-311++G(d,p) + ZPE	-224.691843	-224.690170	-224.690433
6-311++G(d,p) + Gibbs	-224.718066	-224.715729	-224.716840
6-311++G(2d,2p)	-224.810483	-224.807382	-224.808808
6-311++G(2d,2p) + ZPE	-224.746106	-224.745313	-224.745076
6-311++G(2d,2p) + Gibbs	-224.772353	-224.770789	-224.771596
B3PW91			
6-311G(d,p)	-225.251228	-225.249118	-225.249563
6-311G(d,p) + ZPE	-225.187176	-225.187144	-225.186832
6-311G(d,p) + Gibbs	-225.213440	-225.212584	-225.215987
6-311++G(d,p)	-225.257892	-225.258980	-225.259314
6-311++G(d,p) + ZPE	-225.195706	-225.197188	-225.196647
6-311++G(d,p) + Gibbs	-225.222038	-225.222642	-225.224168
6-311++G(2d,2p)	-225.264921	-225.265235	-225.266006
6-311++G(2d,2p) + ZPE	-225.203449	-225.203373	-225.202811
6-311++G(2d,2p) + Gibbs	-225.229716	-225.228818	-225.229476
B3LYP			
6-311G(d,p)	-225.338647	-225.336415	-225.336926
6-311G(d,p) + ZPE	-225.274887	-225.274756	-225.274333
6-311G(d,p) + Gibbs	-225.301154	-225.300205	-225.301714
6-311++G(d,p)	-225.349899	-225.348296	-225.348635
6-311++G(d,p) + ZPE	-225.286387	-225.286837	-225.286263
6-311++G(d,p) + Gibbs	-225.312691	-225.312302	-225.313668
6-311++G(2d,2p)	-225.356856	-225.354777	-225.355486
6-311++G(2d,2p) + ZPE	-225.293187	-225.293210	-225.292608
6-311++G(2d,2p) + Gibbs	-225.319465	-225.318664	-225.319303

Generally, from Table 5.2 the calculation with different basis sets shows that if we include zero point energy (ZPE) in the calculation, the total energy will increase by between 0.064 H to 0.065 H in all methods. This criteria is also shown with the inclusion of Gibbs Free energy, between 0.038 H to 0.039 H. Looking at the effect of basis set, the diffuse functions also play an important part as these create a better chance of overlapping basis function between atoms. The diffuse function lowered the energy by 0.015 H in MP2, 0.0067 H in B3PW91 and 0.113 H in B3LYP. This shows that diffuse function gave similar quality in MP2 and B3LYP method. The increase of polarisation function, from (d,p) to (2d,2p) will increase the flexibility of the basis set polarisation in heavy atoms (C,O,N) and also d function of the atomic orbitals. Thus, giving a lower minimised electronic energy. In comparing the energy between the methods used, MP2 gave the largest effect (0.05 H) than both density functional method, with minimal difference from 0.006 H to 0.007 H. Overall, by including polarisation and diffuse function, the total energy can be decreased by 0.07 H in MP2 and 0.01 H to 0.02 H in density functional methods. The difference between the conformer structures will be discussed in detail later on, in this thesis.

Geometry of each conformer is shown in Table 5.3. A comparison of different methods used in the previous optimisation is shown and compared (Table 5.3 (a)). Bonds lengths, bond angles and dihedral angles were measured for these structures.

The geometries of urea were first calculated and compared with the literature between C_2 conformers. Only calculated data for one half of the urea molecule is shown here because it is symmetrical in bond lengths and bond angle. Our calculation with large basis sets proved to be similar to the literature (Table 5.3 (a)) with the largest difference being 0.002Å for bond length between B3LYP 6-311++G method of (2d,p) and (2d,2p) basis sets. The B3LYP 6-311++G(2d,2p) calculation also provide molecular geometry of the same quality as from MP2 with the same basis set, where there is less then 1% difference in bond lengths and bond angles. However, there is a slight difference in the dihedral angles. Summary of geometry for C_{2v} and C_s is given in Table 5.3 (b).

Table 5.3: (a) C_2 Symmetry Isomer

C_2 Symmetry Isomer				
	BLYP 6-311+G(d,p)[a]	B3LYP 6-311++G(2d,p)[a]	B3LYP 6-311++G(2d,2p)	MP2 6-311G++(2d,2p)
Bond Lengths (Å)				
CO	1.229	1.216	1.216	1.219
CN	1.401	1.386	1.385	1.388
NH	1.017	1.009	1.007	1.005
NH'	1.017	1.009	1.007	1.005
Bond Angles (degrees)				
NCO	123.0	122.9	122.9	123.2
NCN	114.0	114.3	114.3	113.6
HNC	113.3	113.1	113.1	112.4
H'NC	118.4	117.9	118.0	116.7
HNH'	114.5	114.5	114.6	114.0
Dihedral Angles[b] (degrees)				
$NCNH^2$	28.9	29.3	29.0	31.2
180° - [$NCNH^1$]	12.9	13.3	13.4	14.3
180° - [$OCNH^2$]	28.9	29.3	29.0	31.2
$OCNH^1$	12.9	13.3	13.4	14.3
$NCNH^4$	-28.9	-29.3	-29.0	-31.2
180° - [$NCNH^3$]	-12.9	-13.3	-13.4	-14.3
180° - [$OCNH^4$]	-28.9	-29.3	-29.0	-31.2
$OCNH^3$	-12.9	-13.3	-13.4	-14.3

[a] See reference [3]

[b] For H atom designations, refer to Figure 5.7

Table 5.3 : (b) C_{2v} and C_s urea geometries

Isomer	C_{2v}		C_s	
	B3LYP 6-311++G(2d2p)	MP2 6-311G++(2d,2p)	B3LYP 6-311++G(2d2p)	MP2 6-311G++(2d,2p)
Bond Lengths (Å)				
CO	1.219	1.222	1.217	1.215
CN	1.373	1.374	1.381	1.393
NH	1.002	1.000	1.005	1.009
NH'	1.002	0.999	1.005	1.009
Bond Angles (degrees)				
NCO	122.4	122.6	122.4	123.7
NCN	115.3	114.8	115.1	112.6
HNC	117.0	116.9	114.3	115.2
H'NC	123.8	123.5	120.6	111.7
HNH'	119.2	119.6	116.1	113.3
Dihedral Angles[a] (degrees)				
$NCNH^2$	0.0	0.0	23.0	34.3
180° - [$NCNH^1$]	0.0	0.0	11.0	14.7
180° - [$OCNH^2$]	0.0	0.0	18.9	34.3
$OCNH^1$	0.0	0.0	15.1	14.7
$NCNH^4$	0.0	0.0	23.0	34.3
180° - [$NCNH^3$]	0.0	0.0	11.0	14.7
180° - [$OCNH^4$]	0.0	0.0	18.9	34.3
$OCNH^3$	0.0	0.0	15.1	14.7

[a] For H atom designations, refer to Figure 5.7

The C=O bond length between conformers show that there is not much difference between these three, which is about 1.22 Å. However, C-N bond of C_{2v} are shorter than C_2 and C_s conformation, with the sequence in MP2 calculation starting from the shortest bond to be 1.37 Å < 1.38 Å < 1.39 Å. In C_{2v}, where the structure is similar to Figure 5.3 (b), shows that semi double bond characteristics between C and N may exist in this structure. Mulliken charge of MP2 method (Table 5.4) also shows that carbon atom in C_{2v} is more positively charged (0.76) and the nitrogen atoms are more negatively charged (-0.046) compared to C_s

and C_2. This might induce a shorter bond length in the C_{2v} conformer structure. O atom in the C_{2v} structure is the most negatively charged between the three conformers. This high polarisation may cause the C-O bond to be single bonded as suggested in Figure 5.3 (b) and also would suggest that stronger inter-molecular hydrogen bonds can be formed with planar urea. On the other hand, the Mulliken charge distribution between C_s and C_2 shows that the carbon in C_2 is more positively charged (by 0.02) and its nitrogen atoms are more negatively charged (0.014). This explains the opposite pyramidalization of C_2 structure being much lower in energy than C_s.

Table 5.4: MP2 Mulliken charge distribution where atomic charge is summed into heavy atoms

	C_s	C_{2v}	C_2
C	0.678	0.760	0.697
O	-0.648	-0.669	-0.640
N	-0.015	-0.046	-0.029
N	-0.015	-0.046	-0.029

The dihedral angle used to define the structure is based on the degree of NH_2 bending relative to the N-CO-N plane (Figure 5.7) so the values quoted for the angle are small. Here, we can see that the dihedral angle played a significant part in describing the structure of these three conformers. Dihedral angles were labelled accordingly as shown in Figure 5.7 where it shows the direction of NH_2 pyramidal shape, whether they are on the same sides (positive), opposite (negative) or in the plane. We notice that C_2 and C_{2v} have planar heavy atoms (N-CO-N), while in C_s the heavy atoms are not quite planar with the dihedral angle of 4° using B3LYP and 5° using MP2 method. This observation was also reported by Spoliti *et al* [3] from his geometry comparison. This may happen because of the lone pair force in NH_2 pyramidal at the same side of the molecule which creates the slight torsion to the N-CO-N plane. Generally, NH_2 degree of planarity for C_s is greater than C_2 (in B3LYP method), due to it's lower torsion angles comparing the structures presented in Table 5.3 (b) where the average (NCNH and OCNH, B3LYP method) torsion angle for C_s is 17° and C_2 is 21.2°.

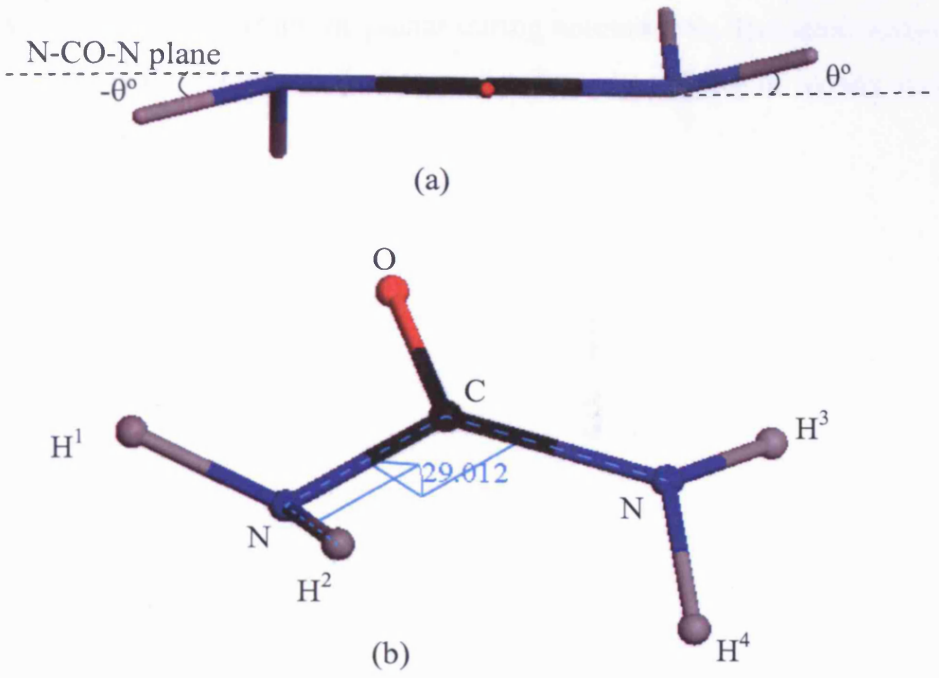


Figure 5.7: (a) Dihedral angle defined relative to N-CO-plane in C_2 conformer. (b) Atom labels in dihedral measurement.

Apart from analysing these three urea conformers structure, we substituted the oxygen atom in urea structure with CH_2 group trying to obtain a reference structure (ethane-1,1-diamine) for which there is no electronegative atom, mimicking resonance in structure Figure 5.3 (a). The initial structures were taken from the optimised urea structures from all three conformers, substituting oxygen atom with CH_2 and re-optimised using density functional B3LYP and MP2 method. After optimisation, generally, the structure were expanding with degree of NH_2 pyramidal broadened out and for the case of C_2 and C_s, it is rotating along C-N bond. Table 5.5 below shows detailed structure and torsion angles of the optimised structure for all three conformers.

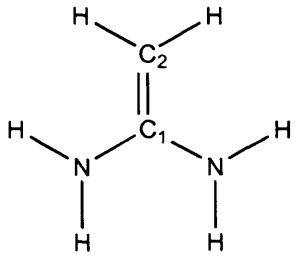
The expanding in size in the optimised structure is due to the bond lengthening for most of the bonds. For example, in C_2 conformer, C=C bond increases in 0.125Å, which is 10.3% of C=O bond because the later is more polar, and in CN bond it increases by 0.016Å (1.2%) in a symmetrical manner. This is about the same for C_{2v} conformer. However, for optimised C_s conformer, it is a little bit different where lengths of bonds are not symmetrical. This can be explain by the unevenly distribution of Mulliken charge (Table 5.6) in nitrogen atoms of both

sides. Further effect of this, causes CN bond of C_s conformer rotates with NH^1 moving downwards and NH^3 upwards from the planar during optimisation. The same occurrence also happened in C_2 but with symmetrical charge distribution resulting in having similar bond length after optimisation.

Table 5.5 : Geometry structures of ethane-1,1-diamine

Isomer	**C_2**		**C_{2V}**		**C_s**	
Method	B3LYP 6-311++G(2d2p)	MP2 6-311G++(2d,2p)	B3LYP 6-311++G(2d2p)	MP2 6-311G++(2d,2p)	B3LYP 6-311++G(2d2p)	MP2 6-311G++(2d,2p)
Bond Lengths (Å)						
CH	1.079	1.077	1.078	1.076	1.079	1.077
CC	1.341	1.344	1.348	1.350	1.343	1.345
CN	1.401	1.403	1.382	1.382	1.395,1.405	1.398,1.406
NH	1.008	1.007	1.000	0.998	1.005,1.008	1.004,1.007
NH'	1.011	1.009	1.000	0.998	1.005,1.012	1.005,1.009
Bond Angles (degrees)						
HCH	117.9	118.8	118.3	119.1	118.2	119.0
CCH	121.0	120.6	120.8	120.4	121.0, 120.8	120.6, 120.3
NCC	124.4	124.5	123.3	123.3	123.3, 123.8	123.2, 123.9
NCN	111.2	111.3	113.4	113.3	112.7	112.7
HNC	113.4	112.7	119.8	119.7	116.1, 112.8	114.8, 112.1
H'NC	113.1	112.3	122.0	121.8	116.3, 113.9	115.0, 113.1
HNH'	111.3	110.8	118.2	118.5	114.0, 110.6	112.9, 110.0
Dihedral Angles (degrees)						
$NCNH^2$	46.3	47.5	0.0	0.0	8.6	12.0
180° - [$NCNH^1$]	5.7	6.4	0.0	0.0	33.2	34.5
180° - [$CCNH^2$]	46.3	47.5	0.0	0.0	3.8	6.5
$CCNH^1$	5.7	6.4	0.0	0.0	38.0	40.1
$NCNH^4$	-46.3	-47.5	0.0	0.0	56.1	58.1
180° - [$NCNH^3$]	-5.7	-6.4	0.0	0.0	3.8	3.1
180° - [$CCNH^4$]	-46.3	-47.5	0.0	0.0	51.3	52.5
$CCNH^3$	-5.7	-6.4	0.0	0.0	1.6	2.5
NCCH	9.4	9.2	0.0	0.0	5.413, 8.774	4.4, 8.8

Table 5.6: MP2 Mulliken charge distribution of ethane-1,1-diamine where atomic charge is summed into heavy atoms and H charge is 0.



	C_s	C_{2v}	C_2
C1	0.358	0.443	0.430
C2	-0.275	-0.266	-0.294
N	-0.033	-0.092	-0.068
N	-0.049	-0.092	-0.068

Comparing this newly optimised structure with earlier urea molecule, we know that with lower polarisibility or atom charge, there is a possibility of NH_2 becoming more pyramidal. This is also increase the flexibility of CN bond where it has less semi-double bond characteristic. We also understand that in lower atomic charge, C_s conformer will have the tendency to restructure itself to be in C_2 conformation where this is the state of lowest energy conformation.

Energy difference of the *ab initio* calculated urea conformers is compared and discussed. From these three conformers, it was found that C_{2v} to be a saddle point which have two imaginary frequencies (Table 5.7), while C_2 and C_s to be featuring as stable conformers. A proper comparison was made based of energy calculated previously and the results are shown in Table 5.6. Energies are quoted relative to that of the C_2 conformation, where it is known having the lowest energy conformation among all three structures.

Table 5.7: Energy difference between urea conformers relative to C_2 conformation of urea conformers.

Method/Basis sets	Energy Difference (kJ mol^{-1})	
	C_{2v} - C_2	C_s - C_2
Literature (MP2 6-311G(d,p)) with ZPE[a]	7.11	5.15
MP2		
6-311G(d,p)	12.85	6.73
6-311G(d,p) + ZPE	6.23	4.91
6-311G(d,p) + Gibbs E	8.11	4.27
6-311++G(d,p)	11.49	5.03
6-311++G(d,p) + ZPE	4.39	3.70
6-311++G(d,p) + Gibbs E	6.14	3.22
6-311++G(2d,2p)	8.14	4.40
6-311++G(2d,2p) + ZPE	2.08	2.70
6-311++G(2d,2p) + Gibbs E	4.11	1.99
aug-cc-pVTZ[b]	5.95	3.93
6-311++G(d,p)[b]	n/a	5.02
6-311++G(3df,3pd)[b]	5.65	3.83
B3PW91		
6-311G(d,p)	5.54	4.37
6-311G(d,p) + ZPE	0.08	0.90
6-311G(d,p) + Gibbs E	2.25	-6.69
6-311++G(d,p)	-2.86	-3.73
6-311++G(d,p) + ZPE	-3.89	-2.47
6-311++G(d,p) + Gibbs E	-1.59	-5.59
6-311++G(2d,2p)	-0.83	-2.85
6-311++G(2d,2p) + ZPE	0.20	1.68
6-311++G(2d,2p) + Gibbs E	2.36	0.63
aug-cc-pVTZ[b]	3.66	2.92

6-311++G(3df,3pd)[b]	4.04	3.10
B3LYP		
6-311G(d,p)	5.86	4.52
6-311G(d,p) + ZPE	0.34	1.45
6-311G(d,p) + Gibbs E	2.49	-1.47
6-311++G(d,p)	4.21	3.32
6-311++G(d,p) + ZPE	-1.18	0.33
6-311++G(d,p) + Gibbs E	1.02	-2.57
6-311++G(2d,2p)	5.46	3.60
6-311++G(2d,2p) + ZPE	-0.06	1.52
6-311++G(2d,2p) + Gibbs E	2.10	0.43

[a]From reference 3

[b]From reference 15

The results describes the comparison of energy stability between each conformers compared to C_2 structure at different method and basis sets. MP2 method gave similar result with the literature [3] at 6-311G(d,p) and 6-311++G(d,p) basis sets. The difference between the literature is about 0.9 to 0.7 kJ mol^{-1} which can be considered as very small error difference. The addition of ZPE and Gibbs energy in the structure calculation increases the energy differences.

In the DFT method, B3PW91 gave poor result where in 6-311G(d,p) basis sets, the energy difference suggested that C_2 structure is higher in energy than C_{2v} and C_s, which is not true. However, this phenomena was explained by Dobrowolski [15] where he mentioned that even if relatively large basis sets are used, the shape of the most stable urea conformer is the subject of controversy: first, one can see that the result is a function of the method applied; second, the differences in energy are small enough to fall within the computational error. Therefore, at this stage of calculations, a conclusion that both urea forms are likely to be observed can be derived only. In B3LYP, small basis sets predicts similar energy difference compared to larger basis sets, 6-311G(d,p) and 6-311++G(2d,2p). This determines that C_2 in most cases of calculation is the lowest energy compared to C_{2v} and C_s.

As a conclusion, even though there are three distinct isomer structures in urea molecule, there is a possibility that these structure can easily change into other isomers as well due to small difference in energy difference. We could also imply that the planar urea structure, C_{2v} as second order saddle point connecting two pair of equivalent non-planar minima (C_s and C_2). The figure below (Figure 5.8) sketches the urea molecule conformation relative to its energy. The energies shown are based on Gaussian calculation method of MP2/6-311G(d,p) with corrected zero point energy.

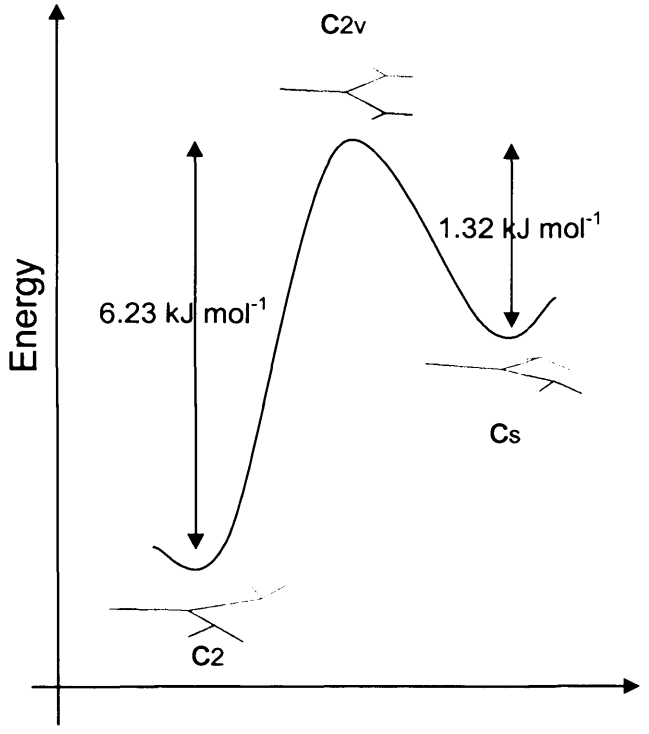


Figure 5.8: Energy difference of urea isomers at MP2/6-311G(d,p) + ZPE

As for the frequency calculation (Table 5.7), only C_{2v} structure exhibits the transition state with imaginary frequency for out of plane distortions. This implies that from all three conformations, C_{2v} is the transition state conformation, with two imaginary frequencies. Two negative vibrational frequencies for planar urea were also found in the literature [4] with the same vibration movement. The first transition state shows the vibration on the planar molecule with nitrogen atoms moving across the plane in opposite directions to one another similar to C_2. The second frequency represents vibration of the N and O atoms up and down the plane in the same direction approaching structure similar to C_s isomer. This shows that C_{2v} is a transition state where planar which will then minimise into C_2 and C_S structure.

Table 5.8: Imaginary frequency between different methods of calculation in C_{2v} conformation

Basis Set	Imaginary frequency	Frequency (cm^{-1})	
		1	2
MP2			
6-311G**	2	-542	-459
6-311++G**	2	-628	-554
6-311++G(2d,2p)	2	-483	-397
B3PW91			
6-311G**	2	-389	-264
6-311++G**	2	-378	-247
6-311++G(2d,2p)	2	-408	-316
B3LYP			
6-311G**	2	-397	-274
6-311++G**	2	-381	-252
6-311++G(2d,2p)	2	-404	-310

5.4 Urea Dimer

5.4.1 Basis Sets Effectiveness with Basis Sets Superposition Error (BSSE) Correction

Earlier, three methods of calculation (MP2 and two hybrid DFT functionals) were used in finding the energy of the urea monomer. In this section, calculations were carried out on a urea dimer (C_{2v}) structure to find the best method and basis sets size and whether to impose BSSE correction for use in building a urea crystal.

The energy of urea dimer were observed. This is to confirm on which basis sets to be used for calculating the urea clusters. Optimisation was done using B3LYP DFT 'tight' method enforcing on the symmetry with different basis sets as shown in Table 5.9. Interaction energy was taken into consideration in this dimer and earlier monomer calculation to compare results of different basis sets and determine on which is the best and most cost effective to be used. The optimisation of dimer also included calculation of its basis sets superposition error.

$$IE_{BSSE\ Corr} = \text{Dimer} - n(\text{monomer}) + BSSE \qquad (5.1)$$

Simulations were done using planar (C_{2v}) conformations. Starting point of the structure used are from crystal urea molecular arrangement, which is shown in Figure 5.9.

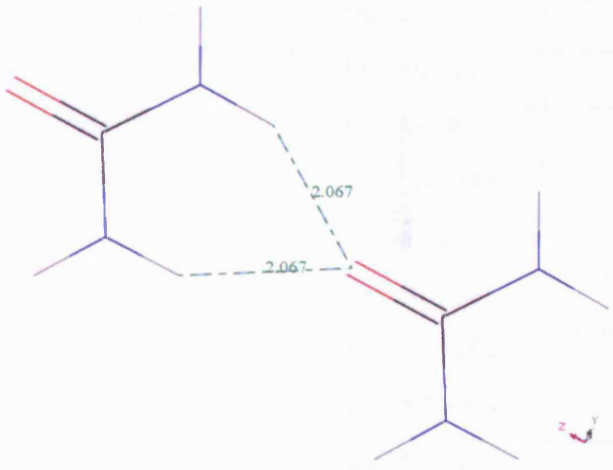


Figure 5.9: Hydrogen bond of the starting point in the urea dimer optimisation

To account for the BSSE, a counterpoise calculation was calculated on the optimised structure. The calculations were done in a larger range of basis sets, which is shown in Table 5.9.

Table 5.9: Basis sets with BSSE correction to urea dimer

B3LYP Basis Sets	Dimer Energy (Hartree)		Interaction Energy (kJ mol^{-1} per molecule)	
	Uncorrected	BSSE Corrected	Uncorrected	BSSE Corrected
6-31G	-450.408562	-450.403936	-26.1	-21.3
6-31+G(d,p)	-450.594440	-450.593690	-18.3	-18.2
6-31++G(d,p)	-450.594577	-450.593819	-17.8	-17.9
6-31G(2d,2p)	-450.573602	-450.568000	-22.7	-16.4
6-31+G(2d,2p)	-450.611634	-450.611145	-17.2	-17.4
6-31++G(2d,2p)	-450.611852	-450.611367	-17.2	-17.4
6-311G	-450.540353	-450.536954	-24.7	-21.4
6-311+G	-450.558106	-450.557099	-21.9	-21.7
6-311++G	-450.558423	-450.557384	-21.9	-21.7
6-311G(d)	-450.664763	-450.661407	-21.3	-17.9
6-311G(d,p)	-450.688896	-450.685549	-21.1	-17.7
6-311+G(d,p)	-450.710002	-450.709602	-17.8	-18.2
6-311++G(d,p)	-450.710138	-450.709713	-17.8	-18.1
6-311G(2d,2p)	-450.702004	-450.698346	-20.4	-16.5
6-311G+(2d,2p)	-450.722322	-450.721906	-16.8	-17.1
6-311G++(2d,2p)	-450.722396	-450.721983	-16.9	-17.1
6-311G(2df,2pd)	-450.718128	-450.714631	-20.3	-16.7
6-311+G(2df,2pd)	-450.737766	-450.737311	-17.0	-17.2
6-311G++(2df,2pd)	-450.737870	-450.737417	-17.0	-17.2

With the corrected basis sets, dimer energy produced are much lower than the previous. The effect of BSSE on binding energy of urea dimer molecule is between 4.8 to 0.2 kJ mol^{-1} and it varies with different basis sets. As larger basis sets were used, the difference between uncorrected and corrected BSSE on urea dimer binding energy becomes smaller.

Adding an extra valence, from 6-31G to 6-311G lowered the interaction energy, by 0.10 kJ mol^{-1}, which is about 0.47%. Diffuse functions, on the other hand, increase the interaction energy by 1.12% of the energy. It shows quite a significant effect with adding one

function, from the minimal basis. However, by adding a second diffuse function to the basis sets, only impacting the energy difference at about 0.02 kJ mol^{-1}, which is only about 0.09%.

Upon adding polarisation on the heavy atom (d), havthe energy difference is 3.56 kJ mol^{-1} and by adding polarisation to the hydrogen (d,p) the energy differs by 3.73 kJ mol^{-1}. This shows a very significant effect to the basis function used. Adding larger polarisation (2d,2p; 2df,2pd) will only have the difference about 1.22 kJ mol^{-1}. This is because, the lower basis function is already approaching its convergence limit.

The convergence limit of this series is about 17.21 kJ mol^{-1} which is calculated using 6-311G++(2df,2pd). However, this calculation takes a long time (8 hours), for only two molecules. Therefore, the basis sets suggested to be used for further investigation is B3LYP/6-311++G(2d,2p) and B3LYP/6-311++G(d,p) due to their cost effectiveness and energy results obtained, or a lower basis which have the same accuracy, 6-31+G(2d,2p).

Upon adding larger basis sets, the BSSE is lowered. However, at a certain basis sets size, BSSE is too small (0.2 kJ mol^{-1}) and can be ignored. This is shown in 6-31+G(2d2p) and 6-31++G(2d2p) basis sets. However, this basis sets are considered as imbalanced basis sets. A more balanced basis sets which have very small BSSE is shown in 6-311++G with difference between the corrected and uncorrected BSSE binding energy at 0.36 kJ mol^{-1}. A larger basis sets that were considered to have good agreement with geometry literature (Table 5.3) and energy converged (Table 5.2) in the earlier discussion is at (6-311++G(2d,2p) where the difference between corrected and uncorrected BSSE is also small, 0.2 kJ mol^{-1}. We decided to use this basis sets type without BSSE applied since it is quite a small value.

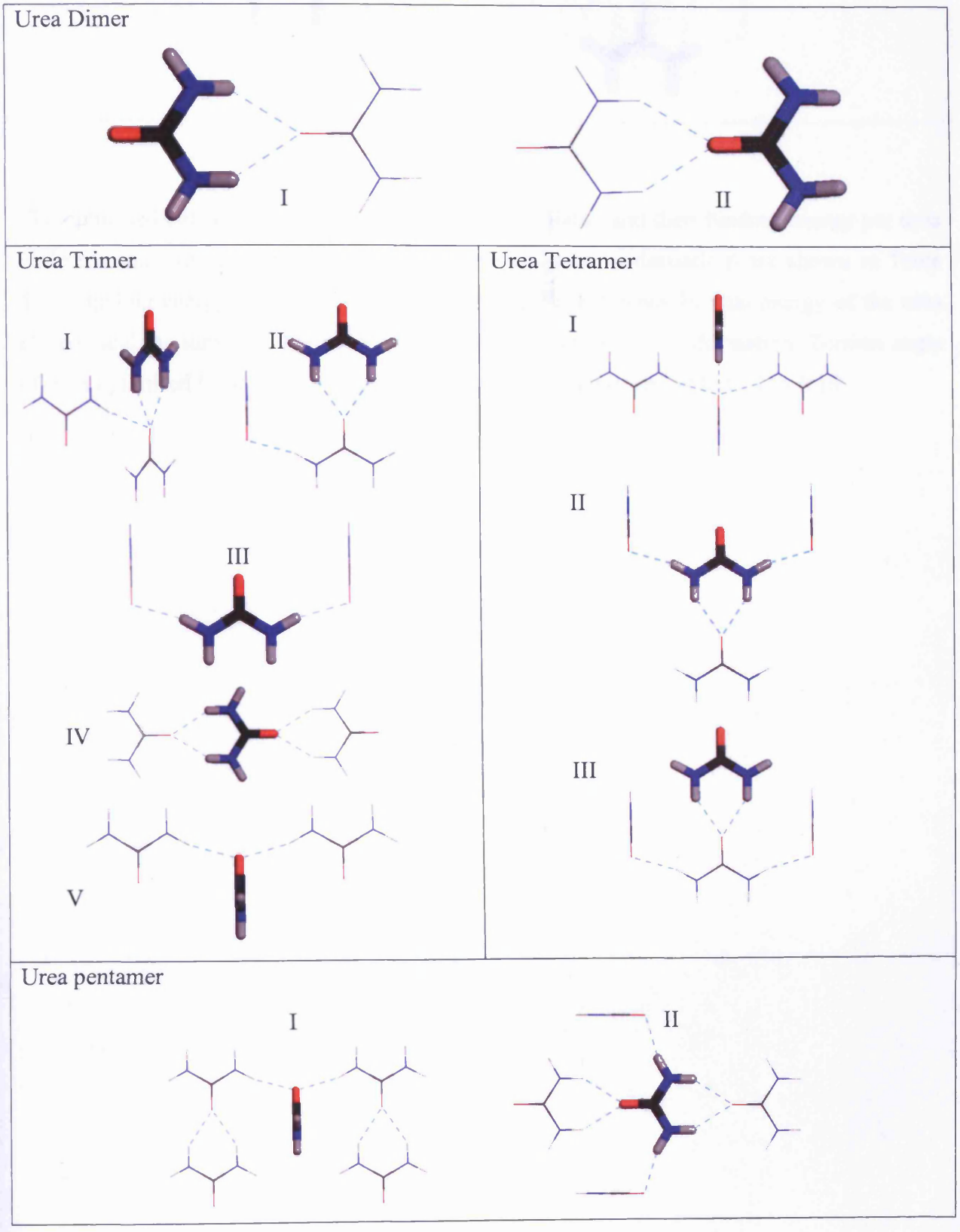
5.5 Urea Oligomers

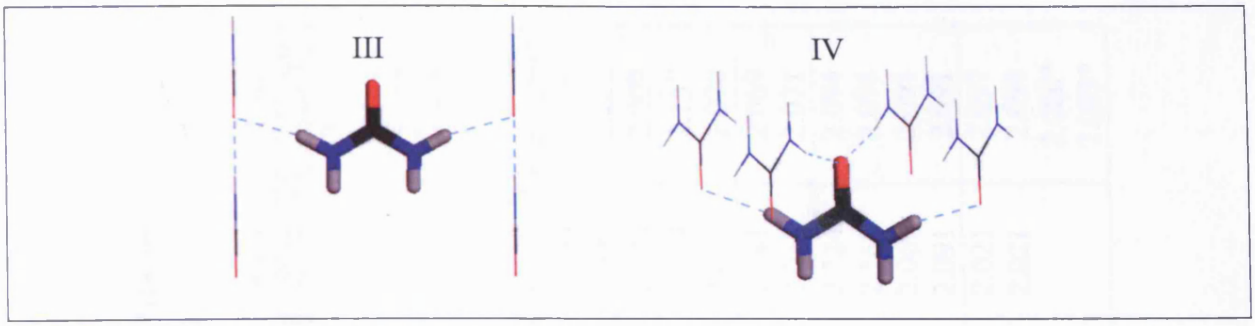
Several works on urea oligomers have been performed previously, but for different purposes [6,9,16]. As shown in Figure 5.6, it was suggested that 14 neighbouring urea molecules have vibration effect to a centre urea molecule. This was also suspected that it also contributed to the formation of flat urea molecule in urea crystal.

Earlier optimisation observation of urea dimer with 'symmetry off' calculation showed that urea will naturally optimise to become urea clusters. However, in this section, we will constrain all urea planar except for the one that will be used for study, whether it will form a flat urea. This urea will be in either C_2 or C_s symmetry. Optimisation will be done using B3LYP method with 6-311++G(2d,2p) basis set calculation.

The construction of urea structure from dimer to pentamer is shown in Figure 5.10. Position of urea C_s and C_2 is illustrated by the bold structure.

Figure 5.10: Initial structure of urea oligomers. Bold urea molecule will be substituted with either C_2 of C_s structure.





The optimised structure of all urea clusters was calculated and their binding energy per urea molecule and hydrogen bond properties (before and after optimisation) are shown in Table 5.10. Binding energy is determined by calculating the difference in total energy of the urea clusters and the same number of isolated urea monomer in C_{2v} conformation. Torsion angle of each optimised C_2 and C_s structure are compared and listed in Table 5.11 to 5.15.

Table 5.10: Energy profiles and hydrogen bond properties of optimised urea oligomers

Structure Type	Energy (Ha)		Binding Energy per mol (kJ mol^{-1})		No. of Total H-bonds	No of Rigid H-bonds	Non-rigid Hbond[a] (Å)			
							Initial	End	Initial	End
Dimer	**C_2**	**C_s**	**C_2**	**C_s**			**C_2**		**C_s**	
I	-450.72091	-450.72097	-9.45	-9.54	2	0	2.148 2.149	2.153 2.152	2.099 2.099	2.132 2.134
II	-450.72097	-450.72212	-9.53	-11.05	2	0	2.102 2.102	2.171 2.171	2.101 2.101	2.157 2.233
Trimer	**C_2**	**C_s**	**C_2**	**C_s**			**C_2**		**C_s**	
I	-676.08263	-676.08264	-10.55	-10.57	3	1	2.149 2.148 2.009	2.090 2.314 2.009	2.099 2.099 2.099	2.116 2.202 2.009
II	-676.09001	-676.09001	-17.02	-17.02	3	1	2.138 2.138	2.155 2.070	2.099 2.099	2.157 2.070
III	-676.07668	-676.07590	-5.35	-4.66	2	0	2.055 2.055	2.063 2.063	2.041 2.038	2.069 2.071
IV	-676.09428	-676.09389	-20.75	-20.41	4	0	2.100 2.081 2.270 2.118	2.080 2.078 2.095 2.098	2.224 2.166 2.069 2.091	2.094 2.094 2.091 2.091
V	-676.09702	-676.09702	-23.15	-23.15	2	0	2.021 2.021	2.095 2.095 2.072* 2.079*	2.021 2.021	2.097 2.094 2.080* 2.089*

Structure Type	Energy (Ha)		Binding Energy (kJ mol^{-1})		No. of Total H-bonds	No of Rigid H-bonds	Non-rigid Hbond[a] (Å)			
							Initial	End	Initial	End
Tetramer	C_2	C_s	C_2	C_s			C_2		C_s	
I	-901.44108	-901.44112	-8.97	-8.99	4	2	2.149 2.148	2.291 2.291	2.099 2.099	2.194 2.193
II	-901.43558	-901.43627	-5.35	-7.74	4	0	2.055 2.055 2.149 2.148	2.021 2.020 2.106 2.106	2.040 2.038 2.099 2.099	2.106 2.106 2.193 2.193
III	-901.45221	-901.45287	-16.27	-16.71	4	2	2.148 2.149	2.037 2.037	2.099 2.099	2.068 2.025
Pentamer	C_2	C_s	C_2	C_s			C_2		C_s	
I	-1126.80569	-1126.80537	-11.24	-11.07	6	4	2.021 2.021	2.019 2.019	2.021 2.021	2.015 2.020
II	-1126.80632	-1126.80628	-11.57	-11.55	6	0	2.055 2.055 2.102 2.102 2.148 2.149	1.995 1.995 2.034 2.034 2.068 2.068	2.038 2.040 2.101 2.101 2.099 2.099	1.990 1.990 2.035 2.035 2.071 2.070
III	-1126.80328	-1126.80374	-9.97	-10.22	6	4	2.055 2.055	2.062 2.062	2.040 2.040	2.100 2.100
IV	-1126.79272	-1126.79265	-4.43	-4.40	4	0	2.021 2.021 2.055 2.055	2.041 2.041 2.026 2.026	2.022 2.021 2.038 2.040	2.040 2.037 2.019 2.018

* Hydrogen bond formed during optimisation

[a] Rigid H-bond is not included in the data

Urea dimer optimisation of both Type I and II resulted in elongation of the hydrogen bonds in the structure. Binding energy for Type I and II of C_2 structures are similar (~9.5 kJ mol^{-1}). However, for C_s structure, Type II gave lower binding energy, which implies a more preferable structure. However, this is also might due to the elongated hydrogen bonds and the large movements urea pyramidal during optimisation which produces this low energy (Figure 5.11).

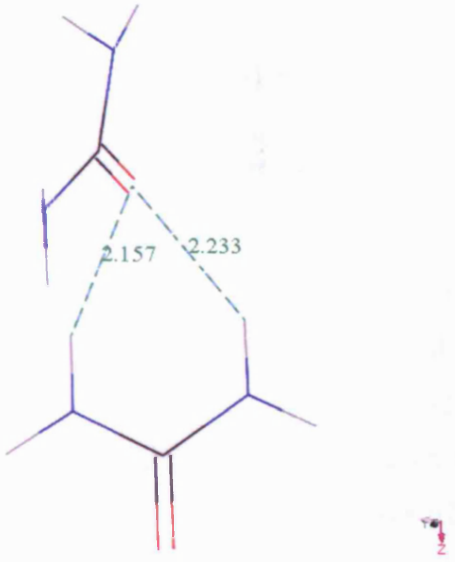


Figure 5.11 Optimised structure of Type II C_s conformers. Note that the optimised conformers have changed to C_2 type conformer.

Table 5.11: Optimised structure of urea monomer of C_2 and C_s isomer

	C_2	C_s
$NCNH^2$	29.0	23.0
180° - [$NCNH^1$]	13.4	11.0
180° - [$OCNH^2$]	29.0	18.9
$OCNH^1$	13.4	15.1
$NCNH^4$	-29.0	23.0
180° - [$NCNH^3$]	-13.4	11.0
180° - [$OCNH^4$]	-29.0	18.9
$OCNH^3$	-13.4	15.1

The original optimised conformer geometry (Table 5.11) is compared with the optimised dimer geometry (Table 5.12). Torsion angle of pyramidal urea decreased after optimisation. This suggests that with increase of urea neighbours and with increasing number appropriate number of hydrogen bonds, the pyramidal urea will eventually becomes a planar molecule. Type II structure of C_s conformer changed its structure to C_2 type of conformation, but with smaller torsion angle.

Table 5.12: Optimised geometry structure of urea dimer

	Type I		**Type II**	
	C_2	**C_s**	**C_2**	**C_s**
$NCNH^2$	17.48	22.30	23.60	-22.85
180° - [$NCNH^1$]	12.49	13.29	13.62	-12.52
180° - [$OCNH^2$]	17.48	18.88	23.60	-22.92
$OCNH^1$	12.49	16.71	13.62	-12.45
$NCNH^4$	-17.47	22.30	-23.60	24.25
180° - [$NCNH^3$]	-12.48	13.25	-13.62	12.93
180° - [$OCNH^4$]	-17.47	18.88	-21.60	24.15
$OCNH^3$	-12.48	16.67	-13.20	13.00

Table 5.13: Optimised geometry structure of urea trimer

	Type I		**Type II**	
	C_2	**C_s**	**C_2**	**C_s**
$NCNH^2$	18.89	18.42	22.64	22.65
180° - $[NCNH^1]$	20.73	10.38	12.09	12.07
180° - $[OCNH^2]$	18.11	14.56	18.46	18.47
$OCNH^1$	21.51	14.23	16.27	16.25
$NCNH^4$	-12.08	26.15	23.00	22.96
180° - $[NCNH^3]$	-14.90	14.04	19.26	19.25
180° - $[OCNH^4]$	-12.87	22.32	18.85	18.82
$OCNH^3$	-14.11	17.87	23.40	23.39
	Type III		**Type IV**	
	C_2	**C_s**	**C_2**	**C_s**
$NCNH^2$	22.60	18.46	8.60	-17.25
180° - $[NCNH^1]$	15.57	10.98	15.45	-10.64
180° - $[OCNH^2]$	22.60	14.68	8.60	-14.44
$OCNH^1$	15.57	14.75	15.45	-13.44
$NCNH^4$	-22.60	18.42	-8.44	-17.24
180° - $[NCNH^3]$	-15.57	10.96	-15.38	-10.41
180° - $[OCNH^4]$	-22.60	14.65	-8.44	-14.44
$OCNH^3$	-15.57	14.73	-15.37	-13.22
	Type V			
	C_2	**C_s**		
$NCNH^2$	15.87	-15.80		
180° - $[NCNH^1]$	10.66	-10.62		
180° - $[OCNH^2]$	15.88	-15.82		
$OCNH^1$	10.65	-10.60		
$NCNH^4$	-15.88	15.90		
180° - $[NCNH^3]$	-10.64	10.70		
180° - $[OCNH^4]$	-15.87	15.88		
$OCNH^3$	-10.64	10.72		

In urea trimer structures, all five types of clusters were compared. Among all five structures, Type V is the most preferable structure, followed by Type IV. In Type V structure, two new hydrogen bonds were formed during optimisation. This is due to the twisting of pyramidal urea molecule which then formed two new hydrogen bonds with its neighbouring urea. Each side of NH^x (designated 1 and 3) acted as a hydrogen bond donor (Figure 5.12). Optimisation of C_s structure also ended as a C_2 conformation type structure where this is a more stabilized structure compared to C_s. The addition of hydrogen bond resulted in lowering the binding energy of this structure (23.15 kJ mol^{-1}). This might be due to the resonance hydrogen bond in urea ribbon where it enhances the role of the well known resonance structures of urea. Geometry of optimised structure in Type V are both about 15.9° and 10.6° degree to planarity (Table 5.13).

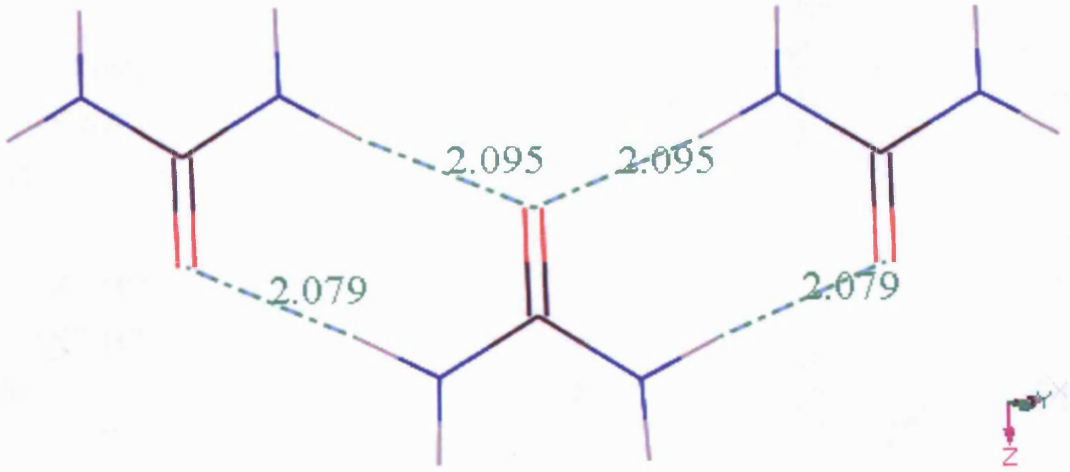


Figure 5.12: Optimised Type V trimer urea structure. The structure formed is similar to a urea ribbon type of structure.

In Type IV structure, four hydrogen bonds were also involved during optimisation, two on oxygen atom which acted as a hydrogen bond acceptor and one at each NH^x (designated 2 and 4) as hydrogen bond donor. Type IV can also be described as a linear type of urea structure. After optimisation, the linear structure didn't change except for the planarity of pyramidal urea. Pyramidal urea becomes more planar where this effect mostly in C_2 structure conformer. The degree of NH_2 pyramidal structure decreased by 70% in C_2 conformer and at

25% in C_s where this causes the hydrogen bond length to shortened (Table 5.13). In C_2 conformer, the structure flipped its pyramidal direction but kept the C_2 type of conformer.

In Type III trimer urea structure, the energy differs at quite a large value than other structure (~-5 kJ mol^{-1}) for both type of conformation because during optimisation, the structure tried to twist itself to be parallel to the neighbouring structure. However, this was restricted due to the two hydrogen bonds attached to the molecule. This also explains why the hydrogen bond lengths became longer from the initial structure.

Table 5.14: Optimised geometry structure of urea tetramer

	Type I		**Type II**	
	C_2	**C_s**	**C_2**	**C_s**
$NCNH^2$	16.02	24.69	9.72	20.77
180° - [$NCNH^1$]	16.78	10.40	20.82	14.51
180° - [$OCNH^2$]	16.02	21.16	9.72	17.23
$OCNH^1$	16.79	13.92	20.82	18.06
$NCNH^4$	-16.01	24.69	-9.72	20.75
180° - [$NCNH^3$]	-16.79	10.40	-20.82	14.51
180° - [$OCNH^4$]	-16.02	21.17	-9.72	17.21
$OCNH^3$	-16.78	13.92	-20.82	18.05

	Type III	
	C_2	**C_s**
$NCNH^2$	9.18	24.74
180° - [$NCNH^1$]	21.11	16.15
180° - [$OCNH^2$]	9.18	20.91
$OCNH^1$	21.12	19.97
$NCNH^4$	-9.18	24.91
180° - [$NCNH^3$]	-21.12	11.15
180° - [$OCNH^4$]	-9.18	21.07
$OCNH^3$	-21.12	14.99

In urea tetramer, Type III is more energy favourable (16.27 kJ mol^{-1}) then the other two structures. This happens because during optimisation, urea pyramidal molecule twisted its arrangement to be parallel to its side structure (Figure 5.13). However, in other structures (Type I and II) which have more hydrogen bonds attached to the pyramidal urea didn't produced such low energy result because the positioning of the restricted neighbouring structures. In Type II structure, the movement of the pyramidal urea is restricted due to its attachment with four hydrogen bonds to its neighbour.

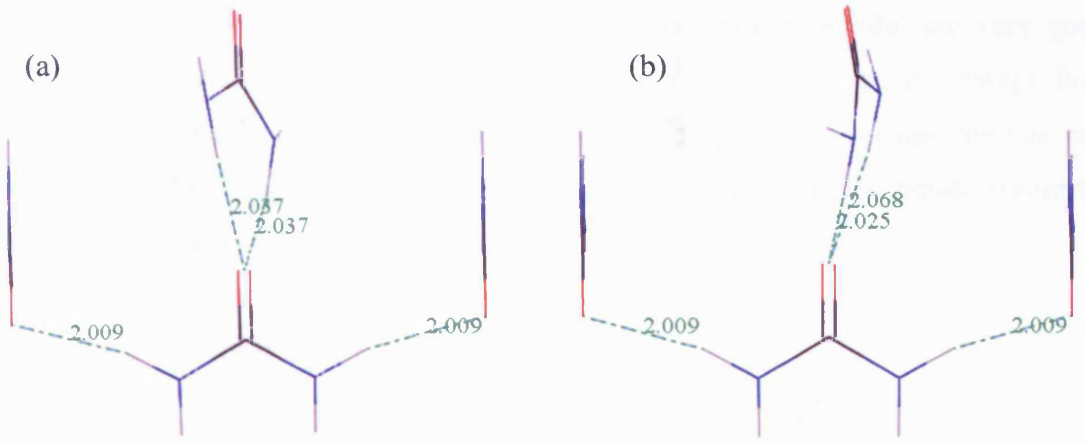


Figure 5.13: Optimised urea tetramer structure of Type III (a) C_2 and (b) C_s conformer.

As for geometry of urea pyramidal, all structures show that the pyramidality is lowered to a more planar-like structure (Table 5.14). We suggested that if it is not due to the hydrogen bonds, this may be due to the intermolecular forces between urea molecules which induce the planar-type change of structure. Type II and III shows that the optimised structure minimised the pyramidal angle, especially for C_2 conformer.

Pentamer urea structures show quite a drastic improvement on the planarity of urea pyramidal (Table 5.15). This especially occurs in C_s starting conformers of Type IV where the average degree of planarity is 0.3°. The structure of C_s Type IV structure eventually turned to a C_2 structure type after optimisation (Figure 5.14). This may be due to the unstability of C_s structure compared to C_2 where C_2 is more symmetrical in geometry and of charge distribution. In terms of binding energy, Type IV structures (both C_2 and C_s) are quite small that other structure type.

In Type I and II pentamer structure, the planarity of urea pyramidal are very good, with average at 4.26° and 4.76° respectively. Both shows that the structure are energy favourable (11-11.5 kJ mol^{-1}) compared to other structure type. This might be because the electrons were fully localised due to the urea structure arrangements and hydrogen bonds connected to it, especially in Type II structure.

Figure 5.14: Urea planar structure (bold) of optimised C_s Type IV structure.

Table 5.15: Optimised geometry structure of urea pentamer

	Type I		**Type II**	
	C_2	**C_s**	**C_2**	**C_s**
$NCNH^2$	19.19	5.77	5.00	6.15
$180° - [NCNH^1]$	9.53	2.60	10.34	2.97
$180° - [OCNH^2]$	12.38	4.61	5.00	4.99
$OCNH^1$	7.15	3.76	10.34	4.11
$NCNH^4$	12.38	5.97	-5.00	6.37
$180° - [NCNH^3]$	7.15	2.71	-10.34	3.47
$180° - [OCNH^4]$	12.38	4.81	-5.00	5.22
$OCNH^3$	7.15	3.87	-10.34	4.63
	Type III		**Type IV**	
	C_2	**C_s**	**C_2**	**C_s**
$NCNH^2$	17.63	22.45	12.38	0.54
$180° - [NCNH^1]$	19.91	14.86	7.15	0.29
$180° - [OCNH^2]$	17.63	18.33	12.38	0.51
$OCNH^1$	19.91	18.98	7.15	0.32
$NCNH^4$	-17.62	22.47	-12.38	-0.22
$180° - [NCNH^3]$	-19.91	14.88	-7.15	-0.12
$180° - [OCNH^4]$	-17.62	18.35	-12.38	-0.25
$OCNH^3$	-19.91	19.00	-7.15	-0.09

5.5.1 Analysis on the planarity of urea

In this section, we will discuss on the effect of number of hydrogen bonds and urea molecule structures to the planarity of urea conformer. Degree of planarity from each structure type was plotted against number of hydrogen bonds as shown in Figure 5.15 and 5.16.

Figure 5.15: Degree of planarity of C_s conformer against different number of hydrogen bond. (0° indicate planar urea structure)

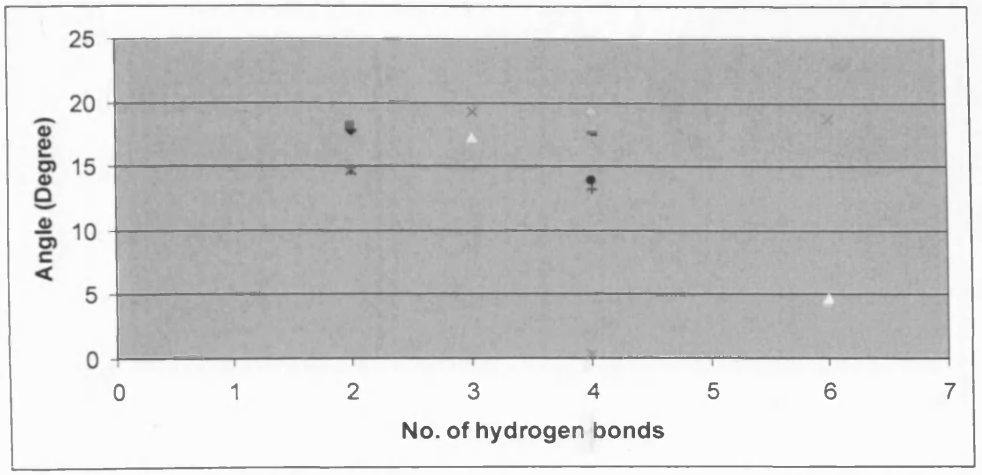
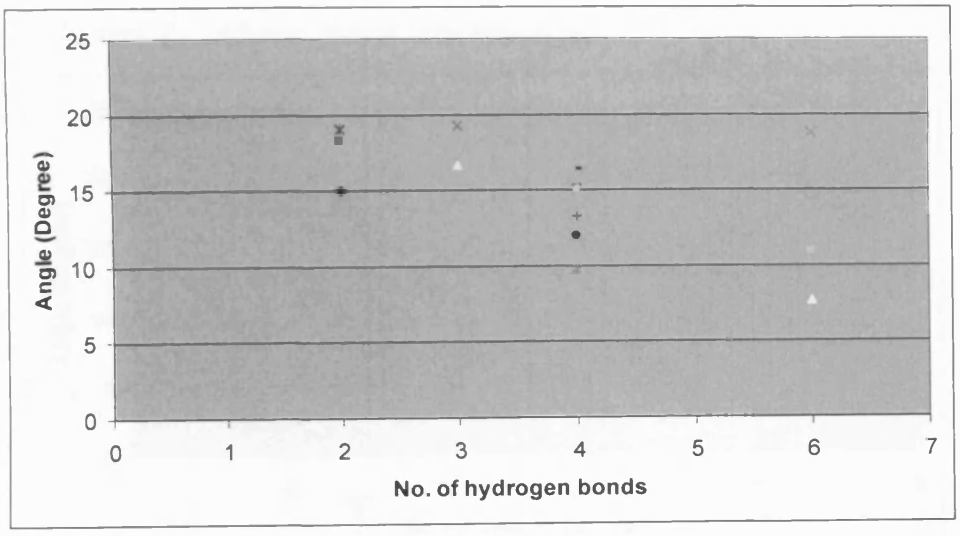


Figure 5.16: Degree of planarity of C_2 conformer against different number of hydrogen bond. (0° indicate planar urea structure)



Both plots demonstrate that the urea structure becomes more planar with increasing number of hydrogen bonds. However, this must also followed by the correct structure arrangements, where degree of planarity will become more effective if the hydrogen bonds are connected directly to the pyramidal molecule (Table 5.10).

Figure 5.17: Degree of planarity of C_2 conformer against different number of urea molecule clusters. (0° indicate planar urea structure)

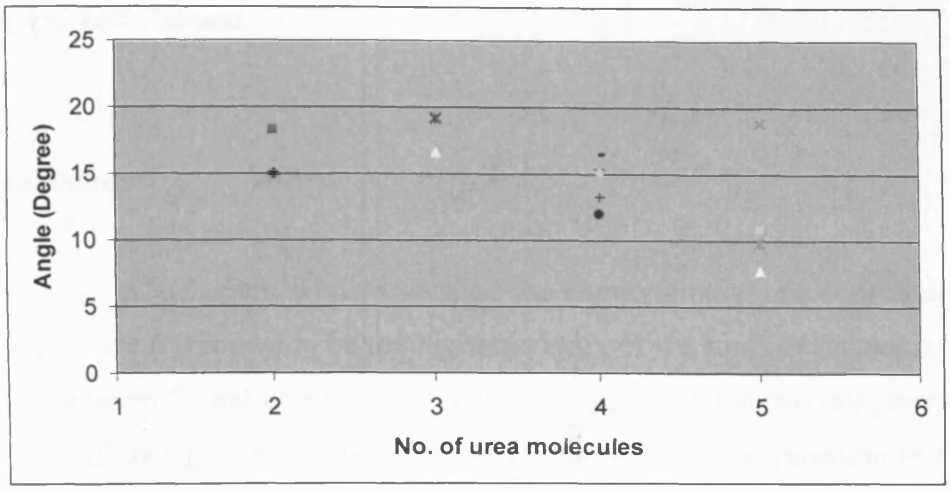


Figure 5.18: Degree of planarity of C_s conformer against different number of urea molecule clusters. (0° indicate planar urea structure)

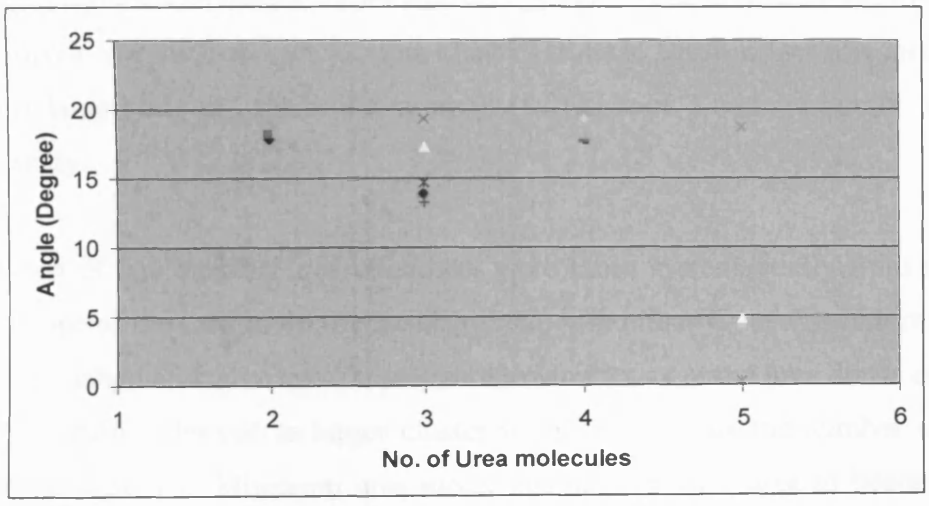


Figure 5.17 shows that the degree of planarity of C_2 conformer increases as the number of urea molecule clusters increases. On the contrary, C_s conformer does not show this type of responses (Figure 5.18). This might be due to the structure arrangements of the urea clusters and also based on the number of hydrogen involves in the optimisation.

From the data and discussion presented in Section 5.5, we can conclude that the barrier to planarisation is systematically lowered by the hydrogen bond formation and application of external electric fields (number of neighbouring molecules). These are the most important

factors in achieving planarisation of urea molecule. However, the existence or formation of hydrogen bonds will indicate lower binding energy in the crystal system, as manifest in Type V of urea trimer optimisation.

5.6 Conclusion

In the first part of this chapter, we investigated the energy conformation of urea monomer, C_2, C_{2v}, and C_s. Urea C_{2v} shown to be the highest energy of the conformers and is a transition conformation between C_2 and C_s structure. This is confirmed by the two imaginary frequency found in C_{2v} confirmation. The higher imaginary frequency shows transition to C_2 structure while the smaller frequency shows transition to C_s structure. Geometry of each urea conformer were also discussed and are similar to the conformation obtained in the literature.

In the second part of this study, we investigate the best basis sets suitable for use in larger urea molecule clusters. The results shows that energy converged at B3LYP/6-311++G(2d,2p) and we decided to use this basis set for urea clusters studies. BSSE effect was also tested and found that at large basis set, the value is small (0.2 kJ mol^{-1}) will be ignore for the next calculation study.

In the last part of this chapter, urea oligomers were build systematically from urea crystal structure and one of the urea molecule is substituted with either C_2 or C_s conformation. This is to study the change of degree upon approaching planarity as in the urea dense crystal. Urea planarity was finally achieved as larger cluster were used and as the number of hydrogen bonds involved increased. Minimum urea molecules needed for a urea to become planar is 5 urea with 4 hydrogen bonds.

5.7 References

(1) King, S. T. *Spectrochimica Acta* **1972**, *28A*, 165-175.

(2) Notario, R.; Castaño, O.; Herreros, M.; José-Luis; Abboud, M. *Journal of Molecular Structure: THEOCHEM* **1996**, *371*, 21-29.

(3) Spoliti, M.; Pieretti, A.; Bencivenni, L.; Sanna, N. *Electronic Journal of Theoretical Chemistry* **1997**, *2*, 149-159.

(4) Masunov, A.; Dannenberg, J. J. *Jounal of Physical Chemistry A* **1999**, *103*, 178-184.

(5) Godfrey, P. D.; Brown, R. D.; Hunter, A. N. *Journal of Molecular Structure* **1997**, *Volumes 413-414*, 405-414

(6) Lee, C.; Stahlberg, E. A.; Fitzgerald, G., 1995; Vol. 99; pp 17737-17741.

(7) Ishida, T.; Rossky, P. J.; Castner, E. W. *The Journal of Physical Chemistry B* **2004**, *108*, 17583-17590.

(8) Masunov, A.; Dannenberg, J. J., 2000; Vol. 104; pp 806-810.

(9) Spoliti, M.; Perrone, G.; Bencivenni, L.; Pieretti, A.; Grandi, A.; Ramondo, F. *Journal of Molecular Structure: THEOCHEM* **2005**, *756*, 113-126.

(10) Rousseau, B.; Keuleers, R.; Desseyn, H. O.; Geise, J. J.; Alsenoy, C. V. *Chem. Phys. Lett.* **1999**, *302*, 55-59.

(11) Bleckmann, P.; Thibud, M. *Journal of Molecular Structure* **1990**, *219*, 7-112.

(12) Brown, R. D.; Godfrey, P. D.; Storey, J. *J. Mol. Spect.* **1975**, *58*, 445-450.

(13) Engkvist, O.; Price, S. L.; Stone, A. J. *Phys. Chem. Chem. Phys.* **2000**, *2*, 3017-3027.

(14) Docherty, R.; Roberts, K. J.; Saunders, V.; Black, S.; Davey, R. J. *Faraday Discussion* **1993**, *95*, 11-25.

(15) Dobrowolski, J. C.; Koos, R.; Sadlej, J.; Mazurek, A. P. *Vibrational Spectroscopy* **2002**, *29*, 261-282

(16) Skurski, P.; Simons, J. *The Journal of Chemical Physics* **2001**, *115*, 10731-10737.

Chapter 6

Conclusions

6.1 Conclusions

DMA approach has been used to model the crystal structures of a range of single molecules and co-crystals, using experimental data for lattice energies, heats of sublimation and unit cell parameters as tests of this approach. For molecules containing only first-row atoms, this worked well with predictions of lattice energies inside the range of experimental data and small deviations from observed crystal parameters. However, for sulfur-containing compounds the performance was considerably worse, which was assigned to problems in the Lennard-Jones parameters for S...S and S...H interactions. Optimisation of these parameters remains an unsolved problem at this stage, and is an area for further research.

New geometries of the non-experimentally observed co-crystals were constructed by modification of known structures. DMAREL approach was used to calculate interaction energies for different polymorphs of these to rationalise the fact that some structures are not observed. The method has succeeded in generating the missing urea/1,8-dho anti-parallel urea ribbon type of co-crystal. The parallel urea ribbon diol co-crystal series is an interesting part for future research to compare both type of polymorph. New polymorph of CA based on the structure of TTCA was also predicted, but the reverse model was less successful. This work suggests that an experimental search for new polymorphs of CA might be a fruitful avenue of research. However, in the absence of reliable potentials for sulfur, predictions for TTCA are not considered to be highly reliable.

Ab initio studies were used for building urea clusters in understanding the formation of planar urea (C_{2v}) in solid crystal from its pyramidal structure (C_2 and C_s) in gas phase. Planar structure was successfully obtained upon increasing the number of neighbouring planar urea mimicking the structure of urea dense crystal. C_s shows a higher probability in becoming planar compared to C_2 conformation. In the near future, simulation of a urea with all eight hydrogen bonds should also be look into and this can be compared to the urea dense crystal.

Appendixes

Appendix 1

Initial pair potential used

Pair Potential	Min R (Å)	Min E (kJ mol^{-1})	lambda
CA CA	3.8905	0.3875	13.5
HY HY	3.3666	0.0414	13.5
NI NI	3.6986	0.6260	13.5
CA NI	3.7946	0.4925	13.5
CA HY	3.6285	0.1267	13.5
NI HY	3.5326	0.1610	13.5
HY HP	3.3666	0.0413	13.5
CA HP	3.2494	0.1267	13.5
NI HP	2.7498	0.1610	13.5
HP HP	3.3666	0.0414	13.5
OX OX	3.6096	0.3347	13.5
CA OX	3.7501	0.3601	13.5
OX NI	3.6541	0.4577	13.5
OX HY	3.4881	0.1177	13.5
OX HP	2.7558	0.1177	13.5

Initial pair potentials in DMAREL format.

```
BUCK CA  CODA  CA  CODA
2343.571908  0.288185 25.067891  0.000000 30.000000
ENDS
 BUCK CA  CODA  OX  CODA
2177.858694  0.277785 18.685030  0.000000 30.000000
ENDS
 BUCK CA  CODA  HP  CODA
766.272414  0.240696 2.782333  0.000000 30.000000
ENDS
 BUCK CA  CODA  NI  CODA
2978.604296  0.281081 27.429370  0.000000 30.000000
ENDS
 BUCK OX  CODA  OX  CODA
2024.241335  0.267377 13.810956  0.000000 30.000000
ENDS
 BUCK OX  CODA  NI  CODA
2768.136418  0.270674 20.327177  0.000000 30.000000
ENDS
 BUCK OX  CODA  HP  CODA
711.841067  0.204133 0.961781  0.000000 30.000000
ENDS
 BUCK NI  CODA  NI  CODA
3786.002617  0.273970 29.895933  0.000000 30.000000
ENDS
 BUCK NI  CODA  HP  CODA
973.716328  0.203688 1.298512  0.000000 30.000000
ENDS
 BUCK HP  CODA  HP  CODA
250.384198  0.249377 1.124509  0.000000 30.000000
ENDS
 BUCK HY  CODA  HY  CODA
250.384198  0.249377 1.124509  0.000000 30.000000
```

```
ENDS
 BUCK CA   CODA   HY   CODA
766.272414  0.268777 5.394521  0.000000 30.000000
ENDS
 BUCK NI   CODA   HY   CODA
973.716328  0.261674 5.837222  0.000000 30.000000
ENDS
 BUCK OX   CODA   HY   CODA
711.841067  0.258377 3.954791  0.000000 30.000000
ENDS
 BUCK HP   CODA   HY   CODA
249.779405  0.249377  1.12179  0.000000 30.000000
ENDS
```

Appendix 2: Creating DMAREL input file

Creating dmain file

1) Create a geom file using neighbour (dave_neigh)
2) Create multipoles file (punch file)
3) Create standard bonds file

Run using neighbour (neigh)

1) Insert geom. File
2) Insert standard bond file
3) Choose option 2 for input file
4) Maximum required inter-molecular contact: 2.0
5) Input c-vector: 1.0 (values in Å)
6) Do you want to standardise bond lengths to hydrogen: Yes
7) Do you want to force any molecules planar? : Yes if contain file contain urea. Put atom molecules which represent urea. If no urea, select No.
8) Do you wish to insert any bond centre sites (Y/N) ?: No
9) Do you have a punch file yet: Yes
10) Did you use Gamess or Cadpac (G/C)?: C (punch file format used)
11) Insert punch file name
12) Insert dmain input file if you wish to change the input name
13) Insert axis definition file
14) Besides creating an input file for DMAREL, neighbours also created 4 atom connectivity files labelled 'filename.nnl, *.mac, *.ccl and *.nem

Appendix 3

GDMA input file

Title "Thiourea13 in Gaussian 03 MP2 for multipoles"

```
FILE turea13.fchk
Angstrom
Multipoles
 Limit 4
 Punch turea13_full.pun
Start

Finish
```

Appendix 4

Awk scripts for shrinking urea/diol co-crystal

Example of awk script (shrink_cryst)

```
#/*****************************************************************************************/
#/*** Car file line for atom data looks like: *********************************************/
#/*** C1     4.781918903   0.926484195   12.519415745 XXXX 1     xx     C  0.000  ***/
#/*****************************************************************************************/


BEGIN { blat =  24.5240 }
#/*** Work out which segment of crystal this atom is in from y (b) co-ordinate ***/

NF == 9 && $3 < blat/3              { print }
NF == 9 && $3 > blat/3 && $3 < 2*blat/3 {  $2 += -2.811015
                            $3 += -4.818266
                            $4 +=  0.576182
                             printf("%-5s %14.9f %14.9f %14.9f GROU 1     %2s     %2s  0.000\n"
,$1,$2,$3,$4, $7, $8)
                            }
NF == 9 && $3 > 2*blat/3            {
                             $2 += -2.811015*2
                             $3 += -4.818266*2
                             $4 +=  0.576182*2
                             printf("%-5s %14.9f %14.9f %14.9f GROU 1     %2s     %2s  0.000\n"
,$1,$2,$3,$4, $7, $8)
                            }
NF !=9 {print}
```

Appendix 5

Sulfur potential converted from α-S_8 of Lennard Jones potential

	Min R (Å)	Min E (kJ mol^{-1})	lambda
CA CA	3.8905	0.3875	13.5
HY HY	3.3666	0.0414	13.5
SU SU	4.4541	0.3707	13.5
NI NI	3.6986	0.6260	13.5
CA SU	3.7946	0.3790	13.5
CA NI	3.7946	0.4925	13.5
SU NI	3.6986	0.4817	13.5
CA HY	3.6285	0.1267	13.5
NI HY	3.5326	0.1610	13.5
SU HY	3.5326	0.1239	13.5
HY HP	3.3666	0.0413	13.5
CA HP	3.2494	0.1267	13.5
SU HP	3.5326	0.1239	13.5
NI HP	2.7498	0.1610	13.5
HP HP	3.3666	0.0414	13.5

In Buckingham potential format

```
BUCK  CA  CODA  CA  CODA
 2343.571908  0.288185 25.067892  0.000000  30.000000
ENDS
BUCK  HY  CODA  HY  CODA
 250.384199  0.249378  1.124509  0.000000  30.000000
ENDS
BUCK  SU  CODA  SU  CODA
 2241.966726  0.329933 54.000273  0.000000  30.000000
ENDS
BUCK  NI  CODA  NI  CODA
 3786.002618  0.273970 29.895933  0.000000  30.000000
ENDS
BUCK  CA  CODA  SU  CODA
 2292.164524  0.281081 21.108084  0.000000  30.000000
ENDS
BUCK  CA  CODA  NI  CODA
 2978.604296  0.281081 27.429370  0.000000  30.000000
ENDS
BUCK  SU  CODA  NI  CODA
 2913.286679  0.273970 23.004586  0.000000  30.000000
ENDS
BUCK  CA  CODA  HY  CODA
 766.272415  0.268778  5.394522  0.000000  30.000000
ENDS
BUCK  NI  CODA  HY  CODA
 973.716328  0.261674  5.837223  0.000000  30.000000
ENDS
BUCK  SU  CODA  HY  CODA
 749.338218  0.261674  4.492123  0.000000  30.000000
ENDS
BUCK  HY  CODA  HP  CODA
 249.779406  0.249378  1.121793  0.000000  30.000000
```

```
ENDS
BUCK  CA  CODA  HP  CODA
 766.272415  0.240696  2.782334  0.000000  30.000000
ENDS
BUCK  SU  CODA  HP  CODA
 749.338218  0.261674  4.492123  0.000000  30.000000
ENDS
BUCK  NI  CODA  HP  CODA
 973.716328  0.203689  1.298512  0.000000  30.000000
ENDS
BUCK  HP  CODA  HP  CODA
 250.384199  0.249378  1.124509  0.000000  30.000000
ENDS
BUCK  OX  CODA  OX  CODA
 2024.241336  0.267378 13.810957  0.000000  30.000000
ENDS
BUCK  OX  CODA  CA  CODA
 2178.062242  0.277781 18.685282  0.000000  30.000000
ENDS
BUCK  OX  CODA  NI  CODA
 2768.353842  0.270674 20.328774  0.000000  30.000000
ENDS
BUCK  OX  CODA  HP  CODA
 711.841067  0.204133  0.961781  0.000000  30.000000
ENDS
BUCK  OX  CODA  HY  CODA
 711.925589  0.258378  3.955261  0.000000  30.000000
ENDS
BUCK  OX  CODA  SU  CODA
 2130.324323  0.298656 28.227924  0.000000  30.000000
ENDS
```

Appendix 6

Awk script to change Buckingham potential to DMAREL format

```
BEGIN { eV_to_J= 1.602192E-19
        kcalmol_to_kjmol= 4.184
        Hartree_to_J = 4.3598E-18
        avos_number= 6.02205E23
        kilo= 1000.0
        Boltz_const= 1.38066E-23
        bohr = 0.5291772
        num_pots= -1
      }

$1 !~ /#/ { num_pots++
        label1[num_pots]=$1
        label2[num_pots]=$2
        R[num_pots]= $3
        E[num_pots]= $4
        L[num_pots]= $5
        printf("Read : %s %s %10.6f %10.6f %10.6f\n",label1[num_pots],
                label2[num_pots], R[num_pots], E[num_pots], L[num_pots])
       }

END { printf("Finished\n Parameters converted from E, R, lambda to ABC:\n")

   /*** Put in combining rules assuming only Homo atom potentials given ***/

    num_homo_pots = num_pots

    for (ipot = 0; ipot <= num_homo_pots; ipot++)
      {
        for (jpot = ipot+1; jpot <= num_homo_pots; jpot++)
         {
           num_pots++
           E[num_pots] = sqrt(E[ipot]*E[jpot])
           R[num_pots] = 0.5*(R[ipot]+R[jpot])
           L[num_pots] = L[ipot]
           label1[num_pots] = label1[ipot]
           label2[num_pots] = label2[jpot]
         }
      }

   for (ipot = 0; ipot <= num_pots; ipot++)
     {
      E[ipot]= (kilo*E[ipot]/eV_to_J)/avos_number

      A[ipot]= 6 * E[ipot]*exp(L[ipot])/(L[ipot]-6)

      B[ipot]= R[ipot]/L[ipot]

      r2 = R[ipot]*R[ipot]
      C[ipot]= L[ipot]* r2 * r2 * r2 * E[ipot]/(L[ipot]-6)

    printf("BUCK  %s  CODA  %s  CODA \n %10.6f %10.6f %10.6f  0.000000  30.000000\n",
            label1[ipot], label2[ipot], A[ipot], B[ipot], C[ipot])
    printf("ENDS\n")
     }

  }
```

Appendix 7

Sulfur potential from final 2

Atom	Min R (Å)	Min E (kJ mol^{-1})	λ (no unit)
O - O	3.6096	0.3347	13.5
C – C	3.8905	0.3875	13.5
H – H (non-polar)	3.3666	0.0414	13.5
H – H (polar)	3.3666	0.0414	13.5
N – N	3.6986	0.6260	13.5
S – S	3.9541	1.1075	13.5

In Buckingham potential format

```
BUCK  OX  CODA  OX  CODA
 2024.241336  0.267378 13.810957  0.000000  30.000000
ENDS
BUCK  CA  CODA  CA  CODA
 2343.571908  0.288185 25.067892  0.000000  30.000000
ENDS
BUCK  HY  CODA  HY  CODA
 250.384199  0.249378  1.124509  0.000000  30.000000
ENDS
BUCK  HP  CODA  HP  CODA
 250.384199  0.249378  1.124509  0.000000  30.000000
ENDS
BUCK  NI  CODA  NI  CODA
 3786.002618  0.273970 29.895933  0.000000  30.000000
ENDS
BUCK  SU  CODA  SU  CODA
 6698.079711  0.292896 78.966559  0.000000  30.000000
ENDS
BUCK  OX  CODA  CA  CODA
 2178.062242  0.277781 18.685282  0.000000  30.000000
ENDS
BUCK  OX  CODA  HY  CODA
 711.925589  0.258378  3.955261  0.000000  30.000000
ENDS
BUCK  OX  CODA  NI  CODA
 2768.353842  0.270674 20.328774  0.000000  30.000000
ENDS
BUCK  OX  CODA  SU  CODA
 3682.190900  0.280137 33.230671  0.000000  30.000000
ENDS
BUCK  CA  CODA  HY  CODA
 766.024395  0.268781  5.393222  0.000000  30.000000
ENDS
BUCK  CA  CODA  NI  CODA
 2978.719419  0.281078 27.428262  0.000000  30.000000
ENDS
BUCK  CA  CODA  SU  CODA
 3961.998416  0.290541 44.500630  0.000000  30.000000
ENDS
BUCK  HY  CODA  HP  CODA
 250.384199  0.249378  1.124509  0.000000  30.000000
ENDS
BUCK  HY  CODA  NI  CODA
```

```
 973.629925  0.261674  5.836705  0.000000  30.000000
ENDS
BUCK  HY  CODA  SU  CODA
 1295.026379  0.271137  9.607742  0.000000  30.000000
ENDS
BUCK  HP  CODA  SU  CODA
 1295.026379  0.271137  9.607742  0.000000  30.000000
ENDS
BUCK  NI  CODA  SU  CODA
 5035.766806  0.283433 48.750692  0.000000  30.000000
ENDS
BUCK  CA  CODA  HP  CODA
 766.272415  0.240696  2.782334  0.000000  30.000000
ENDS
BUCK  NI  CODA  HP  CODA
 973.716328  0.203689  1.298512  0.000000  30.000000
ENDS
BUCK  OX  CODA  HP  CODA
 711.841067  0.204133  0.961781  0.000000  30.000000
ENDS
```