2014

Computational study of boron binding environments and their role in some biological systems

Chamila Chandima De Silva
Iowa State University

Computational study of boron binding environments and their role in some biological systems

by

Chamila Chandima De Silva

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Physical Chemistry
Program of Study Committee:
Thomas A. Holme, Major Professor
Mark S. Gordon
Theresa L. Windus
Edward Yu
Arthur Winter

Iowa State University

Ames, Iowa

2014

DEDICATION

This work is dedicated to my grandmother, Charlotte Premalatha.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to express my utmost gratitude to my research advisor, **Professor Thomas Holme**, first for accepting me in to his group and secondly for all the support and the advices he has given me through out my graduate studies. Also, I really appreciate the freedom he gave for me to do the research and the guidance to keep me in the correct path. Professor Holme showed in example how to work under stressed conditions and how to look at the positive side in every situation has been very inspirational to me. Also, I really appreciate his sense of humor and I have been fortunate and grateful to have Professor Tom Holme as my research advisor.

I thank my program of study committee members for their guidance over the years. I would like to thank **Professor Mark Gordon** and **Professor Theresa Windus** for encouraging me to do graduate studies. I really appreciate **Professor Gordon** for his continuous encouragement through out the years. Also I would like to thank **Professor Windus** for showing how to be a great teacher by example, I have learnt a lot by just observing you, thank you. I thank **Professor Edward Yu** for adding an experimental perspective towards my research and his encouragement through out the years. I also thank **Professor Art Winter** for his insightful comments, which helped me to keep my research on the correct path.

I would like to thank **Dr. Joseph Burnett**, for all the guidance, encouragement and for lending an ear through out my time at Iowa State University as a Teaching Assistant. A special also goes for making me used to the term 'take it easy'.

Also I would like to thank all the Professors and Instructors for whom I have worked as a teaching assistant. I learnt a lot by watching how you teach; work with students and teaching assistants.

A special acknowledgement goes to **Raja De Alwis,** the teacher who showed me wonders in learning.

I also thank to the administrative staff, **Renee Harris**, and **Lynette Edsall**, for their continued assistance in navigating through the paper work.

In addition, I would like to thank **Professor Krishna Kumar**, **Katie Kumar**, **Professor James Cornette**, and **Carolyn Cornette** for their encouraging guidance. I gratefully thank to friends, **Marilee Hall** and **Kumari Rannulu**, for their positive comments and for being an inspirational figures.

I would like to thank my best friends **Sandunil Meepe** and **Bhagya Panadura** for their unchanging friendship from the day we became friends to today.

A special thank to my mother **Srimathi Liyanage** for being an inspirational figure and for supporting me for education in every possible way. Fondly remember my father **Somapala Liyanage**, for his blessings and my brother **Kapila Liyanage** and his family for the positive thoughts. A special thank to my aunt **Gna Liyanage** for the happy memories of my childhood.

Finally, I thank my husband and my best friend **Nuwan** for his support and encouragement through out my graduate studies. My daughter, **Amelia** and son **Jaden**, gets special hugs for being patient through out the final months of my graduate studies and for trying to be happy under all the circumstances

CHAPTER 1. GENERAL INTRODUCTION

General Overview

Boron, a fairly simple atom in the first row of periodic table, makes remarkable dative bonds, which have inspired the curiosity of researchers for decades. Boron has three valence electrons and when in compounds the boron atom is often sp^2 hybridized and has a planar structure. An empty p-orbital, which has the capability of accepting an electron pair, makes boron a good Lewis acid and once the lone pair is gained the structure around the boron becomes tetrahedral. While Boron−Nitrogen, B−N dative bonds have been extensively studied throughout the years, other boron dative bonds such as B−O and B−P also have shown remarkable characteristics.

Boron-Nitrogen containing compounds such as ammonia borane,[1,2] metal amidoboranes,[3] and organoboron boron compounds[4] have shown to be promising hydrogen storage materials. These compounds release hydrogen upon heating, which, can be used as a clean energy source. Thermolysis of amino-boron compounds is an exothermic process. Once the initial hydrogen is released it will gives out enough energy for a self-sustained hydrogen release.[5,6] But this ongoing exothermic reactions increase the temperature of the system and can even cause breaking of boron-nitrogen bond, forming low molecular weight byproducts such as borone, ammonia and borozine, that can pollute the proton exchange membrane fuel cells (PEM fuel cells). This type of

technological utility adds to the interest in carrying out studies to understand the B−N bond energies of different boron nitrogen compounds. Knowing the dissociation bond energies for B−N is useful in controlling the reaction temperatures of thermolysis process.

Boron neutron capture therapy (BNCT) is a powerful form of radiotherapy, which incorporate ^{10}B-containing compounds into tumor cells, followed by the irradiation of tumor/cancer cells with thermal neutrons.[7,8] Subsequent to absorbing the neutron, the resulting ^{11}B is unstable and decays into Li and ^{4}He, a high-energy reaction, which destroys tumor cells without damaging as many of the surrounding healthy cells as other forms of chemotherapy. The most important requirement for the boron neutron-capture therapy is the high and selective accumulation of boron in tumor cells, where the boron-containing molecule should adhere to both boron and to the targeted cancer cell. Finding boron compounds with high selectivity, water solubility and low toxicity in high concentrations is a major problem in the advancement of treatment using BNCT.[9] Therefore studying of boron binding environments is essential in the field of radiotherapy. Phosphorous, which plays an important role in human body as phosphates is a good example of selective boron binding for treatment of cancer cells. In addition to its part as an essential building block in DNA and RNA and as an energy transporter in the form of nucleoside di- and triphosphate, it contributes strongly to the strength and integrity of the bone skeleton. Bisphosphonates, used as a treatment for diseases like osteoporosis and rheumatoid arthritis, can be strongly adsorbed to hydroxylapatite

crystals. Hydroxylapatite crystals are substances that are found in increased quantities in bone cancer cells. Phosphorous makes dative bonds with boron. With an extended study of the strength of the boron-phosphorous bond strength; bisphosphonates might be used as a boron carrier for boron neutron capture therapy.[10,11] In addition to phosphorous it has been found that both oxygen and nitrogen containing molecules also acts as selective boron carriers for BNCT.

The pharmacological uses of boron compounds have also been known for decades. In animal cells at pH values that are present, almost all-natural boron exists as boric acid, which forms molecular additive compounds with many biological compounds such as, amino and hydroxy acids, nucleotides and carbohydrates through the formation of electron donor-acceptor interactions. Recent findings of boron analogous of amino acids and their derivatives express anti-inflammatory,[12] antineoplastic,[13] and hypolipidemic[14] properties. Most available dietary boron supplements use boron chelated with amino acids or hydroxy acids (citric acid, aspartic acid or glycine) in combination with vitamins. However the molecular structure of these boron chelates is poorly understood. It is essential to understand the dative bond strength of a number of boron containing bonds, including B−N, since this bond strength can govern the possible structures for drug design for various diseases to aide in understanding its function in affected sites.

Boric acids and borate anions can be combined with organic compounds to form molecules that contain B−O−C bonds (alcohols and carboxylic acids) so they are capable of forming organic esters.

$$H_3BO_3 + 3R(OH) \rightleftarrows B(OR)_3 + 3H_2O$$

As shown in the equation above, boric acid reacts with an primary alcohol R(OH) in a reversible reaction gives out borate di-ester and water molecules. Both boric and boronic acids can form either neutral or anionic esters depending on the pH of the system. Diol binding by boron acids is favored at high pH values, whereas the esterification of boron by hydroxycarboxylic acids is favored at low pH ranges. At high pH values boron acids can also form an anionic, tetrahedral diester with a couple of diols or with a diol and a relevant divalent ligand. The reversible reaction of esterification of boronic acids has been used in the health industry for developing blood sugar monitoring techniques for diabetics[15] and the same reaction with boric acid is used in the study of plant cell wall biosynthesis where the borate diol cross-linked to connect two side chains of Rhamnogalacturonan II (RG-II) macromolecule which controls the growth of the plant cell wall.

Boron-based glucose receptors for incorporation as sensors in blood sugar monitors for diabetics have developed during the past two decades.[16-18] The relative affinity of boronates for diols in most carbohydrates follows the order of cis-1,2-diol > cis-1,3- diol >> trans-1,2 diol. Therefore it is clear that certain monosaccharides have an intrinsically higher affinity for boron acids.[15] It has been found that the boron-binding site

for a monosaccharide depends on both the type of boron acid and the type of monosaccharide. For an example, boronates has the ability to bind to the 1,2-diol and trans-4,6- diol of glucose in its hexopyranoside form,[17] boronic acid has an affinity for the furanoside form of free hexoes[18] and galactopyranoside has an affinity for cis-3,4-diol and trans-4,6-diol.

Rhamnogalacturonan II (RG-II) is a structurally complex pectic polysaccharide, which is conserved in the plant cell wall despite the evolutionary variation of plants. RG-II contains homogalacturonan backbone composed of at least eight 1→4 linked α-d-galacturonic acid residues. Four different complex oligo glycosyl side-chains named side chain A, side chain B, side chain C and side chain D that contain twelve different glycosyl residues are attached to this backbone.[19] RG-II may exist as a dimmer, which contributes to the strength of the plant cell wall. Covalently cross-linking two side chains by a borate diester lead to RG-II dimerization. The boron needed for the borate binding is absorbed by the plant from the soil solution in the form of either boric or borate acids. The boro-diester reaction occurs in between two α-d-apoise (3-C-hydroxymethyl-D-erythrose) monomers in the side chain A of RG-II structure. The borate esterification is believed to be happened in cis-2,3 diol position.[20] Structure of RG-II and its functions has been studied for more than four decades; still the structure and the functions of RG-II are not completely understood. Therefore the study of borate cross-linking carbohydrates can give some insights in to this complex macromolecule.

Boron is acts as a Lewis acid abduct over a range of fields. The comparable size of boron atom to carbon atom and the strength of dative bonds and the ability to undergo reversible reactions makes a boron a good candidate for the fields of energy, biology, medicine and plant cell biology. As a whole, study of boron binding environments can give insights for many unanswered questions.

Dissertation Organization

This dissertation is comprised of six chapters: Chapter 1 gives the theoretical background of *ab initio* methods, which are relevant to the computational methodologies employed in the following chapters. Chapter 2 provides an extensive study of B−N and B−P bonds in heterocyclic systems. Chapter 3 investigates the existence of cyclic−dimer and cyclic−monomer and linear forms of 2-Aminoethoxydiphenyl Borate in solution. Chapter 4 illustrates a study of modeling hydrolyzation reaction of anhydrous borate to boric acid. Chapter 5 describes the CHARMM all-atom empirical force field parameterization of boro−diester carbohydrates. Chapter 6 provides the general conclusions of this dissertation work.

Theoretical Background

The inadequacy of classical physics to explain the observations of ultraviolet catastrophe and photoelectric effect has led to the development of quantum mechanics in

the twentieth century. Because of the wave particle duality of matter, the physical behavior of an electron can be described using a wavefunction. In quantum mechanics Schrödinger equation[21-26] is used to solve the wavefunction, which describe the state of the system at any given time by the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Phi(r,R,t)=\hat{H}\Phi(r,R,t) \tag{1}$$

where $\hat{H}$ is the Hamiltonian, the energy operator. Φ is the wavefunction, which depends on the electronic coordinates, r, nuclear coordinates, R, and time, t, $\hbar$ is Planck's constant, $h/2\pi$, i is the square root of -1.

For the systems that the potential do not vary with the time, the wavefunction can be written as a product of spatial part (r and R) and the time part (t) and the time-independent Schrodinger equation can be derived as,

$$\hat{H}\Phi(r,R)=E\Phi(r,R) \tag{2}$$

where, E is the total energy of the system.

The Hamiltonian operator $\hat{H}$ has two kinetic energy terms; the kinetic energy of electrons (T_e) and the kinetic energy of nuclei (T_n) and three potential energy terms; the electron-electron repulsion (V_{ee}), the nuclear-nuclear repulsion (V_{nn}), and the electron-nuclear attraction (V_{en}).

$$\hat{H}=T_e(r)+T_n(R)+V_{ee}(r)+V_{nn}(R)+V_{en}(r,R) \tag{3}$$

The more elaborate Hamiltonian in atomic units can be given as,

$$\hat{H} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \quad (4)$$

In Eq. (4), ∇_i^2 and ∇_A^2 are the Laplacian operators for i^{th} electron and A^{th} nucleus. The M_A is the ratio of the mass of nucleus A to the mass of an electron, and Z_A is the atomic number of nucleus A. The distance between electron i and j is r_{ij}, where $r_{ij} = | r_{ij} | = | r_i - r_j |$. The distance between nucleus A and B is R_{AB}, where $R_{AB} = | R_{AB} | = | R_A - R_B |$. The distance between electron i and nucleus A is r_{iA}, where $r_{iA} = | r_{iA} | = | r_i - r_A |$.

Electronic coordinates parametrically depend on the nuclear coordinates. Therefore Eq. (4), which describes the energy of a whole system is not possible to solve without approximations. The Born-Oppenheimer approximation[27] is used to separate the nuclear and electronic motion. This approximation rests on the fact that nuclei are much heavier than electrons and which cause electrons to move much faster than nuclei. As a result, the position of the nuclei can assume to be fixed in space with the motion of electrons and motion of electrons and nuclei can be separated. The result of this assumption caused the nuclear kinetic energy (T_n) goes to zero and nuclear-nuclear repulsion (V_{nn}) term equal to a constant (For each R value, V_{nn} is just a constant which shifts the eigenvalues only by some constant amount). This assumption reduces the Schrödinger equation to second-order differential equation that only depends on the electronic degrees of freedom. Therefore, Eq. (3) can be written as,

$$\hat{H} = \hat{T}_e(r) + \hat{V}_{ee}(r) + \hat{V}_{en}(r,R) \quad (5)$$

Hense, the the electronic time-independent Schrödinger equation can be written as

$$\hat{H}_{elec}\Phi_{elec} = E_{elec}\Phi_{elec} \tag{6}$$

The $\hat{H}_{elec}$ is the electronic Hamiltonian, which depends parametrically on nuclear coordinates (R) and E_{elec} is the total electronic energy of the system. Φ_{elec} is the electronic wavefunction which parametrically depends on nuclear coordinates and describes the motion of electrons in the system.

The total energy of the system with in the Born-Oppenheimer approximation is taken by adding E_{elec} to V_{nn}.

$$E_{total} = E_{elec} + V_{nn} \tag{7}$$

The Eq. (7) describes the potential energy surface (PES), a function of total energy of a system versus the nuclear coordinates. In the electronic Hamiltonian ($\hat{H}_{elec}$), the T_e and V_{en} depends on coordinates of one electron but V_{ee} depends on coordinates of two electrons and makes the solving of electronic Hamiltonian impossible for systems with more than one electron. As a result more approximations are needed to solve the electronic Hamiltonian for systems with more than one electron.

The most simplest approximation is the Hartree-Fock approximation where the explicit electron-electron interactions, V_{ee} in Eq. (5) is replaced with an averaged interaction, $\nu^{HF}(i)$ and the time independent Schrödinger equation is solved in a self-consistent manner. The replacement of two electron repulsion term, V_{ee} with one electron

potential, $v^{HF}(i)$ allow the separation of variables and hence solve the time independent Schrödinger equation. This approximation is known as the self-consistent field (SCF) Hartree-Fock (HF) method.[28-31] The Hartree-Fock potential, $v^{HF}(i)$ is defined as

$$v^{HF}(1) = \sum_{i=1}^{N} \left[\hat{J}_i(1) - \hat{K}_i(1) \right] \tag{8}$$

In Eq. (8) $\hat{J}$ is the Coulomb operator and $\hat{K}$ is the exchange operator. The one electron Fock operator can be written as

$$\hat{F} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 \; - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} v^{HF}(i) \tag{9}$$

The Fock operator can be expressed as a sum of one-electron operators

$$\hat{F} = \sum_{i=1}^{N} \hat{F}(i) = \sum_{i=1}^{N} \left[-\frac{1}{2} \nabla_i^2 \; - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + v^{HF}(i) \right] \tag{10}$$

Hartree-Fock equations can be written as a sequence of coupled one-electron eigenvalue problems.

$$\hat{F}(i)\chi(i) = \varepsilon(i)\chi(i) \tag{11}$$

The eigenfunction of the one-electron Fock operator is called molecular orbital, $\chi(i)$. The Fock operator is a sum of one-electron operators. Therefore the total solution for the system is the product of one-electron functions, called Hartree product Ψ^{HP}.

$$\Psi^{HP}(1,2,3,\ldots,N) = \chi(1)\chi(2)\chi(3)\cdots\chi(N) \tag{12}$$

The Hartree product is not an anti-symmetric wavefunction therefore it does not obey the Pauli exclusion principle. This problem can be fixed by expressing the wavefunction in a compact form of a Slater determinant.[32]

$$\Psi^{HF} = \left|\chi(1)\chi(2)\chi(3)\cdots\chi(N)\right\rangle \tag{13}$$

The molecular orbitals used in Slater determinants has a spatial ψ_i and a spin (α or β) components. The spatial component of the molecular orbital is constructed as a linear combination of atomic orbitals, ϕ_μ (LCAO),

$$\psi_i = \sum_{\mu=1}^{N} C_{\mu i}\phi_\mu \tag{14}$$

where, $C_{\mu i}$ are the LCAO coefficients. The atomic orbitals are comprised of Gaussian functions or contracted Gaussian functions. To get an exact molecular orbital, one should use a complete expansion of atomic orbitals. This is not possible because in Hibert space, complete basis set goes to infinite. Therefore a truncated basis set is used.

The one-electron operators given in Eq. (8), $\hat{J}$ and $\hat{K}$ act on spin orbitals via the following equations:

$$\hat{J}_b(1)\chi_a(1) = \left[\int dx_2 \chi_b^*(2) r_{12}^{-1} \chi_b(2)\right]\chi_a(1) \tag{15}$$

$$\hat{K}_b(1)\chi_a(1) = \left[\int dx_2 \chi_b^*(2) r_{12}^{-1} \chi_a(2)\right]\chi_b(1) \tag{16}$$

The Hartree-Fock equations define the energy of the spin orbitals χ_a as ε_a; Hartree-Fock equations are optimized iteratively until the self-consistency is obtained. The

wavefunction and the energy are unknown at the beginning and a guess orbital is used. Using the guess orbitals, Fock matrix is formed and diagonalized and the energy eigenvalues for the orthogonal molecular orbitals are obtained. Using the molecular orbitals obtained in the above process a new guess at the density is formed and repeats the process until the energy and wavefunctions are converged. Hartree-Fock method uses variational principle. Therefore the calculated Hartree-Fock energy, E_{HF} is always an upper bound to the exact non-relavistic energy, E_{exact}.

$$E_{exact} \leq E_{HF} \tag{17}$$

The difference in energy between the exact energy and the Hartree-Fock energy is called the correlation energy, E_{corr}.

$$E_{exact} = E_{HF} + E_{corr} \tag{18}$$

The correlation energy arises from the replacement of electron-electron interaction term, V_{ee} in Eq. (9) by the Hartree-Fock potential $\nu^{HF}(i)$. Even though electron-electron correlation is a small value it is necessary to obtain the accurate energies of chemical systems. In order to capture correlation energy accurately many methods have been developed.

Many-body perturbation theory (MBPT) is a common way to include electron correlation to a system at a reasonable cost. In this method a small perturbation, ν is added to the independent particle H_0.

$$\hat{H} = \hat{H}_0 + \lambda V \tag{19}$$

λ is a parameter that can take the values 0 and 1. The energy and the wave function can be expanded in a series as,

$$\Psi_{MPn} = \Psi_0 + \sum_{i=1}^{n} \lambda^i \Psi_i \tag{20}$$

$$E_{MPn} = E_0 + \sum_{i=1}^{n} \lambda^i E_i \tag{21}$$

The nth order perturbation theory adds higher order energy corrections to give better energies and wave functions that represent the system. Møller-Plesset[33] perturbation theory at n=2, (MP2), founds to account for ~80%-90% of total correlation energy. Eq.22 gives the second order energy correction using MP2.

$$E_2 = \sum_D \frac{\left|\left\langle \Psi_D \middle| V \middle| \Psi_{HF} \right\rangle\right|^2}{E_D^{(0)} - E_{HF}^{(0)}} \tag{22}$$

The summation in Eq. 22 runs over all double excitations from occupied to unoccupied orbitals Ψ_D. The denominators correspond to the zeroth-order energy differences between each doubly excited determinant and the HF determinant.

Even though MP2 is more computationally expensive (order of $O(N^5)$) than Hartree-Fock approximation (order of $O(N^4)$) it is still more feasible for many chemical systems of medium and large scale. The disadvantage of perturbation theory is that it is not variational and cannot guarantee that the resulting energy is always above or equal to the exact energy. Also as the order of the perturbation increases, the perturbation expansion itself often does not converge.

References

(1) Marder, T. B. *Angew. Chem* **2007**, *46*, 8116.

(2) Hess, N. J. S., G. K.; Hartman, M. R.; Daemen, L. L.; Proffen, T.; Kathmann, S. M.; Mundy, C. J.; Hartl, M.; Heldebrant, D., J.; Stowe, A. C.; Autrey T. *J. Phys. Chem. A* **2009**, *113*, 5723.

(3) Xiong, Z. Y., C. K.; Wu, G.; Chen, P.; Shaw, W.; Karkamkar, A.; Autrey, T. et.al., *Nat. Mater.* **2008**, *7*, 138.

(4) Campbell, P. G. Z., L. N.; Grant, D., J.; Dixon, D. A.; Liu, S., H. *J. AM. Chem. Soc.* **2010**, *132*, 3289.

(5) Geanangle, R., A.; Wendlandt, W., W. *Thermochim. Acta* **1985**, *86*, 375.

(6) Hu, M., G.; Geanangle, R., A.; Wendlandt, W., W. *Thermochim. Acta* **1978**, *23*, 249.

(7) Hawthrone, M., F., *Angew. Chem. Intl. Ed.* **1993**, *32*, 950.

(8) Valliant, J. F., Guenther, K., J., King, A., S., Morel, P., Schaffer, P., Sogbein, O., O., Stepheson, K., A. *Coord. Chem. Rev.* **2002**, *232*, 173.

(9) Hawthrone, M., F., Lee, M., W., *J. neurooncol.* **2003**, *62*, 33.

(10) Sparidans, R. W. T., I., M.; Talbot, S. *Pharm. World. Sci.* **1998**, *20*, 206.

(11) Stresing, V. D., F.; Benzaid, I.; Monkkoen, H.; Clezardin, P. *Cancer Lett.* **2007**, *257*.

(12) Hall, I. H. S., C. O., McPhail, A. T.;Wisian-Neilson, P., Das., M. K.; Harchelroad jr., Spielvogel, B. F. *J. Pharmacol. Sci.* **1980**, *69*.

(13) Hall, I. H. G., C. J.;, McPail, A., T.; Morse, Hassett, K.; Spielvogel, B. *J. Pharmacol Sci.* **1985**, *74*.

(14) Hall, I. H. W. J., W. L.; Gilbert, C. J.;, McPail, A., T.; Spielvogel, B. *J. Pharmacol Sci.* **1984**, *73*, 973.

(15) Pappin, B. M., J. K.; Houston, T. A.; .

(16) Mader, H. S. W., O. S.; *Microchim. Acta* **2008**, *162*.

(17) Ferrier, R. J. P., R. *J. Chem. Soc.* **1965**, 7429.

(18) Eggert, H. F., J.; Mortin, C.; Norrild, J. C.; *J. Org. Chem.* **1999**, *64*, 3846.

(19) Vidal, S. W., T.D.P.; Pellerin,P., York, W. A., O'Neill, M.A. Glushka, A.G Darill, P Albersheim *Carbohydrate Research* **2000**, *326*

(20) O'Neil M. A.; Ishii T.; Albersheim, P. D. A. G. *Annu. Rev. Plant. Biol.* **2004**, *55*, 109.

(21) Schrödinger, E. *Naturwissenschaften* **1926**, *14*, 664.

(22) Schrödinger, E. *Annalen Der Physik* **1926**, *79*, 361.

(23) Schrödinger, E. *Annalen Der Physik* **1926**, *79*, 489.

(24) Schrödinger, E. *Annalen Der Physik* **1926**, *79*, 734.

(25) Schrödinger, E. *Annalen Der Physik* **1926**, *80*, 437.

(26) Schrödinger, E. *Annalen Der Physik* **1926**, *81*, 109.

(27) Born, M.; Oppenheimer, R. *Annalen Der Physik* **1927**, *84*, 0457.

(28) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 89.

(29) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 111.

(30) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 426.

(31) Fock, V. *Physik* **1930**, *61*, 126.

(32) J. C. Slater, *Phys. Rev.* **1929**, *34*, 1293.

(33) Moller, C.; Plesset, M. S. *Phys. Rev. Lett.* **1934**, *46*, 618.

CHAPTER 2. COMPUTATIONAL STUDIES OF DATIVE BOND CONTAINING HETEROCYCLIC RING STRUCTURES

A paper published in *Computational and Theoretical Chemistry*

Chamila C. De Silva and Thomas A. Holme

Abstract

Ab initio calculations are reported for several related heterocyclic compounds, each of which contains two dative bonds when they self-dimerize. Thus, these molecules are nominally dimers that contain either a boron-carbon-nitrogen (BCN) or boron-carbon-phosphorous (BCP) segment. Molecules with this motif have been found experimentally to have several unusual properties that may be related to a “multi-polar framework" that results from charge separation associated with the two dative bonds. Structures obtained from full geometry optimizations without symmetry constraints, dative bond energies and charge distributions for four multipolar molecules are reported, the BCN-BCN dimer and the BCN-BCP dimer with and without carboxylation of one boron atom. Comparisons to single dative bond, self-cyclized monomers and the role of ring strain in molecular stabilities are also discussed.

I. Introduction

Experimental work with heterocyclic ring compounds containing two boron nitrogen dative bonds was reported over 30 years ago.[1-4] Since initial work by Miller and Muetterties, a number of these compounds have been synthesized, with a variety of permutations of the heteroatoms and substituents on the ring system.[3,4] The general class of molecules shows an interesting feature with dimers formed by two Lewis acid-base adducts, or dative bonds. These compounds, form heterocycles of the form BCNBCN, and have been labeled Multipolar Framework Heterocycles.[4]

Experimental studies of compounds that contain B–N dative bonds have been fairly limited. The prototype molecule for this class of compound BH3NH3 has received substantial attention,[5-8] and the B–N analog of benzene, borazine, likewise has attracted attention.[9] Bartlett and coworkers have studied the BCN fragment with an eye towards comparing it to its all carbon analog.[10] Despite their rich synthetic history and the intriguing nature of the multi-polar framework hypothesis, the specific class of dimer compounds for which we are reporting calculations has been the subject of only one calculation study, by Hseu using semi-empirical procedures.[11]

From the perspective of a model system for computational investigation, this category of molecules presents several positive attributes. First, the previously noted concept of a multipolar heterocyclic system provides an interesting theoretical challenge.

Because the dative bond forms due to electron donation from the nitrogen (or phosphorous) to the boron, the charge separation is commonly designated with a positive nitrogen and negative boron, $^{\delta+}N–B^{\delta-}$. A canonical perspective of this bond using electronegativity arguments, however, would predict the negative charge on the nitrogen atoms. Because charge separation plays a role in the way molecules are solvated or act in biochemical environments, such a model system may provide an important template for study of these effects.

Second, as a theoretical target, this set of molecules provides interesting questions. The determination of atomic charges poses a difficult challenge for electronic structure calculations, and this system admits particular difficulties for consideration. Because the system is rather large in its experimentally stabilized forms[2-4] there is little ability to utilize large basis sets simultaneously with computationally expensive models for inclusion of electron correlation. Thus, the calculations presented here provide insight at levels of theory that may be possible in relatively large molecular systems. Finally, in principle this class of molecules allows the exploration of Lewis acid-base reaction chemistry in a unimolecular system.

In this study two related heterocyclic compounds are considered. The first category are those molecules composed of like monomers, such as 1,1,4,4-Tetramethyl-1,4-diazonia-2,5-diboratacyclohexane, (Figure 1-a). This system will generally be designated as BCNBCN. The second category includes heterocycles composed of

different monomers, such as 1,1,1,4-Tetramethyl-1-azonia-4-phosphonia-2,5-diboratacyclohexane, (Figure 1-b). The basic designation of this system will be BCNBCP throughout. The carboxylic acid derivatives of both types of rings have also been synthesized and are therefore calculated. Thus, Figure 1-c shows 1,1,4,4-Tetramethyl-1,4-diazonia-2,5-diboratacyclohexane-2-carboxylic Acid, designated (COOH) BCNBCN and Figure 1-d shows 1,1,4,4-Tetramethyl-1-azonia-4-phosphonia-2,5-diboratacyclohexane-2-carboxylic acid, designated (COOH) BCPBCN. In addition to these ring structures with two dative bonds, dimerization may also occur via the formation of a single dative bond (referred to as a linear dimer, hereafter). This category of dimer is also investigated for both BCNBCN and BCNBCP in order to help elucidate the role of ring formation relative to that of dative bond formation.

In order to estimate energetics of dative bond formation, the monomers from which these dimers form have also been calculated, in both a cyclic and open (referred to as linear, hereafter) form. For these smaller systems the relative role of ring strain may compete energetically with dative bond formation. To estimate the role of ring strain an additional set of cyclic monomers with increasing number of carbon atoms in the cycle were also calculated. Thus additional monomers such as $BH_2CH_2CH_2N(CH_3)_2$ (designated BCCN) and $BH_2CH_2CH_2CH_2N(CH_3)_2$ (designated BCCCN) were calculated in both linear and cyclic form. These systems may form 8 member and 10 member rings by dimerization and comparisons of energies of species in this set provide insight into the relative roles of ring-strain and dative bond strength in the overall stability of this

category of molecule. Calculations reveal that the cyclic monomers of all three cases are lower in energy than the linear monomer, indicating that the stabilization of the dative bond formation (for self-cyclization) exceeds ring strain concerns even in the case of a three-member ring. Not surprisingly the relative added stability from self-cyclization of the monomers increases with the increasing number of atoms in the linear monomers.

II. Computational Details

All calculations were performed using the GAMESS[12,13] electronic structure code, and the molecules were visualized with MacMolPlt.[14] All molecules were fully optimized in gas phase with C1 symmetry, using 2nd order Moller-Plesset perturbation theory (MP2)[15] and the 6-311G(d,p)[16] basis set. All MP2 calculations utilize a frozen-core approximation. Zero point energy was calculated for all the molecules. Conformational searches for all heterocyclic rings were carried out using the Complete Rotation from the Evaluation of Potential Energy Surface (CREPES) program[17] at MP2/6-311G(d,p) level and the lowest energy conformers were optimized. Positive definite Hessian calculations confirm the reported structures are true minima.

Dative bond energies were estimated by comparing the energy of the entire molecule with that of the fragments derived from breaking both dative bonds simultaneously. Linear structures were obtained for each heterocyclic molecule by breaking each dative bond. Atomic charges were estimated using the Generalized

Atomic Polar Tensors method (GAPT)[18] which is derived from the trace of the dipole derivative tensor, calculated numerically. Ring Strain Energy for monomers and some dimers were calculated using a method given by Dudev and Lim.[19] In this method Ring Strain Energy (E_{RS}) of a cyclic molecule is calculated relative to its acyclic counterpart containing the same number of heavy atoms.

III. Results and Discussion

The primary system of interest arises from the dimerzation of boron-carbon-nitrogen/phosphorous systems into six-member rings. Four such cyclic dimer structures are depicted in Figure 1. Each has a chair conformation that is the lowest energy conformer. For the BCNBCN cyclic dimer the lowest energy conformer is a regular chair and other structures show a slightly twisted chair conformation with a twisted angle varying from 1 to 4 degrees. In the experimentally synthesized structure (COOH)BCPBCN, carboxylic acid is attached to boron that forms a dative bond with nitrogen rather than phosphorous, because the carboxyl group originates on the BCP monomer. Calculations have also been carried out for (COOH)BCNBCP where the carboxylic acid is attached to boron bonded to phosphorous. The energy of this structure is 0.41kcal/mol less than that of the experimentally reported structure. Nonetheless, because this energy difference is very small, indeed smaller than the likely accuracy of the calculation methods applied here, for subsequent discussions, the experimental (COOH)BCPBCN structure will be used for this system.

As noted earlier, self-cyclization of the monomers provides an important set of heterocycles to consider as well. In order to understand effects of ring strain in these systems, additional monomers were considered in this study, beyond those that have been experimentally reported. Figure 2 shows the basic structures of the cyclized monomers whose geometries and energies have been calculated.

Structural and energy information is summarized for the molecules calculated in this work, in a series of tables. Key structural features for the molecules are provided in Table 1. In this table, C' and C" represent the carbons in the axial and equitorial methyl groups. C"' is the carbon next to boron in BCN, BCCN and BCCCN monomers. In the systems that include carboxy-boranes, B' designates the boron with attached carboxylic group. N' is the nitrogen next to the boron with carboxylic group. Finally, for the BCNBCP system B" is the boron that forms the dative bond to phosphorous.

The structural parameters, such as the bond lengths tabulated here, are generally in line with those seen for B–N, B–P, B–C, C–N, etc.[20] Limited experimentally derived structural data is available for these systems. Only for the single case of the BCNBCN dimer (Figure 3 includes labeling information for this structure) has the x-ray structure has been published.[22] A comparison of calculated and experimental structural information for BCNBCN molecule is provided in Tables 2 and 3 based on atom definitions provided in Figure 3.

Agreement for bond distances for BCNBCN molecule is generally good, though calculations show shorter B–H bond distances, for example. Calculated bond angles (Table 3) are also generally in good agreement. A decrease in angles from 109.5° of N3-B4-H12, N3-B4-H11, H7-C2-H8 and C16-N3-C15 (Figure 3) is common for both calculated and experimental structures. For the rest of the molecule the slight difference in lengths and in angles in experimental data from that of calculated data can emerge from intermolecular repulsions due to close packing of the molecule.

Both experimental and calculated torsional angles of BCNBCN are shown in Figure 4. Calculated structures tend to deviate somewhat less from the dihedral ideal (60°) than was observed experimentally. Nonetheless, from the perspective of experimental verification of the calculations, these comparisons suggest the calculations are doing reasonably well in terms of predicting structure.

The relative energies of the various molecules are provided using the system with fewer dative bonds as the reference, and chosen as the zero of energy. In Table 4 the energies of the monomers are provided while dimer energies are given in Table 5.

Trends observed here are not unexpected. Focusing first on the dimers, when the multipolar framework contains only nitrogen, the stabilization derived from forming two dative bonds is greater than when a P atom is swapped for N. In addition, BCNBCN and

BCNBCP are each stabilized less when compared to their respective carboxylic acid derivatives (COOH)BCNBCN and (COOH)BCPBCN where the electron withdrawing carboxylic acid is bonded to a boron atom. For all four structures the stabilization energy for the linear dimer is notably less than for the cyclic dimer, essentially revealing the strength of the second dative bond. The experimental observations of BCN cyclic structures[21] are consistent with these findings.

Looking at the energy information in Table 4 a little more closely reveals some additional insights into the nature of the dative bonding. Even for these monomers, with three heavy atoms, the cyclic monomer of BCN is notably more stable than that of BCP, mimicking the results for the dimer systems. Noting the relative energies of B–N bonds and B−P bonds in the linear dimers is also interesting. BCPBCN (B–P dative bond) is 3.5 kcal/mol higher in energy than BCNBCP (B–N dative bond). This effect appears to be smaller than for the self-cyclization case where the formation of the B–P dative bond is putatively ~14 kcal/mol less stable, the increase energy in cyclic form is attributed to the higher ring strain of cyclic BCP monomer than that of cyclic BCN monomer – noted in Table 6. This set of observations tends to corroborate the importance of ring strain when the P atom is incorporated into the cyclic monomer.

An intriguing experimental hypothesis of these systems is that the formation of two dative bonds to form the dimer will lead to a multipolar framework, that is that the 6-atom heterocycle has alternating charges because of the electron donor/acceptor nature of

dative bonds. To consider this hypothesis further, charges (calculated using GAPT) and charge differences are given in Tables 7 and 8. The heavy atoms included in the data presented are B, C in the COOH group, and the N or P in each molecule. Table 7 first shows charge differences found in linear dimers. When the system includes the carboxyl group, B' designates the boron with attached carboxylic group and B" is the boron next to phosphorous in BCNBCP. N' is the nitrogen next to the boron with carboxylic group. In BCNBCN structure B' and N' are the atoms making the dative bond.

Considering the linear dimers, perhaps the most surprising aspect is that the formation of the dative bonds results in a fairly modest change in atomic charge for any of the atoms involved when a B-N dative bond is formed in these systems. While about 0.5 unit change in charge is observed when a B-P bond is created, the charge shift between boron and nitrogen in the B-N case is negligible. This observation is largely due to the fact that the much more electronegative nitrogen atom has significant electron density in the monomers, and continues to have that enhanced charge density in the dimer. While small, there is a slight tendency for the amount of charge transfer from nitrogen to boron to increase when the boron is carboxylated, thus the electron withdrawing character of COOH appears to bolster the electron density transfer from nitrogen enough that the boron gains greater electron density.

When the second dative bond forms, to yield the heterocycle, similar trends in charge distributions arise as seen in Table 8. For the BCNBCN system, the amount of net

charge transfer is relatively small in either case. When a P atom is introduced, the charge transfer at that dative bond is larger than the B-N bond in the multipolar framework, but not quite as large as when the B−P bond forms the linear dimer. Perhaps somewhat surprisingly, for BCNBCP, the presence of the phosphorous also leads to slightly larger changes in the B–N dative bond charge distributions for the carboxylated system. Nonetheless, the most important observation related to charge in these systems is that the traditional synthetic view, which essentially designates formal charges on the atoms in the multipolar framework, is not observed. In so far as quantum mechanical calculations utilize canonical wavefunctions, this is to be expected.

In BCNBCN structure, cyclic BCN monomer is lower in energy than that of linear monomer. Despite the bond angles around 60 degrees (Figure 2-a) and dihedral angles between methyl group and hydrogen atoms in boron and carbon are 5.4 and 3.9 degrees, the BCN cyclic monomer is more stable than its linear monomer. When the same structures are calculated with the methyl groups replaced by hydrogen atoms and these cyclic BCN monomers are not more stable than the linear dimers, which implies that electron donor methyl groups strengthens the B-N dative bond. This relatively high stability of BCN cyclic monomers may explain an unusual experimental observation of irreversible decomposition of BCNBCN dimer at temperatures above 160 ^{0}C.[19] The stability of cyclic BCN monomer (with methyl groups) is contrary to the stability of structures made by atoms with similar atomic radii, such as cyclopropane. To investigate the role of cyclization by dative bond formation versus ring strain further, dimers of

BCCN and BCCCN were also optimized. Energies relative to uncyclized monomers are presented in Table 9.

The cyclic monomer of BCCN (Figure 2-b) has ring bond angles of around 90 degrees and is roughly 18 kcal/mol more stable than the linear monomer. Which once again means that despite any ring strain, a boron nitrogen dative bond is apparently strong enough to hold the molecules in a cyclic structure. Furthermore, by considering all the cyclic monomers summarized in Tables 4, 5 and 9, it is apparent that the stability of cyclic monomers increases with the increasing number of atoms in the system, at least until the angles between the atoms reaches around 109 degrees. The calculated ring strain of BCN, BCCN and BCCCN monomers and dimers (Table 6) follow the expected ring strain for cyclic molecules of 3 to 10 atoms. Thus, the overall stability of heterocyclic compounds formed by dative bonds is largely governed by the competition between the strength of the dative bond and the ring strain the system must accommodate.

IV. Conclusions

This paper presents *ab initio* calculations of a range of heterocyclic compounds that contain boron. Because of the Lewis acid characteristic of three-coordinate boron atoms, a number of different intramolecular and intermolecular processes may give rise to the formation of cyclic systems via the formation of a dative bond. The relative

stabilities of these various possibilities have been determined. In many cases, the competition between the energy lowering observed from dative bond formation and ring strain associated with the formation of the cyclic system is a close such that in some cases the strained ring is more stable than an open structures and visa versa. Putative charge distributions, predicated on formal charge style analysis that suggest the electron donating atom becomes positively charged is not seen for these systems, but rather the canonical orbital bonding in which electronegativity best describes charge build-up appears to be the operative paradigm.

References

(1) B.F Spielvogel, R. F. Bartton, and C. G. Moreland, Cyanoborane chemistry. I. Preparation of a novel series of macrocyclic cyanoboranes *J. Am. Chem. Soc.*, **1972**, *94*, 8597-8598.

(2) N.E. Miller and E.L. Muetterties, Chemistry of Boranes. XVI. A New Heterocyclic Boron Compound *Inorg. Chem.*, **1964**, *3*, 1196-1197.

(3) N.E. Miller, Borane isonitriles and carboxylates. Synthesis and characterization of the 2-carboxylic acid derivative of 1,1,4,4-tetramethyl-1,4-diazonia-2,5-diboratacyclohexane *Inorg. Chem.*, **1991**, *30*, 2228-2231.

(4) S.W. Dougherty and N.E. Miller, Synthesis and chemistry of 1,1,4,4-tetramethyl-1-azonia-4-phosphonia-2,5-diboratacyclohexane, a novel multipolar framework heterocycle *Inorg. Chem.*, **1993**, *32*, 5889-5893.

(5) Y. Mo, L. Song, W.Wuand Q. Zhang, Charge Transfer in the Electron Donor−Acceptor Complex BH3NH3 *J. Am. Chem. Soc.*, **2004**, *126*, 3974-3982.

(6) J.S. Binkley, L.R. Thorne, A Theoritical study of the properties of BH3NH3. *J. Chem. Phys.*, **1983**, *79*, 2932-2940.

(7) L. R. Thorne, R. D. Suenram, F. J. Lovas, Microwave spectrum, torsional barrier, and structure of BH3NH3. *J. Chem. Phys.*, **1983**, *78*, 167-171.

(8) S. Trudel, D. F. R. Gilson, High-Pressure Raman Spectroscopic Study of the Ammonia−Borane Complex. Evidence for the Dihydrogen Bond. *Inorg. Chem.*, **2003**, *42*, 2814-2816.

(9) T. Wideman and L.G. Sneddon, Convenient Procedures for the Laboratory Preparation of Borazine. *Inorg. Chem.*, **1995**, *34*, 1002-1003.

(10) R.B. Kaner, J. Kouvetakis, C.E. Warble, M.L. M.L. Sattler and N. Bartlett, Boron-carbon-nitrogen materials of graphite-like structure. *Material Research Bulletin*, **1987**, *22*, 399-404.

(11) T.H. Hseu, MO and Vibrational Normal Coordinate Calculations of 1,1,4,4,-Teteamethyl-1,4-Diazonia-2,5-Diboratacyclohexane. *J. Mol. Struct.* **1979**, *53*, 121-127.

(12) M. S. Gordon and M. W. Schmidt, Theory and Applications of Computational Chemistry: The First Forty Years; Elsevier: Amsterdam, The Netherlands, 2005.

(13) M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, **1993**, *14*, 1347-1363.

(14) B. M. Bode and M. S. Gordon, MacMolPlt: a graphical user interface for GAMESS. *J. Mol. Graph. Modell.*, **1998**, *16*, 133-138.

(15) C. Møller and M. S. Plesset, Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.*, **1934**, *46*, 618-622.

(16) R. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, Self-consistent molecular orbital methods. A basis set for correlated wave functions. *J. Chem. Phys.*, **1980**, *72*, 650-654.

(17) M. P. ver Haag, ab initio investigation of boron binding, PhD Thesis, Iowa State University, 2012.

(18) J. Cioslowski, A new population analysis based on atomic polar tensors. *J. Am. Chem. Soc.*, **1989**, *111*, 8333-8336.

(19) T. Dudev and C. Lim, J. Am. Ring strain Energies from ab Initio Calculations. *J. Am. Chem. Soc.*, **1998**, *120*, 4450-4458.

(20) F.H. Allen, O. Kennard, D.G. Watson, L. Brammer, A.G. Orpen and R. Taylor, Tables of bond lengths determined by X-ray and neutron difraction. Part 1. Bond lenghts in organic compounds. *J. Chem. Soc. Perkin Trans. II*, **1987**, *12*, S1-S19.

21. N.E. Miller, M.D. Murphy and D. L. Reznicek, Thermal Properties of Dimethylaminomethylborane Cyclic Dimer. *Inorg. Chem.*, **1966**, *5*, 1832-1834.

22. T.H. Hseu, L.A. Larson, 1,1,4,4-Tetramethyl-1,4-diazonia-2,5-diboratacyclohexane. Crystal structure and a new preparation. *Inorg. Chem.*, **1975**, *14*, 330-334.

Table 1. Key structural features of the main cyclic structures

Molecule	BCNBCN	BCNBCP	(COOH)-BCNBCN	(COOH)-BCPBCN	BCN Monomer	BCCN Monomer	BCCCN Monomer
R_{B-N}	1.620	1.633	1.620	-	1.610	1.681	1.648
$R_{B'-N}$	-	-	1.612	1.624	-	-	-
$R_{B"-P}$	-	1.926	-	1.923	-	-	-
R_{B-C}	1.623	1.639	1.622	1.631	-	-	-
$R_{B-C"'}$	-	-	-	-	1.589	1.644	1.637
$R_{B"-C}$	-	-	-	-	-	-	-
$R_{B'-C}$	-	-	1.624	1.639	-	-	-
$R_{B"-C}$	-	1.632	-	-	-	-	-
R_{C-N}	1.508	1.513	1.504	-	1.501	1.509	1.491
$R_{C-N'}$	-	-	1.511	1.517	-	-	-
R_{P-C}	-	1.800	-	1.800	-	-	-
$R_{B'-COOH}$	-	-	1.621	1.624	-	-	-
$R_{C"'-C}$	-	-	-	-	-	1.534	1.557
R_{C-C}	-	-	-	-	-	-	1.533
$R_{N-C'H3}$	1.483	1.482	1.483	-	1.463	1.473	1.480
$R_{N-C"H3}$	1.481	1.484	1.482	-	1.463	1.468	1.475
$R_{N'-C'H3}$	-	-	1.488	1.488	-	-	-
$R_{N'-C"H3}$	-	-	1.487	1.490	-	-	-
$R_{P-C'H3}$	-	1.818	1.818	1.818	-	-	-
$R_{P-C"H3}$	-	1.820	1.819	1.819	-	-	-
A_{CNB}	109.0	112.8	108.9	-	61.3	85.9	102.2
$A_{CN'B'}$	-	-	107.1	110.7	-	-	-
A_{CPB}	-	107.8	-	107.2	-	-	-
A_{BCN}	115.1	-	-	-	-	-	-
$A_{BC"'N}$	-	-	-	-	62.7	-	-
$A_{BCN'}$	-	-	115.6	117.7	-	-	-
$A_{B'CN}$	-	-	114.9	-	-	-	-
$A_{B"CN}$	-	117.0	-	-	-	-	-
A_{BCP}	-	110.1	-	-	-	-	-
$A_{B'CP}$	-	-	-	110.4	-	-	-
$A_{BC"'C}$	-	-	-	-	-	86.4	106.1
$A_{C"'CN}$	-	-	-	-	-	95.3	-
$A_{C"'CC}$	-	-	-	-	-	-	106.3
A_{CCN}	-	-	-	-	-	-	106.1
A_{NBC}	109.7	111.7	109.5	-	-	-	-
$A_{NB'C}$	-	-	110.6	-	-	-	-
$A_{N'B'C}$	-	-	-	112.5	-	-	-
$A_{NB"C}$	-	106.9	-	-	-	-	-
$A_{NBC"'}$	-	-	-	-	56.0	85.1	99.8
A_{PBC}	-	-	-	106.6	-	-	-

Table 2. Experimental and Calculated Bond lengths for the cyclic BCNBCN structure, based on atom definitions provided in Figure 3

Bond	Calculated Value (Å)	Experimental Value (Å)
C2−B1	1.623	1.609 (2)
N3−B4	1.620	1.615 (1)
N3−C2	1.508	1.511 (1)
N6−C18	1.483	1.482 (2)
N6−C17	1.481	1.484 (2)
B4−H12	1.213	1.168 (8)
B4−H11	1.213	1.152 (7)
C2−H7	1.096	1.064 (8)
C2−H8	1.100	1.035 (8)
C18−H28	1.095	0.987 (10)
C18−H29	1.091	1.025 (9)
C18−H30	1.087	1.034 (9)
C17−H25	1.095	1.022 (9)
C17−H26	1.090	1.029 (10)
C17−H27	1.092	0.979 (10)

Table 3. Experimental and Calculated Bond angles for the BCNBCN cyclic structure, based on atom definitions provided in Figure 3

Angle	Calculated Value (degrees)	Experimental Value (degrees)
B4−N3−C2	109.0	108.2 (1)
B4−N3−C15	111.5	111.8 (1)
B4−N3−C16	108.9	109.5 (1)
C2−N3−C15	110.1	110.4 (1)
C16−N3−C15	107.6	107.5 (1)
N3−B4−H12	105.4	104.1 (4)
N3−B4−H11	107.1	106.7 (4)
N3−B4−C5	109.7	110.6 (1)
C5−B4−H12	113.1	113.3 (3)
C5−B4−H11	109.9	109.7 (4)
H11−B4−H12	111.5	112.2 (5)
N3−C2−H8	107.3	107.1 (4)
N6−C5−H9	105.0	102.8 (4)
N3−C2−B1	115.1	115.4 (1)
B1−C2−H8	109.7	111.5 (4)
B1−C2−H7	112.6	113.0 (4)
H7−C2−H8	106.7	106.1 (7)
N3−C15−H21	109.6	110.1 (6)
N3−C15−H20	109.2	110.0 (5)
N3−C15−H19	108.3	105.7 (5)
H21−C15−H20	109.1	108.8 (8)
H19−C15−H21	110.2	111.0 (8)
H20−C15−H19	110.3	111.3 (7)
N3−C16−H24	108.5	106.5 (5)
N3−C16−H23	108.9	106.8 (6)
N3−C16−H22	109.4	109.3 (6)
H22−C16−H23	110.4	111.7 (8)
H23−C16−H24	109.5	110.0 (9)
H24−C16−H22	110.2	112.3 (9)

Table 4. Energies of monomers relative to the energy of the linear monomer in each molecule

Molecules without COOH group		Molecules with COOH group	
Monomer	Relative Energy (kcal/mol)	Monomer	Relative Energy (kcal/mol)
BCN Linear Monomer	0.00	(HOOC)BCN Linear Monomer	0.00
BCN Cyclic Monomer	−16.22	(HOOC)BCN Cyclic Monomer	−20.96
BCP Linear Monomer	0.00	(HOOC)BCP Linear Monomer	0.00
BCP Cyclic Monomer	−2.50	(HOOC)BCP Cyclic Monomer	−5.22

Table 5. Energies of dimers relative to the energy of linear dimer in each molecule

Molecules without COOH group		Molecules with COOH group	
Dimer	Relative Energy (kcal/mol)	Dimer	Relative Energy (kcal/mol)
BCNBCN Linear Dimer	0.00	(HOOC)BCNBCN Linear Dimer	0.00
BCNBCN Cyclic Dimer	−51.88	BCNB(COOH)CN Linear Dimer	−9.13
		(HOOC)BCNBCN Cyclic Dimer	−59.79
BCPBCN Linear Dimer	0.00	(HOOC)BCPBCN Linear Dimer	0.00
BCNBCP Linear Dimer	−3.51	BCNB(COOH)CP Linear Dimer	−11.36
BCNBCP Cyclic Dimer	−47.90	(HOOC)BCPBCN Cyclic Dimer	−55.06

Table 6. Calculated ring strains

Molecule	Ring Strain Energy (kcal/mol)
Cyclopropane (3 atoms)	26.88
BCN Cyclic Monomer (3 atoms)	14.57
BCP Cyclic Monomer (3 atoms)	27.99
BCCN Cyclic Monomer (4 atoms)	13.04
BCCCN Cyclic Monomer (5 atoms)	0.65
BCN Cyclic Dimer (6 atoms)	−21.09
BCP Cyclic Dimer (6 atoms)	−17.11
BCCN Cyclic Dimer (8 atoms)	13.81
BCCCN Cyclic Dimer (10 atoms)	3.29

Table 7. Charges and changes in charge for linear dimer formation

Molecule	BCNBCN		BCPBCN		BCNBCP	
	charge	Δ charge	charge	Δ charge	charge	Δ charge
B	0.56	−0.03	0.64	0.04	0.62	0.01
B’	0.58	−0.01	-	-	-	-
B”	-	-	0.19	−0.41	0.57	−0.03
N	−0.61	0.01	−0.64	−0.01	−0.55	0.07
N’	−0.53	0.09	-	-	-	-
P	-	-	1.05	0.53	0.53	0.00

Molecule	(COOH)BCNBCN		BCN(COOH)BCN		(COOH)BCPBCN		BCN(COOH)BCP	
	charge	Δ charge	charge	Δ charge	charge	Δ charge	charge	Δ charge
B	0.58	−0.01	0.57	−0.03	-	-	-	-
B’	0.56	−0.02	0.54	−0.05	0.65	0.00	0.58	−0.07
B”	-	-	-	-	0.10	−0.50	0.57	−0.02
N	−0.53	0.08	−0.61	0.00	-	-	-	-
N’	−0.61	0.01	−0.51	0.12	−0.65	−0.03	−0.53	0.09
P	-	-	-	-	1.08	0.55	0.53	0.01
C (COOH)	−0.23	0.02	−0.33	−0.09	−0.24	0.02	−0.35	−0.09

Table 8. Charges and changes in charge for cyclic dimer formation

Molecule	BCNBCN		BCNBCP		(COOH)BCNBCN		(COOH)BCPBCN	
	charge	Δ charge	charge	Δ charge	charge	Δ charge	charge	Δ charge
B	0.52	−0.04	0.56	−0.08	0.51	−0.06	-	-
B’	-	-	-	-	0.48	−0.08	0.51	−0.15
B”	-	-	0.14	−0.42	-	-	0.13	−0.44
N	−0.54	0.07	−0.54	0.10	−0.53	0.09	-	-
N’	-	-	-	-	−0.52	0.09	−0.52	0.13
P	-	-	0.98	0.45	-	-	1.00	0.47
C (COOH)	-	-	-	-	−0.32	−0.09	−0.33	−0.09

Table 9. Energies of monomers and dimers relative to the total energy of linear monomers in each molecule for extended BCCN and BCCCN

Energy relative to linear monomers		Energy relative to linear dimers	
Monomer	Relative Energy (kcal/mol)	Monomer	Relative Energy (kcal/mol)
BCCN Linear Monomer	0.00	BCCN Linear Dimer	0.00
BCCN Cyclic Monomer	−17.75	BCCN Cyclic Dimer	−16.98
BCCCN Linear Monomers	0.00	BCCCN Linear Dimer	0.00
BCCCN Cyclic Monomer	−30.15	BCCCN Cyclic Dimer	−27.50

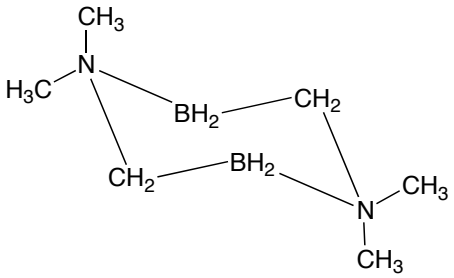


Figure 1-a

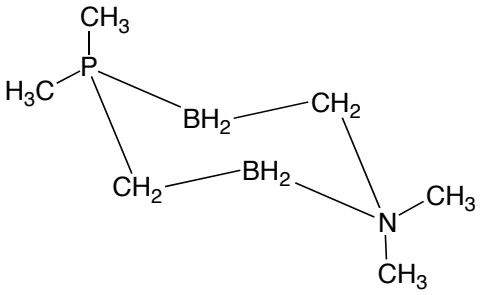


Figure 1-b

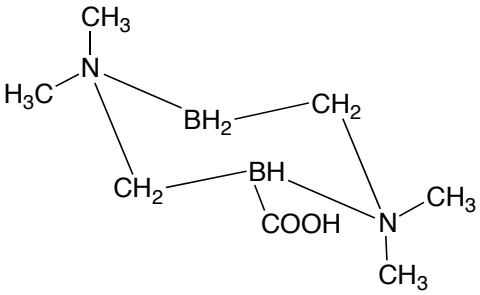


Figure 1-c

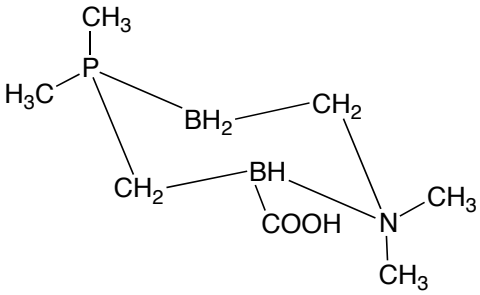


Figure 1-d

Figure 1. Basic structures for all the cyclic dimer systems.

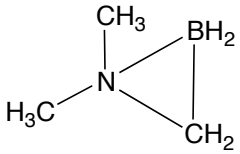


Figure 2-a

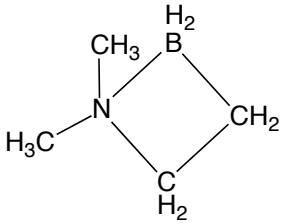


Figure 2-b

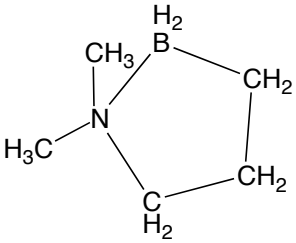


Figure 2-c

Figure 2. Basic structures for self-cyclized monomers.

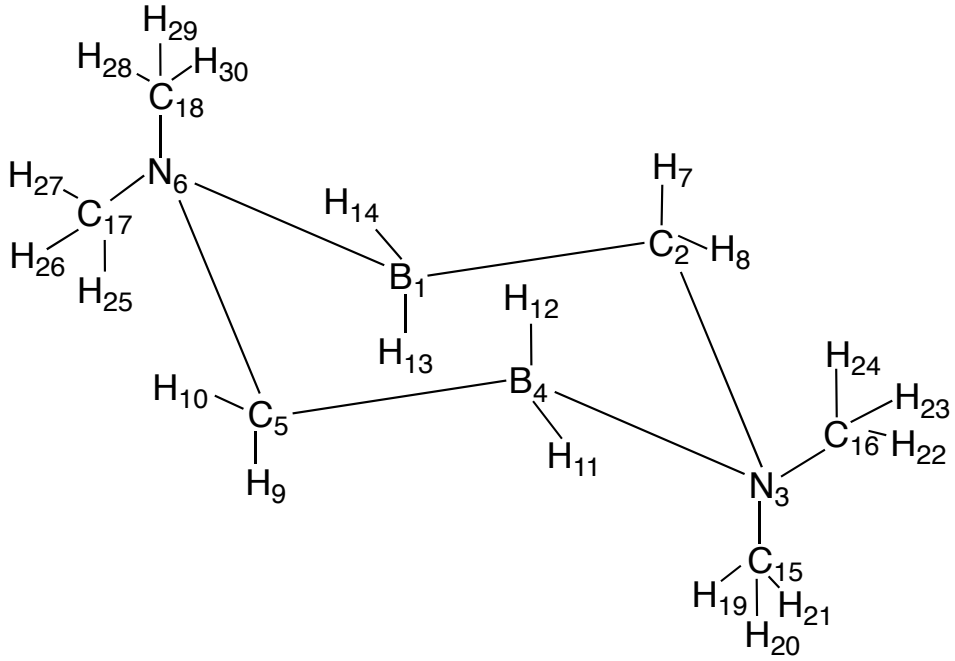


Figure 3. BCNBCN cyclic structure with atom designation for structural data.

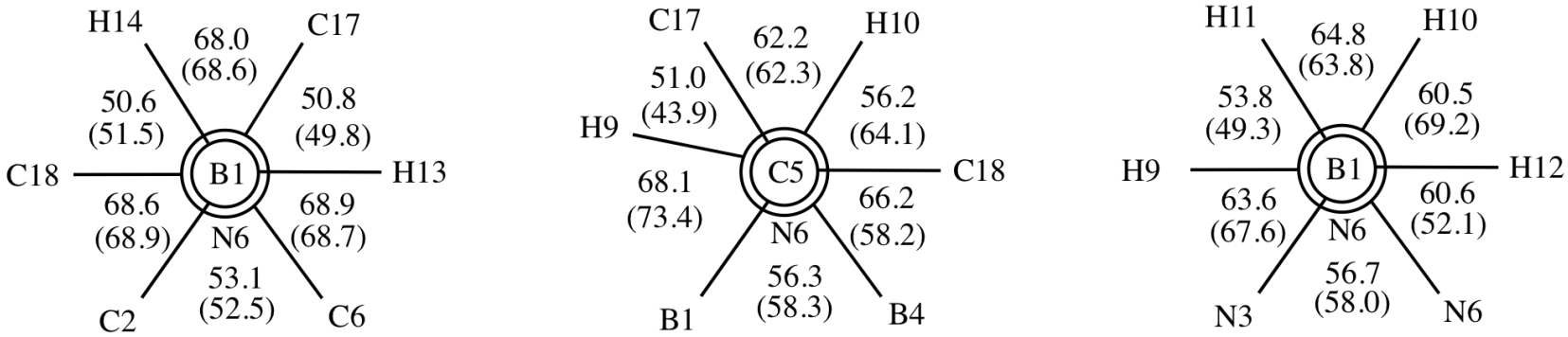


Figure 4. Torsional angles of BCNBCN ring conformation. Experimental values are given in parentheses.

CHAPTER 3. MASS SPECTRAL AND THEORETICAL INVESTIGATION OF 2-AMINOETHOXYDIPHENYL BORATE IN SOLUTION

A paper submitted to *Physical Chemistry Chemical Physics*

Chamila C. De Silva, Marsha A. Collins and Thomas A. Holme

Abstract

2-Aminoethoxydiphenyl borate (2-APB) is a reliable blocker of store-operated Ca^{2+} entry in cells. 2-APB is a five membered ring with boron nitrogen dative bond, which has a probability of dimerizing. Ab-initio calculations done on this molecule have been limited to monomer structure and only a few experimental studies mention the existence of a dimer. In this study mass spectroscopy has been used to provide evidence of the dimerization of 2-APB in water. The peak for the dimer, formed by two dative B-N bonds is present only in electrospray mass spectra. Harder ionizing mass spectral techniques show only monomers and smaller fragments. Fragmentation patterns in the electrospray experiments show not only monomers and dimers but also additional cyclic compounds and are explained by ab initio calculations done on both gas and solvent phase on monomer, dimer and fragment geometries. Thus 2-APB can exist as both monomer and a dimer structure in solvent phase and the free energy of the dimer structure is about 2 kcal/mol lower than cyclic-monomer.

I. Introduction

Determining a molecular structure represents a key component of understanding the pharmacological mechanism of any physiologically active molecule. To the extent that many pharmacores derive their activity from interaction with proteins, a clear picture of the structure of a molecule is vital. This prerogative becomes even more compelling when affected enzymes are unidentified or the structure of the enzyme active site is unknown. Calcium ion movement in cells represents one physiological process where small molecule pharmacores show a selective effect, but the mechanism for both the effect and its selectivity remain unknown.

The movement of intraluminal calcium ions is important for many cellular processes conduct by Endoplasmic Reticulum (ER) including protein folding, vesicle trafficking,[1,2] release of stress[3] signals and apoptosis.[4] In addition to these, the ER is well known for its function as an agonist-sensitive Ca^{2+} store and sink.[5] Inositol 1,4,5-trisphosphate receptors ($InsP_3R$) are calcium channels in the ER membrane. IP_3Rs release calcium from ER[6] when they are activated by the second messenger inositol 1,4,5-trisphosphate (IP_3). Because of the role of ER as a source of Ca^{2+} it is important that the cell maintain the Ca^{2+} levels in the ER from falling too low. Refilling of Ca^{2+} in to the ER is central process for all eukaryotic cells and is done by store-operated calcium channels

(SOCCs) in the plasma membrane. This process is modeled as capacitative Ca^{2+} entry or store-operated Ca^{2+} entry model where calcium pools are refilling by extracellular Ca^{2+} in the presence of IP_3R receptors and closing the pathway from Ca^{2+} pools to cytosol when the agonist IP_3R is removed.[7]

Alterations to the activity of Ca^{2+} signaling pathways to generate inappropriate Ca^{2+} responses which are either too high or too low can leads to many diseases.[8] Cardiac hormones like ET-1 (endothelin-1) or α-adrenergic agents in atrial cells generate IP_3, and as a result IP_3/Ca^{2+} signaling pathway increase the force of contraction of the atrial cells which leads to irregular heart beat or atrial arrhythmias.[9] The Ca^{2+} hypothesis of Alzheimer's disease has argued that a permanent elevation of Ca^{2+} concentration due to the increase in the activity of IP_3R in neuronal cells cause the erasure of memory.[10]

Development in the research of store-operated Ca^{2+} influx has been slowed down by the lack of relatively specific inhibitors of the underlying Ca^{2+} channels.[11] 2-Aminoethoxydiphenyl borate (2-APB) is a membrane permeable agent use as a calcium inhibitor in cells. It has been found that 2-APB inhibits IP3 receptors to block the increase in intracellular calcium concentration in rabbit atrial cells.[12] By contrast, another study has found that 2-APB inhibits SOCC but not through a mechanism involving IP3. Rather binding of 2-APB to either the channel protein or an associated regulatory protein is implicated in rat liver cells.[13] Also 2-APB appears to be an inconsistent inhibitor of

$InsP_3R$ but can universally serve as a SOCC inhibitor.[11,14] While there are debates about whether 2-APB blocks IP_3R or SOCC channel, studies show that its effect is extending beyond these areas. 2-APB has been found to inhibit Ca^{2+} influx through TRPM2 channels in Rat Dorsal Root Ganglion neurons in response to oxidative stress,[15] inhibits sarco/endoplasmic reticulum Ca^{2+}- ATPase SERCA pumps,[16] inhibits Ca^{2+} efflux from mitochondria,[17] in the absence of other stimuli 2-aminoethoxydiphenyl borate activates heat-activated cation channels, TRPV1, TRPV2, and TRPV3.[18] Moreover it blocks electrical coupling and inhibits voltage-gated K^+ channels in guinea pig arteriole cells.[19]

The structure of 2-APB has been proposed in the literature to be either a monomer or a dimer. Knowledge of its precise structure in biological media is important because the biological activity of a pharmacophore is often determined by its three dimensional placement. The X-ray crystal structure data of 2-APB predicts its existence as a monomer with dative bond between the nitrogen and the boron forming a five-membered boroxazolidine ring.[20,21] An NMR study has predicted the stable five membered heterocyclic structure.[22] An open-chain monomer structures with and with out positively charged amines were suggested.[23,24] A dimerized form of 2-APB has also been suggested.[25,24] An *ab initio* study has been reported for the structure of the five membered ring with the phenyl groups replaced by methyl groups[26] and conformational studies of the linear monomer structure have also been reported,[27] but no *ab initio* studies have been reported on the full dimer structure of 2-APB. This study reports *ab initio* results for both monomer and dimer structures of 2-APB in solvent phase with all the conformational

studies. Mass spectral data with several ionization methodologies for 2-APB are also presented, to provide a comparison of calculated structures to experimental data.

II. Computational Details

2-Aminoethoxydiphenyl borate was obtained from Aldrich chemical. The electronic ionization (EI) and chemical ionization (CI) mass spectra were acquired using a Hewlett-Packard 5985 Mass Spectrometer equipped with a temperature programmable Direct Insertion Probe (DIP). Solid samples were introduced directly into spectrometer via the DIP and heated from 50-350^0 C at 30^0 C/minute, while scanning a mass range of 100-800 amu. Data were acquired and processed using HP Chemstation software. The electrospray mass spectra were acquired using an Agilent 1100 LC/MSD. This sample was introduced via flow injection in 100% aqueous solution.

All *ab initio* calculations were performed using the GAMESS,[28,29] electronic structure code, and the molecules were visualized with MacMolPlt.[30] All the Electronic structure calculations were carried out at Moller-Plesset perturbation theory (MP2) level[31] with 6-311G(d,p)[32] basis set as this combination has found to give better energies comparable to experimental values for systems with B-N dative bond in a previous study.[27] Conformational searches for all heterocyclic rings were carried out using the Complete Rotation from the Evaluation of Potential Energy Surface (CREPES)[33] program. To find the multiple low-lying conformers in all three structures, one phenyl

group was rotated with respect to the rotation of other phenyl groups. In linear and cyclic dimer structures additional rotations were done by rotating the linear and cyclic structure at two different points in addition to the rotations of phenyl groups. For all the above conformations single point energies were calculated using MP2/6-311G(d,p) (PCM)[34] level and the lowest energy conformations were optimized using at MP2/6-311G(d,p) (PCM) with water as the solvent. Calculations designed to confirm the stability of fragments that occur in the mass spectrum were done in gas phase with the same level of theory and the same basis set.

The free energy of dissociation of the cyclic monomer and the dimerization energies with respect to linear monomers and cyclic monomers were calculated using the equation

$$G = -RT \ln\left(\frac{\sum P_{\text{products}}}{\sum P_{\text{reactants}}}\right) \tag{1}$$

The Boltzmann distributions of conformers were generated using the equation

$$P_x = \left(\frac{e^{-\frac{E_x}{kT}}}{\sum_i e^{-\frac{E_i}{kT}}}\right) \tag{2}$$

E_i is the electronic energy with Zero point energy correction at 298.15K for the i^{th} conformer.

To estimate the influence of explicit solvent molecules, calculations were carried out with the addition of 60 Effective fragment potential (EFP) waters in the case of linear and cyclic monomers and 120 EFP waters in the case the dimer structure. Potential minimum structures for the water clusters were obtained via a Monte Carlo simulated annealing method.[35,36] These solvent configurations were then used to optimize using MP2/6-311G(d,p) with the whole cluster embedded in a continuum solvent modeled by PCM to get the final structures. Obtaining numerical hessians with EFP calculations in GAMESS is prohibitively time consuming for systems of the size calculated here, so the internal vibration frequencies of the fragments were not available. Therefore all the energies calculated are electronic energies without zero point energy corrections. For the water clusters used, the largest components of these corrections would arise from the water, and would change quite modestly for the various configurations, so the impact of this limitation on the energy differences can be expected to be small.

III. Results and Discussion

The electron ionization (EI) spectrum for 2-APB is presented in Figure 1. The peak assignments are listed in Table 1. The absence of peaks at masses greater than 255 amu either indicates that 2-APB does not exist as a dimer or that the fragmentation caused by EI is extensive enough that there is no molecular peak for the dimer. Assuming that the former is true, the fragmentation pattern does not indicate the presence of one monomer structure over another. A fragment containing B-N would be indicative of the

monomer ring structure. However, as O and NH_2 possess similar masses, several fragments would present as equivocal peaks. These include the phenyl-boron-oxygen and phenyl-boron-amide peaks at 105 amu, and the phenyl, phenyl-boron-oxygen and phenyl, phenyl-boron-amide peaks at 182 amu.

To reduce the fragmentation of the sample, and possibly detect a peak at a mass greater than that of the monomer, a softer ionization method was employed, specifically chemical ionization (CI). This spectrum is shown in Figure 2 and peak assignments are given in Table 2. While the softer ionization does lead to less extensive fragmentation relative to what was seen for the EI spectra the results about the existence of the dimer remain inconclusive in this spectrum as well. Thus, with chemical ionization the mass spectrum indicates that the dimer structure is either not present or fragmented via the CI ionization process.

An even softer method of ionization, electrospray ionization (ESI), was used to further reduce the fragmentation of the sample. While ESI imparts sufficient energy to fragment the observed molecules, the experimental parameters are such that such fragmentation is commonly limited by kinetics, so that larger molecules can be observed unfragmented[36]. The spectrum resulting from the ESI method and the peak assignments are presented in Figure 3 and Table 3. The peaks indicate the presence of the monomer at 225 amu, the dimer at 451 amu, a peak of the dimer with sodium which is a common

feature in ESI mass spectrometry, and a fragment at 287.2 amu. This fragment is new, and a proposed reaction that leads to this fragment is provided in Figure 4.

The assignments of the dimer and fragment peaks are confirmed by the isotopic breakdown shown in Figure 5, which indicates that the dimer peak (Figure 5(a)) contains two boron atoms and the fragment observed at 287.2 amu contains only one boron atom. The mass of this molecule is consistent with a fragmentation where a boron with two phenyl is lost and then second fragment of boron with two phenyl groups is formed. Importantly, these predicted isotopic abundance patterns match the intensities observed in the electrospray mass-spectrum shown in Figure 3 for both the dimer and this newly proposed molecule derived from fragmentation.

To further characterize the results of these mass spectroscopic experiments, quantum mechanical calculations were carried out to determine both structural and energetic information about the 2-APB system. The structures of three important molecules including some key structural features are provided in Figure 6.

Bond lengths and angles for the 2-APB monomer may be compared with those obtained from crystal structures using X-ray crystallography.[21] Overall, the RMS deviation for bond distances is 0.016 Å. Much of this deviation is the result of experimental C-C bond distances in the benzene rings being shorter than the 1.40

distances obtained in these calculations, and this distance is commonly accepted for benzene. For bond angles the RMS deviation is 0.55°, and the majority of this discrepancy arises in three angles, in the ring, which for the calculations here are between 1.5 and 2.5 degrees smaller than for the crystal structure. Calculated 2-APB cyclic monomer in solvent medium maintains the same envelope structure as the structure obtained by X-ray crystallography and the structures given by previous theoretical calculations,[27] where in all cases the same CH_2 group coming out of plane. Gibbs free energy for the dative bond dissociation in 2-APB cyclic monomer is 14.07 kcal/mole, which suggests that 2-APB is more stable as a cyclic monomer than a linear monomer (Figure 6). The available experimental data also does not support an existence of a linear structure.[25]

The first method used to include the effect of solvent is the PCM continuum model. In these calculations, the ten-member ring of 2-APB dimer takes a chair-boat-chair like conformation. Calculated B-N bond distances for this dimer (Figure 6) are 1.661 Å and 1.663 Å and bond angles in OBN are 109.7 and 109.6 degrees. Conformer searches were performed using CREPES[33] to identify the low-lying conformers for this system. Over 7800, 10-member dimer conformers and 700 linear conformers and 400 cyclic conformers were tested via this process to search for the lowest energy ones. The energies of the lowest energy conformers are given in Table 4. The conformational structures of cyclic dimer, cyclic monomer and linear monomer are labeled as D, C and L respectively. The primary structural difference between the different conformers were the

composite rotations of four phenyl groups in all structures. In addition to that in 10-member ring conformational structures were found by rotating around the oxygen-oxygen axis and the carbon-carbon axis (carbon next to nitrogen) and in the linear structure conformations were taken by rotating CH_2 and NH_2 groups around the axis of oxygen-carbon. In the cyclic monomer CH_2 group was rotated using carbon and oxygen as the axis. All the energies are given with respect to the energy of the lowest cyclic dimer, D1. In this comparison, energy of the dimer structure is compared to the total energy of two monomer structures. All the energies given here are electronic energy with the zero point energy correction.

Energetically, dimerization lowers the free energy in solvent phase with respect to the linear monomers by -30.22 kcal/mole. Note, however, that taking the two linear monomers and allowing them to self-cyclize lowers the free energy by 28.14 kcal/mole. The dimerization free energy in solvent with respect to the cyclic monomer is -1.93 kcal/mole. Since the free energy difference between 2 cyclic monomers and the cyclic dimer are fairly low, it is reasonable to predict that both the dimer and cyclic monomer exist in solvent at ambient temperatures.

Because the utility of 2-APB arises from use in living systems, the role of non-covalent interactions are potentially important. The 2-APB molecule has several candidate sites for hydrogen bonding with water molecules, including both N–H and O–

H bonds and nitrogen and oxygen lone pairs. In order to adjudicate the possible effect of hydrogen bonding between water molecules in the solvent and the 2-APB molecules explicit solvent molecules must be added. To this end, calculations were carried out with 60 effective fragment potential (EFP) waters with further solvent accounted via PCM for both linear and cyclic monomers. To allow energy comparisons the same calculation requires 120 EFP waters to be added within PCM for the dimer structure. Minimal structures obtained from Monte Carlo simulated annealing method were optimized using MP2/6-311G(d,p) PCM to get the final structures. As noted earlier, computational limitations for hessians of these large systems means that for this comparison electronic energies are used without zero point energies corrections.

The dative bond dissociation energy in 2-APB cyclic monomer with 60 EFP waters is 0.48 kcal/mol higher than the dative bond dissociation energy with PCM. One might expect the bond dissociation energy would go down with all the possible hydrogen bonding the linear monomer could make but the orientation of the oxygen and nitrogen in the minimal structure of linear monomer does not favor for the maximum number of hydrogen bonding. Therefore we looked at different conformations with the ability to make hydrogen bonding. We used 20 different linear monomer conformation energies used in Table 5 in a bootstrap estimate of variance using STATA, Data Analysis and Statistical Software Programe[35.] In this program we changed the number of replications from 20 to 100000 and observed the mean value of energy with 95% Confident interval (Table 6).

The average energy for the linear monomer conformation taken from the bootstrap estimate of variation was -697.1564 Hartree. The cyclic monomer structure with 60 EFP waters makes the expected four hydrogen bonds; as a result additional statistical calculations were not necessary for the cyclic monomer. When we use the average energy we obtained from the bootstrap variance, bond dissociation energy becomes is 8.38 kcal/mol higher than that of with PCM. Considering the structures of above 20 linear conformations and the calculated value of bootstrap average it is clear that the linear monomer does not easily make the five hydrogen bonds, as we would expect. The position of oxygen atom which is expected to make 2 hydrogen bonds is placed in between two phenyl groups which resists the water in the solvent to make hydrogen bonds with oxygen as expected. Also in the lowest energy structure one of the hydrogens in the NH2 is stabilized over the electron cloud of one of the phenyl groups and is not in the region of making possible hydrogen bonds. Also we calculated the distances between all the oxygen atoms on EFP waters on the linear structure, in which we found that out of 60 EFP waters approximately 80 oxygen-oxygen bond distances are below 4.00 angstroms. This concludes that making hydrogen bonds in between solvent is more favorable than making hydrogen bonds with the solute. In conclusion, for a large molecule with higher degree of freedom the ability of making hydrogen bonds does not depend only on the number of atoms capable of making hydrogen bonds but also on the other factors like flexibility of the molecule and the groups attached to it. The dimerization energy of cyclic monomer and linear monomer are 4.23 kcal/mol and 21.00 kcal/mol more stable than that of with out EFP waters.

These free energy differences are fairly modest, so an additional experiment to determine the molecular mass of 2-APB by observation of the freezing point depression was conducted. Although only a small amount of material needs to dissolve, 2-APB was observed to be not soluble enough in water or reagent-grade cyclohexane to affect the freezing point of these solvents. However, a solution containing 1.0870g of 2-APB and 50.00ml of reagent grade benzene was observed to freeze at 5.20 ^{0}C. This is a freezing point depression of 0.32 ^{0}C, indicating a 2-APB molar mass of 390 g/mol. This molar mass is greater than that of monomer structure and less than that of dimer structure, indicating the existence of equilibrium between both structures in benzene solution. The resulting equilibrium constant of this solution is 1.4, slightly favoring the dimer. Hence in solution 2-APB can exists in equilibrium with its monomer and dimer structure.

The fragment observed at 287.2 amu in ES spectrum is identified in the reaction scheme suggested in Figure 4. The nine-member ring is a stable minimum on the potential energy surface of the system because the Hessian matrix is positive-definite. This experiment appears to be the first report, experimentally or theoretically, of this molecule. Nonetheless, the calculated reaction enthalpy in gas phase for the fragmentation reaction shown in Figure 4, is +103.51 kcal/mole and in solvent phase it is +104.52 kcal/mole. The free energies of fragmentation are also large and positive; $\Delta_{gas}G$ = +91.55 kcal/mol and $\Delta_{solv}G$ = +91.05 kcal/mole. These large, positive free energies for

both reactions reveal that the somewhat exotic nine-member ring observed due to fragmentation is possible because of the high-energy conditions associated with mass spectrometry.

IV. Conclusions

This paper reports experimental and theoretical investigations into the structrure of the 2-APB molecule. Because this molecule acts as a selective inhibitor of capacitative calcium uptake in cells, it's structure in physiological solutions is likely to be important in terms of understanding it's observed activity. A key question is whether or not the molecule dimerizes via the formation of two dative bonds. Evidence for dimerization is found in three ways. First, and foremost, a dimer peak is found using electrospray mass spectrometry. This technique uses a particularly soft ionization method, and under these conditions the dimer is observable. Electron ionization and chemical ionization mass-spectra do not show the dimer peaks. Second, while the molecule shows limited solubility in water or cyclohexane, solution concentrations capable of showing freezing point depression are achievable in benzene as a solvent. Results from this work lead to a molar mass that is consistent with equilibrium between monomers and dimers.

Finally, *ab initio* calculations have also been carried out with full optimization at the MP2 level of theory. The results of these calculations show that the ten-member ring

resulting from dimerization via the formation of two dative bond has a considerably lower free energy than the separated, but open chain monomers and slightly lower energy than the self-cyclizing monomers.

Taken in total, the results of this study suggest that it is possible that 2-APB has an equilibrium between monomers and dimers in the relevant physiological solutions. Thus, either form of the system could be responsible for the molecular dynamics that lead to the pharmacological action. For example, if the dimer were the specific form that inhibits calcium uptake by binding to a key protein, the results presented here would allow for this mode of action. Once bound, dimers would be removed from the equilibrium, so the mode of action might actually drive formation of the dimers. There has been no structural hypothesis put forward about the nature of the molecular interactions inherent in 2-APB activity, and this study suggests that such a hypothesis could be consistent with the basic thermodynamics of the system with either monomer or dimer indicated as the active species.

References

(1) Beckers, C. J. M.; Balch, W. E., Calcium and GTP: Essential Components in Vesicular Trafficking between the Endoplasmic Reticulum and Golgi Apparatus. *J. Cell. Bio.* **1989**, *108*, 1245-56.

(2) Gorelick, F. S.; Shugrue, C., Exiting the endoplasmic reticulum. *Mol. Cell. Endocrinol.* **2001**, *177*, 13-18.

(3) Kaufman, R. J., Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. *Genes. Dev.* **1999**, *13*, 1211-1233.

(4) Ferri, K. F.; Kroemer, G., Organelle-specific initiation of cell death pathways. *Nat. Cell Biol.* **2001**, *3*, E255-E263.

(5) Berridge, M. J., The endoplasmic reticulum: a multifunctional signaling organelle. *Cell. Calcium.* **2002**, *32*, 235-249.

(6) Akimzhanov, A. M.; Boehning, D., IP3R function in cells of the immune system. *WIREs Membr. Transp. Signaling* **2012**, *1*, 329-339.

(7) Putney Jr, J. W., A model for receptor-regulated calcium entry. *Cell. Calcium.* **1986**, *1*, 1-12.

(8) Berridge, M. J., Calcium signalling remodelling and disease. *Biochem. Soc. Trans.* **2012**, *40*, 297-309.

(9) Kockskamper, J.; Zima, A. V.; Roderick, H. L.; Pieske, B.; Blatter, L. A.; Bootman, M. D., Emerging roles of inositol 1,4,5-trisphosphate signalling in cardiac myocytes. *J. Mol. Cell. Cardiol.* **2008**, *45*, 128-147.

(10) Cheung, K. H.; Shineman, D. M. l., M.; ; Ca´rdenas, C.; Mei, L.; Yang, J.; Tomita, T.; Iwatsubo, T.; Lee, V. M.; Foskett, J. K., Mechanisms of Ca2+ disruption in Alzheimer's disease by presenil in regulation of InsP3 receptor channel gating. *Neuron* **2008**, *58*, 871-883.

(11) Parekh, A. B.; Putney, J. J. W., Store-Operated Calcium Channels. *Physiol. Rev.* **2005**, *85*, 757-810.

(12) Wacker, M. J.; Kosloski, L. M.; Gilbert, W. J. R.; Touchberry, C. D.; Moore, D. S.; Kelly, J. K.; Brotto, M.; Orr, J. A., Inhibition of Thromboxane A2-Induced Arrhythmias and Intracellular Calcium Changes in Cardiac Myocytes by Blockade of the Inositol Trisphosphate Pathway. *J. Pharmacol. Exp. Ther.* **2009**, *331*, 917-924.

(13) Gregory, R. B.; Rychkov, G.; Barritt, G. J., Evidence that 2-aminoethyl diphenylborate is a novel inhibitor of store- operated Ca2+ channels in liver cells, and acts through a mechanism which does not involve inositol trisphosphate receptors. *Biochem. J.* **2001**, *354*, 285-290.

(14) Bootman, M. D.; Collins, T. J.; Mackenzie, L.; Roderick, H. L.; Berridge, M. J.; Peppiatt, C. M., 2-Aminoethoxydiphenyl borate (2-APB) is a reliable blocker of

store-operated Ca2+ entry but an inconsistent inhibitor of InsP3-induced Ca2+ release. *FASEB J.* 2002, *16*, 1145-1150.

(15) Nazıroğlu, M.; Özgül, C.; Çelik, Ö. Ç., B; ; Sözbir, E., Aminoethoxydiphenyl borate and flufenamic acid inhibit Ca2+ influx through TRPM2 channels in rat dorsal root ganglion neurons activated by ADP-ribose and rotenone.. *J. Membr. Biol.* **2011**, *241*, 69-75.

(16) Bilmen, J. G.; Wootton, L. L.; Godfrey, R. E. S., O. S.;; Michelangeli, F., Inhibition of SERCA Ca2+ pumps by 2-aminoethoxydiphenyl borate (2-APB). *Eur. J. Biochem.* **2002**, *269*, 3678-3687.

(17) Prakriya, M.; Lewis, R. S., Potentiation and inhibition of Ca2+ release-activated Ca2+ channels by 2-aminoethyldiphenyl borate (2-APB) occurs independently of IP3 receptors. *J. Physiol.,* **2001**, *536*, 3-19.

(18) Hu, H. Z.; Gu, Q. H.; Wnag, C. B.; Colton, C. K.; Tang, J. S.; Kinoshita-Kawada, M.; Lee, L. Y.; Wood, J. D.; Zhu, M. X., 2-aminoethoxydiphenyl borate is a common activator of TRPV1, TRPV2, and TRPV3. *J. Biol. Chem.* **2004**, *279*, 35741-35748.

(19) Ma, K. T.; Guan, B. C.; Yang, Y. Q.; Nuttall, A. L.; Jiang, Z. G., 2-Aminoethoxydiphenyl borate blocks electrical coupling and inhibits voltage-gated K+ channels in guinea pig arteriole cells. *Am. J. Physiol. Heart. Circ. Physiol.* **2011**, *300*, H335-H346.

(20) Rettig, S. J.; Trotter, J., Crystal and Molecular Structure of B,B- Diphenylboroxazolidine (2-AminoethylDiphenylborinate), Ph2BO(CH2)2NH2. *Can. J. Chem.* **1973**, *51*, 1288-1294.

(21) Rettig, S. J.; Trotter, J., Crystal and molecular structure of B,B-bis(p-tolyl)bsroxazolidine and the orthorhsmbic form of B,B-diphenylboroxazolidine. *Can. J. Chem.* **1976**, *54*, 3130-3141.

(22) Farfan, N.; Castillo, D.; P.;, J.-N.; Contreras, R.; Szentpaly, L., Through-bond Modulation of N+B Ring Formation shown by N M R and X-Ray Diffraction Studies of Borate Derivatives of Pyridyl Alcohols. *J. Chem. Soc. Perkin Trans. II* **1992**, 527-532.

(23) Maruyama, T.; Kanaji, T.; Nakade, S.; Kanno, T.; Mikoshiba, K., 2APB, 2-Aminoethoxydiphenyl Borate, a Membrane-Penetrable Modulator of Ins(l,4,5)P3-Induced Ca2+ Release. *j. Biochem.* **1997**, *122*, 498-505.

(24) van Rossum, D. B.; Patterson, R. L.; Ma, H.; Gill, D. L., Ca2+ Entry Mediated by Store Depletion, S-Nitrosylation, and TRP3 Channels Printed in U.S.A. *J. Biol. Chem.* **2000**, *122*, 28562-28568.

(25) Dobrydneva, Y.; Blackmore, P., 2-Aminoethoxydiphenyl Borate Directly Inhibits Store-Operated Calcium Entry Channels in Human Platelets. *Mol. Pharm.* **2001**, *60*, 541-552.

(26) Hopfl, H.; Galvan, M.; Farfan, N.; Santillan, R., Ab initio study of substituted 2-aminoethylborinates. *J. Mol. Struct. (Theochem)* **1997**, *427*, 1-13.

(27) LeTourneau, H. A.; Birsch, R. E.; Korbeck, G.; Radkiewicz-Poutsma, J. L., Study of the Dative Bond in 2-Aminoethoxydiphenyl Borate at Various Levels of Theory: Another Poor Performance of the B3LYP Method for B-N Dative Bonds. *J. Phys. Chem. A* **2005**, *109*, 12014-12019.

(28) Gordon, M. S.; Schmidt, M. W., Theory and Applications of Computational Chemistry: The First Forty Years. *Elsevier: Amsterdam, The Netherlands* 2005.

(29) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347-1363.

(30) Bode, B. M.; Gordon, M. S., MacMolPlt: a graphical user interface for GAMESS. *J. Mol. Graph. Modell.* **1998**, *16*, 133.

(31) Møller, C.; Plesset, M. S., Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618-622.

(32) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A., Self-consistent molecular orbital methods. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650-654.

(33) Var Hagg, M. P.; Holme, T. A., ab initio investigation of boron binding. *PhD Thesis, Iowa State University* 2012.

(34) Tomasi, J.; B.;, M.; R., C., Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999-3093.

(35) StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

(36) Gabelica, V.; DePauw, E. Internal energy and fragmentation of ions produced in electrospray sources. *Mass. Spect. Rev.* **2005**, *24*, 566-587.

Table 1. Peak assignments for EI mass spectrum of 2-APB

m/z (amu)	Peak Assignment
78	phenyl group
91	Phenyl-BH_3
105	Phenyl-boron-oxygen or phenyl-boron-amide
148	monomer less one phenyl
163	boron and two phenyl groups
182	phenyl, Phenyl−boron-oxygen or phenyl, Phenyl−boron-amide
224	2-APB monomer

Table 2. Peak assignments for CI mass spectrum of 2-APB

m/z (amu)	Peak Assignment
105	phenyl-boron-oxygen or phenyl-boron-amide
119	phenyl-boron-oxygen-methyl or phenyl-boron-amide-methyl or boron- phenyl, oxygen, amide
148	monomer less one phenyl
165	boron and two phenyl groups
183	Phenyl, Phenyl-boron-oxygen or Phenyl, Phenyl-boron-amide
226	2-APB monomer

Table 3. Peak assignments for ES mass spectrum of 2-APB

m/z (amu)	Peak Assignment
225	2-APB monomer
287	New dimer fragment – identity hypothesized here
451	2-APB dimer
473	2-APB dimer + Na^+

Table 4. Relative energies (in kcal/mol) of linear (L), cyclic monomer (C) and dimer (D) structures

Structure	Relative Energy	Structure	Relative Energy	Structure	Relative Energy
D1	0.0	C1	1.9	L1	30.1
D2	0.2	C2	2.2	L2	31.0
D3	0.8	C3	3.1	L3	31.4
D4	4.7	C4	3.5	L4	31.6
		C5	3.5	L5	31.7
				L6	31.8
				L7	31.9

Table 5. Energies of 20 different linear monomer conformations

Linear +60 EFP waters	Energy (Hartree)	Linear +60 EFP waters	Energy (Hartree)
LE1	−697.1633	LE11	−697.1413
LE2	−697.1640	LE12	−697.1523
LE3	−697.1640	LE13	−697.1407
LE4	−697.1599	LE14	−697.1566
LE5	−697.1678	LE15	−697.1547
LE6	−697.1651	LE16	−697.1384
LE7	−697.1607	LE17	−697.1414
LE8	−697.1690	LE18	−697.1636
LE9	−697.1583	LE19	−697.1522
LE10	−697.1636	LE20	−697.1511

Table 6. Bootstrap estimate of variance of mean energy

Replications	Observed Coeff.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf. Interval]	
20	−697.1564	0.0019622	−3.60E+05	0	−697.16	-697.1526
100	−697.1564	0.0022099	−3.20E+05	0	−697.16	-697.1521
1000	−697.1564	0.0021103	−3.30E+05	0	−697.16	-697.1523
5000	−697.1564	0.0020822	−3.30E+05	0	−697.16	-697.1523
10000	−697.1564	0.0021006	−3.30E+05	0	−697.16	-697.1523
100000	−697.1564	0.0020992	−3.30E+05	0	−697.16	-697.1523

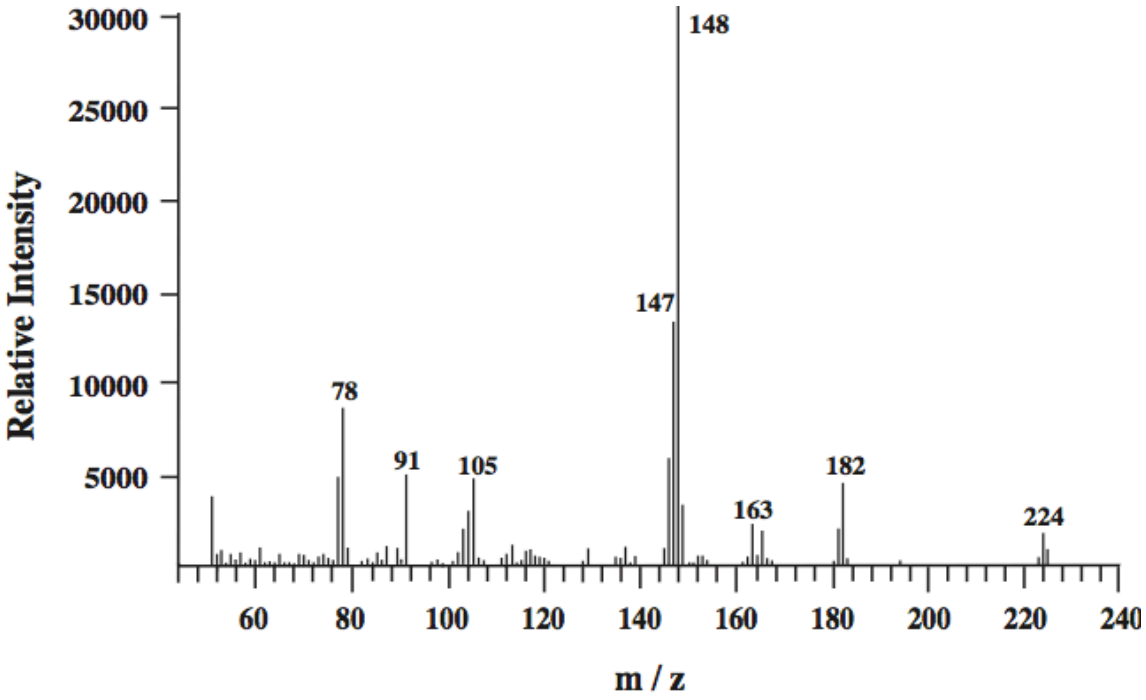


Figure 1. Electron Ionization mass spectrum of 2-APB.

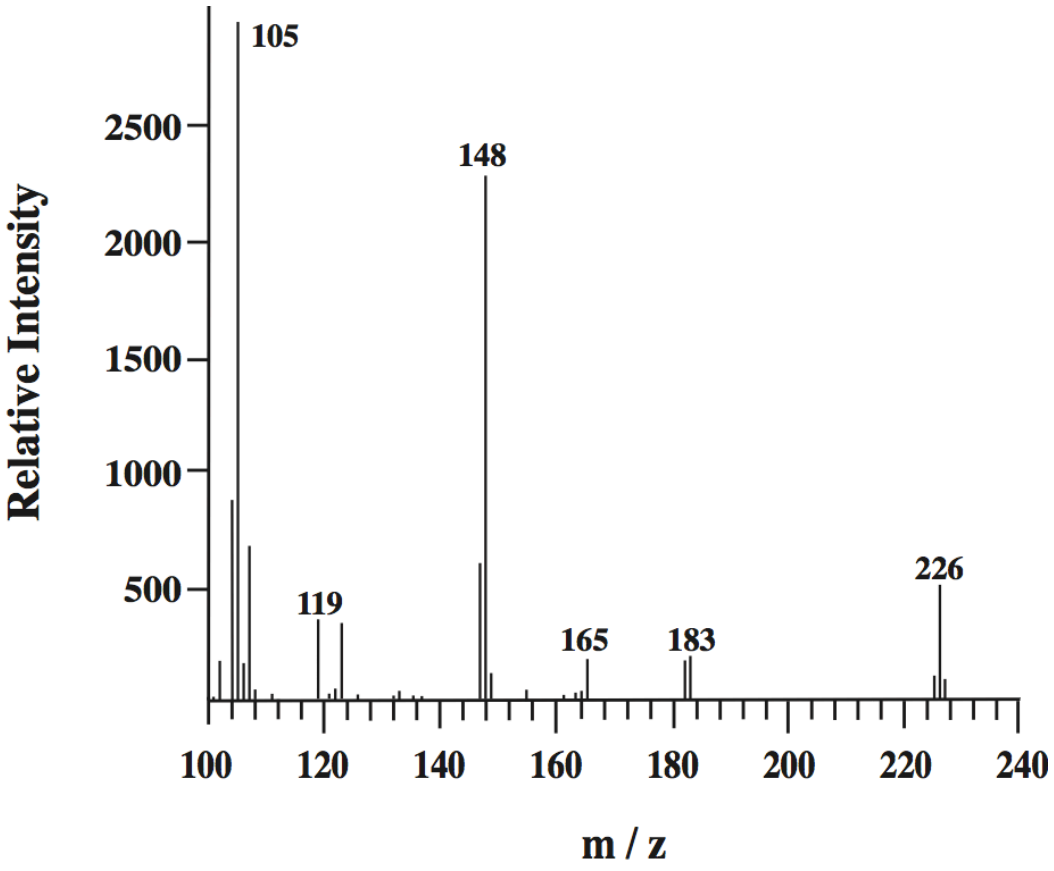


Figure 2. Chemical Ionization mass spectrum of 2-APB.

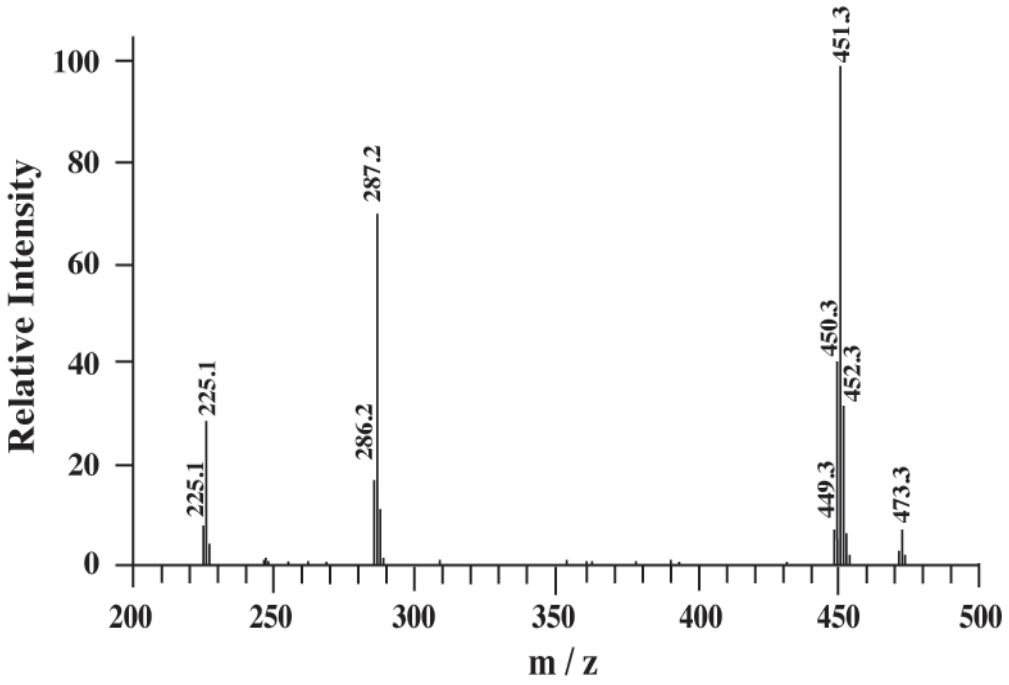


Figure 3. Electrospray (ES) mass spectrum of 2-APB.

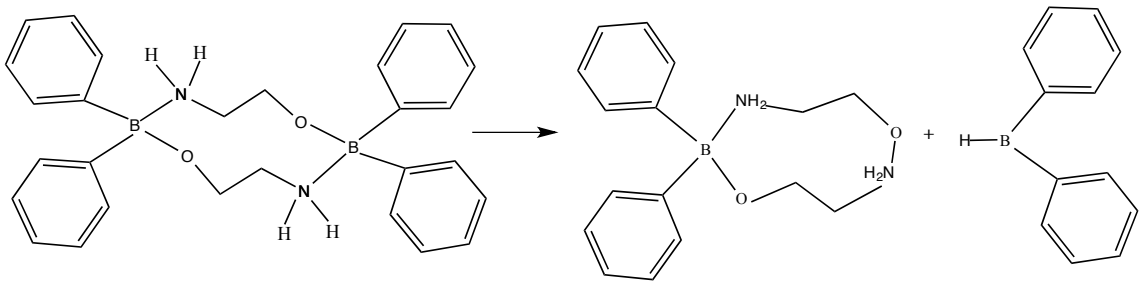


Figure 4: Fragmentation reaction that gives rise to 9-member ring with a peak at 287 amu.

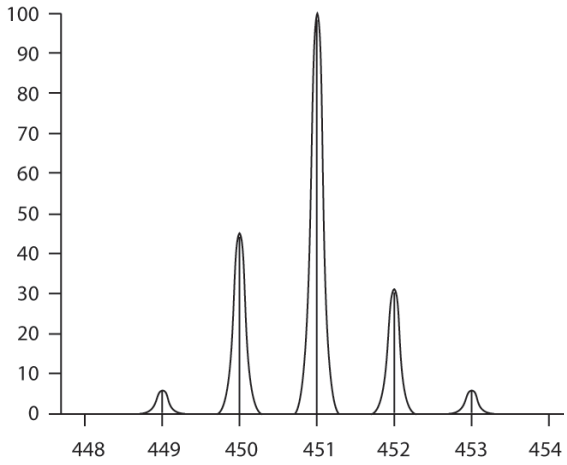


Figure 5(a)

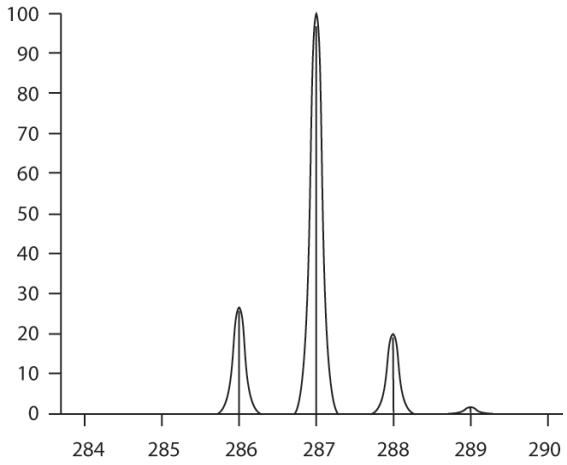


Figure 5(b)

Figure 5. Calculated isotopic peak intensities for a 2-APB dimer peak (a) and for the proposed 9-member ring fragment (b).

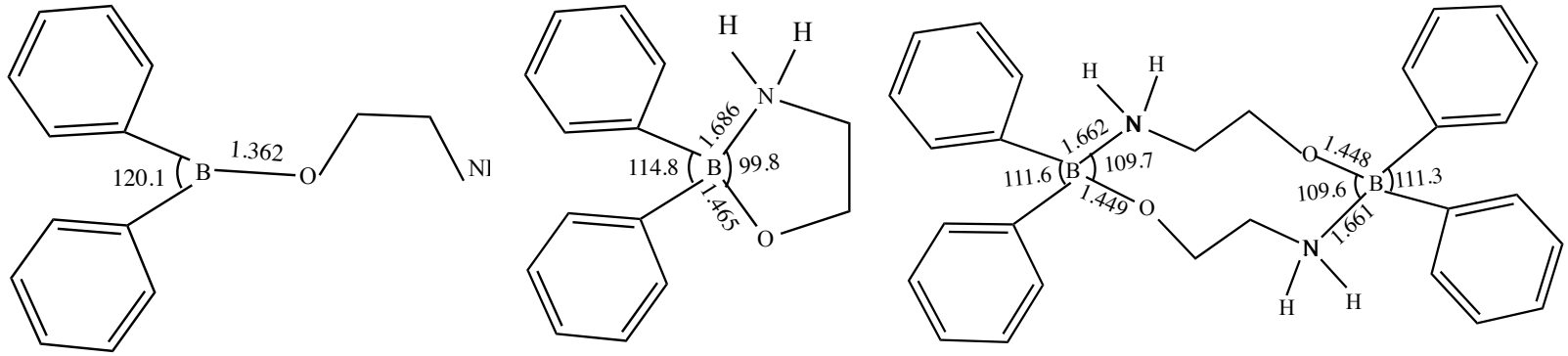


Figure 6: Structures, with key structural details, for monomers and dimer of 2-APB.

CHAPTER 4. THEORETICAL STUDY OF HYDROLIZATION OF B_2O_3

A paper to be submitted to *The Main Group Chemistry*

Chamila C. De Silva and Thomas A. Holme

Abstract

In order to quantify the conversion of anhydrous boric acid to aqueous boric acid model studies were carried out. The polymeric structure of B_2O_3 is theoretically modeled as interconnected ribbons of BO_3^{3-} repeating units (monomers). The hydrolyzation reaction barrier heights are predicted for consecutive bond breaking of three B−O bonds in the central BO_3^{3-} unit of the B_2O_3 structure at RHF/6-31G(d,p) and MP2/6-31G(d,p)//RHF/6-31G(d,p) levels. The hydrolyzation reaction barrier heights are lower when two water molecules are involved than for a single water molecule transition states. The successive barrier heights for the hydrodrolysis of B_2O_3 with two waters are predicted as 9.43, 12.31, and 17.30 kcal/mol at MP2/6-31G(d,p)//RHF/6-31G(d,p) level relative to their reactant complexes.

I. Introduction

Hydrolyzation of non-metal oxides such as carbon, nitrogen and sulfur has been studied extensively using both experimental and theoretical methods.[1-5] Among the non-metals boron is an interesting element, which has a wide variety of borate minerals that occur naturally[6] and are synthetically accessible. Despite the number of various borates, the reaction between borates and water has not theoretically studied yet and experimental studies have been limited to one.[7]

Boron is an important micronutrient element for plants and animals. The role of boron in plant reproduction,[8] cellular membrane functions[9] and most widely the role in cell-wall formation[10,11] has been studied. It is important to study the hydrolyzation of boric oxide since plants respond only to the boron activity in soil solution.[12] Boric acid is available to soils naturally during rock weathering[13] and by adding fertilizers containing boron.[14] Boron is up taken by plants roots as boric acid ($B(OH)_3$) and or borate $\left(B(OH)_4^-\right)$.[15-17] Availability of boron to plants as boric acid depends on various factors such as pH, texture, moisture and temperature of the soil. For example, high pH, coarse textured soils, dry soils and low soil temperature decrease the boron availability to plants.[12] Unlike other nutrient elements the concentration difference in soils between the scarce and toxic level of boron is very small.[18] In order to add the precise amount of boron to soil without making it too deficient or too toxic it is important to understand the reaction of boron in fertilizers (synthetic fertilizers or natural minerals) with water.

A wide variety of Borate minerals occur in geological environments.[19-22] Fertilizers used for boron deficiency are sodium tetra borate ($Na_2B_4O_7.5H_2O$), sodium tetra borate decahydrate (borax) ($Na_2B_4O_7.10H_2O$), hydrous calcium borate[14] (colemanite) $Ca_2B_6O_{11}.5H_2O$, disodium octaborate tetrahydrate[23] (solubor) $Na_2B_8O_{13}\cdot 4H_2O$ and boron humate (contains humic acid and anhydrous boric acid (B_2O_3)). Boron exists in these compounds and in many other naturally existing soil minerals in the form of triangular, $B\phi_3$ and tetrahedral, $B\phi_4$ coordination (ϕ is an unspecified ligand, either oxygen or hydroxyl).[6,24]

While there are certainly multiple possible pathways for the hydrolyzation reaction of borate, it can be argued that their overall energies are likely to be similar. Therefore it is possible to identify one example of this reaction rather than model all the different minerals in order to observe the hydrolyzation reaction between BO_3^{3-} and water. Therefore in this study a single representation model was constructed for boron containing minerals. Boron trioxide anion (BO_3^{3-}) is the repeating unit of the B_2O_3 polymer. In nature such a polymer of B_2O_3 is comprised of repeating BO_3^{3-} units in forms of infinitely long ribbons, which are connected to its neighboring ribbons by sharing oxygen atoms. This study uses a truncated model of the anhydrous boric acid, B_2O_3 that includes thirteen BO_3^{3-} structural units. This model provides a site for hydrolyzation that would be present at the surface in many soil minerals.

II. Feature of the Model and Chemistry

The proposed model (Figure 1) has an array of repeating units containing boron and oxygen atoms. Each oxygen in the anhydrous borate (model) makes two single bonds that connect to two boron atoms, except for the outer oxygen atoms of the model that are capped with hydrogen atoms to truncate the structure. Each boron atom in the model has sp^2 hybridization and has an empty p-orbital. This empty p-orbital is oriented perpendicular to the surface of planer BO_3 unit therefore has the ability to accept a pair of electrons hence these boron atoms can act as Lewis acid. A water molecule that is introduced to this system, can interact either to the oxygen in the model where the lone pair of electrons in the oxygen acts as a H-bond donor or to the boron, where boron acts as a Lewis acid and makes a dative bond with the oxygen in the water (Figure 1-b). Breaking any three consecutive B−O bonds around a boron atom in the model can contribute to a transition from anhydrous borate to solvated boric acid. This study only focuses on breaking away one BO_3 unit to produce one boric acid. The boron atom that provides the best model of a surface boron to break is central boron atom in the anhydrous borate, highlighted in Figure 1. An equal environment of boron and oxygen atoms surrounds the central boron atom. Boron containing minerals in soil also has an equal environment of boron and oxygen atoms. Therefore the central boron atom represents the best place for the hydrolyzation reaction. Technically; one or many water

molecules can be used in the hydrolyzation. In this study either 1 or 2 water molecules have been used to break a B−O bond (Figure 2-a, b).

When two water molecules are brought to the surface of the anhydrous borate, there is a possibility of forming three different reactant complexes, $BO_3(H_2O)_2$, where two water molecules are physisorbed to the anhydrous borate, $BO_3(H_2O) + (H_2O)$, where one water molecule is physisorbed to anhydrous borate and the other water molecule is infinitely apart, $B_2O_3 + (H_2O)_2$, where the two water molecules make a dimer but infinitely apart from the anhydrous borate. In the reference state of reactants both water molecules and anhydrous borate are infinitely apart, $B_2O_3 + 2H_2O$ (Figure 3 a-d). The first option, $BO_3(H_2O)_2$, is energetically more favorable than the other possible complexes and therefore forms a reaction complex for the hydrolyzation reactions studied here.

The transition states describing the breaking of the first of three B−O bonds is shown in Figure 4. In the transition state two water molecules are concertedly transferring their hydrogen-bonding protons to the proton acceptor oxygen atoms, while a new B−O bond is being formed between the newly generated hydroxyl group and the boron atom of anhydrous boron.

The product complexes of the first two B−O bond breaking result in a partially hydrolyzed borate and a water molecule, where as breaking the third bond results a boric

acid molecule and a water molecule. In all three-bond breakings all the water molecules and the boric acid molecule (in the third bond breaking) make hydrogen bonds to the oxygen atoms in the anhydrous borate (Figure 5).

In the overall reaction anhydrous borate (model) and four water molecules are combined for a three consecutive B−O bond breaking reaction (Figure 5). The products of this reaction are one molecule of boric acid, a partially hydrolyzed borate (model) and a second water molecule, which assists the proton transfer of the final transition state. This model only shows the breaking of three B−O bonds to make one boric acid but with the presence of many water molecules the same procedure would be repeated to produce many boric acids.

III. Computational Details

All calculations were performed using the GAMESS[25,26] electronic structure code, and the molecules were visualized with MacMolPlt.[27] All the Electronic structure calculations were carried out using restricted Hartree Fock (RHF) calculation with 6-31G(d,p) basis set and a single point energy calculations on the stationary states (reactants and on transition states and products) were carried with MP2/6-31G(d,p) level. Geometry optimizations were carried on B_2O_3 structure and on all reactants and products in C1 symmetry. Hessian calculations were performed on the optimized structures to confirm the true minima (all positive force constants).

Transition states of all three reactions were identified by a Hessian calculation showing one imaginary frequency followed by a saddle point calculations to optimize the transition state structures. Intrinsic Reaction Coordinate (IRC) calculations were performed to confirm that the located transition state structures connect the expected reactants and products.

All the calculations were performed in a solvent model using Polarizable Continuum Model (PCM)[29] with water as the solvent.

IV. Results and Discussion

As noted earlier, the model polymer structure of B_2O_3 is truncated into three interconnected BO_3^{3-} ribbons[30] (Figure 1). The ribbons consisted of five, five, and three BO_3^{3-} units. Hydrogen atoms bonded to oxygen cap the ends of the ribbons. In each ribbon, the BO_3^{3-} units all adopt a planar configuration (local C_3h symmetry). Boron containing minerals have both 4-coordinated and 3-coordinated boron atoms. 4-coordinated boron has a hybridization of sp^3. There is no empty p-orbitl to accept a lone pair from the oxygen atom of water molecule. Therefore 4-coordinated boron cannot act as a Lewis acid and it does not under go hydrolyzation as readily as the 3-coordinated borates. For this reason 4-coordinated boron is not included in to the anhydrous borate

model. The B−O bond distances and O−B−O bond angles were compared with the available experimental data and the results are presented in Table 1. The calculated bond lengths are normally shorter than experimental values. This should be expected since the calculated geometries are taken from Hartree-Fock calculations, where electron correlation is taken as an average and hence bond lengths are shorter than experimental values. The bond angles are in acceptable agreement with experimental values. In the modeled B_2O_3 structure the central BO_3^{3-} unit is taken as the place to undergo the hydrolization reaction (Figure 1), because it is the best location in the model.

Two different studies were carried out using either one or two quantum water molecules to find the hydrolyzation transition state (Figure 2 a, b) for the initial breaking of a B–O bond. Saddle points were obtained for the attack of he B−O bond of BO_3^{3-} unit by either one or two waters. Because a different number of water molecules are involved, the physisorbed reaction complex energies are different. Therefore, energies are compared to an initial state with waters infinitely removed from the borate model. The attack of reaction with one water molecule, which is a 4-centered transition state, has a activation barrier of 19.67 kcal/mol at MP2/6-31G(d,p)// RHF/6-31G(d,p) level with respect to the isolated reactants. The transition state for hydrolization with two water molecules is six-centered, where two protons transfer simultaneously. and in this two water molecule system it is –11.87 kcal/mol with respect to the isolated reactants. Note that, when compared to the reaction complex, this second barrier is 9.43 kcal/mole above the energy of that complex. The large energy difference between the 4-centered and 6-

centered transitions states suggests the single water molecule path is not energetically competitive. Therefore for subsequent hydrolization reactions, transition state for one water molecule attacks were not sought and the breakings of 2^{nd} and 3^{rd} B−O bonds were modeled with two water molecules. The function of the second water molecule is to assist in the proton transfer from the first water molecule, which is attacking the boron atom. In this way the net effect is to transfer a proton to one of the three oxygen atoms, which already attached to boron.

Henceforth, all the reported energies of the consecutive bond breakings are reported with respect to the energy of isolated reactants made with initial anhydrous borate structure and four water molecules. First transition state obtained at -11.87 kcal/mol. Once again, this value is above the starting point of physisorbed waters on the surface of anhydrous borate. This reactant state was determined by carrying out an optimization after determining 256 steps in the Intrinsic Reaction Coordinate (IRC) path of first transition state heading towards reactants. The reaction complex found in this way indicates that the two water molecules are physisorbed to B_2O_3 polymer. In fact we observed this formation of reactant complexes in all three-bond breakings (Figure 5). In addition to reactant complexes we determined product complexes (Figure 7) using the same IRC-based method, where the extra water molecule of hydrolyzation makes hydrogen bonds to the B_2O_3 polymer. This implies that in all three consecutive bond breakings reactants and products may not acheive an infinite separation imediately but rather produce a more stable complex with water molecules.

All the transition states are shown in Figure 6 and the geometries of each transition state is given in Table 2. The energy profiles for all three consecutive bond breakings are given in Figure 8. Considering the first B−O bond breaking as an isolated event (once an exothermic reaction complex is formed) shows an endothermic reaction (+7.46 kcal/mol) with a 9.43 kcal/mol barrier. Two possible transition state geometries can be identified for the hydrolyzation of the second B−O bond; (1) breaking the second B−O bond in the central BO_3, or (2) breaking a B−O bond to an adjacent BO_3. Both transition states yield the same product complex (Figure 7). A solvation reaction occurs in the central B−O bond is an exothermic reaction with a large energy difference (−9.51 kcal/mol) and a comparably large activation barrier (12.3 kcal/mol), where as the when reaction happens in the adjacent B−O, is an endothermic reaction with a small energy difference (+2.36 kcal/mol) and relatively small activation barrier (8.71 kcal/mol). Since the overall reaction includes energy liberated from the association of water molecules with the surface, both thermodynamic and kinetic factors must be considered. As a result, both of these reactions are energetically favorable. The third B−O bond breaking involves the highest activation energy (17.3 kcal/mol) where as the first and the second B−O bond breaking energies are on average 7.15 kcal/mol less than the third. However, when the overall reaction, from reactant complex of first bond breaking to the product complex of the third bond breaking is considered the result is an exothermic reaction (−22.14 kcal/mol).

For each sequential bond breaking, the reactant and product complexes were determined by tracing the IRC path from the transition state. With many possible ways for water molecules to associate with the surface, these states are not going to be identical, so the sequential modeling essentially builds in a step where the system changes between the product state of one bond-breaking to the reactant state of the next. This transferring from a product complex to a reactant complex is always an energetically favorable process (Figure 8). In the vicinity of many water molecules rearranging of product complex to give the reactants complex can be done though many different paths. Therefore, identifying a single saddle point between the product and reactant complexes would not be very helpful in establishing the activation energy so no saddle points were calculated. Since the availability of abundant free water molecules in the solution should not contribute to a high-energy saddle point, the reaction from product complex to reactant complex should happen in a low energy profile. In Figure 8, energies from a product complex to a reactant complex of the next bond to break are connected using dashed lines to infer that there can be additional energy steps in between the given states. Also for the same reason reaction profiles from isolated reactants/products to reactant/product complexes are joined using dashed lines.

The final isolated products are much higher in energy than its product complex, but this is only in the perspective of bond energies. When the final products goes from complex to isolated products the attractive physisorbtion interactions must be overcome so the internal energy is greater. In an experimentally observed system there would

always be additional water molecules to solvate the products. Thus, the motivation for calculating isolated products is only to obtain an overall comparison for the three bond-breaking reactions to the initial, isolated reactants. The overall reaction is largely exothermic and the breaking of B−O bonds results an increase in the number of degree of freedom, the entropy of the overall reaction increases. Therefore thermodynamically the overall reaction is favorable. The rate-determining step can be predicted to be the third B−O bond breaking which has a high barrier. In our model system we considered only two explicit water molecules with continuum solvation. However in the soil solution there can be many water molecules involve in the hydrolyzation reaction and any of the barrier heights could possibly be lowered.

It is important to know that we used the terms breaking 1^{st}, 2^{nd} and 3^{rd} are used in this study is with respect to the anhydrous borate model we used. But when concerning all the possible borate fertilizers and minerals the 2^{nd} bond breaking we used in this study can be the first bond breaking in the relevant environment (for example: in Borax breaking the first bond is similar to breaking the 2^{nd} bond in this study).

V. Conclusions

We have modeled the solvation of anhydrous borate and the energy profile is in agreement with the observation of dissolving borate fertilizers and minerals on water. The theoretical study of the hydrolyzation reaction in B_2O_3 polymer have been presented,

using MP2/6-311G(d,p)//RHF/6-311G(d,p). The transition states of all three B−O bond breaking of the hydrolyzation reactions were found in B_2O_3:$2H_2O$ with PCM water. The optimized geometries of the corresponding reactants and products and their associated complexes are also presented. The overall reaction is thermodynamically favorable. The rate determining step, the third B−O bond breaking, has high barrier. Therefore the reaction is kinetically less favorable. However when consider a soil solution with many water molecules around the hydrolyzation reaction, the barrier height can be possibly lowered.

References

(1) Nguyen M. T., H. T. *J. Am. Chem. Soc.* **1983**, *106*, 599.

(2) Wojcik M. J., B. M., Ford T. A. *Chem. Phys. Lett.* **2001**, *348*, 126.

(3) Lemke, K. H.; Seward, T. M. *J. Geophys. Res.-Atmos.* **2008**, *113*, 2156.

(4) Lovejoy, E. R.; Hanson, D. R.; Huey, L. G. *J. Phys. Chem.* **1996**, *100*, 19911.

(5) Meijer E. J., S. M. *J. Phys. Chem. A* **1998**, *102*, 2893.

(6) Schindler, M., Hawthorne, F. C. *Can. Mineral.* **2001**, *39*, 1243.

(7) Kim Y. Y., L. J. T., Park S. H. *J. Phys. Chem. Solids.* **1999**, *60*, 969.

(8) Dell, B.; Huang, L.; Bell, R. W. *Boron in plant reproduction*, 2002.

(9) Goldbach, H. E.; Wimmer, M. A. *J. Soil Sci. Plant Nutr.* **2007**, *170*, 39.

(10) Matoh, T.; Kobayashi, M. *Boron function in plant cell walls - Research progress since **1997***, 2002.

(11) O'Neill M. A., Albersheim P., Darvill A. G. *Annu. Rev. Plant Biol.* **2004**, *55*, 109.

(12) Goldberg, S. *Plant and Soil* **1997**, *193*, 35.

(13) Nable R. O. , B. u. G. S., Paull J. G. *Plant and Soil* **1997**, *193*, 181.

(14) Saleem M., K. Y. M., Fauziah Ishak Y. M. , Samsuri A. W. *Commun. Soil Sci. Plant Anal.* **2011**, *42*, 293.

(15) Jiao, X. Y., Zhu, Y. G., Jarvis, B. C., Quick, W. P., Christie, P. *J. Plant Nutr.* **2005**, *28*, 351.

(16) Hu H. N., B. P. H. *Plant and Soil* **1997**, *193*, 49.

(17) Power P. P., W. W. G. *Plant and Soil* **1997**, *193*, 1.

(18) Reisenauer, H. M. W., L. M.; Hoeft, R. G. *Soil Test. Plant Anal.* **1973**, *173*, 173.

(19) Grew, E. S. *Rev. Mineral.* **1996**, *33*, 387.

(20) London, D. M., G.B., VI; Wolf, M. B. **1996**, *33*, 299.

(21) Dingwell, D. B., Pichavant, M., Holtz, F. **1996**, *33*, 331.

(22) Smith, G. I., Medrano, M.D. **1996**, *33*, 263.

(23) Gowariker S., Dhanorkar M., Paranjape K., Borlaug N. 2009.

(24) Grice J. D., B. P. C., Hawthorone F. C. *Can. Mineral.* **1999**, *37*, 731.

(25) Schmidt, M. W. B., K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(26) Gordon, M. S.; Schmidt, M. *Theory and Applications of Computational Chemistry: The First Forty Years*; Elsevier: Amsterdam, The Netherlands, 2005.

(27) Bode, B. M.; Gordon, M. S. *J. Mol. Graph. Modell.* **1998**, *16*, 1347.

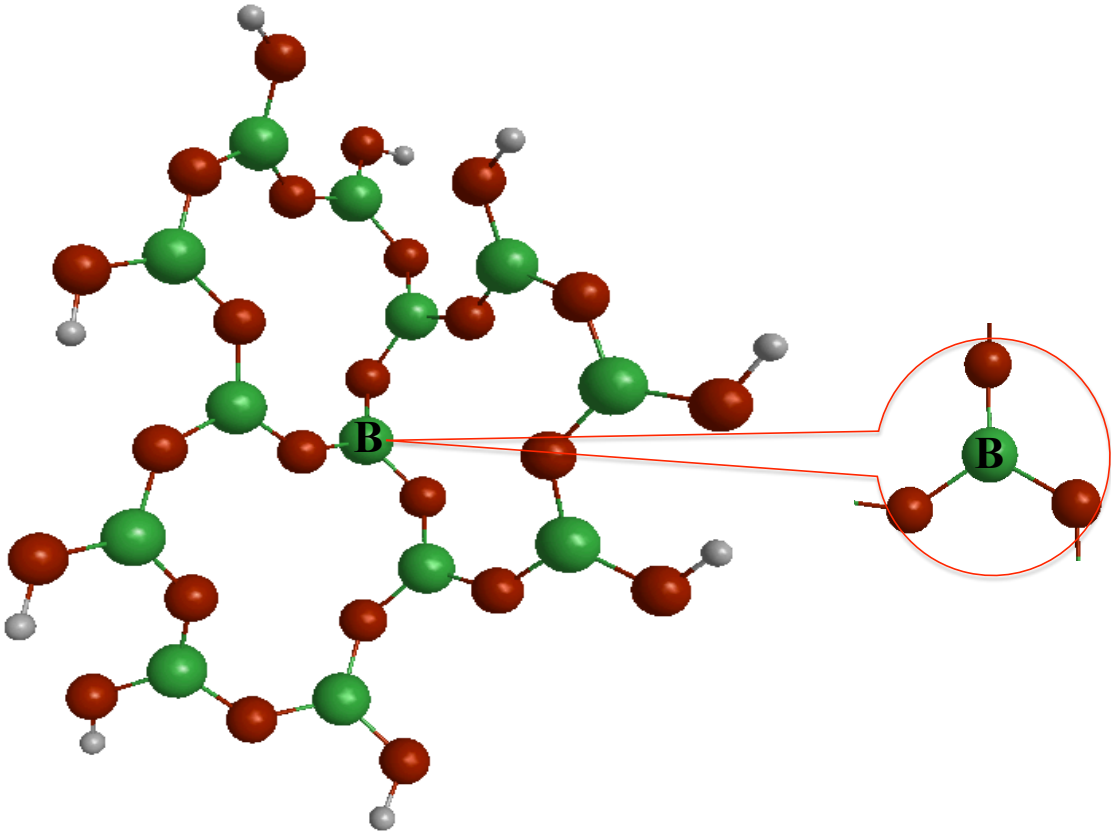
(28) S., M. R. *J. Chem. Phys.* **1955**, *23*, 1833.

(29) J. Tomasi, B. M., R.Cammi *Chem. Rev.* **2005**, *105*, 2999.

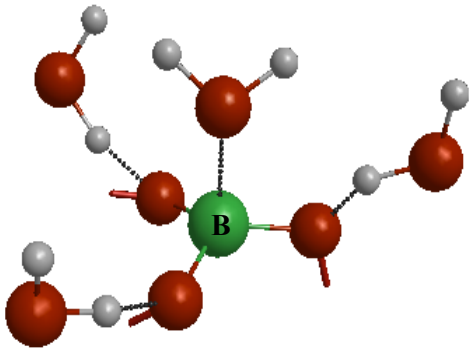
(30) Gurr, G. E.; Montgome.Pw; Knutson, C. D.; Gorres, B. T. *Acta Crystallographica Section B* 1970, *B 26*, 906.

(32) P. V. Ramachandran, P. D. Gagare Inorg. Chem. **2007**, 46, 7810

(33) N. Morimoto Mineralogical Journal, **1956**, 2, 1

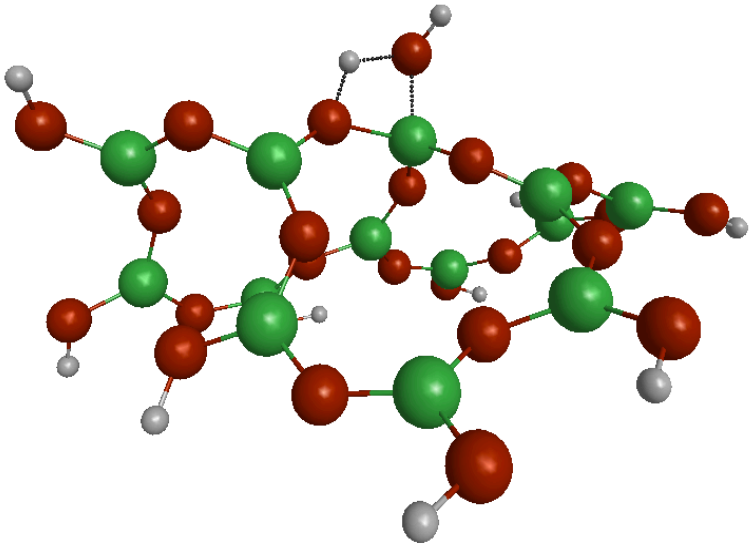


(a) Truncated polymer structure of B_2O_3

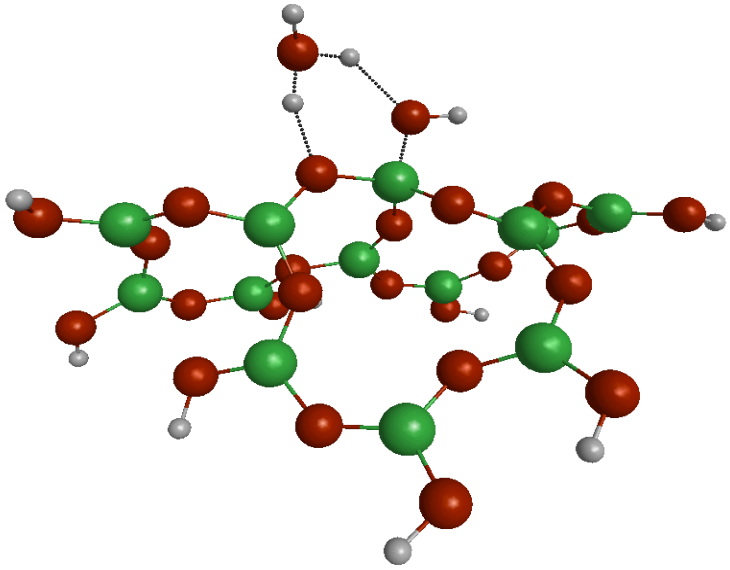


(b) Anhydrous borate with water molecules

Figure 1. (a) Truncated polymer structure of B2O3 with the magnified central boron where the hydrolyzation reaction happens. Green spheres represent boron atoms and the red spheres represents oxygen atoms. (b) All possible interactions of the anhydrous borate with water molecules, for simplicity, only the central BO3 unit is shown.

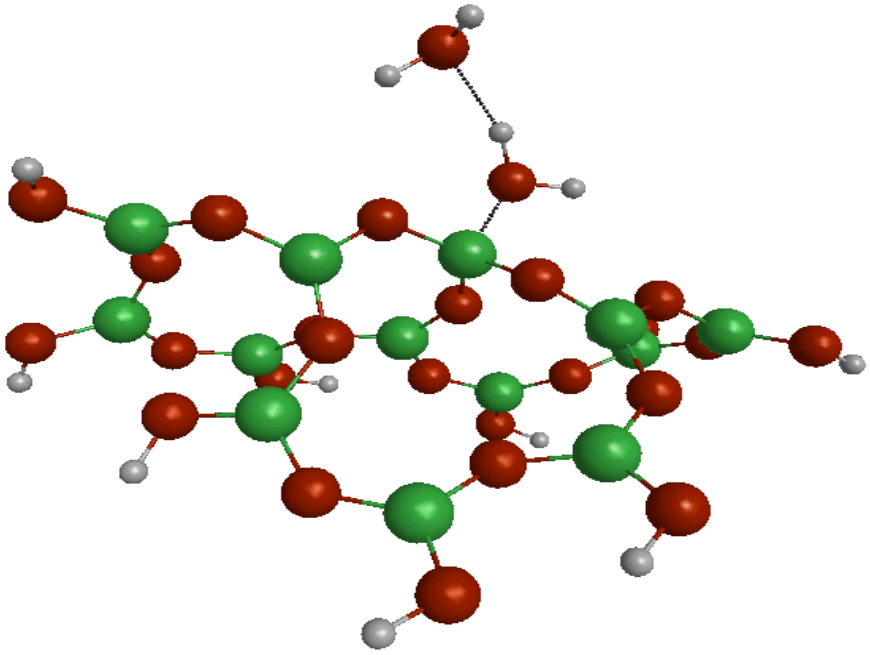


(a) Hydrolyzation of anhydrous borate with one water molecule

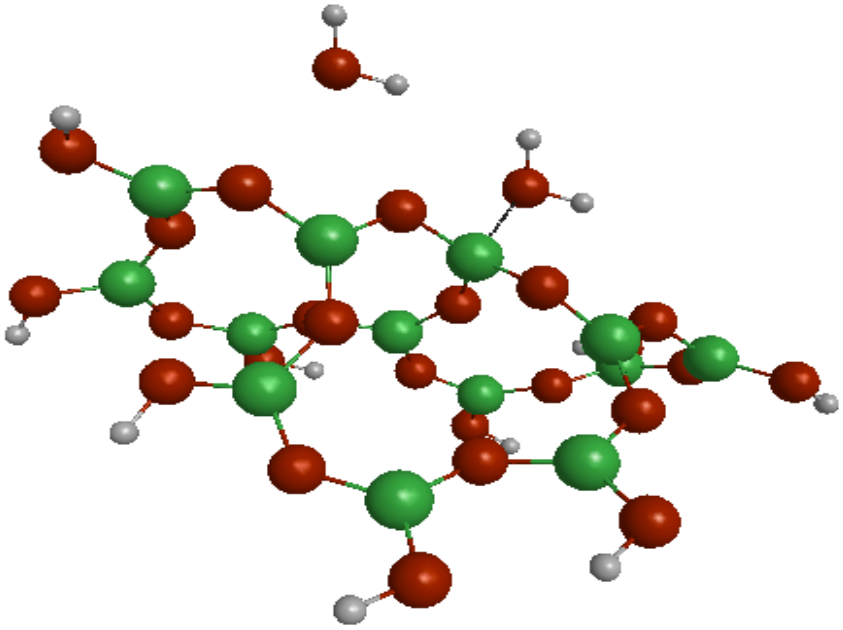


(b) Hydrolyzation of anhydrous borate with two water molecules

Figure 2. (a) Hydrolyzation of anhydrous borate with one water molecule. (b) Hydrolyzation of anhydrous borate with two water molecules

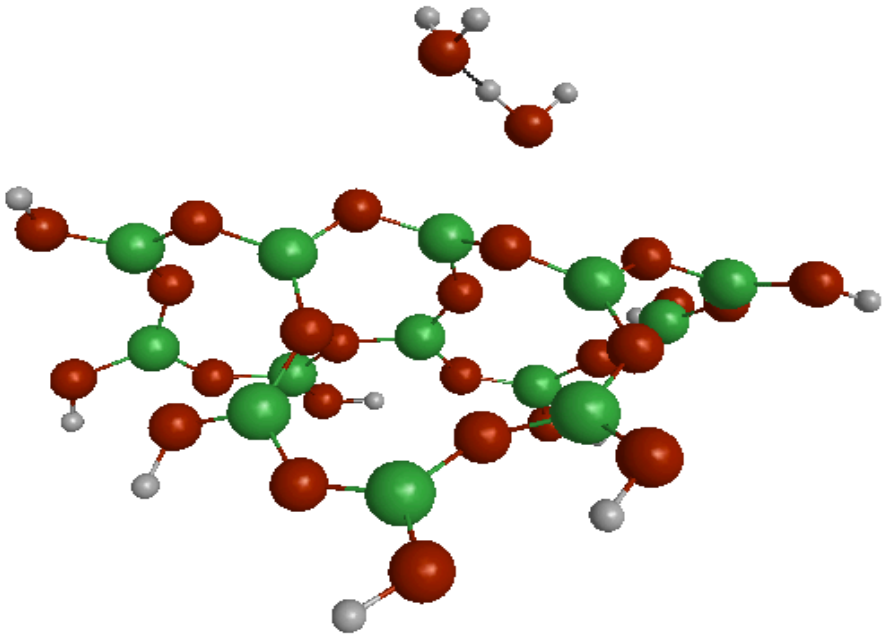


(a) $BO_3(H_2O)_2$

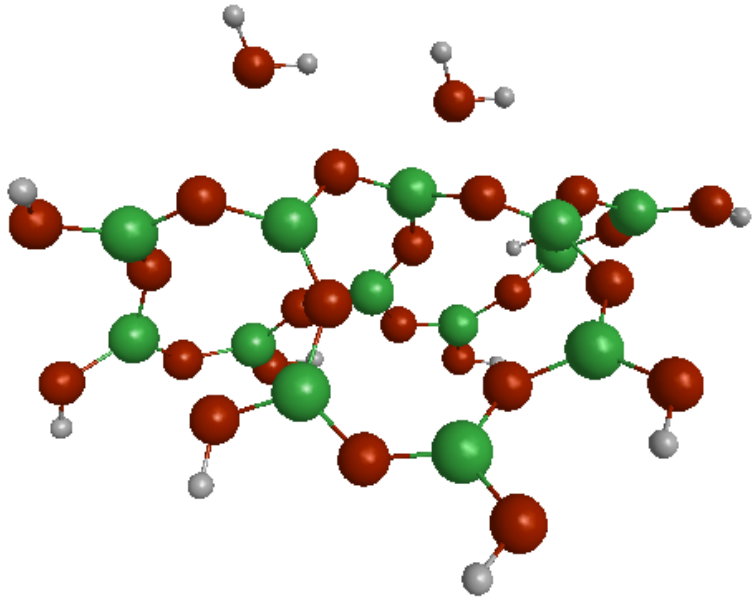


(b) $BO_3(H_2O)$

Figure 3. Possible reactant complexes, (a) $BO_3(H_2O)_2$ (b) $BO_3(H_2O) + (H_2O)$ (c) $B_2O_3 + (H_2O)_2$ (d) $B_2O_3 + 2H_2O$



(c) $B_2O_3 + (H_2O)_2$



(d) $B_2O_3 + 2H_2O$

Figure 3. (continued)

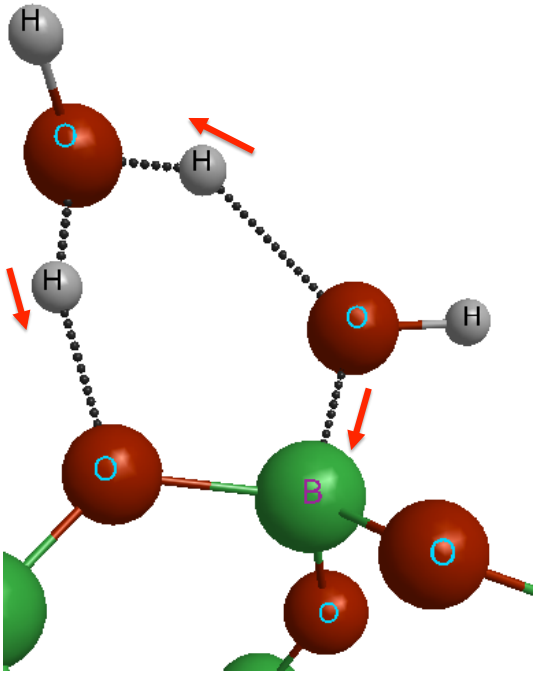
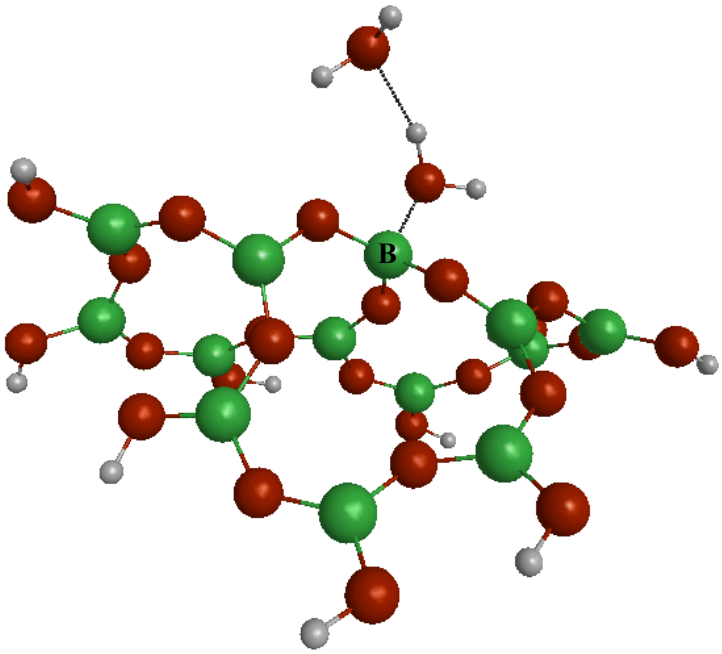
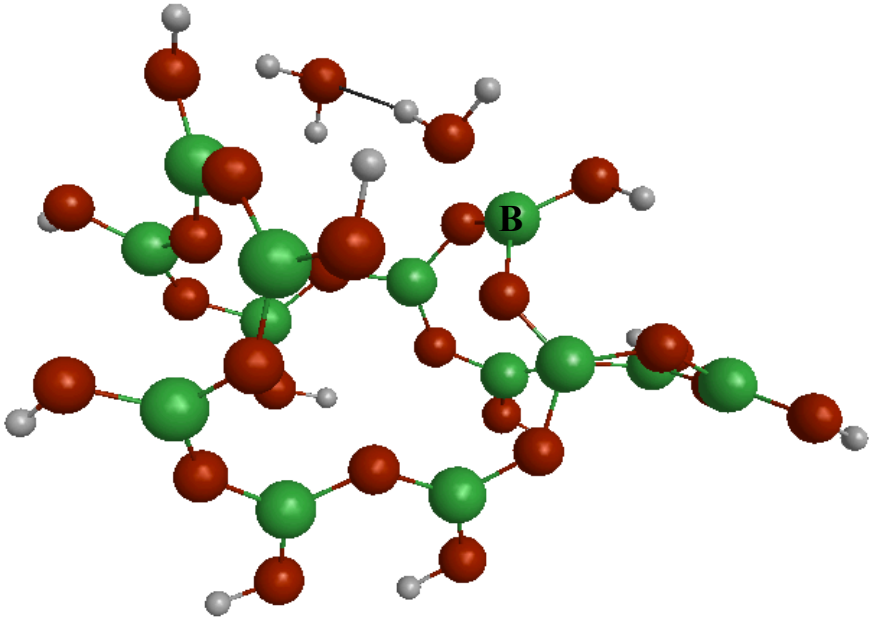


Figure 4. Transition state two water molecules are concertedly transferring their hydrogen-bonding protons to the proton acceptor oxygen atoms. Reaction coordinate vectors are shown in red arrows.

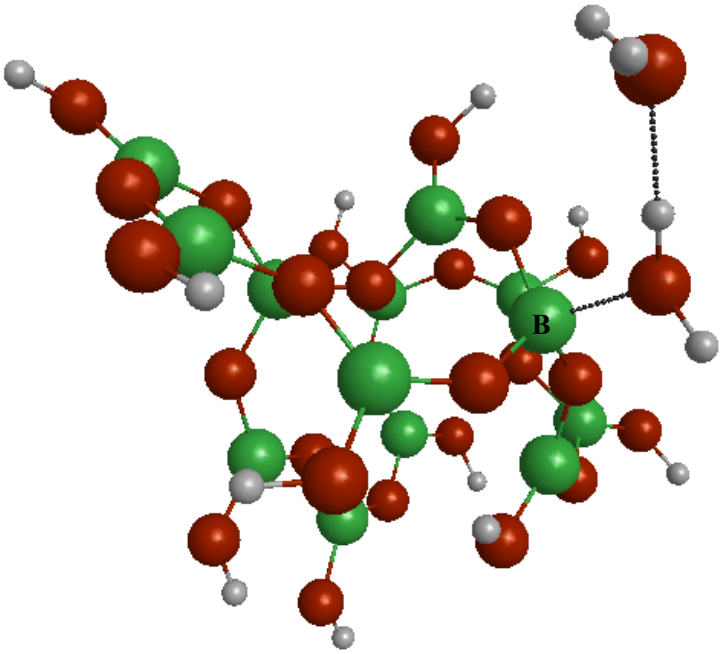


(a) 1st bond breaking

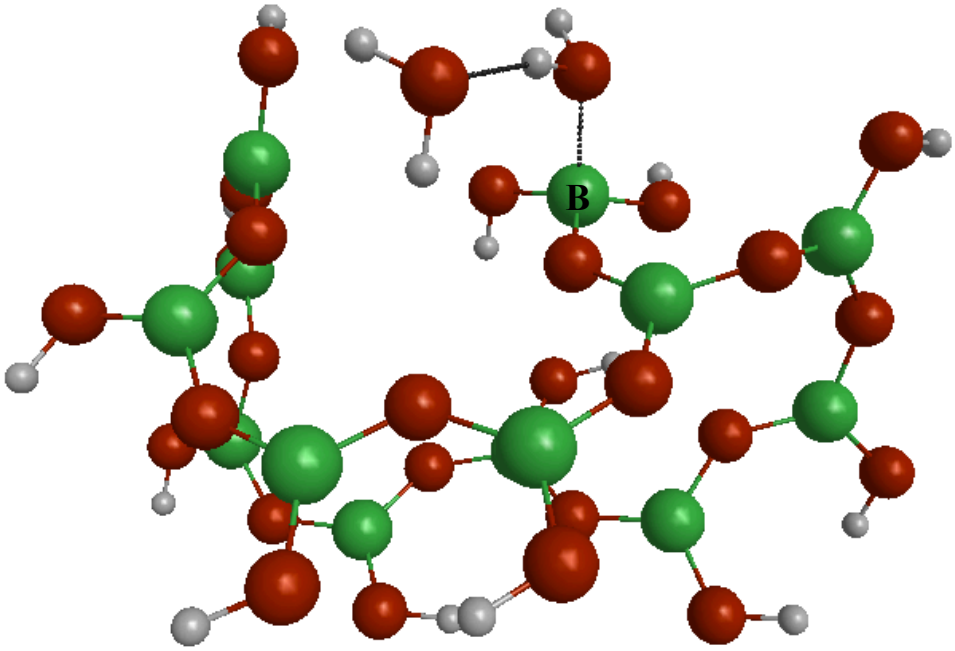


(b-1) 2nd bond breaking at central boron atom

Figure 5. Reactions complexes for (a) 1st bond breaking, (b-1) and (b-2) for second bond breaking at central and adjacent boron atoms and (c) 3rd bond breaking.

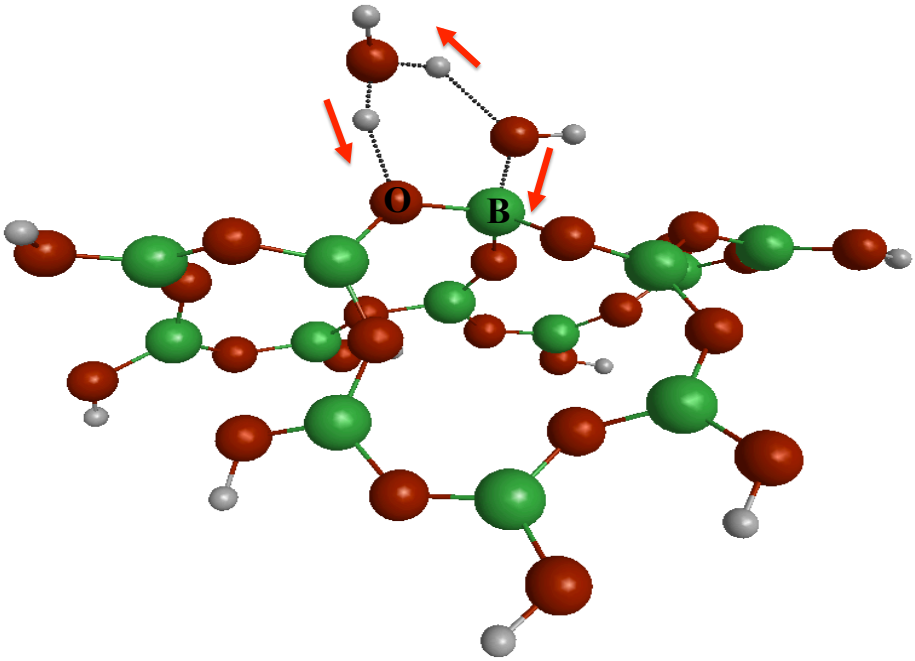


(b-2) 2nd bond breaking at adjacent boron atoms

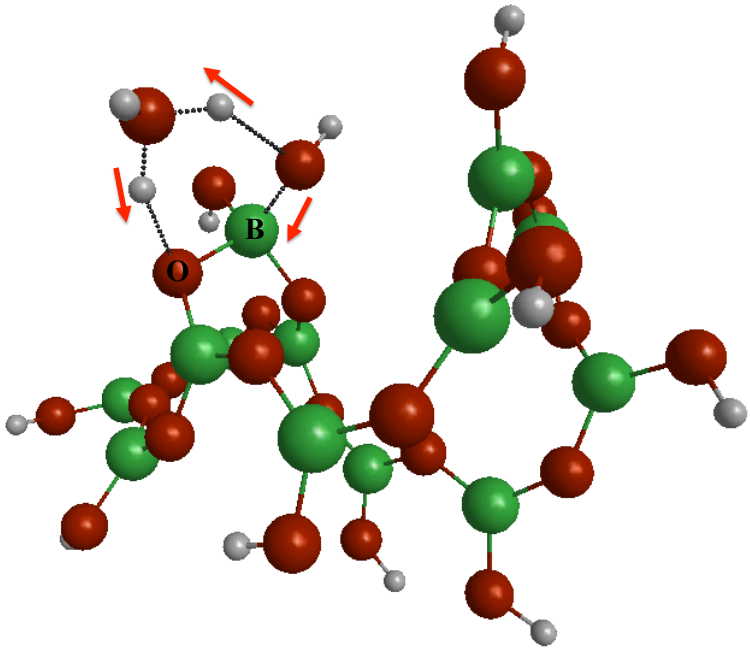


(c) 3rd bond breaking

Figure 5. (continued)

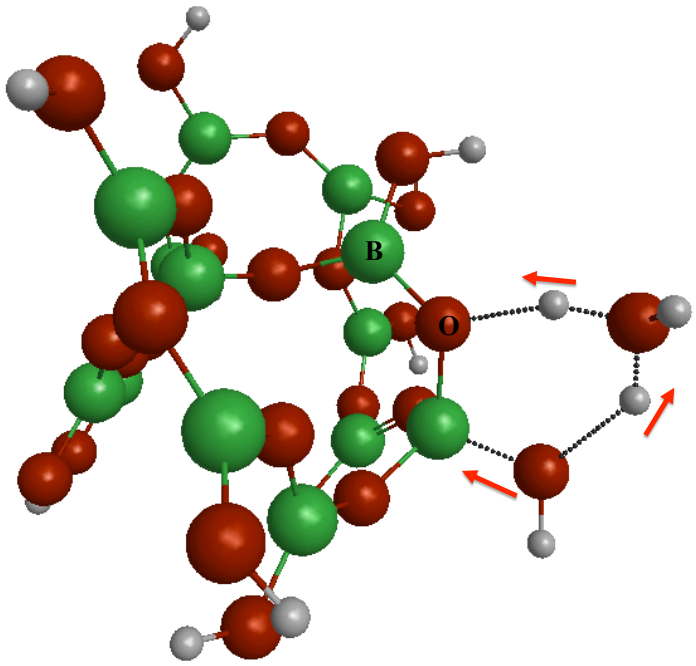


(a) 1st bond breaking

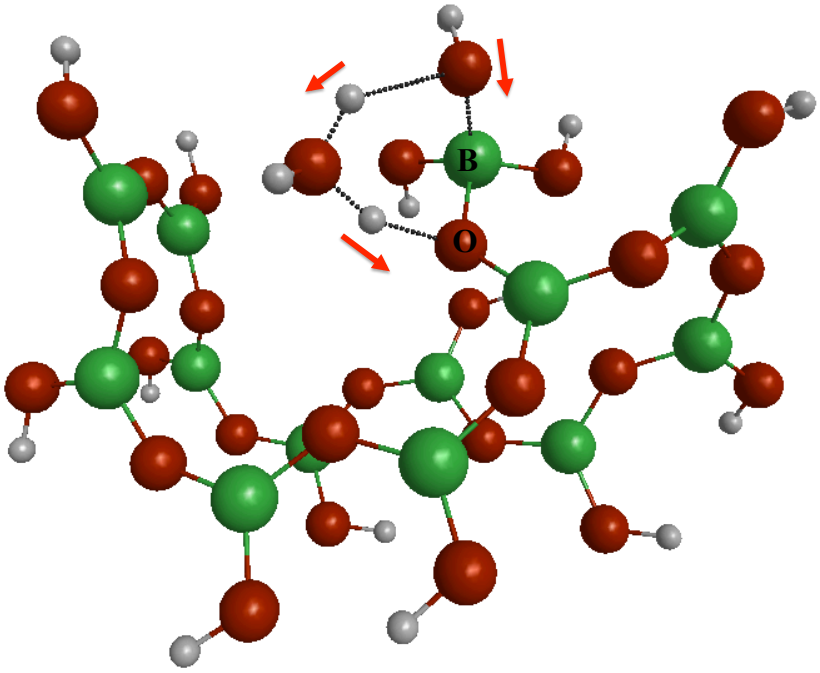


(b-1) 2nd bond breaking at central boron atom

Figure 6. Transition States for (a) 1st bond breaking, (b-1) and (b-2) for second bond breaking at central and adjacent boron atoms and (c) 3rd bond breaking. Boron and Oxygen atoms involve in B−O bond breaking is marked on the figures and the arrows shows the movement of proton transfer.

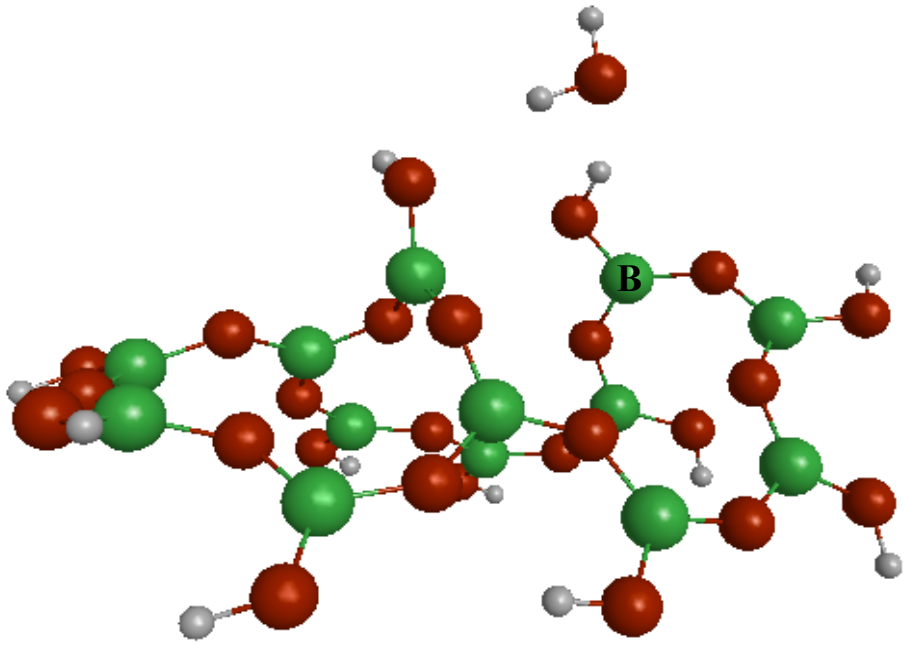


(b-2) 2nd bond breaking at adjacent boron atoms

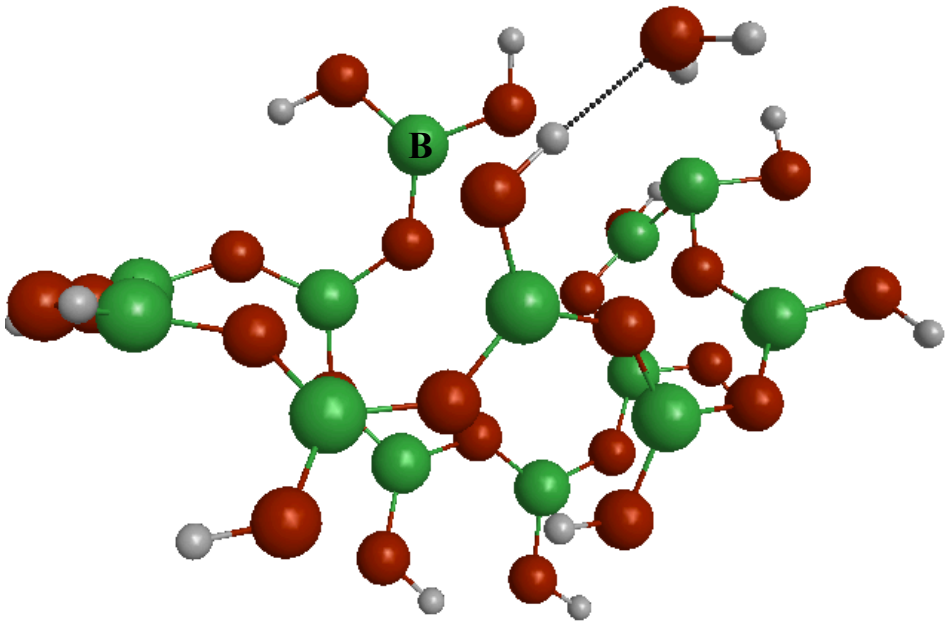


(c) 3rd bond breaking

Figure 6. (continued)

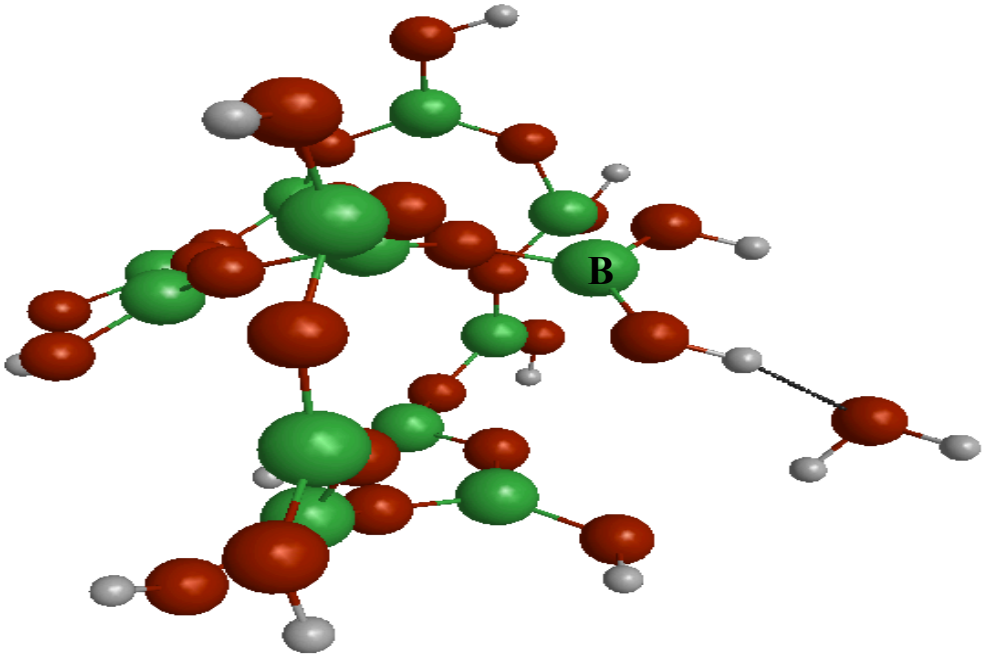


(a) 1st bond breaking

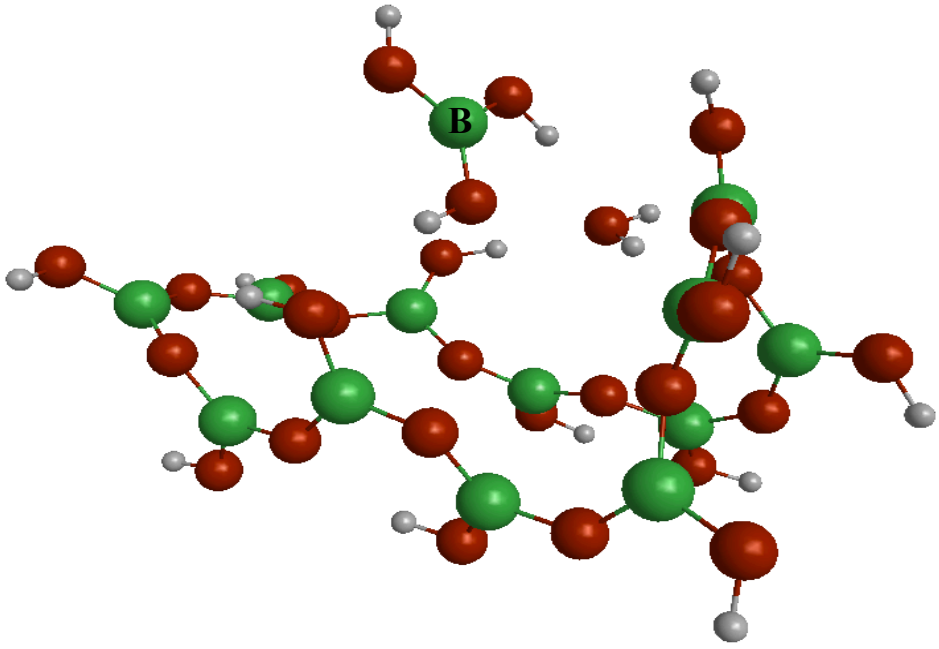


(b-1) 2nd bond breaking at central boron atom

Figure 7. Product Complexes for (a) 1st bond breaking, (b-1) and (b-2) for second bond breaking at central and adjacent boron atoms and (c) 3rd bond breaking.



(b-2) 2nd bond breaking at adjacent boron atoms



(c) 3rd bond breaking

Figure 7. (continued)

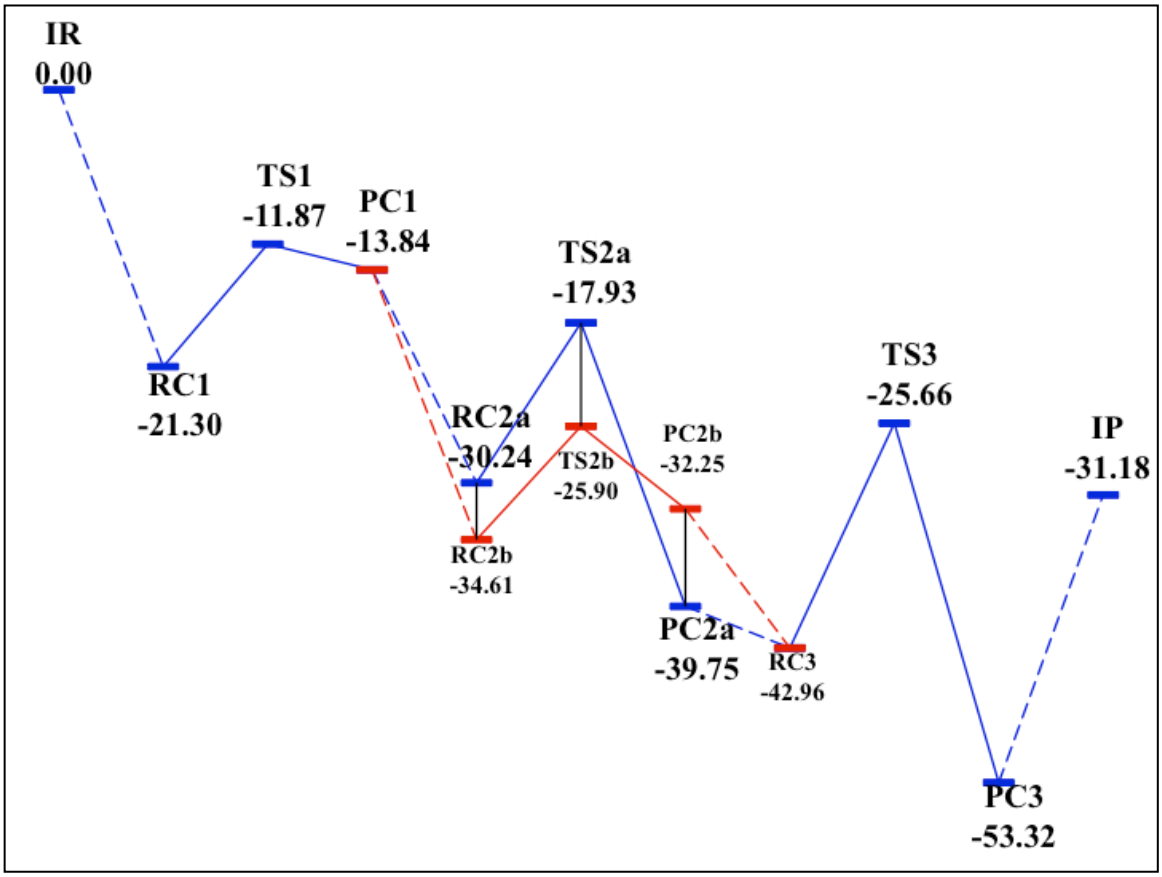
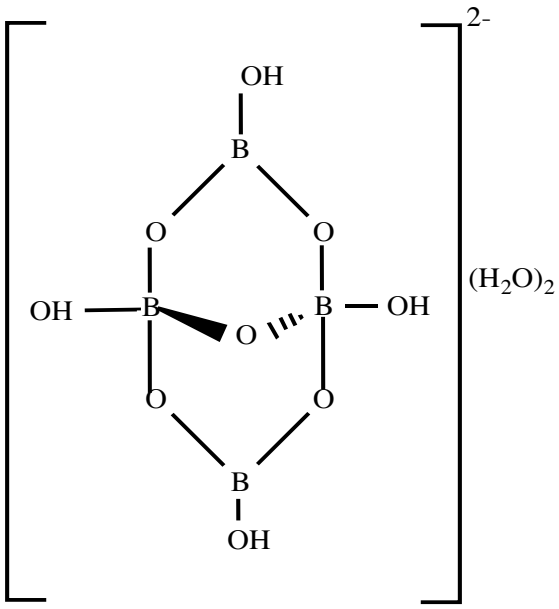
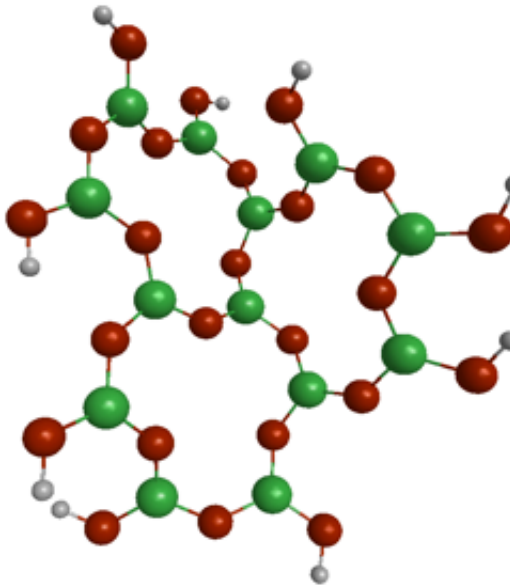


Figure 8. Energy Profile For three consecutive B−O bond breaking. All the energies are given with respect to the isolated reactants, IR (B_2O_3 polymer + $4H_2O$) in units of kcal/mol. The labels given in the graph are described as follows with the associated Figure.

IR − Isolated reactants (truncated polymer structure B2O3 (Figure 1) and four water molecules
RC1 − Reactant Complex for first bond breaking – Figure 5-(a)
TS1 − Transition State for first bond breaking – Figure 6-(a)
PC1 − Product Complex for first bond breaking – Figure 7-(a)
RC2a − Reactant Complex for second bond breaking at central boron – Figure 5-(b-1)
TS2a − Transition State for second bond breaking at central boron – Figure 6-(b-1)
PC2a − Product Complex for second bond breaking at central boron – Figure 7-(b-1)
RC2b − Reactant Complex for second bond breaking at adjacent boron – Figure 5-(b-2)
TS2b − Transition State for second bond breaking at adjacent boron – Figure 6-(b-2)
PC2b − Product Complex for second bond breaking at adjacent boron – Figure 7-(b-2)
RC3 − Reactant Complex for third bond breaking – Figure 5-(c)
TS3 − Transition State for third bond breaking – Figure 6-(c)
PC3 − Product Complex for third bond breaking – Figure 7-(c)
IP − Isolated products (Boric acid, a water molecule and resulting B2O3 polymer)



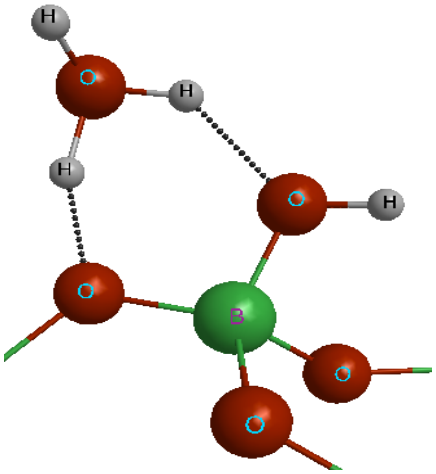


(a) **(a)**

Table 1. Experimented and calculated bond lengths and angles for 3-coordinated boron in boron containing minerals with the model compound. Figure (a) shows the structure of Borax, an example for an experimental structure and figure (b) is the computed structure.

	Average Experimental	Calculated
B−O Bond Distance (Å)	1.37*	1.36
O−B−O Bond Angle (degrees)	117-122 *	117-122

*The avarage value from References 31-33.



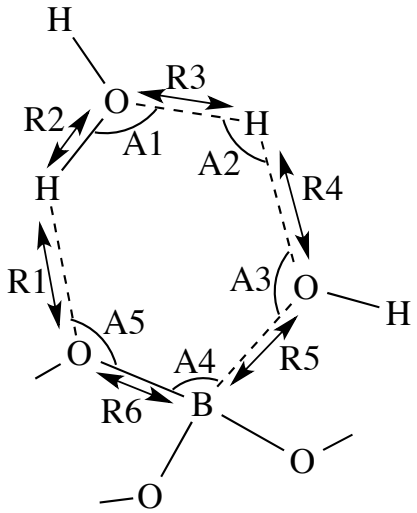


Table 2. Bond distances and angles at the transition states. Bond lengths (Angstrom) and bond angles (degree).

B–O Bond	R1	R2	R3	R4	R5	R6	A1	A2	A3	A4	A5
First	1.43	1.04	0.98	1.67	1.46	1.51	97.0	137.2	114.7	103.1	110.3
Second (a)	1.48	1.02	0.99	1.69	1.46	1.56	98.9	137.6	107.4	104.5	101.9
Second (b)	1.48	1.02	0.98	1.68	1.46	1.51	98.3	135.6	115.1	104.1	109.2
Third	1.38	1.06	0.97	1.84	1.46	1.55	101.0	132.1	101.0	104.2	99.8

Second (a) and Second (b) are the 2nd bond breaking at the central B and at the adjacent B respectively.

CHAPTER 5. PARAMETERIZATION OF BORO-DIESTER CARBOHYDRATES IN THE CHARMM ALL-ATOM EMPIRICAL FORCE FIELD

Manuscript in Preparation

Chamila C. De Silva and Thomas A. Holme

Abstract

An extension of the bio-molecular CHARMM all-atom empirical force field parameters is described for modeling boron complexes of carbohydrates in which the boron is bound to the carbohydrate through boro-diester linkage. The model is developed to be consistent with the CHARMM all-atom carbohydrates force field, and the existing parameters for pyranose and furanose sugars were transferred from carbohydrate force fields to develop new boro-carbohydrate parameters. The additional parameterization is based on MP2/6-31G(d,p) geometries, solute-water interaction energies and torsional potentials. The optimized geometries are reported for a set of galactose, fucose, mannose, glucose, xylose, apoise and fructofuranose boron complexes. The model satisfactorily reproduces the structures of thirteen boro-carbohydrate complexes within 0.03 Å accuracy for bond lengths and 3 degrees accuracy for bond angles. The torsional barriers are well reproduced, within 0.6 kcal/mol.

I. Introduction

Rhamnogalacturonan II (RG-II) is a predominant pectic polysaccharide present in the primary cell wall of higher plants. RG-II along with the other pectic polysaccharides, homogalacturonan (HG) and rhamnogalacturonan I (RG-I) contributes to the growth, mechanical strength and physical properties of plant cell wall.[1-3] RG-II has been isolated from the cell walls of variety of plant types including monocots, dicots and gymnosperm. Common feature that all these plants cell walls share is that the structure of RG-II in all these plant types are same.[4-6]

RG-II contains homogalacturonan backbone composed of at least eight 1→4 linked α-d-galacturonic acid residues. Four different complex oligo glycosyl side-chains (A, B, C and D) that contain twelve different glycosyl residues[7] are attached to this backbone.[8] The three polysaccharides, HG, RG-I and RG-II are covalently linked to one another to form a pectic macromolecule. Furthermore covalent and non-covalent cross-linking of some glycosyl residues in these macromolecules form a three-dimensional pectic network, which coexists in the primary wall with a network, composed of non-covalent cellulose micro-fibrils.[9-10] The mechanical strength of the primary cell wall is given by the interactions within and between these networks. Therefore when the cell wall needs to expand, these interactions need to be modified allowing the necessary spacing and flexibility for cell growth.

RG-II exists predominantly as a dimer that is covalently cross-linked by a borate diester.[11-13] These borate di-esters are formed by the binding of boron in between two α-D-apoise monomers (3-C-hydroxymethyl-D-erythrose) in side chain A of RG-II. Boron cross-linking in plant cell wall plays an important role in the reinforcing strong cell walls. Therefore boron becomes an essential micronutrient for plants. It has been shown that boron deficient plants are brittle and show lack of growth.[7,14-16] Divalent metal cations such as Mg^{+2}, Ca^{+2}, Sr^{+2}, Ba^{+2} and Pb^{+2} have also been reported along with the presence of boron cross linked RG-II dimer.[13-17] It has been found that the presence of the divalent metal cations stabilize the boron cross-linking in plant cell walls, however the reason for this reaction is not strongly proved yet.[18]

Structure of RG-II and its functions has been studied for more than four decades; still the structure and the functions of RG-II are not completely understood. The glycosyl sequence of RG-II appears to be an evolutionary conserved structure, staying essentially the same in all higher plants.[19] Therefore this pectic polysaccharide structure must have a fundamental structural role in the plant cell wall.[19-20] Understanding the structure and its possible functions other than regulating the cell wall growth is still incomplete. The majority of the experimental work suggests that boron makes a single di-ester in RG-II dimer, where it cross-links side chain A in both RG-II polysaccharides. Meanwhile other experimental groups have reported of two boron di-esters in RG-II dimer, where boron should cross linked to side chain A and B respectively.[21] Therefore it is important to

understand the complete description of possible borate cross-linking of RG-II with all the side chains and the backbone.

The structure of the RG-II dimer has been analyzed in many experimental studies but computational studies on that are limited to one.[20] The interest of the study reported here is to model sugars present in the borated cross-linked RG-II structure to build basics for modeling studies of the chemistry of boro-diester formation in plants and other biological systems.

Despite rapid development in computational modeling and simulations of large biological molecules still confronts a high computational cost for accurate results. The most efficient and commonly used technique for reliable computational results is the use of techniques based on empirical force fields. Empirical force fields consist of a potential energy function U(R) with number of adjustable parameters. The use of simplified models to calculate the potential energy of a system, as a function of its three dimensional structure allows for computational efficiency sufficient to study large molecular systems. To be broadly useful, force fields with their parameter set should describe an entire class of the molecules to be modeled.

Boron cross-linking in the RG-II macromolecule is essential for the growth and strength of the plant cell. Therefore studying of boron ester bonding sites and their structural energies are important in the understanding of borate cross-linking. The use of

molecular mechanics methods is likely the best approach for this macromolecule. But the only obstacle here is that the CHARMM force fields currently do not have parameters for boro-diester carbohydrates. Therefore the first step of modeling boro-diesters should be getting parameters for boro-diester carbohydrates. The aim of this work is to introduce boron as an atom type to CHARMM force fields. Boron exists as three coordinate and four coordinate compounds, Here the CHARMM 38 force field for carbohydrates is extended to do calculate three-coordinate boro-carbohydrates. Both furanose and pyranose borated carbohydrate compounds were studied. All bonded and non-bonded parameters for carbohydrates were initially transferred from the hexopyranose[22] and furanose[23] force fields.

II. Computational Details

In general, force field parameters are optimized using the accurate optimization of the intermolecular (non-bonded) parameters, reproduction of experimental target geometries, conformational properties and vibrational spectra. But experimental data are not available for boro-diester carbohydrates; therefore a well-known approach of replacing experimental data with QM data was used. QM calculations of sufficiently high level of theory with or with out empirical scaling factors has proven to give values that are comparable to experimental data[24] for the purpose of parameterization. Here empirical force fields calculations were performed using the program CHARMM.[25] Quantum mechanical calculations were performed using GAMESS.[26-27] Boro-diester

carbohydrates used in this parameter optimization are divided in to two groups, training set and test set. The training set is implemented to build up a model by using the values in the training set to optimize the force field parameters. The test set is used to validate the model built using training set. If the parameters obtained in the training set were unable to validate the test set then the training set is iteratively re-optimized until its parameters validate both test and training sets.

The potential energy function used in the program CHARMM is shown in eq. (1)

$$U(R) = \sum_{bonds} K_b\left(b - b_0\right)^2 + \sum_{angles} K_\theta\left(\theta - \theta_0\right)^2 + \sum_{dihedrals} K_\chi\left(1 + \cos\left(n\chi - \delta\right)\right) + \sum_{impropers} K_\varphi\left(\varphi - \varphi_0\right)^2 + \sum_{nonbonded} \varepsilon_{ij}\left\{\left[\left(\frac{R\min_{ij}}{r_{ij}}\right) + \left(\frac{R\min_{ij}}{r_{ij}}\right)^6\right] + \left(\frac{q_i q_j}{4\pi e r_{ij}}\right)\right\} \quad (1)$$

Equation (1) includes the bond length, b the valence angle, θ, the dihedral or torsion angle, χ, the improper angle, φ, and the distance between atoms i and j, r_{ij}. Parameters being optimized in the present work include, the bond force constant and equilibrium distance, K_b and b_0, respectively, the valence angle force constant and equilibrium angle, K_θ and θ, respectively, the dihedral force constant, multiplicity and phase angle, K_χ, n and δ, respectively and the improper force constant and equilibrium improper angle, K_φ and φ_0 respectively. Above parameters are called the internal parameters. The non-bonded or interaction parameters between atoms i and j were also optimized, including

partial atomic charges, q_i, and Lennard-Jones (LJ) well depth, ε_{ij}, and minimum interaction radius, $R_{min_{ij}}$, for the interacting atoms via combining rules. In CHARMM, ε_{ij} values are obtained via the geometric mean, $\varepsilon_{ij} = \text{sqrt}(\varepsilon_i * \varepsilon_j)$, where ε_i is the atomic softness of the i^{th} atom, and $R_{min_{ij}}$ via the arithmetic mean, $R_{min_{ij}}=(R_{min_i}+R_{min_j})/2$, where R_{min_i} is the van der Waals radii of the i^{th} atom. In CHARMM parameterization procedures, it is customary to do all the *ab initio* structure calculations in gas phase.[28] Therefore in all MM calculations, the dielectric constant, e, is set to one in all calculations, corresponding to the permittivity of vacuum.

Conformation study for borated sugars

First, sugar molecules (with out borate ester) were investigated doing a composite rotation using the Complete Rotation from the Evaluation of Potential Energy Surface (CREPES)[34] program. Here each functional group (OH & CH_3) was rotated around its axis in steps of 60/90 degrees with respect to the rotation of other functional groups. Single point energies of each of these rotations were calculated at RHF/6-31G(d,p). A total of 1500-1800 single point energy conformations for each sugar molecule were analyzed and were grouped according to the rotation angle of the first rotation in the composite rotation. From each of these groups minimum energy structures were taken and optimized at MP2/6-31G(d) in gas phase. The lowest energy structure obtained from these optimizations was used to as the sugar to make borate esters.

An optimized structure of boric acid was added to the sugars where they can make borate monoesters. Depending on the structure of the sugars there are three possible places where borate esters can be formed. When borate ester formed on the C1 and C2 of the sugar, it is called 1,2-carbohydrate and when the borate ester is formed on C2, C3 and C3, C4 then those sugars are called 2,3-carbohydrate and 3,4-carbohydrate respectively (Figure 1). Once the borate ester is built, the planer structure of borate group was rotated angles from -60° to +60° in steps of 5°and the single point energy of each rotation was calculated at RHF/6-31G(d,p). The lowest energy conformation was then optimized at MP2/6-31G(d) in gas phase and used as the QM structure.

There are 8 molecules in the training set and 5 molecules in the test set as shown in Figure 2. These molecules are named according to the sugar and the place of carbon atoms where borate di-ester binds. The molecules in the training set are 1,2-boro-diol-alpha-D-Galactose (1,2-aGal), 1,2-boro-diol-alpha-D-Xylose (1,2-aXyl), 1,2-boro-diol-alpha-Fructofuranose (1,2-aFrucfur), 2,3-boro-diol-alpha-D-Galactose (2,3-aGal), 2,3-boro-diol-alpha-L-Fucose (2,3-aFuc), 3,4-boro-diol-alpha-D-Galactose (3,4-aGal), 3,4-boro-diol-alpha-L-Fucoose (3,4- aFuc) and 3,4-boro-diol-alpha-D-Mannose (3,4-aMan). The test contains of 1,2-boro-diol-beta-D-Fucose (1,2-bFuc), 2,3-boro-diol-alpha-Fructofuranose (2,3-aFrucfur), 2,3-boro-diol-alpha-D-Mannose (2,3-aMan), 2,3-boro-diol-alpha-D-Apoise (2,3-aApi) and 3,4-boro-diol-alpha-L-Rhamnose (3,4-Ram).

To maintain the transferability of parameters with other CHARMM bio-molecular force fields, previously established protocol for CHARMM force fields to parameterize borated carbohydrates[29,30] were followed.

Force field parameterization is an iterative process where all the parameters in the force field are independent[30] and need to be optimized in a self-consistent manner. However, according to previous studies, using the following order to optimize each parameter often allows the parameterization to converge in one or two iterations.[24]

(i) Partial atomic charges: Hydrogen bonding water-solute pair interaction energies and distances were calculated for charge optimization using the following procedure. α-D-galactose with borate ester formed at C1 and C2, 1,2-boro-diol-alpha-D-Galactose (1,2-aGal) was taken as the model compound for the solute. Solute geometry was optimized in gas phase at MP2/6-31G(d) level[31]. A water molecule with the internal geometries identical to the TIP3P[32] water model was used to construct solute-solvent pairs as shown in Figure 3. In pair **a-i** and **a-ii** the hydroxyl hydrogen in the borate is the hydrogen-bond donor and in all the other pairs hydroxyl oxygen in the borate is the hydrogen-bond acceptor. In **a-i** and **a-ii** the O–H bond vector is constrained to lie on the bisector of H–O–H with B–O–O_{water}–H dihedral angles of 180° and 90°. In **b-i**, **b-ii** and **b-iii**, where water is the donor the O–H bond vector is constrained to lie on the bisector of B–O–H with B–O–O_{water}–H dihedral angles of and 0°, 90° and 180°. In pairs c-i and c-ii one hydrogen atom of the water molecule is in the B–O_{ring}–C plane and the other

hydrogen is located below and above the plane, respectively. For pairs **e-i** and **e-ii** the arrangement of water molecule is same as **c-i** and **c-ii** except that the plane of B–O_{ring}–C is opposite to the O_{ring} of **c-i** and **c-ii.** For pairs **d-i** and **d-ii**, hydrogen of the water molecule is 120° above and below, the B–O_{ring}–C plane respectively. The possibility of a water molecule making two hydrogen bonds with the oxygens in B–O_{ring}–C and B–O–H was modeled in **f-i**. In both QM and MM, the solute water interaction energy was determined by optimizing the distance between the sugar and the water molecule with all the other degrees of freedom constrained. GAMESS, the QM program that was used, has a limitation on the number of constrained internal coordinates, therefore for both QM and MM this constrained energy was calculated manually by mapping with a series of single point energies, the distance at 0.25Å increments. The QM water interaction energy was calculated at RHF/6-31G(d) level and MM calculations were calculated in gas phase with no truncation of non-bonded interactions. Both QM and MM interaction energy ($E_{interaction}$) was computed using the equation, $E_{interaction} = (E_{solute+water} - E_{solute} - E_{water})$.

Charge was optimized by adjusting charges to reproduce QM minimum interaction energies and distances between the solute and TIP3P water. Parameterization using QM structures has well known limitations so the well-established CHARMM additive force field empirical scaling rules[29,30] were used. In this method the MM distance (R_{MM}) is calculated as, $R_{MM} = R_{QM} - 0.2$ Å and the MM interaction energy (E_{MM}) is given by the expression 1.16* E_{QM}, where E_{QM} is the QM interaction energy.

(ii) Optimization of the equilibrium bond lengths and valence angle parameters to reproduce the QM values of the training set. These QM values are obtained from optimizing the structures of the lowest constraint dihedral angle (see below) at MP2/6-31G(d). The same geometry was optimized in MM with Newton-Raphson minimization (nrap) method.

(iii) Optimization of bond, angle, improper and dihedral force constants need to be done using scaled MP2/6-31G(g) vibrational spectrum. This step is a part of future work for this study and will require determining ways to construct a test set that has experimental validation of at least some of the structures. Such experimental studies are not currently available.

(iv) Following the established procedures[33,34] for parameter optimization of the dihedral angles, the parameters for the dihedral angle O_{ring}–B–O–H was calculated as follows. In QM the target dihedral was scanned at 15° intervals from 0°-360° and a constrained optimization was carried out at each angle with dihedral angle constrained at MP2/6-31G(d). Single point energy of the each constrained-optimized geometry was calculated at MP2/cc-pVTZ level of theory. Constrained optimization on MM was done at the same dihedral angles as QM with the constrained dihedral force constant of 1000 kcal/mol.

During the parameterization of borated esters, the above-mentioned steps for each parameterization were iterated until the best values for each parameter were obtained (converged).

III. Results

The initial topology information and initial parameters for the pyranose and furanose rings were directly transferred from the existing CHARMM 38, top_all36_carb.rtf and par_all36_carb.prm files. The parameters for $BO_2(OH)$ ester functional form were the new parameters to assign and these were assigned by the analogy to existing parameters. The parameters belonging to the carbohydrate ring are the initial parameters, which are already optimized. Some of these initial parameters, parameters belong to the atoms where the boro-diester connects to the carbohydrate, so they needed to be re-optimized to reproduce the geometrical data.

The atom types introduced in this study are listed in Table 1. The set of new parameters for the training/test sets of molecules is provided in Table 2. Here boron is introduced as a new atom-type. Typically experimental data would be required to optimize the LJ parameters, but due to the lack of borated sugar experimental data, LJ parameters (ε_{ij} and $R_{min_{ij}}$) were started with guess values and iteratively optimized with the other parameter optimization. Because the ultimate aim of this study is to model RG-II pectic polysaccharide and characterize possible borate binding sites, the LJ parameters

optimized by experimental values such as heats of evaporation, crystal lattice and pure solvent molecular volume are likely not essential for the ultimate modeling goal of studying cross linking in plant cell walls. Charges of four atoms (B, O_{ring} , O_{ring} , O_{OH} , H) in the borate-diol ester were optimized using the interaction energy and hydrogen-bond distance of water–solute interaction. Optimization of improper parameters, K_{imp} and φ_0 were essential to keep the planar geometry of borate ester, BO_2OH. To specify a total of 13 molecules, the parameter set includes 144 new torsional definitions, with each having parameters, K_{χ}, n and δ, where δ will be either 180° or 0°. 12 bond parameters and 31 angle parameters have also been introduced. The QM, MM bond lengths and the differences between bond lengths and that of bond angles are listed in tables 4 to 16. The overall root mean square deviations (RMSD) for bond lengths and angles over the entire training/test set is given in table 17. The torsional energy profiles are the plots of potential energy versus the torsional angle. The potential energy for each boro-carbohydrate was calculated with respect to the lowest potential energy of that boro-carbohydrate. For sugars with mono-boro-diesters the only torsional angle that change the conformational energy is the O_{ring}–B–O–H dihedral angle. The dihedral parameters from the hydroxyl hydrogen in sugars are directly transferred from earlier work.[23,33] The torsional energy graphs for all the molecules for both MM and QM are shown in Figures 3 to 15. The barrier heights and RMSD for each system are listed in Table 18.

IV. Discussion

The set of new parameters and re-optimized old parameters used in this study are listed in Table 2. These parameters along with existing CHARMM parameters[23,33] for carbohydrate were used to parameterize boro-ester sugars.

When using QM geometries for comparisons in CHARMM, up to 0.03 Å respective deviations in bond length and up to 3 degrees respective deviations in bond angles are acceptable. The boron-oxygen bond distances (Figure1) of B−O_H (in B-O_H-H), where oxygen is the hydroxyl oxygen, is consistently smaller than that of *ab initio* value with an average deviation of -0.01 Å and standard deviation of 0.00 Å. In the two B−O_{ring} (in B−O_{ring}−C) bonds, where oxygen binds to two carbons, the average and standard deviations are 0.00, 0.02 Å and -0.01, 0.02 Å. Here the major contribution to the average is given by B-O_{ring} distance where O_{ring} is connected to furanose carbohydrate. The O−H bond length in B−O−H is 0.02 Å lower than that of *ab initio* value. B−O in B−O−H is in a very good fit through out the training set with 0.00 Å for both average and standard deviations of change in distance. The C−C distance (C1−C2, C2−C3 or C3−C4 as shown in Figure1), where each carbon binds to oxygen in B− O_{ring} −C, needed to be re-optimized. Here the difference in C−C bond distance with respect to its *ab initio* value deviated in both negative and positive direction with an absolute average of 0.03 Å and a standard deviation of 0.03 Å. Also the bond distance C−$O_{carb.ring}$, where oxygen is in the pyranose or furanose structure, B−C−$O_{carb.ring}$ (Figure 1−a) was re-optimized with an

absolute average of 0.02 Å in both bond deviation and standard deviation. The average relative deviation of bond lengths in all 13 molecules (Table 4-13) is with in 0.03 Å, therefore the optimized parameters gives bond lengths with in the acceptable values for bond lengths.

The new angle parameters added for this study are discussed below. The three bond angles around planner $(BO_3)^{3-}$ deviates as follows. O_{ring}−B−O_{ring} angle and one of the O_{ring}−B−O_H angles are smaller than the corresponding *ab initio* values. Where as the other O_{ring}−B−O_H angle has a higher bond angle than the *ab initio* counterpart. The maximum and the minimum deviations in the bond angles around the planner boron are 3.90 degrees and 0.20 degrees and absolute averages of 1.37, 2.50 and 2.9 degrees and a common standard deviation of 1.00 degrees. The bond angle, B−O_H−H equally deviates in both positive and negative directions and has a good absolute average of 0.12 degrees. One of the C−O_{ring}−B angles deviates significantly in the range 0.00 degrees to 4.40 degrees with an absolute average of 2.26 degrees and a standard deviation of 1.34 degrees. The other C−O_{ring}−B angle deviates in between the maximum and minimum of 5.50 degree and 0.60 degrees. In both of these cases the maximum bond angle is higher than the general accepted value of 3 degrees. Attempts to lower these angles decreased the quality of the overall parameter fit. Therefore these angles are considered to be the best values for C−O_{ring}−B angle. The O_{ring}−C−$O_{carb.ring}$ is in good agreement with the *ab initio* values with maximum and minimum values of 0.20 degrees and 2.30 degrees. The two C−C−O_{ring} angles are in good agreement with the *ab initio* values except for the type

of structures shown in Figure 1-c, where the angles are slightly above 3 degrees. The absolute average and standard deviations for the two C−C−O_{ring} bonds are 2.08, 1.14 degrees and 1.21, 1.26 degrees respectively. The acceptable individual deviation for parameterizations in bond angle using QM optimized structures is 3 degrees. When considering the above optimized bond angles it is clear that all the bond angles deviations are not with in 3 degrees. This discrepancy in deviation can be explained using the structures in training and test sets. The parameters optimized in this study cover two types of sugar ring conformations (pyranose and furanose) with three different places for boron ester to bind (C1−C2, C2−C3 and C3−C4) in each conformation. It is clear that the attempt here is to fit a minimum number of parameters over a large range of molecular types. Considering the main idea of force field parameter optimization is to use minimum number of parameters with the appropriate level of accuracy for the application of interest, parameters obtained in here gives satisfactory results. Also note that the average root mean square deviation of bond angles across 13 molecules is 2.5 degrees (Table 17), which tells that as an overall the deviation in bond angle is a reasonable value for a system with a range of molecular types.

The average difference in conformational energy for dihedral angle of O_{ring}−B−O_{OH}−H varies from positive 0.30 kcal/mol to negative 0.28 kcal/mol resulting a zero average difference across 13 molecules (Table 18). The absolute average difference in conformational energy is 0.13 kcal/mole and the average standard deviation for all the molecules is 0.63 kcal/mol. Also referring to torsional energy profiles given from Figure

3 to Figure 15 it is clear that the discrepancy in most of the conformational energies occurs in the high-energy regions and the low energy regions are in good agreement with the QM values. The future work of this study involves MD simulations of sugar molecules in plant cell wall at room temperature. At room temperatures high-energy conformations would not be populated therefore the discrepancy in high-energy conformations can be neglected.

The partial atomic charges were optimized so that they produce the minimum interaction energy difference and minimum hydrogen−bond distance difference between QM and MM values. The interaction energies and distances for both QM and MM are provided in Table 3. The molecular mechanics solute−water interaction energies and distances are in close agreement with HF target data when water molecule acts as both hydrogen bond donor and acceptor. The average interaction energy error is 0.11 kcal/mol and average interaction distance error is −0.08 Å, both these values confirm that the overall solvation of the borate-ester is reasonable.

Because of the lack of experimental data for borated sugar compounds, the Lennard-Jones parameters were optimized in the iterative process until reasonable values for geometries and charges are obtained. To get the proper Lennard-Jones parameters for boron, we will suggest a future project needs to be pursued where; Lennard-Jones parameters are optimized for boron in boron nitrogen systems. The boron nitrogen systems are suggested due to the availability of experimental data.

The topology and full parameter files are attached as Supplement 1 and 2 respectively.

V. Conclusions

We have presented the development of CHARMM force field for boron carbohydrate complexes. This model based in part on fitting to geometries, solute-water interaction energies and torsional potentials at MP2/6-31G(d,p). The parameters were validated by its ability to reproduce the electronic structure geometries and torsional barriers. The average root mean square deviation of bond length is 0.02 Å and for the bond angle the deviation is 2.5 degrees. The torsional barriers are well reproduced within 0.6 kcal/mol. These values are within the acceptable accuracy of CHARMM force fields. Therefore this force field parameterization can be used to study the RG-II interactions in the plant cell wall with atomistic simulation methods.

References

(1) Jarvis, M. C.; McCann, M. C. *Plant Physiology and Biochemistry* **2000**, *38*, 1.

(2) Jarvis, M. C. *Plant Cell and Environment* **1984**, *7*, 153.

(3) Chormova, D.; Messenger, D. J.; Fry, S. C. *Plant Signal Behav* **2014**, *9*, e28169.

(4) Alan G. Darvill, M. M., and Peter Albersheim *Plant Physiol.* **1978**, *62*.

(5) Jerry R. Thomas1, A. G. D., Peter Albersheim *Carbohydrate Research* **1989**, *185*.

(6) Jerry R. Thomas, M. M., Alan G. Darvill, and Peter Albersheim *Plant Physiol.* **1987**, *83*.

(7) Goldbach, H. E. *Journal of Trace and Microprobe Techniques* **1997**, *15*, 51.

(8) S Vidal, T. D., P Williams, P Pellerin, W.S York, M.A O'Neill, J Glushka, A.G Darill, P Albersheim *Carbohydrate Research* **2000**, *326*

(9) McCann, M. C.; Roberts, K. *Journal of Experimental Botany* **1994**, *45*, 1683.

(10) Carpita, N. C.; Gibeaut, D. M. *Plant Journal* **1993**, *3*, 1.

(11) Kobayashi, M.; Matoh, T.; Azuma, J. *Plant Physiology* **1996**, *110*, 1017.

(12) Ishii, T.; Matsunaga, T. *Carbohydrate Research* **1996**, *284*, 1.

(13) Oneill, M. A.; Warrenfeltz, D.; Kates, K.; Pellerin, P.; Doco, T.; Darvill, A. G.; Albersheim, P. *Journal of Biological Chemistry* **1996**, *271*, 22923.

(14) Loomis, W. D.; Durst, R. W. *Biofactors* **1992**, *3*, 229.

(15) Matoh, T.; Ishigaki, K.; Ohno, K.; Azuma, J. *Plant and Cell Physiology* **1993**, *34*, 639.

(16) Brown, P. H.; Bellaloui, N.; Wimmer, M. A.; Bassil, E. S.; Ruiz, J.; Hu, H.; Pfeffer, H.; Dannel, F.; Romheld, V. *Plant Biology* **2002**, *4*, 205.

(17) Pellerin, P.; Doco, T.; Vidal, S.; Williams, P.; Brillouet, J. M.; Oneill, M. A. *Carbohydrate Research* **1996**, *290*, 183.

(18) Kobayashi, M.; Nakagawa, H.; Asaka, T.; Matoh, T. *Plant Physiology* **1999**, *119*, 199.

(19) Ishii, T.; Matsunaga, T.; Pellerin, P.; O'Neill, M. A.; Darvill, A.; Albersheim, P. *Journal of Biological Chemistry* **1999**, *274*, 13098.

(20) Perez, S.; Rodriguez-Carvajal, M. A.; Doco, T. *Biochimie* **2003**, *85*, 109.

(21) Matoh, T.; Kawaguchi, S.; Kobayashi, M. *Plant and Cell Physiology* **1996**, *37*, 636.

(22) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. *JOURNAL OF COMPUTATIONAL CHEMISTRY* **2008**, *29*, 2543.

(23) Hatcher, E. G., O.; MacKerell, A.D.; *JOURNAL OF PHYSICAL CHEMISTRY B* **2009**, *113*, 12466.

(24) Vanommeslaeghe, K. H., E.; Acharya, C.; Kundu,S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr. A. D.; *Journal of Computational Chemistry* **2010**, *31*, 671.

(25) B.R. Brooks, C. L. B. I., A.D. Mackerell Jr, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, J.M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus *J. Comput. Chem.* **2009**, *30*, 1545.

(26) Schmidt, M. W. B., K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., *J. Comput. Chem.* **1993**, *14*, 1347.

(27) Gordon, M. S. S., M. W., *Elsevier: Amsterdam, The Netherlands* **2005**.

(28) Mackerell, A. D. J. *J. Comput. Chem.* **2004**, *25*.

(29) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wio´rkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*.

(30) Foloppe N.; MacKerell, A. D. J. *Journal of Computational Chemistry* **2000**, *21*.

(31) Møller, C. P., M. S., *Phys. Rev.* **1934**, *46*.

(32) Jorgensen W. L.; Chandrasekhar, J. M., J. D.; Impey R. J.; Klein M. L.; *J. Chem. Phys.* **1983**, *79*, 926.

(33) Guvench, O. G., S.N.; Kamath, G.; Brady, J.W.; Venable, R.M.; Pastor, R.W.; Mackerell, A.D. *JOURNAL OF COMPUTATIONAL CHEMISTRY* **2008**, *29*, 2543.

(34) Guvench, O. G., Mackerell, A.D. Jr. *J. Phys. Chem. A* **2006**, *110*.

(35). Var Hagg, M. P.; Holme, T. A., ab initio investigation of boron binding. *PhD Thesis, Iowa State University* 2012.

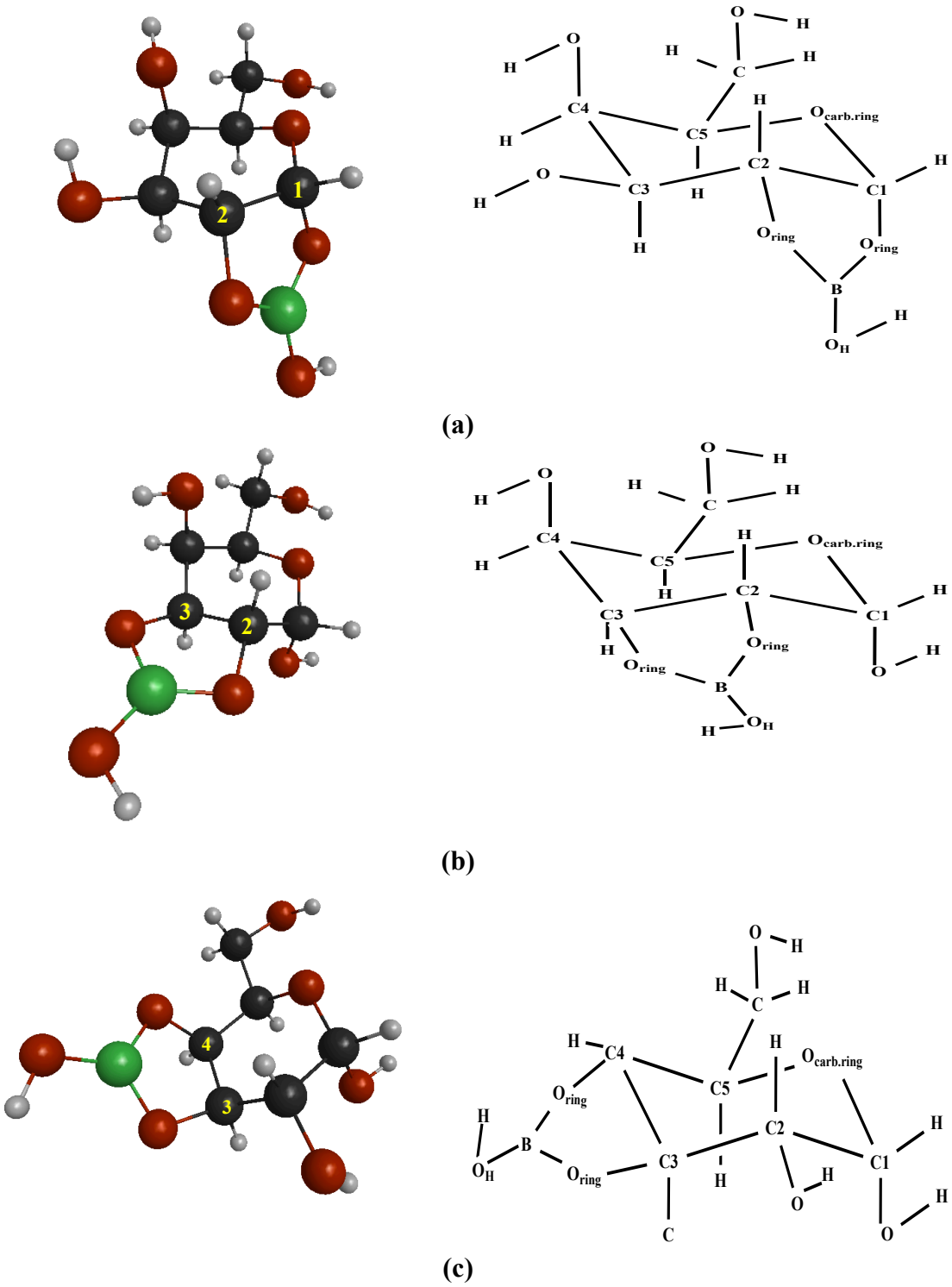
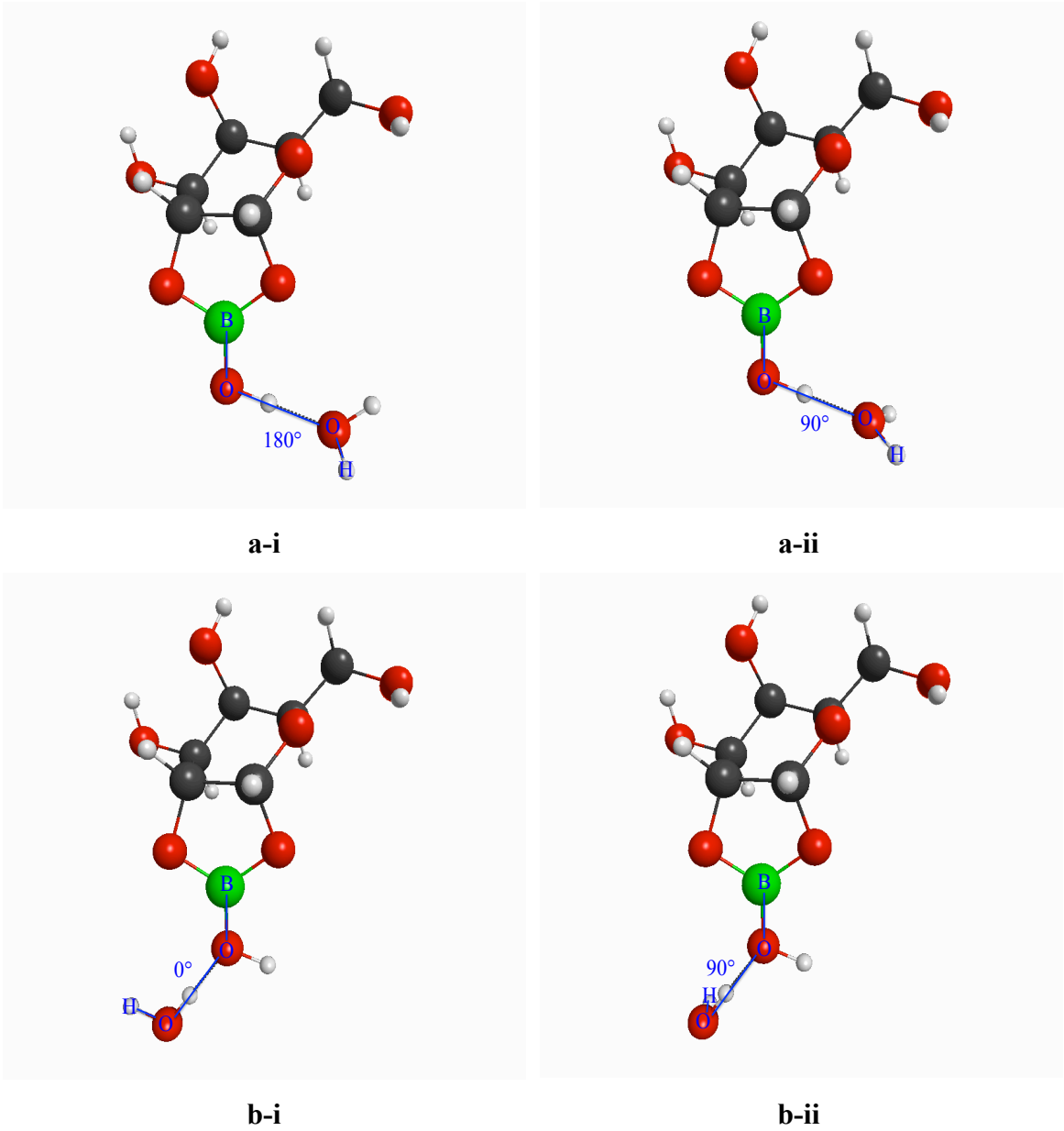


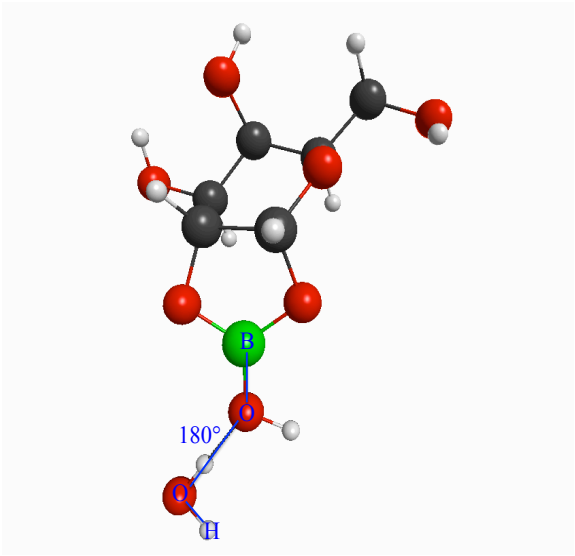
Figure 1. Carbohydrate boron-diester binding sites(α-d-galactose is taken as an example carbohydrate). (a) 1,2−carbohydrate, (b) 2,3−carbohydrate (c) 3,4−carbohydrate



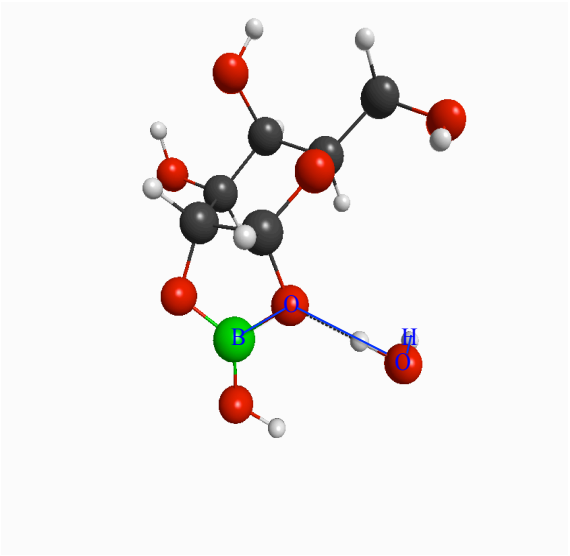
a-i **a-ii**

b-i **b-ii**

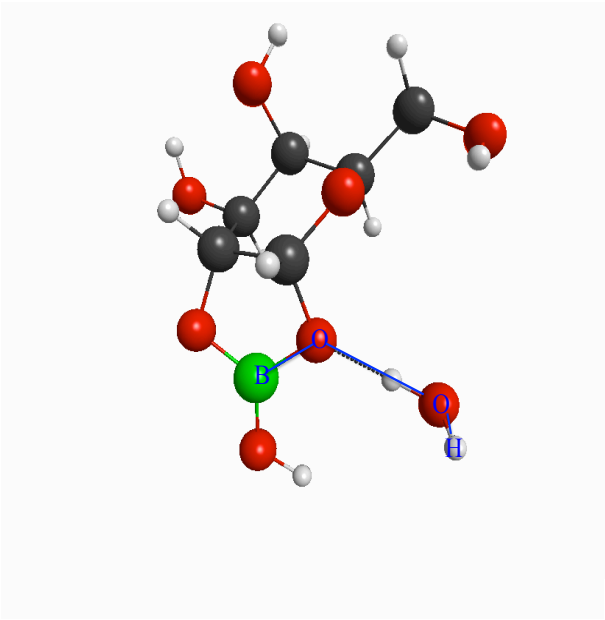
Figure 2. Schema a−f, water1−2−boro−diester−alpha−d−galactose interactions



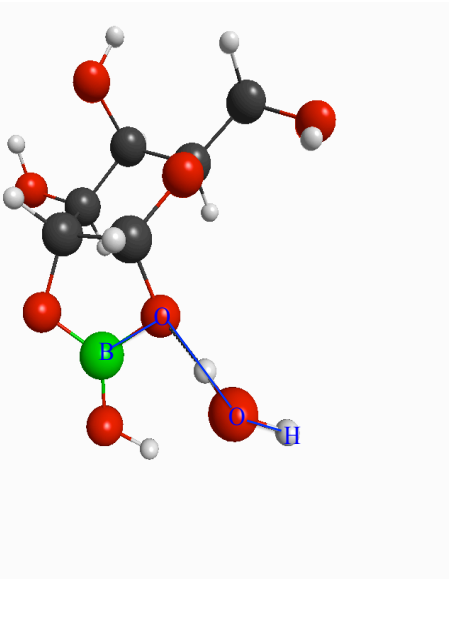
b-iii



c-i

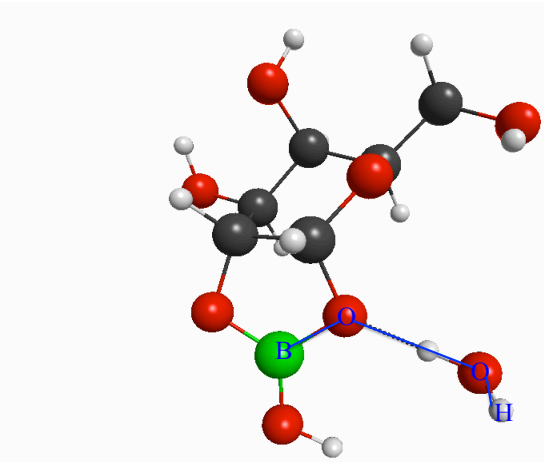


c-ii

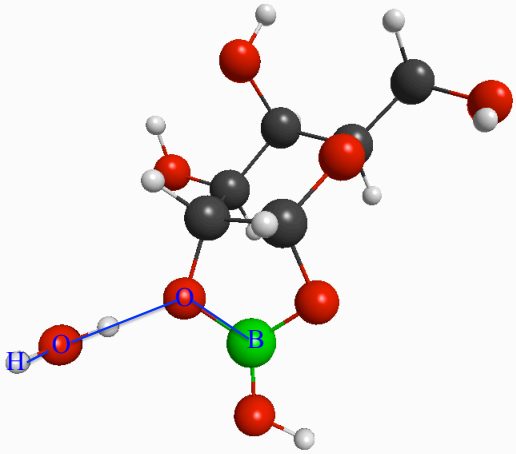


d-i

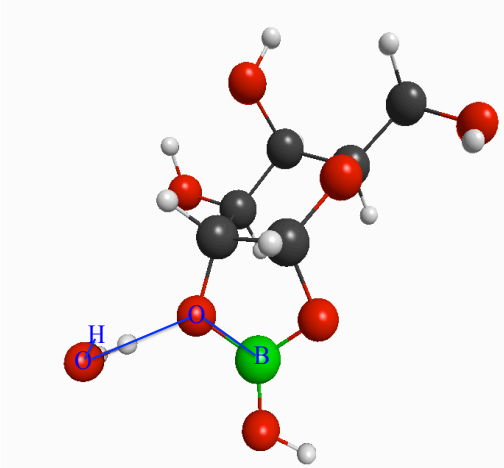
Figure 2. (continued)



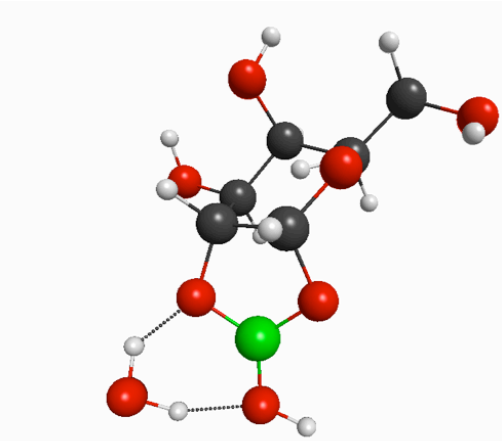
d-ii



e-i



e-ii



f-i

Figure 2. (continued)

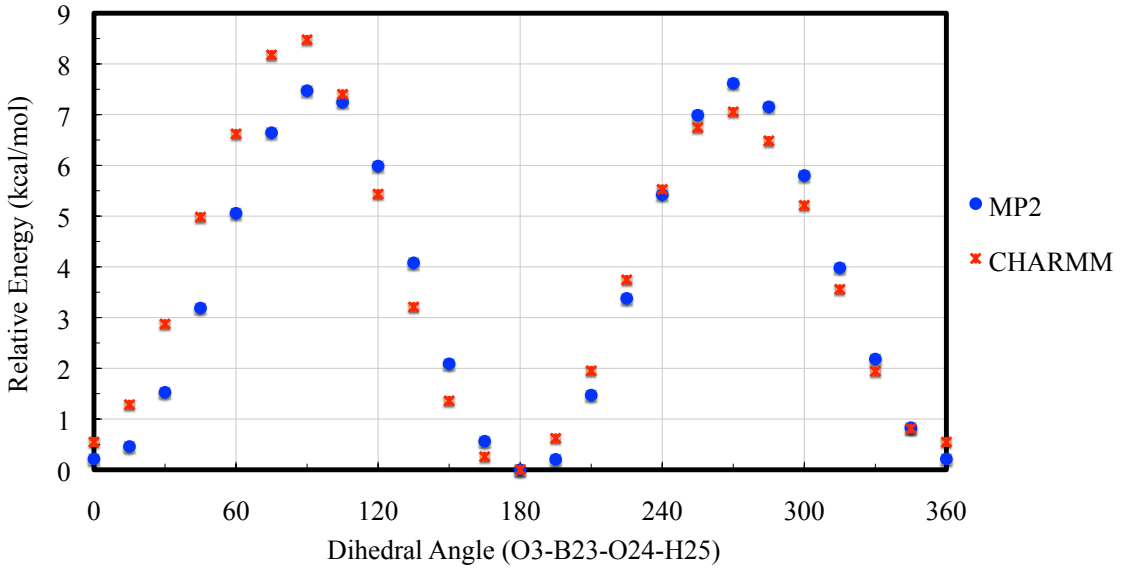


Figure 3. Torsional Profile for Molecule 1,2−aGal

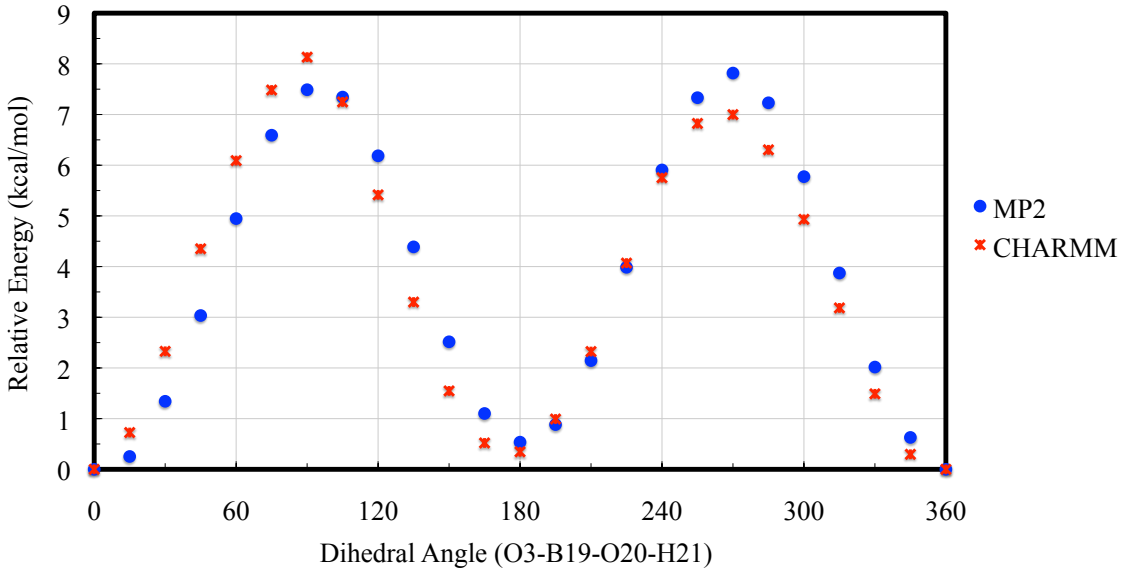


Figure 4. Torsional Profile for Molecule 1,2−aXyl

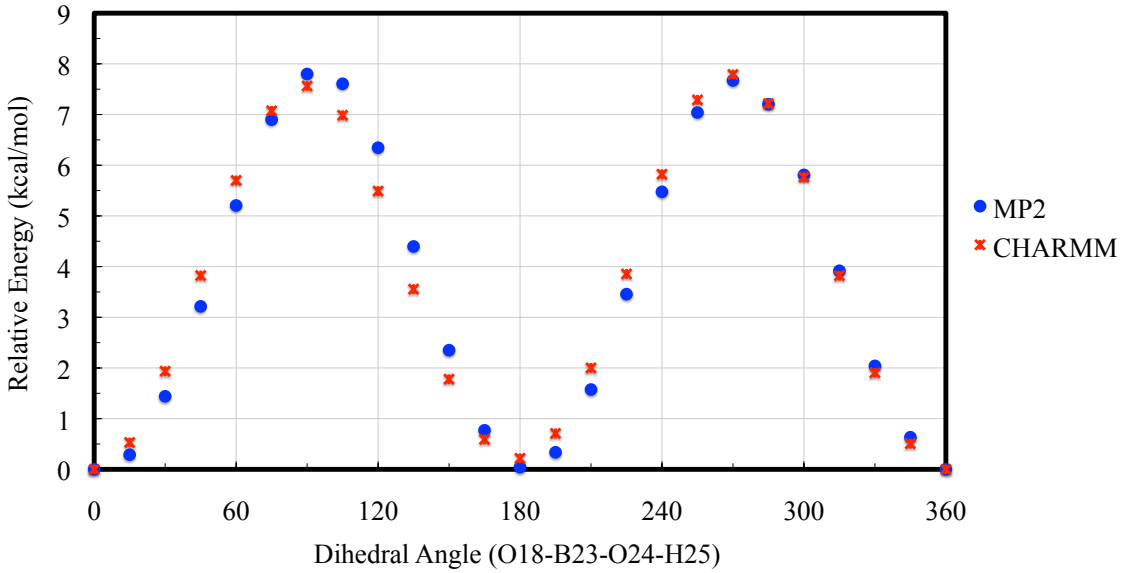


Figure 5. Torsional Profile for Molecule 1,2−aFrucfur

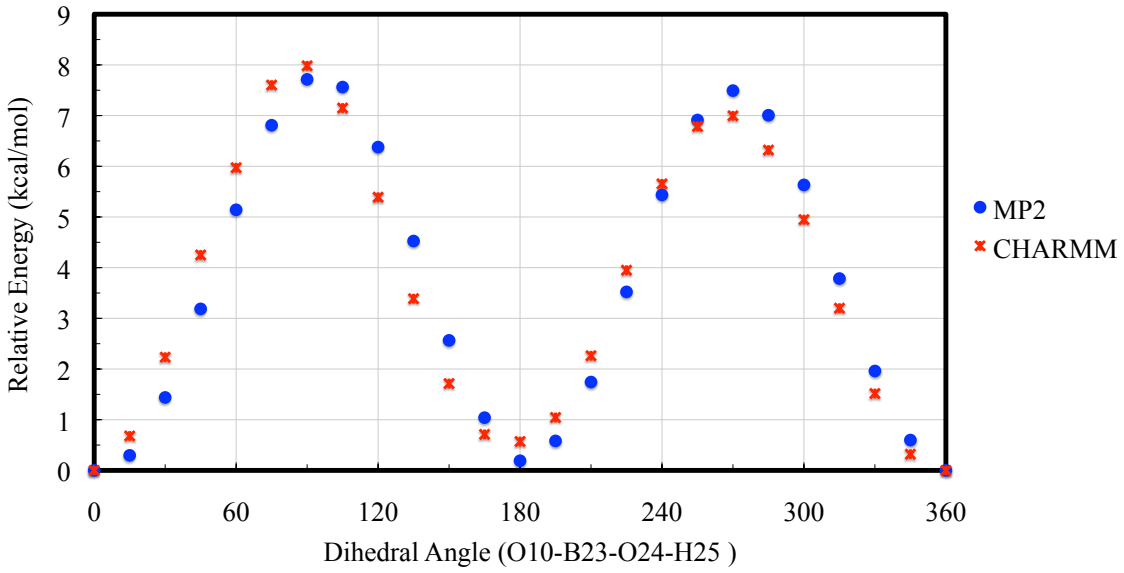


Figure 6. Torsional Profile for Molecule 2,3−aGal

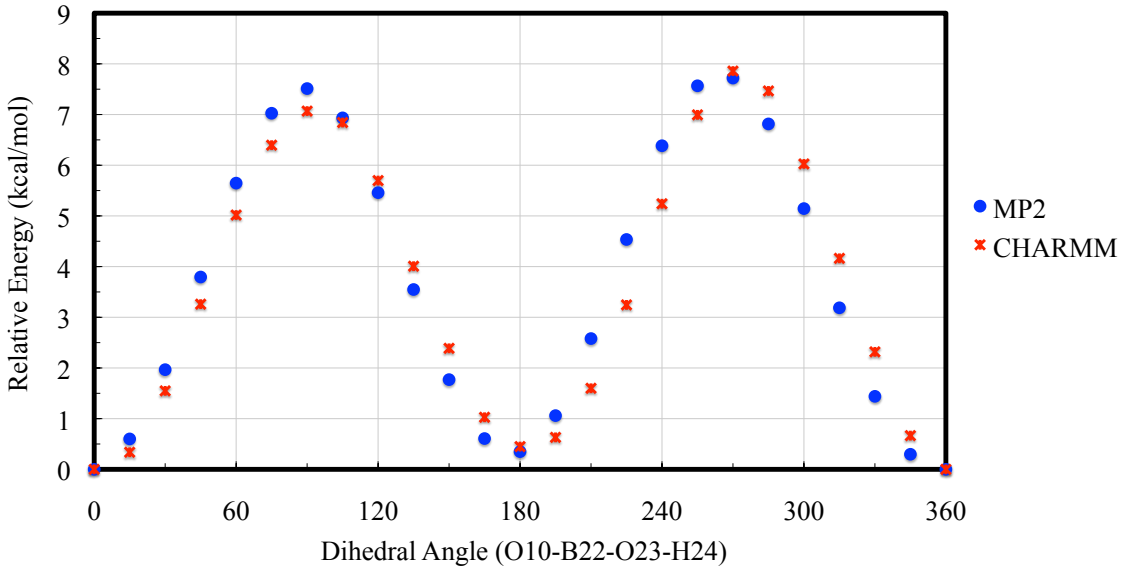


Figure 7. Torsional Profile for Molecule 2,3−aFuc

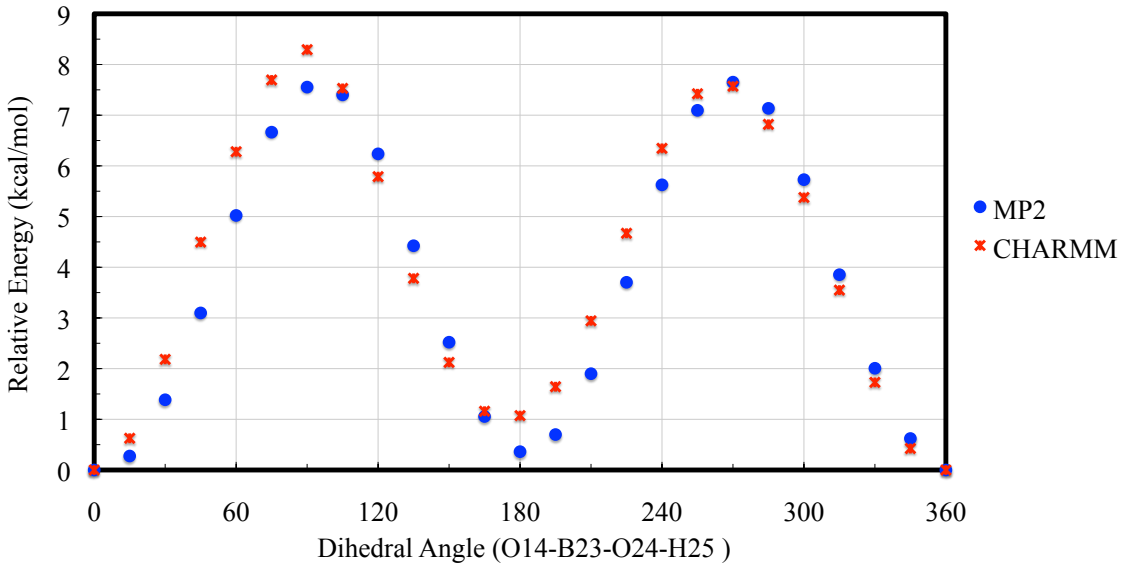


Figure 8. Torsional Profile for Molecule 3,4−aGal

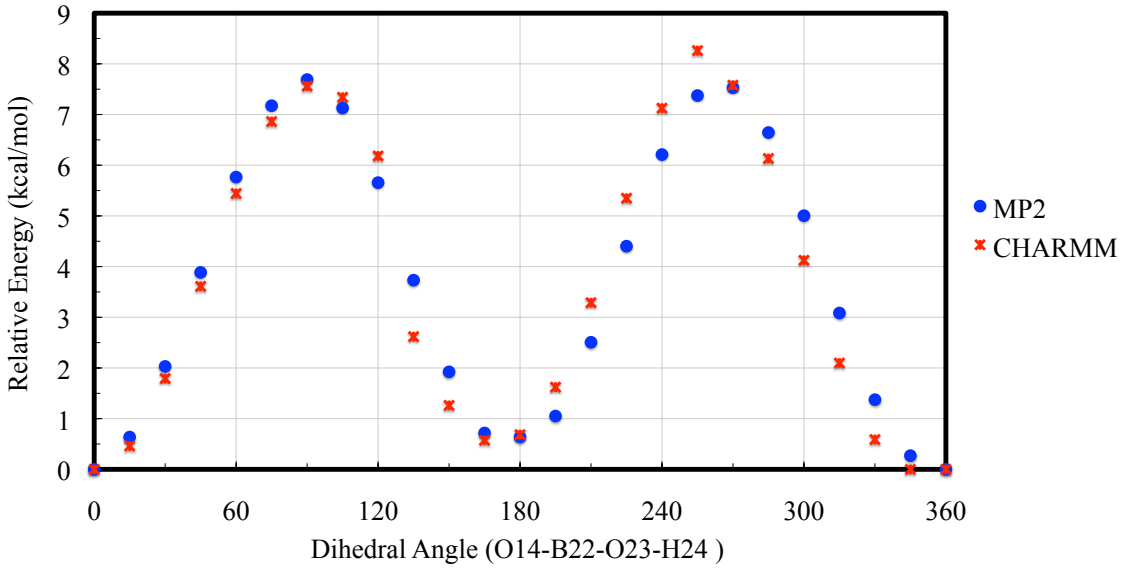


Figure 9. Torsional Profile for Molecule 3,4− aFuc

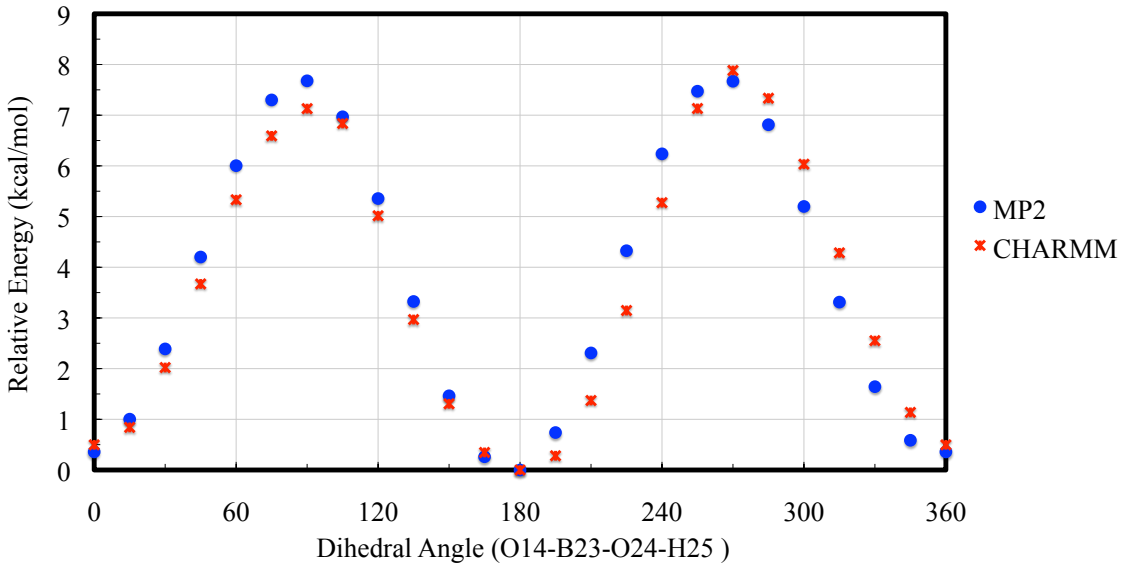


Figure 10. Torsional Profile for Molecule 3,4−aMan

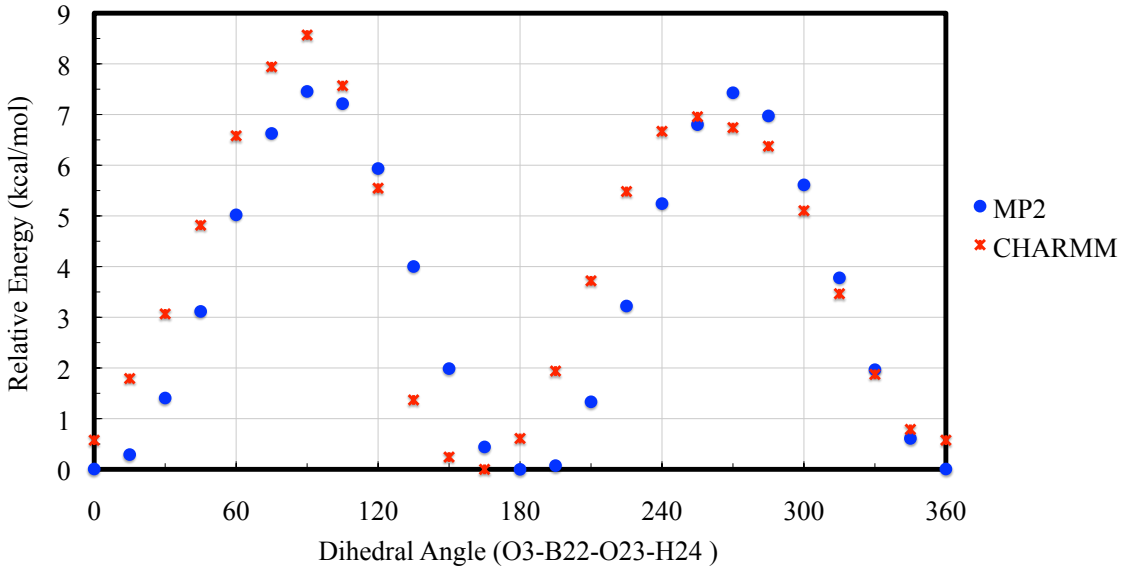


Figure 11. Torsional Profile for Molecule 1,2−bFuc

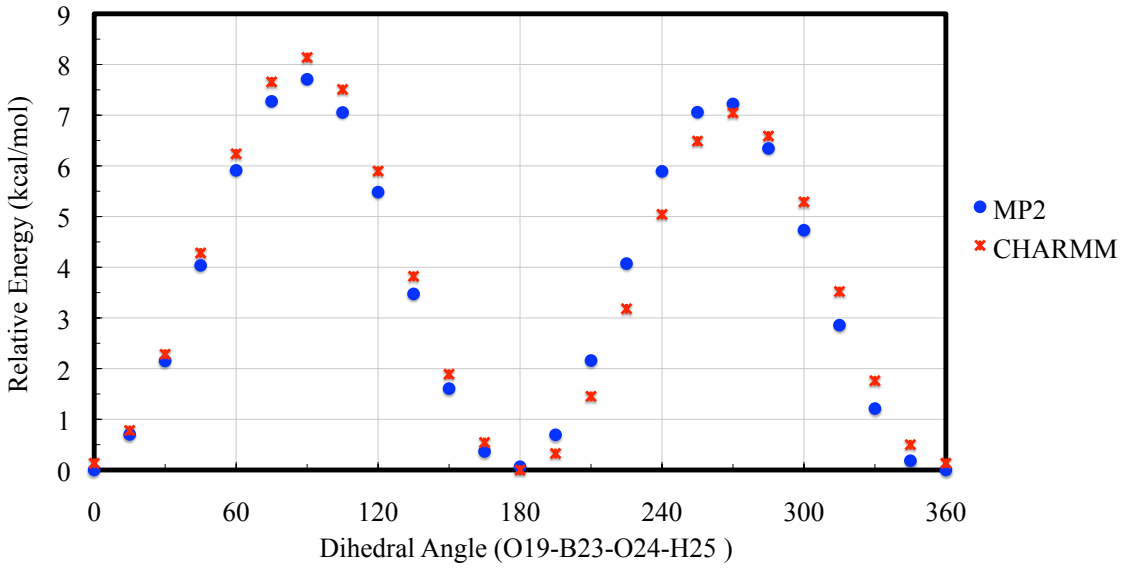


Figure 12. Torsional Profile for Molecule 2,3−aFrucfur

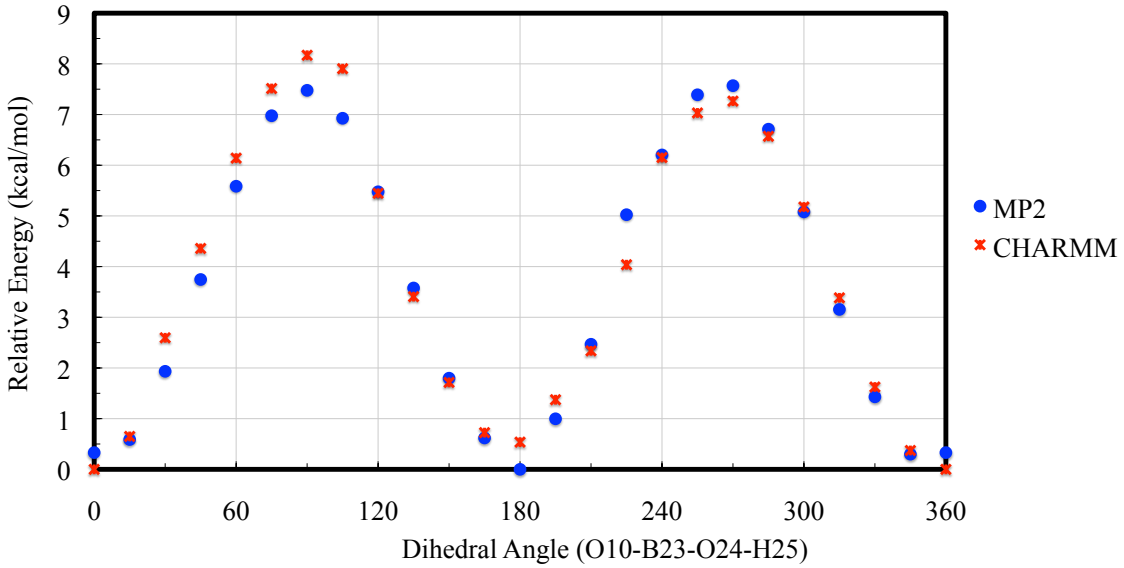


Figure 13. Torsional Profile for Molecule 2,3−aMan

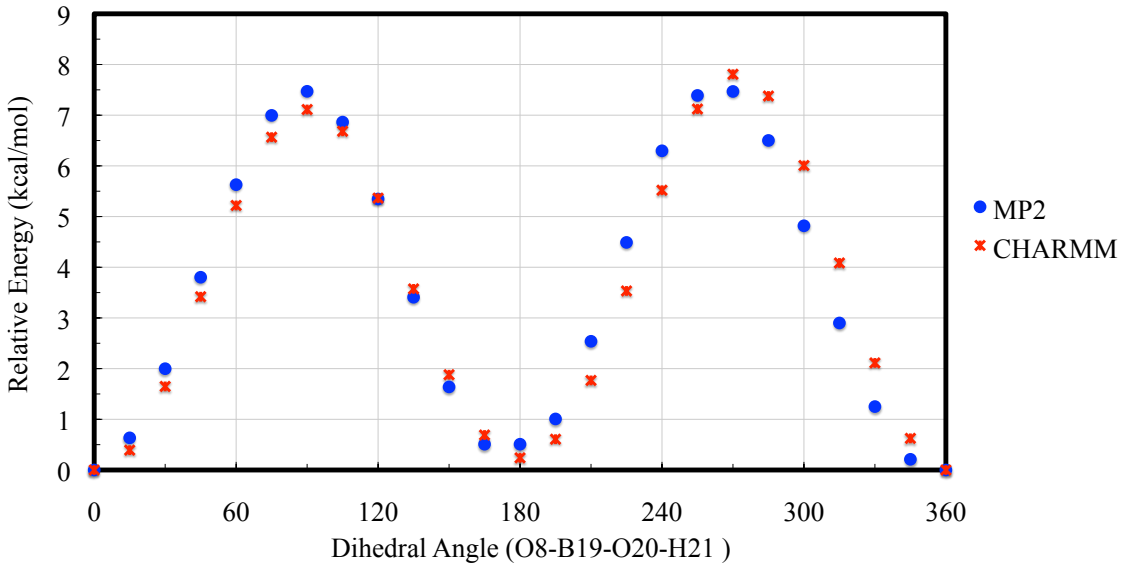


Figure 14. Torsional Profile for Molecule 2,3−aApi

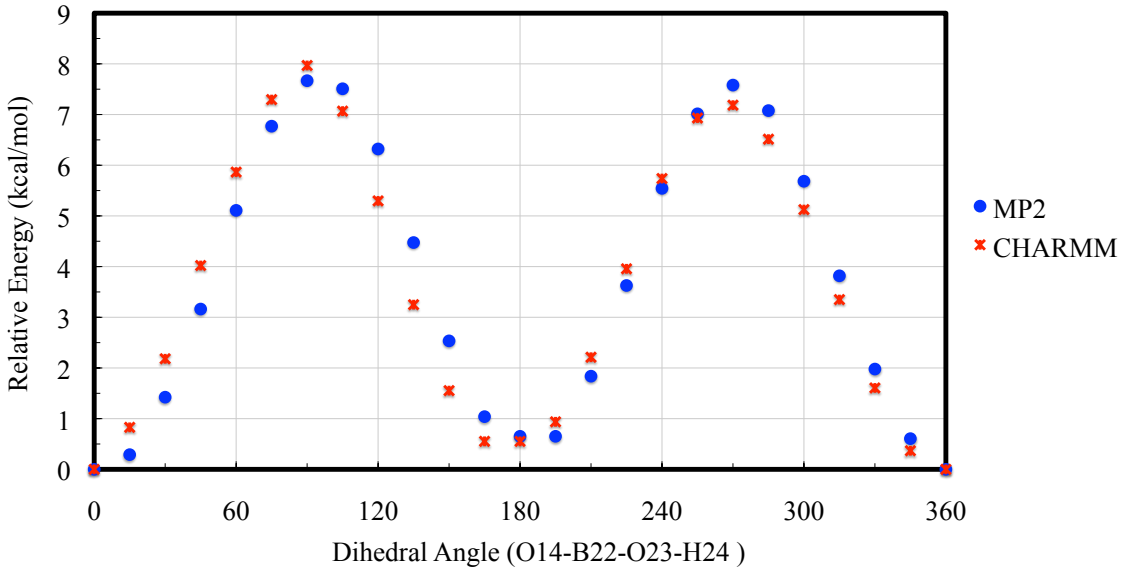


Figure 15. Torsional Profile for Molecule 3,4−aRam

Table 1. New Atom Definitions Needed for CHARMM Parameterization

Atom Number	Atom Type	Mass	Description
510	BO3	11.003	3 coordinated boron (sp2 hybridization) in boric bind to carbohydrate
512	OB1	15.9994	Oxygen in B−O−H
513	HOB3	1.008	Hydrogen in B−O−H
514	OB2	15.9994	Oxygen in B−O−C−O (C−O in sugar ring)

Table 2. CHARMM Parameters for Boron Di-esters in Sugars

Bonded Parameters

Stretching Parameters

Atom I	Atom J	K_θ(kcal/(mol*Å^{-2})	b_0(Å)
OB1	HOB3	545.0	0.960
BO3	OB1	230.0	1.400
BO3	OB2	500.0	1.380
BO3	OB3	500.0	1.220
OB2	CC3151	410.0	1.395
OB2	CC3051	410.0	1.500
OB2	CC3162	428.0	1.440
OB2	CC3161	410.0	1.440
CC3151	CC321	222.5	1.490
CC3153	HCA2	307.0	1.100
CC3161	*CC3161*	*250.5*	*1.550*
CC3162	*OC3C61*	*400.0*	*1.400*
CC3161	*CC3162*	*222.5*	*1.560*
CC3151	*CC3151*	*210.0*	*1.508*

The re−optimized parameters from carbohydrate.prm is given in italic.

Bending Parameters

Atom I	Atom J	Atom K	K_θ (kcal/(mol*rad^{-2})	$\theta(^0)$
BO3	OB1	HOB3	53.0	115.00
BO3	OB2	CC3162	50.0	115.00
BO3	OB2	CC3161	50.0	120.00
BO3	OB2	CC3051	50.0	109.00
BO3	OB2	CC3151	50.0	109.00
OB2	BO3	OB1	100.0	130.00
OB2	BO3	OB2	70.0	125.00
OB2	CC3151	CC321	75.7	110.10
OB2	CC3151	CC3153	75.7	110.10
OB2	CC3162	CC3161	100.0	110.00
OB2	CC3162	OC3C61	100.0	112.00
OB2	CC3162	HCA1	55.0	108.89
OB2	CC3161	HCA1	55.0	110.00
OB2	CC3161	CC3161	40.7	120.00
OB2	CC3161	CC3162	70.0	115.10
OB2	CC3161	CC3163	85.0	110.10
OB2	CC3161	CC3263	75.7	120.00
OB2	CC3051	OC3C51	460.0	120.50

Table 2. (continued)

OB2	CC3051	CC321	75.7	110.10
OB2	CC3051	CC3151	55.7	118.10
OB2	CC3151	CC3051	75.7	110.10
OB2	CC3151	HCA1	55.0	108.89
OB2	CC3151	CC3151	65.7	118.10
HCA2	*CC3153*	*HCA2*	*34.5*	*110.10*
HCA2	*CC3153*	*OC3C51*	*45.0*	*109.50*
OC311	*CC321*	*CC3151*	*75.7*	*110.10*
CC3151	*CC3151*	*CC321*	*58.4*	*113.50*
CC3153	*CC3151*	*CC321*	*58.4*	*113.50*
CC3151	*CC3153*	*HCA2*	*33.4*	*110.10*
CC3152	CC3151	OB2	75.7	110.10
CC3151	*CC321*	*HCA2*	*33.4*	*110.10*

The re-optimized parameters from carbohydrate.prm are given in italic.

Torsional Parameters

Atom I	Atom J	Atom K	Atom L	K_χ(kcal/(mol*rad^{-2})	n	χ (^{o})
OB2	BO3	OB1	HOB3	1.60	2	180.0
OB2	BO3	OB1	HOB3	1.15	1	180.0
OB2	BO3	OB1	HOB3	0.50	3	180.0
OB1	BO3	OB2	CC3161	0.76	1	180.0
OB1	BO3	OB2	CC3161	1.25	2	0.0
OB1	BO3	OB2	CC3161	0.48	3	180.0
OB1	BO3	OB2	CC3162	0.76	1	180.0
OB1	BO3	OB2	CC3162	1.25	2	0.0
OB1	BO3	OB2	CC3162	0.48	3	180.0
CC3161	CC3161	OB2	BO3	0.29	1	0.0
CC3162	CC3161	OB2	BO3	1.00	3	0.0
CC3161	CC3162	OB2	BO3	1.00	3	0.0
CC3162	OB2	BO3	OB2	1.24	3	0.0
HCA1	CC3162	CC3161	OB2	0.20	3	0.0
HCA1	CC3162	OB2	BO3	0.30	3	0.0
OB2	CC3162	OC3C61	CC3163	0.41	1	180.0
OB2	CC3162	OC3C61	CC3163	0.89	2	0.0
OB2	CC3162	OC3C61	CC3163	0.05	3	0.0
OB2	CC3162	CC3161	HCA1	0.20	3	0.0
OB2	CC3162	CC3161	OB2	1.24	3	0.0
OB2	CC3162	CC3161	CC3161	1.24	3	0.0
OB2	BO3	OB2	CC3161	1.00	3	0.0

Table 2. (continued)

OC3C61	CC3162	OB2	BO3	1.50	2	0.0
OC3C61	CC3162	OB2	BO3	1.00	3	0.0
OC3C61	CC3162	OB2	BO3	0.20	1	0.0
OC3C61	CC3162	CC3161	OB2	2.75	1	180.0
OC3C61	CC3162	CC3161	OB2	0.26	2	180.0
OC3C61	CC3162	CC3161	OB2	0.10	3	0.0
HCA1	CC3161	OB2	BO3	0.30	3	0.0
OB2	CC3161	CC3161	HCA1	0.24	3	0.0
OB2	CC3161	CC3161	OC311	2.65	1	180.0
OB2	CC3161	CC3161	OC311	0.00	2	180.0
OB2	CC3161	CC3161	OC311	0.13	3	180.0
OB2	CC3161	CC3161	CC3161	0.20	3	180.0
CC3162	CC3161	CC3161	OB2	0.20	3	180.0
OC311	CC3162	CC3161	OB2	2.65	1	180.0
OC311	CC3162	CC3161	OB2	0.00	2	0.0
OC311	CC3162	CC3161	OB2	0.13	3	180.0
CC3163	CC3161	CC3161	OB2	0.20	3	0.0
OB2	CC3161	CC3161	OB2	1.24	3	0.0
CC3163	CC3161	OB2	BO3	0.50	3	180.0
HCA1	CC3163	CC3161	OB2	0.30	3	180.0
OC3C61	CC3163	CC3161	OB2	1.36	1	180.0
OC3C61	CC3163	CC3161	OB2	0.16	2	0.0
OC3C61	CC3163	CC3161	OB2	1.01	3	0.0
OB2	CC3161	CC3163	CC321	0.20	3	0.0
CC3263	OC3C61	CC3162	OB2	0.89	2	0.0
CC3263	OC3C61	CC3162	OB2	0.41	1	180.0
CC3263	OC3C61	CC3162	OB2	0.05	3	0.0
CC3263	CC3161	CC3161	OB2	0.20	3	0.0
CC3263	CC3161	OB2	BO3	0.29	1	0.0
CC3263	CC3161	OB2	BO3	0.62	2	0.0
CC3263	CC3161	OB2	BO3	0.05	3	0.0
HCA2	CC3263	CC3161	OB2	0.14	3	0.0
OC3C61	CC3263	CC3161	OB2	1.36	1	180.0
OC3C61	CC3263	CC3161	OB2	0.16	2	0.0
OC3C61	CC3263	CC3161	OB2	1.01	3	0.0
CC331	CC3163	CC3161	OB2	0.20	3	0.0
OB2	CC3161	CC3163	CC2O2	0.20	3	0.0
OC3C51	CC3051	OB2	BO3	0.50	3	180.0
OC3C51	CC3051	CC3151	OB2	0.65	2	180.0

Table 2. (continued)

OB1	BO3	OB2	CC3051	3.85	2	180.0
OB2	BO3	OB2	CC3051	0.50	3	0.0
CC3051	CC3151	OB2	BO3	0.19	3	0.0
OB2	CC3051	OC3C51	CC3153	0.86	3	0.0
OB2	CC3051	CC321	HCA2	0.14	3	0.0
OB2	CC3051	CC321	OC311	1.72	3	180.0
OB2	CC3051	CC3151	HCA1	0.14	3	0.0
OB2	CC3051	CC3151	OB2	1.64	3	180.0
CC3151	CC3151	CC3051	OB2	0.88	3	180.0
OB2	BO3	OB2	CC3151	0.50	3	0.0
CC3153	CC3151	CC3151	OB2	0.01	1	180.0
CC321	CC3051	OB2	BO3	0.19	3	0.0
CC321	CC3051	CC3151	BO2	0.84	3	0.0
CC3151	CC3051	OB2	BO3	0.19	3	0.0
OB1	BO3	OB2	CC3151	3.85	2	180.0
HCA1	CC3151	OB2	BO3	0.18	3	0.0
OB2	CC3151	CC3151	HCA1	0.14	3	0.0
OC311	CC3151	CC3151	OB2	0.23	3	0.0
CC3151	CC3151	OB2	BO3	0.08	3	0.0
CC321	CC3051	CC3151	OB2	0.84	3	0.0
OC3C51	CC3152	CC3151	OB2	1.26	1	180.0
OC3C51	CC3152	CC3151	OB2	1.27	2	0.0
OC3C51	CC3152	CC3151	OB2	0.53	3	0.0
OC3C51	CC3153	CC3151	OB2	0.14	1	0.0
OC3C51	CC3153	CC3151	OB2	0.70	2	0.0
OC3C51	CC3153	CC3151	OB2	0.18	3	0.0
OC3C51	*CC3153*	*CC3151*	*CC321*	*0.14*	*1*	*0.0*
OC3C51	*CC3153*	*CC3151*	*CC321*	*0.70*	*2*	*0.0*
OC3C51	*CC3153*	*CC3151*	*CC321*	*0.18*	*3*	*0.0*
CC3152	*OC3C51*	*CC3153*	*HCA2*	*0.30*	*3*	*180.0*
CC3152	CC3151	OB2	BO3	0.08	3	0.0
CC3152	CC3151	CC3151	OB2	0.01	1	180.0
CC3152	*CC3151*	*CC3151*	*CC321*	*1.26*	*1*	*180.0*
OC311	CC3152	CC3151	OB2	2.87	1	180.0
OC311	CC3152	CC3151	OB2	0.03	2	0.0
OC311	CC3152	CC3151	OB2	0.23	3	0.0
OB2	CC3152	CC3151	HCA1	0.14	3	0.0
CC3151	*CC3151*	*CC3153*	*HCA2*	*0.55*	*1*	*180.0*
CC3151	*CC3151*	*CC3153*	*HCA2*	*0.55*	*1*	*0.0*

Table 2. (continued)

CC3051	CC3151	CC3151	OB2	2.07	1	0.0
CC3051	CC3151	CC3151	OB2	2.13	2	0.0
CC3051	CC3151	CC3151	OB2	2.71	3	180.0
OC311	CC3051	CC3151	OB2	0.12	1	180.0
OC311	CC3051	CC3151	OB2	1.87	2	180.0
OC311	CC3051	CC3151	OB2	1.64	3	180.0
CC3153	CC3151	OB2	BO3	0.29	1	180.0
CC3153	CC3151	OB2	BO3	0.55	2	180.0
CC3153	CC3151	OB2	BO3	0.08	3	180.0
OB2	CC3151	CC3153	HCA1	0.14	3	0.0
CC321	CC3153	CC3151	OB2	0.76	1	180.0
CC321	CC3153	CC3151	OB2	0.40	2	180.0
CC321	CC3153	CC3151	OB2	0.40	3	180.0
OB2	CC3151	CC3151	OB2	2.87	1	0.0
OB2	CC3151	CC3151	OB2	0.03	2	180.0
OB2	CC3151	CC3151	OB2	0.23	3	180.0
OB2	CC3151	CC3152	HCA1	0.14	3	0.0
CC3151	*CC3151*	*CC321*	*OC311*	*0.01*	*1*	*0.0*
CC3151	*CC3151*	*CC321*	*OC311*	*0.14*	*2*	*0.0*
CC3151	*CC3151*	*CC321*	*OC311*	*0.70*	*3*	*180.0*
CC321	*CC3151*	*CC3151*	*HCA1*	*0.20*	*3*	*0.0*
CC3151	*CC321*	*OC311*	*HCP1*	*0.12*	*1*	*0.0*
CC3151	*CC321*	*OC311*	*HCP1*	*0.42*	*2*	*0.0*
CC3151	*CC321*	*OC311*	*HCP1*	*0.29*	*3*	*0.0*
OB2	CC3051	CC3151	OC311	0.12	1	180.0
OB2	CC3051	CC3151	OC311	1.87	2	180.0
OB2	CC3051	CC3151	OC311	1.64	3	180.0
OB2	CC3151	CC321	OC311	0.07	1	0.0
OB2	CC3151	CC321	OC311	1.99	2	180.0
OB2	CC3151	CC321	OC311	1.72	3	180.0
CC321	CC3151	CC3151	OB2	0.94	1	0.0
CC321	CC3151	CC3151	OB2	1.59	2	180.0
CC321	CC3151	CC3151	OB2	0.84	3	0.0
OB2	CC3151	CC3153	HCA2	0.14	3	0.0
CC321	*CC3151*	*CC3153*	*HCA2*	*0.14*	*3*	*0.0*
CC321	CC3151	OB2	BO3	0.05	1	180.0
CC321	CC3151	OB2	BO3	0.14	2	0.0
CC321	CC3151	OB2	BO3	0.70	3	180.0
CC3153	*CC3151*	*CC321*	*OC311*	*0.01*	*1*	*0.0*

Table 2. (continued)

CC3153	*CC3151*	*CC321*	*OC311*	*0.14*	*2*	*0.0*
CC3153	*CC3151*	*CC321*	*OC311*	*0.70*	*3*	*180.0*
CC3151	*CC3151*	*CC321*	*HCA2*	*0.35*	*3*	*0.0*
OB2	CC3151	CC321	HCA2	0.30	3	0.0
HCA2	*CC321*	*CC3151*	*CC3153*	*0.20*	*3*	*0.0*

The re-optimized parameters from carbohydrate.prm are given in italic.

Improper Torsional Parameters

Atom I	Atom J	Atom K	Atom L	Kφ (kcal/(mol*rad^{-2})	$\varphi_0(^0)$
BO3	OB2	OB1	OB2	80.00	0.0

Non−Bonded Parameters

LJ parameters

Atom I	ε_{ij}	R_{min_i} (Å)
HOB3	−0.0460	0.2245
OB1	−0.1921	1.7650
BO3	−0.0980	1.7000
OB2	−0.1000	1.6500

Charge parameters

Atom I	Charge
BO3	0.70
OB2	−0.44
OB1	−0.70
HOB3	0.42

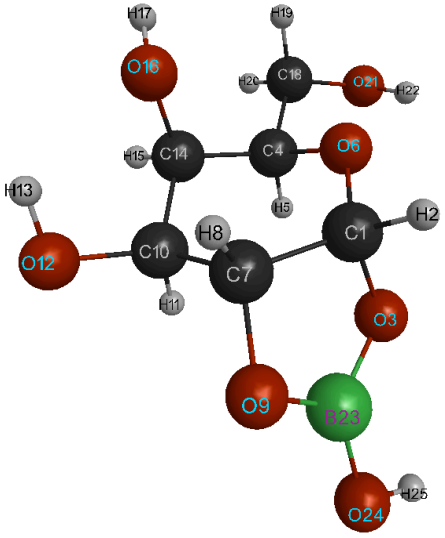
* Charge is given in units of electron charge.

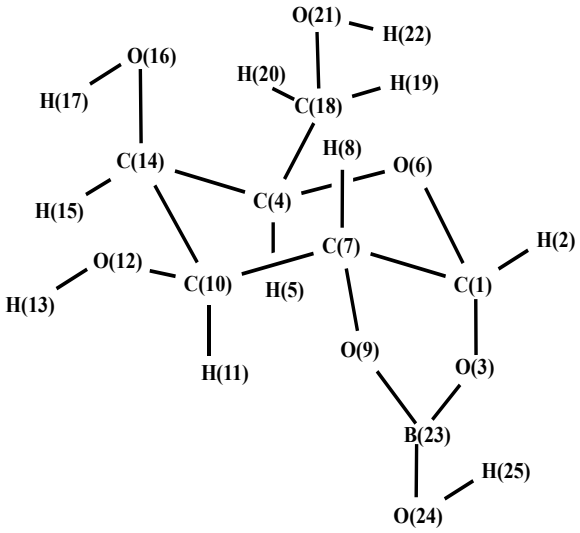
Table 3. Solute−water Interaction Energies and Distances for Borated Carbohydrates

Geometry	*Ab initio*		CHARMM38		Difference	
	E_{min*}	R_{min*}	E_{min}	R_{min}	ΔE	ΔR
ai	−7.57	1.70	−6.85	1.75	0.72	0.05
aii	−5.50	1.80	−5.54	1.75	0.05	−0.05
bi	−4.97	2.05	−5.72	2.00	−0.75	−0.05
bii	−4.10	2.05	−4.63	2.00	−0.53	−0.05
biii	−3.19	2.05	−3.58	2.00	−0.39	−0.05
ci	−4.53	1.80	−3.66	1.75	0.87	−0.05
cii	−3.98	1.80	−3.10	1.75	0.87	−0.05
di	−3.84	1.80	−2.95	1.75	0.89	−0.05
dii	−3.98	1.80	−3.10	1.75	0.87	−0.05
ei	−5.17	2.05	−5.71	1.75	−0.55	−0.30
eii	−3.99	2.05	−3.71	1.75	0.28	−0.30
fi	−5.78	2.30	−6.75	2.25	−0.97	−0.05
Average					0.11	−0.08
Standard deviation					0.72	0.11

* Energies are in kcal/mol and distances are in Å. $R_{min*} = R_{QM} - 0.2$ Å and $E_{min*} = 1.16 * E_{QM}$. See Figure 2 for interaction orientations.

Table 4. Structural Differences Between CHARMM and Ab Initio Results for **1,2−aGal**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	2	1.12	1.10	0.02
1	7	1.56	1.54	0.02
1	6	1.41	1.39	0.01
1	3	1.43	1.44	−0.01
3	23	1.38	1.40	−0.02
4	18	1.51	1.52	0.00
4	6	1.44	1.44	0.00
4	5	1.12	1.10	0.02
4	14	1.52	1.52	0.00
7	8	1.11	1.09	0.02
7	10	1.56	1.53	0.03
7	9	1.44	1.44	0.00
9	23	1.38	1.38	0.00
10	14	1.56	1.52	0.04
10	12	1.42	1.42	0.00
10	11	1.11	1.10	0.02
12	13	0.96	0.98	−0.01
14	15	1.12	1.10	0.01
14	16	1.43	1.43	−0.01
16	17	0.96	0.97	−0.01
18	21	1.43	1.42	0.01

Table 4. (continued)

18	20	1.11	1.09	0.02
18	19	1.11	1.10	0.01
21	22	0.97	0.98	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
			RMS	0.01

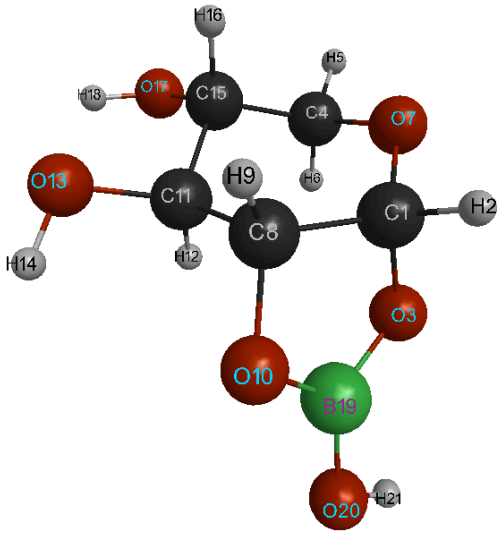
Bond Angles (in degrees)

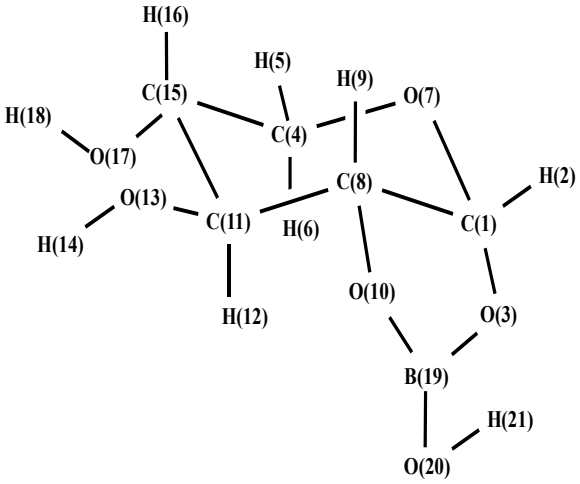
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	7	10	112.4	113.8	−1.4
1	6	4	114.0	115.3	−1.3
1	3	23	99.8	104.8	−5.0
1	7	8	108.9	111.6	−2.7
1	7	9	100.9	103.1	−2.2
2	1	3	108.5	108.4	0.1
2	1	6	107.7	105.0	2.7
2	1	7	108.8	111.6	−2.8
3	23	24	122.7	124.2	−1.5
3	23	9	112.0	113.6	−1.6
3	1	7	101.6	103.1	−1.5
3	1	6	113.1	112.3	0.8
4	14	15	109.1	110.1	−1.0
4	14	10	108.0	110.0	−2.0
4	14	16	112.4	112.6	−0.2
4	18	19	110.2	108.8	1.4
4	18	20	110.0	110.2	−0.2
4	18	21	111.2	110.9	0.3
5	4	18	108.2	108.8	−0.6
5	4	14	108.6	109.4	−0.8
5	4	6	109.1	109.5	−0.4
6	4	18	107.3	104.1	3.2
6	4	14	110.8	109.9	0.9
6	1	7	116.9	116.4	0.5
7	10	12	111.9	111.7	0.2
7	10	14	111.1	110.7	0.4
7	10	11	106.8	108.2	−1.4
7	9	23	106.0	105.3	0.7

Table 4. (continued)

8	7	9	111.6	109.0	2.6
8	7	10	109.3	109.8	−0.5
9	7	10	113.4	109.3	4.1
9	23	24	124.9	122.2	2.7
10	12	13	108.4	104.7	3.7
10	14	16	111.4	104.6	6.8
10	14	15	107.3	108.7	−1.4
11	10	12	107.4	106.3	1.1
11	10	14	107.2	109.0	−1.8
12	10	14	112.2	110.9	1.3
14	16	17	110.6	109.0	1.6
14	4	18	112.8	115.0	−2.2
15	14	16	108.5	110.7	−2.2
18	21	22	108.7	105.2	3.5
19	18	20	108.1	108.8	−0.7
19	18	21	109.3	111.2	−1.9
20	18	21	108.0	106.9	1.1
23	24	25	110.6	110.5	0.1
				RMS	2.2

Table 5. Structural Differences Between CHARMM and Ab Initio Results for **1,2−aXyl**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	8	1.56	1.54	0.02
1	3	1.43	1.45	−0.01
1	2	1.12	1.10	0.02
1	7	1.41	1.39	0.02
3	19	1.38	1.39	−0.01
4	15	1.51	1.52	−0.01
4	5	1.11	1.09	0.02
4	6	1.11	1.10	0.02
4	7	1.43	1.44	−0.01
8	10	1.44	1.45	0.00
8	9	1.11	1.09	0.02
8	11	1.56	1.52	0.04
10	19	1.38	1.39	−0.01
11	15	1.56	1.51	0.05
11	13	1.42	1.43	0.00
11	12	1.12	1.10	0.02
13	14	0.96	0.98	−0.01
15	17	1.42	1.42	0.00
15	16	1.11	1.10	0.01
17	18	0.97	0.98	−0.01
19	20	1.36	1.36	0.00

Table 5. (continued)

20	21	0.96	0.97	−0.02
			RMS	0.02

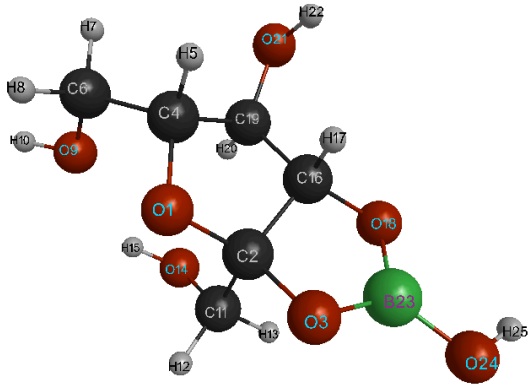
Bond Angles (in degrees)

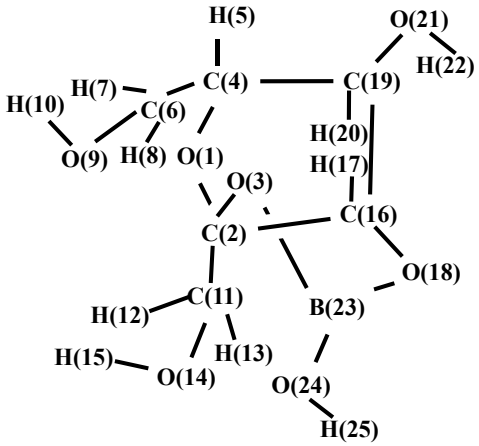
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	3	19	99.8	105.3	−5.5
1	8	11	112.6	112.4	0.2
1	8	10	101.3	103.9	−2.6
1	7	4	113.8	115.3	−1.5
1	8	9	109.3	112.3	−3.0
2	1	7	108.1	105.0	3.1
2	1	8	109.2	112.0	−2.8
2	1	3	108.6	108.0	0.6
3	19	20	122.8	124.6	−1.8
3	1	8	101.6	102.4	−0.8
3	19	10	112.0	113.8	−1.8
3	1	7	113.5	113.3	0.2
4	15	16	109.7	109.3	0.4
4	15	17	110.0	108.1	1.9
4	15	11	107.9	107.9	0.0
5	4	15	109.5	111.1	−1.6
5	4	6	107.4	109.0	−1.6
5	4	7	108.4	105.7	2.7
6	4	7	110.6	110.9	−0.3
6	4	15	111.0	110.3	0.7
7	1	8	115.5	116.1	−0.6
7	4	15	110.0	109.6	0.4
8	11	12	107.2	108.6	−1.4
8	11	13	111.2	112.1	−0.9
8	11	15	111.3	111.4	−0.1
8	10	19	105.6	104.4	1.2
9	8	11	109.6	109.9	−0.3
9	8	10	111.8	109.3	2.5
10	19	20	124.8	121.6	3.2
10	8	11	112.1	108.8	3.3
11	13	14	109.3	107.0	2.3
11	15	16	109.2	109.5	−0.3
11	15	17	111.2	110.7	0.5

Table 5. (continued)

12	11	13	108.3	110.0	−1.7
12	11	15	106.2	107.1	−0.9
13	11	15	112.4	107.5	4.9
15	17	18	107.5	105.4	2.1
16	15	17	108.8	111.3	−2.5
19	20	21	110.6	110.5	0.1
				RMS	2.1

Table 6. Structural Differences Between CHARMM and Ab Initio Results for **1,2−aFrucfur**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	2	1.43	1.42	0.01
1	4	1.45	1.46	−0.01
2	16	1.48	1.50	−0.02
2	11	1.54	1.53	0.00
2	3	1.49	1.45	0.04
3	23	1.39	1.40	−0.01
4	6	1.51	1.51	0.01
4	19	1.55	1.57	−0.02
4	5	1.12	1.10	0.01
6	8	1.12	1.10	0.02
6	7	1.12	1.10	0.02
6	9	1.43	1.43	0.00
9	10	0.97	0.97	−0.01
11	14	1.43	1.41	0.02
11	13	1.11	1.09	0.02
11	12	1.11	1.10	0.01
14	15	0.98	0.98	0.00
16	19	1.52	1.51	0.01
16	18	1.39	1.44	−0.05
16	17	1.13	1.11	0.02
18	23	1.39	1.42	−0.02
19	21	1.41	1.42	−0.01

Table 6. (continued)

19	20	1.11	1.09	0.02
21	22	0.96	0.98	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
			RMS	0.02

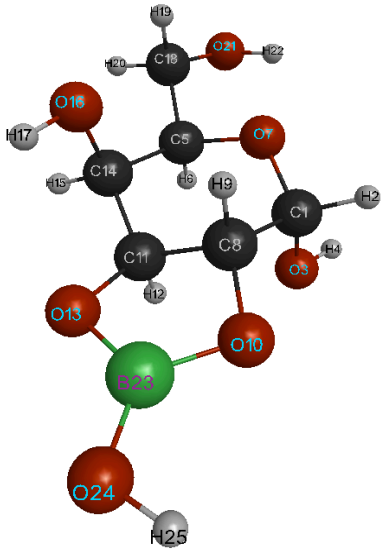
Bond Angles (in degrees)

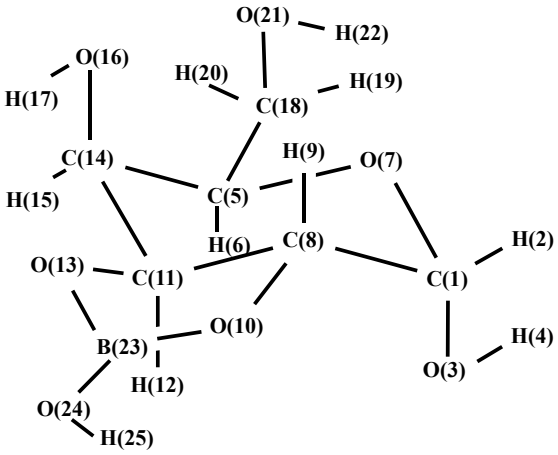
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	4	6	111.4	110.0	1.4
1	2	16	103.2	103.7	−0.5
1	4	19	107.7	108.4	−0.7
1	2	3	121.8	119.5	2.3
1	2	11	108.1	109.3	−1.2
1	4	5	104.9	108.4	−3.5
2	11	14	116.6	115.2	1.4
2	16	17	107.5	105.3	2.2
2	1	4	107.9	105.9	2.0
2	16	19	104.2	103.0	1.2
2	16	18	103.2	102.8	0.4
2	3	23	97.7	100.7	−3.0
2	11	12	108.1	105.0	3.1
2	11	13	109.5	109.6	−0.1
3	2	11	105.0	104.7	0.3
3	23	24	122.9	121.7	1.2
3	23	18	116.2	114.7	1.5
3	2	16	98.9	100.0	−1.1
4	19	20	113.8	112.0	1.8
4	19	16	96.6	93.7	2.9
4	6	9	112.7	109.7	3.0
4	6	7	109.0	107.6	1.4
4	6	8	108.9	108.4	0.5
4	19	21	109.0	111.4	−2.4
5	4	19	109.4	106.5	2.9
5	4	6	107.6	106.5	1.1
6	4	19	115.4	116.8	−1.4
6	9	10	110.1	108.6	1.5
7	6	9	109.0	111.0	−2.0
7	6	8	108.2	109.0	−0.8

Table 6. (continued)

8	6	9	109.0	111.0	−2.0
11	14	15	109.9	108.8	1.1
11	2	16	120.8	120.3	0.5
12	11	13	106.6	107.9	−1.3
12	11	14	108.9	112.8	−3.9
13	11	14	106.8	106.3	0.5
16	18	23	99.2	99.2	0.0
16	19	20	114.2	114.0	0.2
16	19	21	112.0	118.2	−6.2
17	16	18	104.4	105.8	−1.4
17	16	19	104.1	104.3	−0.2
18	23	24	121.0	123.6	−2.6
18	16	19	131.7	133.1	−1.4
19	21	22	109.6	107.9	1.7
20	19	21	110.6	107.1	3.5
23	24	25	110.3	110.4	−0.1
				RMS	2.1

Table 7. Structural Differences Between CHARMM and Ab Initio Results for **2,3−aGal**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	2	1.11	1.10	0.02
1	3	1.40	1.40	0.00
1	8	1.57	1.51	0.06
1	7	1.41	1.43	−0.02
3	4	0.96	0.98	−0.02
5	14	1.54	1.54	0.00
5	6	1.12	1.10	0.02
5	18	1.51	1.52	0.00
5	7	1.45	1.45	0.00
8	11	1.54	1.51	0.03
8	10	1.44	1.44	0.00
8	9	1.11	1.10	0.02
10	23	1.38	1.40	−0.02
11	14	1.55	1.51	0.04
11	13	1.44	1.45	0.00
11	12	1.11	1.10	0.02
13	23	1.38	1.40	−0.02
14	15	1.11	1.10	0.01
14	16	1.42	1.42	0.00
16	17	0.97	0.98	−0.01
18	21	1.43	1.42	0.00

Table 7. (continued)

18	20	1.11	1.09	0.02
18	19	1.11	1.10	0.02
21	22	0.96	0.97	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
			RMS	0.02

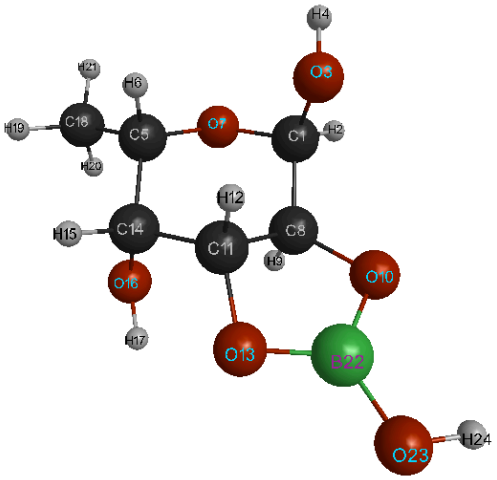
Bond Angles (in degrees)

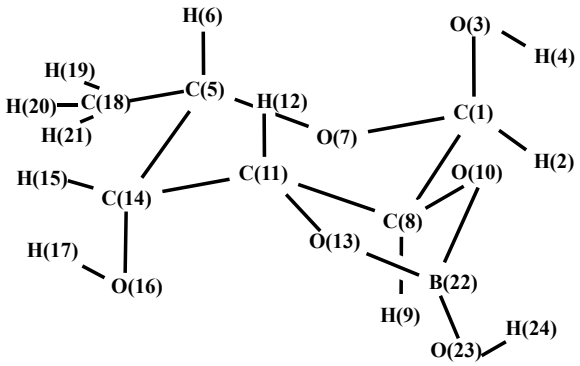
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	8	9	106.2	106.5	−0.3
1	3	4	107.6	107.4	0.2
1	8	10	122.1	119.9	2.2
1	8	11	110.3	109.1	1.2
1	7	5	114.0	116.2	−2.2
2	1	3	108.8	111.8	−3.0
2	1	7	109.0	104.4	4.6
2	1	8	107.6	113.3	−5.7
3	1	7	113.8	113.4	0.4
3	1	8	110.6	109.3	1.3
5	14	11	108.1	105.6	2.5
5	18	20	109.8	110.0	−0.2
5	18	21	111.4	110.5	0.9
5	18	19	110.1	108.8	1.3
5	14	16	112.4	109.7	2.7
5	14	15	109.0	108.2	0.8
6	5	14	107.1	108.1	−1.0
6	5	18	107.3	109.2	−1.9
6	5	7	108.3	108.5	−0.2
7	5	14	114.4	113.7	0.7
7	1	8	106.9	104.5	2.4
7	5	18	107.4	103.9	3.5
8	10	23	103.2	102.6	0.6
8	11	12	108.1	110.2	−2.1
8	11	13	101.8	102.4	−0.6
8	11	14	107.1	109.6	−2.5
9	8	10	107.8	108.6	−0.8
9	8	11	109.1	109.6	−0.5
10	23	13	110.8	114.1	−3.3

Table 7. (continued)

10	8	11	100.8	102.9	−2.1
10	23	24	123.5	124.3	−0.8
11	13	23	106.9	102.6	4.3
11	14	16	110.5	111.5	−1.0
11	14	15	108.6	111.1	−2.5
12	11	13	106.4	108.5	−2.1
12	11	14	106.9	109.6	−2.7
13	23	24	125.4	121.6	3.8
13	11	14	125.5	116.3	9.2
14	5	18	112.0	113.3	−1.3
14	16	17	109.5	105.8	3.7
15	14	16	108.1	110.5	−2.4
18	21	22	108.5	104.9	3.6
19	18	20	107.8	108.6	−0.8
19	18	21	109.4	111.9	−2.5
20	18	21	108.3	106.9	1.4
23	24	25	110.1	110.3	−0.2
				RMS	2.7

Table 8. Structural Differences Between CHARMM and Ab Initio Results for **2,3−aFuc**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	3	1.40	1.41	−0.01
1	7	1.41	1.43	−0.02
1	2	1.11	1.10	0.02
1	8	1.57	1.51	0.06
3	4	0.96	0.98	−0.02
5	7	1.44	1.45	−0.01
5	6	1.12	1.10	0.02
5	18	1.54	1.52	0.03
5	14	1.54	1.54	0.00
8	11	1.54	1.51	0.03
8	10	1.44	1.44	0.00
8	9	1.11	1.10	0.02
10	22	1.38	1.40	−0.02
11	14	1.55	1.51	0.05
11	13	1.44	1.45	0.00
11	12	1.11	1.10	0.02
13	22	1.38	1.40	−0.02
14	15	1.11	1.10	0.01
14	16	1.42	1.42	0.00
16	17	0.96	0.98	−0.01
18	21	1.11	1.09	0.02
18	20	1.11	1.09	0.02
18	19	1.11	1.09	0.01

Table 8. (continued)

22	23	1.36	1.36	0.00
23	24	0.96	0.97	−0.02
			RMS	0.02

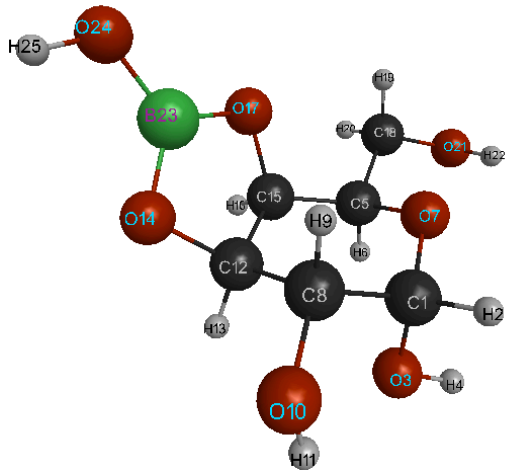
Bond Angles (in degrees)

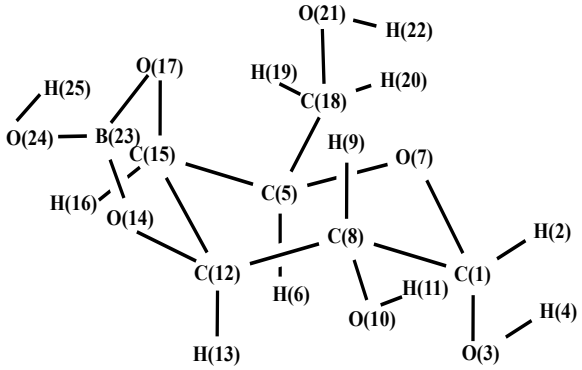
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	7	5	114.0	116.3	−2.3
1	8	10	122.1	120.1	2.0
1	8	9	106.2	106.5	−0.3
1	8	11	110.4	109.0	1.4
1	3	4	107.2	107.0	0.2
2	1	3	108.7	111.5	−2.8
2	1	7	108.8	104.4	4.4
2	1	8	107.6	113.1	−5.5
3	1	8	110.6	109.2	1.4
3	1	7	113.8	113.4	0.4
5	14	16	113.0	109.8	3.2
5	14	15	108.8	108.0	0.8
5	14	11	107.6	105.9	1.7
5	18	19	110.7	110.3	0.4
5	18	21	110.2	109.9	0.3
5	18	20	109.8	109.4	0.4
6	5	18	106.9	110.5	−3.6
6	5	7	108.0	108.4	−0.4
6	5	14	107.4	106.7	0.7
7	5	14	114.9	113.5	1.4
7	5	18	106.4	105.4	1.0
7	1	8	107.2	105.0	2.2
8	10	22	103.1	102.5	0.6
8	11	12	108.0	110.1	−2.1
8	11	13	101.9	102.4	−0.5
8	11	14	106.9	109.5	−2.6
9	8	10	107.8	108.6	−0.8
9	8	11	109.1	109.6	−0.5
10	22	13	110.8	114.2	−3.4
10	8	11	100.8	102.9	−2.1
10	22	23	123.5	124.2	−0.7
11	13	22	106.9	102.5	4.4

Table 8. (continued)

11	14	16	110.4	111.4	−1.0
11	14	15	108.8	111.1	−2.3
12	11	13	106.4	108.4	−2.0
12	11	14	106.9	109.7	−2.8
13	22	23	125.4	121.6	3.8
13	11	14	125.7	116.4	9.3
14	16	17	109.5	105.7	3.8
14	5	18	113.0	112.4	0.6
15	14	16	108.2	110.5	−2.3
19	18	20	108.7	108.9	−0.2
19	18	21	109.0	109.0	0.0
20	18	21	108.3	109.4	−1.1
22	23	24	110.1	110.2	−0.1
				RMS	2.6

Table 9. Structural Differences Between CHARMM and Ab Initio Results for **3,4−aGal**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	8	1.58	1.52	0.06
1	3	1.40	1.42	−0.02
1	7	1.40	1.41	−0.01
1	2	1.11	1.10	0.02
3	4	0.96	0.98	−0.01
5	15	1.52	1.52	0.01
5	7	1.44	1.44	0.00
5	6	1.12	1.10	0.02
5	18	1.52	1.52	0.00
8	12	1.56	1.53	0.03
8	10	1.43	1.42	0.01
8	9	1.11	1.09	0.02
10	11	0.96	0.98	−0.01
12	15	1.55	1.54	0.02
12	14	1.44	1.44	0.00
12	13	1.11	1.09	0.02
14	23	1.37	1.39	−0.02
15	17	1.44	1.45	0.00
15	16	1.11	1.10	0.01
17	23	1.38	1.39	−0.01
18	21	1.43	1.42	0.00
18	20	1.11	1.09	0.02
18	19	1.11	1.10	0.02

Table 9. (continued)

21	22	0.96	0.97	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
			RMS	0.02

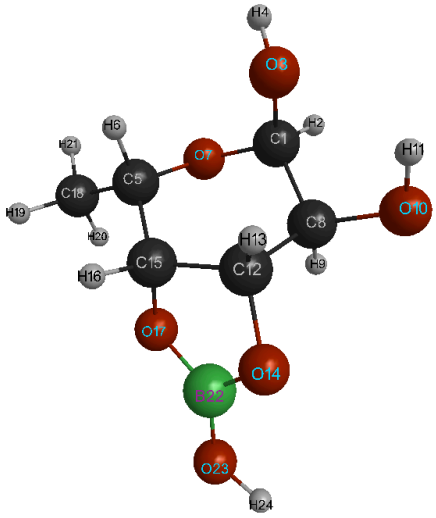
Bond Angles (in degrees)

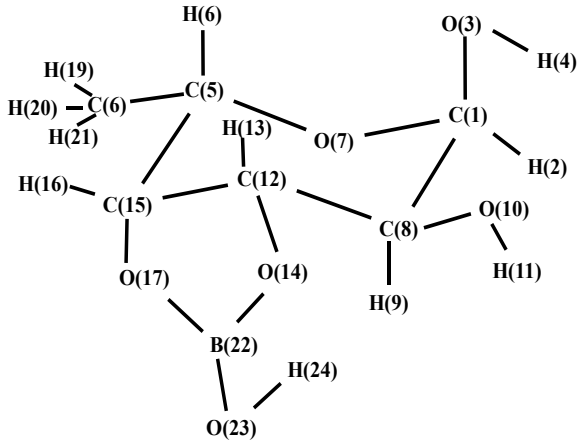
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	3	4	107.2	108.1	−0.9
1	8	10	113.4	110.8	2.6
1	8	12	111.8	110.6	1.2
1	7	5	112.9	113.3	−0.4
1	8	9	105.7	107.8	−2.1
2	1	8	107.4	110.7	−3.3
2	1	7	108.3	105.4	2.9
2	1	3	108.7	111.2	−2.5
3	1	7	115.1	113.0	2.1
3	1	8	109.8	105.5	4.3
5	18	21	111.5	110.3	1.2
5	18	20	109.7	110.1	−0.4
5	18	19	110.1	108.9	1.2
5	15	12	115.7	115.4	0.3
5	15	16	108.8	108.2	0.6
5	15	17	113.3	112.2	1.1
6	5	15	107.5	108.1	−0.6
6	5	18	106.8	109.0	−2.2
6	5	7	109.2	108.7	0.5
7	5	15	114.1	112.6	1.5
7	1	8	107.3	111.3	−4.0
7	5	18	107.5	104.7	2.8
8	10	11	109.0	105.5	3.5
8	12	13	108.6	110.0	−1.4
8	12	14	115.9	109.1	6.8
8	12	15	112.0	112.9	−0.9
9	8	10	107.1	106.5	0.6
9	8	12	107.0	108.6	−1.6
10	8	12	111.4	112.3	−0.9
12	14	23	103.7	105.0	−1.3
12	15	16	108.1	109.9	−1.8

Table 9. (continued)

12	15	17	103.2	103.0	0.2
13	12	14	111.0	108.9	2.1
13	12	15	109.3	112.6	−3.3
14	12	15	99.7	103.0	−3.3
14	23	24	123.7	124.4	−0.7
14	23	17	110.1	113.6	−3.5
15	5	18	111.5	113.6	−2.1
15	17	23	106.6	104.9	1.7
16	15	17	107.2	107.8	−0.6
17	23	24	125.8	122.0	3.8
18	21	22	108.7	105.1	3.6
19	18	20	107.7	108.6	−0.9
19	18	21	109.5	112.0	−2.5
20	18	21	108.3	106.9	1.4
23	24	25	110.0	110.2	−0.2
				RMS	2.4

Table 10. Structural Differences Between CHARMM and Ab Initio Results for **3,4−aFuc**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	7	1.40	1.41	−0.01
1	2	1.11	1.10	0.02
1	8	1.58	1.52	0.06
1	3	1.40	1.42	−0.02
3	4	0.96	0.98	−0.01
5	15	1.53	1.52	0.01
5	7	1.43	1.44	−0.01
5	6	1.12	1.10	0.02
5	18	1.54	1.52	0.02
8	12	1.56	1.53	0.03
8	10	1.43	1.42	0.01
8	9	1.11	1.09	0.02
10	11	0.96	0.98	−0.01
12	15	1.55	1.54	0.02
12	14	1.44	1.45	−0.01
12	13	1.11	1.09	0.02
14	22	1.37	1.39	−0.02
15	17	1.44	1.45	0.00
15	16	1.11	1.10	0.01
17	22	1.38	1.39	−0.01
18	21	1.11	1.09	0.02
18	20	1.11	1.09	0.02

Table 10. (continued)

18	19	1.11	1.09	0.02
22	23	1.36	1.36	0.00
23	24	0.96	0.97	−0.02
			RMS	0.02

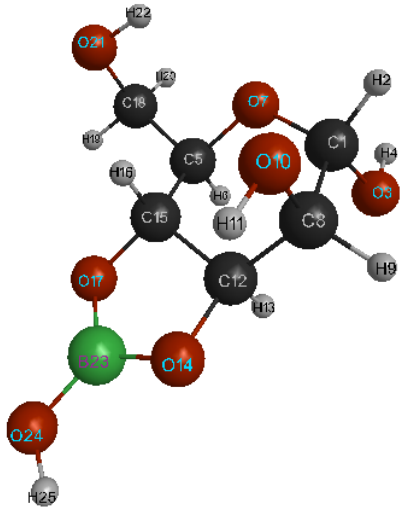
Bond Angles (in degrees)

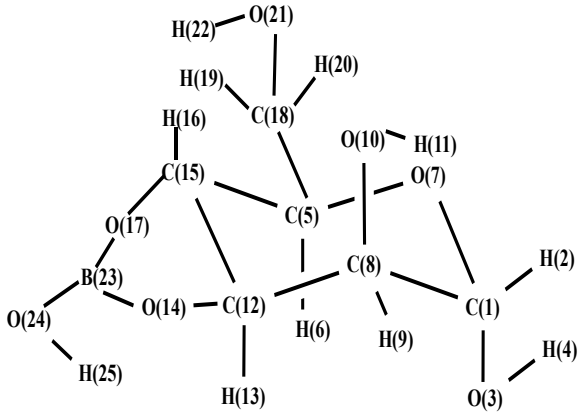
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	8	10	113.4	110.8	2.6
1	8	9	105.6	107.8	−2.2
1	7	5	112.9	113.5	−0.6
1	8	12	111.9	110.5	1.4
1	3	4	106.9	107.7	−0.8
2	1	7	108.1	105.3	2.8
2	1	3	108.6	110.8	−2.2
2	1	8	107.2	110.5	−3.3
3	1	8	109.9	105.3	4.6
3	1	7	115.0	113.1	1.9
5	18	21	110.4	109.8	0.6
5	18	20	109.8	109.5	0.3
5	18	19	110.7	110.3	0.4
5	15	12	115.2	115.6	−0.4
5	15	16	108.7	108.0	0.7
5	15	17	114.0	112.4	1.6
6	5	15	107.7	106.7	1.0
6	5	18	106.5	110.4	−3.9
6	5	7	108.9	108.5	0.4
7	5	15	114.7	112.3	2.4
7	1	8	107.7	111.9	−4.2
7	5	18	106.3	106.3	0.0
8	10	11	108.8	105.3	3.5
8	12	13	108.7	110.0	−1.3
8	12	14	115.9	109.2	6.7
8	12	15	111.8	112.7	−0.9
9	8	10	107.2	106.6	0.6
9	8	12	106.9	108.6	−1.7
10	8	12	111.4	112.3	−0.9
12	14	22	103.6	104.8	−1.2
12	15	16	108.3	109.9	−1.6

Table 10. (continued)

12	15	17	103.0	103.0	0.0
13	12	14	111.1	108.9	2.2
13	12	15	109.4	112.7	−3.3
14	12	15	99.8	103.1	−3.3
14	22	23	123.6	124.2	−0.6
14	22	17	110.2	113.7	−3.5
15	5	18	112.4	112.6	−0.2
15	17	22	106.6	104.9	1.7
16	15	17	107.2	107.8	−0.6
17	22	23	125.8	122.0	3.8
19	18	20	108.5	108.8	−0.3
19	18	21	109.0	109.0	0.0
20	18	21	108.4	109.4	−1.0
22	23	24	109.9	110.1	−0.2
				RMS	2.3

Table 11. Structural Differences Between CHARMM and Ab Initio Results for **3,4−aMan**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	8	1.60	1.54	0.06
1	3	1.40	1.41	−0.01
1	7	1.41	1.43	−0.01
1	2	1.11	1.10	0.02
3	4	0.96	0.98	−0.01
5	15	1.51	1.51	0.00
5	7	1.45	1.46	−0.01
5	6	1.12	1.10	0.02
5	18	1.51	1.52	0.00
8	12	1.55	1.51	0.04
8	10	1.43	1.42	0.01
8	9	1.11	1.10	0.01
10	11	0.96	0.98	−0.01
12	15	1.54	1.51	0.03
12	14	1.44	1.45	−0.01
12	13	1.12	1.10	0.02
14	23	1.38	1.41	−0.03
15	17	1.44	1.44	0.00
15	16	1.11	1.09	0.02
17	23	1.38	1.39	−0.01

Table 11. (continued)

18	21	1.43	1.42	0.01
18	20	1.11	1.10	0.01
18	19	1.11	1.09	0.02
21	22	0.96	0.98	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
			RMS	0.02

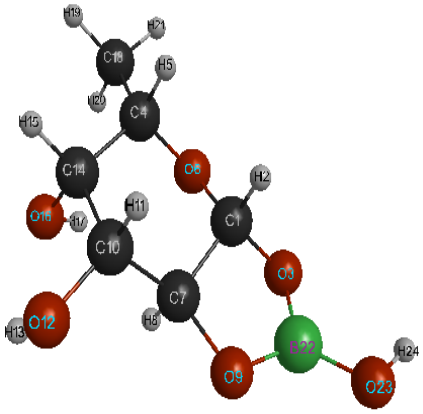
Bond Angles (in degrees)

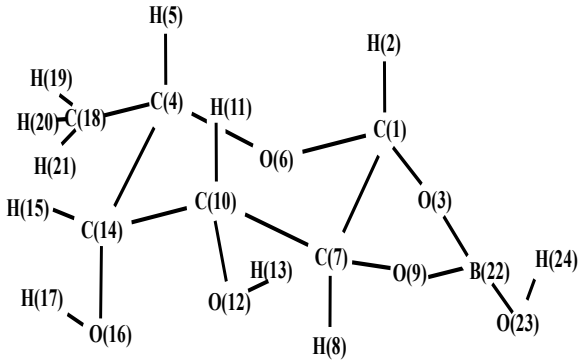
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	3	4	107.1	106.7	0.4
1	8	10	113.1	107.5	5.6
1	8	12	107.1	105.3	1.8
1	7	5	114.4	117.1	−2.7
1	8	9	108.4	108.1	0.3
2	1	8	106.8	109.8	−3.0
2	1	7	108.2	103.5	4.7
2	1	3	108.2	112.4	−4.2
3	1	7	111.0	111.3	−0.3
3	1	8	108.6	105.4	3.2
5	18	21	111.7	111.1	0.6
5	18	20	109.8	109.0	0.8
5	18	19	109.8	109.4	0.4
5	15	12	109.7	109.6	0.1
5	15	16	109.9	107.5	2.4
5	15	17	115.6	118.1	−2.5
6	5	15	109.2	111.5	−2.3
6	5	18	108.0	110.8	−2.8
6	5	7	110.1	110.7	−0.6
7	5	15	108.5	104.4	4.1
7	1	8	113.8	114.6	−0.8
7	5	18	108.1	104.8	3.3
8	10	11	109.7	106.1	3.6
8	12	13	106.7	108.9	−2.2
8	12	14	127.8	116.8	11.0
8	12	15	106.8	110.0	−3.2
9	8	10	108.0	111.3	−3.3
9	8	12	109.0	111.9	−2.9

Table 11. (continued)

10	8	12	111.1	112.3	−1.2
12	14	23	104.1	102.1	2.0
12	15	16	109.9	110.4	−0.5
12	15	17	102.8	102.5	0.3
13	12	14	107.1	108.3	−1.2
13	12	15	108.4	110.4	−2.0
14	12	15	98.9	102.3	−3.4
14	23	24	123.5	123.7	−0.2
14	23	17	110.9	114.1	−3.2
15	5	18	112.9	114.3	−1.4
15	17	23	105.6	102.5	3.1
16	15	17	108.5	108.6	−0.1
17	23	24	125.4	122.2	3.2
18	21	22	108.6	105.3	3.3
19	18	20	108.0	108.7	−0.7
19	18	21	108.6	107.1	1.5
20	18	21	108.9	111.4	−2.5
23	24	25	110.4	110.5	−0.1
				RMS	3.0

Table 12. Structural Differences Between CHARMM and Ab Initio Results for **1,2−bFuc**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	7	1.54	1.51	0.03
1	6	1.40	1.41	0.00
1	3	1.43	1.42	0.01
1	2	1.11	1.11	0.01
3	22	1.38	1.41	−0.03
4	18	1.54	1.52	0.03
4	5	1.12	1.10	0.01
4	6	1.44	1.45	−0.02
4	14	1.53	1.54	0.00
7	8	1.12	1.10	0.02
7	10	1.56	1.51	0.05
7	9	1.44	1.44	0.00
9	22	1.38	1.39	−0.01
10	14	1.57	1.55	0.02
10	12	1.42	1.41	0.00
10	11	1.11	1.10	0.01
12	13	0.97	0.98	−0.01
14	15	1.11	1.10	0.02
14	16	1.42	1.44	−0.01
16	17	0.97	0.98	−0.01
18	21	1.11	1.09	0.02
18	20	1.11	1.09	0.02
18	19	1.11	1.09	0.02
22	23	1.36	1.36	0.00

Table 12. (continued)

23	24	0.96	0.97	−0.02
			RMS	0.02

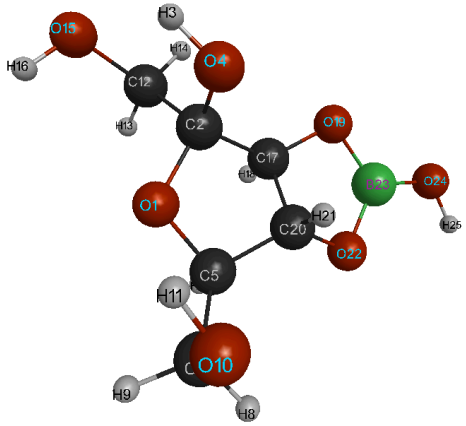
Bond Angles (in degrees)

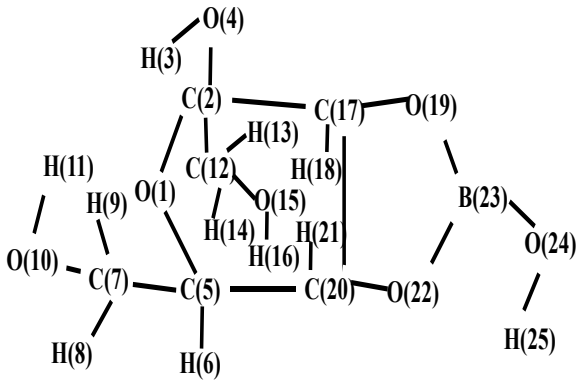
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	3	22	100.2	102.4	−2.2
1	7	10	107.7	109.2	−1.5
1	7	9	101.7	102.2	−0.5
1	6	4	110.3	106.1	4.2
1	7	8	107.0	108.3	−1.3
2	1	3	109.2	108.2	1.0
2	1	7	111.4	111.0	0.4
2	1	6	111.7	108.9	2.8
3	1	6	114.7	114.3	0.4
3	22	23	122.5	123.6	−1.1
3	22	9	112.3	114.0	−1.7
3	1	7	101.4	103.3	−1.9
4	14	16	110.5	110.3	0.2
4	14	15	109.5	109.9	−0.4
4	14	10	109.8	114.2	−4.4
4	18	19	110.6	110.0	0.6
4	18	20	110.8	110.8	0.0
4	18	21	110.1	109.6	0.5
5	4	18	107.0	109.6	−2.6
5	4	14	108.2	108.5	−0.3
5	4	6	109.1	108.5	0.6
6	4	18	107.4	107.0	0.4
6	4	14	111.2	109.4	1.8
6	1	7	107.9	111.2	−3.3
7	10	12	112.1	114.4	−2.3
7	10	14	107.5	105.0	2.5
7	10	11	108.2	110.1	−1.9
7	9	22	105.2	102.5	2.7
8	7	9	105.1	107.7	−2.6
8	7	10	106.1	108.8	−2.7
9	7	10	127.8	119.9	7.9
9	22	23	124.8	122.5	2.3
10	12	13	108.0	104.8	3.2

Table 12. (continued)

10	14	16	111.1	107.8	3.3
10	14	15	108.1	108.5	−0.4
11	10	12	107.7	106.8	0.9
11	10	14	109.0	110.7	−1.7
12	10	14	112.3	109.9	2.4
14	16	17	108.5	106.2	2.3
14	4	18	113.9	113.7	0.2
15	14	16	107.8	105.8	2.0
19	18	20	108.6	108.4	0.2
19	18	21	108.7	109.0	−0.3
20	18	21	108.0	109.0	−1.0
22	23	24	110.5	110.6	−0.1
				RMS	2.3

Table 13. Structural Differences Between CHARMM and Ab Initio Results for **2,3−aFrucfur**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	5	1.46	1.48	−0.03
1	2	1.45	1.47	−0.02
2	17	1.53	1.51	0.02
2	4	1.39	1.40	0.00
2	12	1.53	1.52	0.02
3	4	0.97	0.98	−0.02
5	7	1.51	1.52	−0.01
5	6	1.11	1.10	0.01
5	20	1.51	1.51	−0.01
7	8	1.11	1.10	0.02
7	10	1.43	1.42	0.01
7	9	1.11	1.10	0.01
10	11	0.97	0.98	−0.01
12	15	1.44	1.43	0.01
12	14	1.12	1.09	0.02
12	13	1.11	1.10	0.02
15	16	0.97	0.98	−0.01
17	20	1.46	1.50	−0.04
17	19	1.39	1.43	−0.04
17	18	1.12	1.10	0.01

Table 13. (continued)

19	23	1.40	1.40	0.00
20	22	1.39	1.44	−0.06
20	21	1.12	1.10	0.03
22	23	1.40	1.41	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
			RMS	0.02

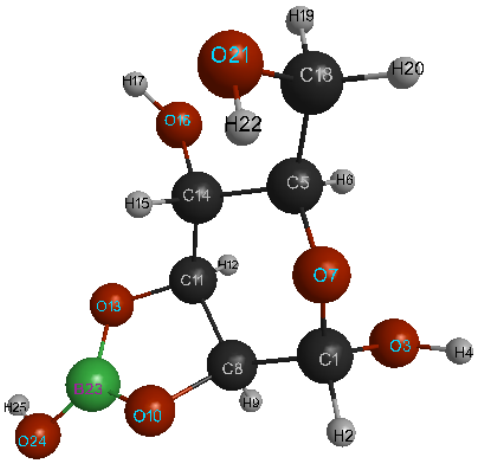
Bond Angles (in degrees)

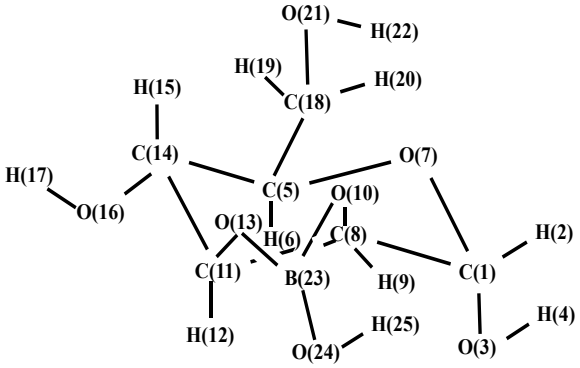
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	2	4	114.4	110.9	3.5
1	2	17	101.1	98.2	2.9
1	5	6	107.8	108.8	−1.0
1	5	20	100.2	100.4	−0.2
1	5	7	110.4	106.3	4.1
1	2	12	106.0	105.5	0.5
2	12	13	109.8	110.2	−0.4
2	17	18	108.9	106.4	2.5
2	12	14	107.5	109.7	−2.2
2	17	19	125.7	127.7	−2.0
2	17	20	99.3	100.4	−1.1
2	12	15	115.3	108.6	6.7
2	1	5	110.6	111.9	−1.3
2	4	3	104.3	104.7	−0.4
4	2	17	113.6	111.7	1.9
4	2	12	108.6	110.3	−1.7
5	20	21	110.3	108.6	1.7
5	20	22	122.3	125.9	−3.6
5	20	17	100.9	101.0	−0.1
5	7	9	109.9	108.7	1.2
5	7	10	111.8	111.2	0.6
5	7	8	109.7	109.9	−0.2
6	5	20	111.4	112.3	−0.9
6	5	7	110.3	111.7	−1.4
7	10	11	109.4	105.9	3.5
7	5	20	116.2	116.3	−0.1
8	7	9	108.0	108.3	−0.3

Table 13. (continued)

8	7	10	108.4	107.1	1.3
9	7	10	108.9	111.6	−2.7
12	2	17	112.9	119.2	−6.3
12	15	16	110.4	105.6	4.8
13	12	14	107.2	109.6	−2.4
13	12	15	109.4	112.1	−2.7
14	12	15	107.2	106.6	0.6
17	19	23	98.0	100.9	−2.9
17	20	21	111.6	110.1	1.5
17	20	22	103.0	101.9	1.1
18	17	19	108.1	108.4	−0.3
18	17	20	111.7	109.2	2.5
19	23	24	123.1	121.5	1.6
19	23	22	115.7	114.7	1.0
19	17	20	102.4	103.4	−1.0
20	22	23	97.5	101.0	−3.5
21	20	22	108.1	108.3	−0.2
22	23	24	121.2	123.8	−2.6
23	24	25	110.2	110.5	−0.3
				RMS	2.4

Table 14. Structural Differences Between CHARMM and Ab Initio Results for **2,3−aMan**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	2	1.11	1.10	0.02
1	3	1.40	1.41	−0.01
1	8	1.58	1.52	0.06
1	7	1.40	1.42	−0.02
3	4	0.96	0.98	−0.01
5	14	1.52	1.52	−0.01
5	6	1.12	1.10	0.02
5	18	1.52	1.52	0.00
5	7	1.44	1.44	0.00
8	11	1.55	1.54	0.01
8	10	1.44	1.44	0.00
8	9	1.11	1.10	0.02
10	23	1.38	1.39	−0.01
11	14	1.56	1.52	0.04
11	13	1.45	1.45	0.00
11	12	1.11	1.09	0.02
13	23	1.37	1.39	−0.02
14	15	1.12	1.10	0.02
14	16	1.42	1.43	0.00
16	17	0.96	0.97	−0.01
18	21	1.43	1.43	0.00

Table 14. (continued)

18	20	1.11	1.10	0.02
18	19	1.11	1.09	0.02
21	22	0.96	0.98	−0.01
23	24	1.36	1.36	0.00
24	25	0.96	0.97	−0.02
				0.02

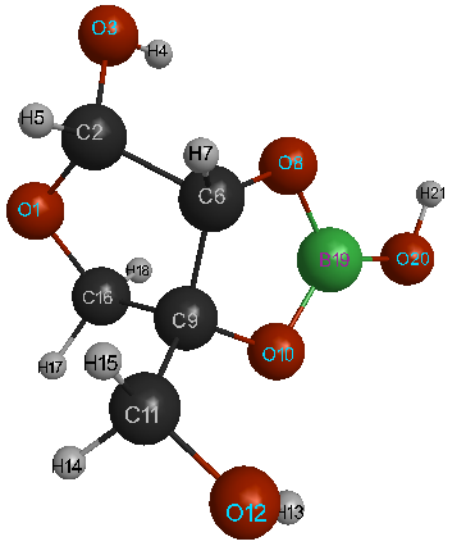
Bond Angles (in degrees)

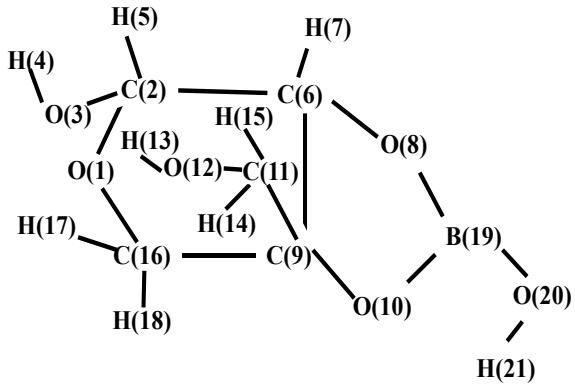
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	8	9	107.3	108.1	−0.8
1	3	4	107.4	106.7	0.7
1	8	10	118.3	110.0	8.3
1	8	11	114.9	115.7	−0.8
1	7	5	112.8	112.4	0.4
2	1	3	108.0	112.1	−4.1
2	1	7	108.3	104.6	3.7
2	1	8	106.6	109.8	−3.2
3	1	7	112.4	111.4	1.0
3	1	8	108.8	105.4	3.4
5	14	11	112.3	110.3	2.0
5	18	20	109.8	109.1	0.7
5	18	21	111.9	111.5	0.4
5	18	19	109.8	109.8	0.0
5	14	16	110.9	111.2	−0.3
5	14	15	108.1	107.5	0.6
6	5	14	108.6	110.4	−1.8
6	5	18	107.2	109.9	−2.7
6	5	7	110.3	110.4	−0.1
7	5	14	109.9	107.8	2.1
7	1	8	112.4	113.6	−1.2
7	5	18	107.2	105.8	1.4
8	10	23	103.8	104.9	−1.1
8	11	12	109.0	113.2	−4.2
8	11	13	102.6	102.8	−0.2
8	11	14	111.1	112.3	−1.2
9	8	10	107.1	108.7	−1.6
9	8	11	108.2	110.8	−2.6
10	23	13	110.1	113.7	−3.6

Table 14. (continued)

10	8	11	100.4	103.4	−3.0
10	23	24	125.6	121.9	3.7
11	13	23	107.1	105.0	2.1
11	14	16	110.6	107.9	2.7
11	14	15	106.2	108.6	−2.4
12	11	13	110.9	109.0	1.9
12	11	14	108.3	109.6	−1.3
13	23	24	124.0	124.4	−0.4
13	11	14	114.7	109.5	5.2
14	5	18	113.5	112.6	0.9
14	16	17	109.0	106.8	2.2
15	14	16	108.5	111.3	−2.8
18	21	22	108.8	105.7	3.1
19	18	20	107.5	108.3	−0.8
19	18	21	108.5	106.7	1.8
20	18	21	109.2	111.4	−2.2
23	24	25	110.0	110.1	−0.1
				RMS	2.6

Table 15. Structural Differences Between CHARMM and Ab Initio Results for **2,3−aApi**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	16	1.43	1.44	−0.01
1	2	1.43	1.42	0.01
2	6	1.54	1.55	0.00
2	3	1.39	1.41	−0.01
2	5	1.11	1.09	0.02
3	4	0.96	0.98	−0.02
6	9	1.49	1.54	−0.05
6	7	1.11	1.09	0.02
6	8	1.40	1.43	−0.02
8	19	1.38	1.39	−0.01
9	10	1.40	1.45	−0.05
9	16	1.53	1.53	0.00
9	11	1.53	1.52	0.01
10	19	1.38	1.39	−0.01
11	15	1.11	1.09	0.02
11	14	1.11	1.10	0.01
11	12	1.43	1.42	0.00
12	13	0.97	0.97	−0.01
16	18	1.12	1.10	0.02
16	17	1.12	1.09	0.02
19	20	1.36	1.36	0.00

Table 15. (continued)

20	21	0.96	0.97	−0.02
			RMS	0.02

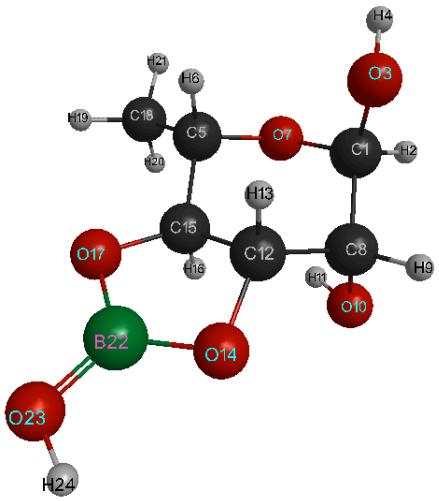
Bond Angles (in degrees)

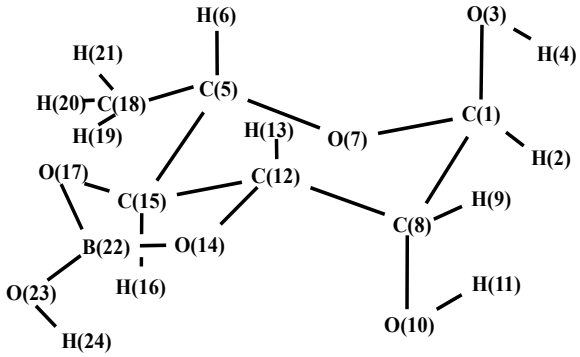
Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	2	3	111.9	113.7	−1.8
1	16	18	108.1	111.6	−3.5
1	16	17	109.4	106.5	2.9
1	2	5	106.6	107.5	−0.9
1	16	9	104.8	106.3	−1.5
1	2	6	106.2	102.3	3.9
2	3	4	106.6	108.2	−1.6
2	6	8	114.7	114.1	0.6
2	1	16	109.6	107.7	1.9
2	6	7	111.5	110.8	0.7
2	6	9	104.0	101.6	2.4
3	2	5	108.9	105.8	3.1
3	2	6	112.2	113.3	−1.1
5	2	6	110.9	114.4	−3.5
6	8	19	102.9	104.7	−1.8
6	9	11	112.2	114.9	−2.7
6	9	16	101.7	103.9	−2.2
6	9	10	106.4	102.0	4.4
7	6	8	108.2	110.1	−1.9
7	6	9	112.9	114.4	−1.5
8	19	20	121.6	124.8	−3.2
8	19	10	115.2	113.5	1.7
8	6	9	105.4	105.7	−0.3
9	11	15	111.8	108.7	3.1
9	16	17	109.8	113.2	−3.4
9	16	18	110.1	110.6	−0.5
9	10	19	103.6	106.2	−2.6
9	11	14	109.2	108.8	0.4
9	11	12	112.5	112.9	−0.4
10	9	11	113.8	109.5	4.3
10	9	16	108.7	111.3	−2.6
10	19	20	123.2	121.6	1.6
11	9	16	113.1	114.5	−1.4
11	12	13	108.5	108.1	0.4

Table 15. (continued)

12	11	14	106.2	112.0	−5.8
12	11	15	109.9	106.2	3.7
14	11	15	106.8	108.0	−1.2
17	16	18	114.2	108.6	5.6
19	20	21	110.5	110.5	0.0
				RMS	2.7

Table 16. Structural Differences Between CHARMM and Ab Initio Results for **3,4−aRam**





Bond Lengths (Å)

Atom I	Atom J	CHARMM	*Ab Initio*	Difference
1	7	1.41	1.43	−0.02
1	2	1.11	1.10	0.02
1	8	1.59	1.54	0.05
1	3	1.40	1.41	−0.01
3	4	0.96	0.98	−0.01
5	15	1.51	1.51	0.00
5	7	1.44	1.46	−0.02
5	6	1.12	1.10	0.02
5	18	1.54	1.51	0.03
8	12	1.56	1.52	0.04
8	10	1.43	1.42	0.01
8	9	1.11	1.09	0.02
10	11	0.96	0.98	−0.01
12	15	1.54	1.51	0.03
12	14	1.44	1.44	0.00
12	13	1.12	1.10	0.02
14	22	1.38	1.40	−0.02
15	17	1.44	1.44	0.00
15	16	1.11	1.10	0.01
17	22	1.38	1.40	−0.02
18	21	1.11	1.09	0.02

Table 16. (continued)

18	20	1.11	1.09	0.02
18	19	1.11	1.09	0.02
22	23	1.36	1.36	0.00
23	24	0.96	0.97	−0.02
			RMS	0.02

Bond Angles (in degrees)

Atom I	Atom J	Atom K	CHARMM	*Ab Initio*	Difference
1	8	10	113.0	109.5	3.5
1	8	9	108.3	110.0	−1.7
1	7	5	114.1	115.4	−1.3
1	8	12	107.6	104.3	3.3
1	3	4	107.2	107.0	0.2
2	1	7	108.5	103.5	5.0
2	1	3	108.1	112.1	−4.0
2	1	8	107.1	109.6	−2.5
3	1	8	109.1	108.0	1.1
3	1	7	112.1	112.5	−0.4
5	18	21	110.2	110.6	−0.4
5	18	20	110.4	110.0	0.4
5	18	19	110.2	109.4	0.8
5	15	12	109.9	109.8	0.1
5	15	16	109.5	108.0	1.5
5	15	17	116.0	118.4	−2.4
6	5	15	109.9	110.2	−0.3
6	5	18	108.3	110.9	−2.6
6	5	7	110.9	109.5	1.4
7	5	15	107.9	104.3	3.6
7	1	8	111.8	111.2	0.6
7	5	18	107.0	107.0	0.0
8	10	11	110.8	106.5	4.3
8	12	13	106.2	107.6	−1.4
8	12	14	129.2	119.3	9.9
8	12	15	106.7	110.1	−3.4
9	8	10	107.2	106.8	0.4
9	8	12	108.2	112.0	−3.8
10	8	12	112.4	114.1	−1.7
12	14	22	104.4	102.6	1.8

Table 16. (continued)

12	15	16	109.8	110.1	−0.3
12	15	17	103.6	102.9	0.7
13	12	14	106.8	107.7	−0.9
13	12	15	108.3	109.3	−1.0
14	12	15	98.3	102.6	−4.3
14	22	23	123.5	124.1	−0.6
14	22	17	110.8	114.3	−3.5
15	5	18	112.9	114.5	−1.6
15	17	22	105.3	102.4	2.9
16	15	17	107.9	107.5	0.4
17	22	23	125.5	121.6	3.9
19	18	20	108.6	108.9	−0.3
19	18	21	108.8	109.2	−0.4
20	18	21	108.5	108.7	−0.2
22	23	24	110.1	110.1	0.0
				RMS	2.7

Table 17. RMS Deviations of Structural Differences Between Molecular Mechanics and Ab Initio for both Training and Test Set

Molecule	Bond Lengths (Å)	Bond Angles (degrees)
1,2−aGal	0.01	2.2
1,2−aXyl	0.02	2.1
1,2−aFrucfur	0.02	2.1
2,3−aGal	0.02	2.7
2,3−aFuc	0.02	2.6
3,4−aGal	0.02	2.4
3,4− aFuc	0.02	2.3
3,4−aMan	0.02	3.0
1,2−bFuc	0.02	2.3
2,3−aFrucfur	0.02	2.4
2,3−aMan	0.02	2.6
2,3−aApi	0.02	2.7
3,4−Ram	0.02	2.7
Average RMSD	0.02	2.5

Table 18. Statistical Results from Torsional Profiles

Molecule	Barrier Difference[a]	RMSD[a]	Figure
1,2−aGal	−0.13	0.77	3
1,2−aXyl	−0.11	0.70	4
1,2−aFrucfur	0.02	0.40	5
2,3−aGal	−0.04	0.62	6
2,3−aFuc	−0.07	0.64	7
3,4−aGal	0.30	0.61	8
3,4− aFuc	0.08	0.60	9
3,4−aMan	−0.28	0.58	10
1,2−bFuc	−0.10	1.24	11
2,3−aFrucfur	−0.04	0.44	12
2,3−aMan	0.44	0.44	13
2,3−aApi	−0.01	0.58	14
3,4−Ram	−0.08	0.57	15
Average	**0.00**	**0.63**	

[a]All the energies are in kcal/mol

CHAPTER 6. GENERAL CONCLUSIONS

In chapter 2, *ab initio* calculations are reported for several related heterocyclic compounds, each of which contains two dative bonds when they self-dimerize. Thus, these molecules are nominally dimers that contain either a boron-carbon-nitrogen (BCN) or boron-carbon-phosphorous (BCP) segment. Molecules with this motif have been found experimentally to have several unusual properties that may be related to a "multi-polar framework" that results from charge separation associated with the two dative bonds. Structures obtained from full geometry optimizations without symmetry constraints, dative bond energies and charge distributions for four multipolar molecules are reported, the BCN-BCN dimer and the BCN-BCP dimer with and without carboxylation of one boron atom. Comparisons to single dative bond, self-cyclized monomers and the role of ring strain in molecular stabilities are also discussed.

In chapter 3, *ab initio* calculations done on this molecule have been limited to monomer structure and only a few experimental studies mention the existence of a dimer. In this study mass spectroscopy has been used to provide evidence of the dimerization of 2-APB in water. The peak for the dimer, formed by two dative B-N bonds is present only in electrospray mass spectra. Harder ionizing mass spectral techniques show only monomers and smaller fragments. Fragmentation patterns in the electrospray experiments show not only monomers and dimers but also additional cyclic compounds and are explained by ab initio calculations done on both gas and solvent phase on monomer,

dimer and fragment geometries. Thus 2-APB can exist as both monomer and a dimer structure in solvent phase and the free energy of the dimer structure is about 2 kcal/mol lower than cyclic-monomer.

In chapter 4, the polymeric structure of B_2O_3 is theoretically modeled as interconnected ribbons of BO_3^{3-} repeating units (monomers). The hydrolyzation reaction barrier heights are predicted for consecutive bond breaking of three B−O bonds in the central BO_3^{3-} unit of the B_2O_3 structure at RHF/6-31G(d,p) and MP2/6-31G(d,p)//RHF/6-31G(d,p) levels. The hydrolyzation reaction barrier heights are lower when two water molecules are involved than for a single water molecule transition states. The successive barrier heights for the hydrodrolysis of B_2O_3 with two waters are predicted as 9.43, 12.31, and 17.30 kcal/mol at MP2/6-31G(d,p)//RHF/6-31G(d,p) level relative to their reactant complexes.

In chapter 5, an extension of the bio-molecular CHARMM all-atom empirical force field parameters is described for modeling boron complexes of carbohydrates in which the boron is bound to the carbohydrate through boro-diester linkage. The model is developed to be consistent with the CHARMM all-atom carbohydrates force field, and the existing parameters for pyranose and furanose sugars were transferred from carbohydrate force fields to develop new boro-carbohydrate parameters. The additional parameterization is based on MP2/6-31G(d,p) geometries, solute-water interaction energies and torsional potentials. The optimized geometries are reported for a set of

galactose, fucose, mannose, glucose, xylose, apoise and fructofuranose boron complexes. The model satisfactorily reproduces the structures of thirteen boro-carbohydrate complexes within 0.03 Å accuracy for bond lengths and 3 degrees accuracy for bond angles. The torsional barriers are well reproduced, within 0.6 kcal/mol.