Summer 2021

Modeling of Argon Bombardment and Densification of Low Temperature Organic Precursors Using Reactive MD Simulations and Machine Learning

Kwabena Asante-Boahen
Missouri State University, Kwabena442@live.missouristate.edu

MODELING OF ARGON BOMBARDMENT AND DENSIFICATION OF LOW

TEMPERATURE ORGANIC PRECURSORS USING REACTIVE MD SIMULATIONS

AND MACHINE LEARNING

A Master's Thesis

Presented to

The Graduate College of

Missouri State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science, Materials Science

By

Asante-Boahen Kwabena

July 2021

MODELING OF ARGON BOMBARDMENT AND DENSIFICATION OF LOW TEMPERATURE ORGANIC PRECURSORS USING REACTIVE MD SIMULATIONS AND MACHINE LEARNING

Physics, Astronomy and Materials Science

Missouri State University, July 2021

Master of Science

Asante-Boahen Kwabena

ABSTRACT

In this study, an important aspect of the synthesis process for a-BxC:Hy was systematically modeled by utilizing the Reactive Molecular Dynamics (MD) in modeling the argon bombardment from the orthocarborane molecules as the precursor. The MD simulations are used to assess the dynamics associated with the free radicals that result from the ion bombardment. By applying the Data Mining/Machine Learning analysis into the datasets generated from the large reactive MD simulations, I was able to identify and quality the kinetics of these radicals. Overall, this approach allows for a better understanding of the overall mechanism at the atomistic level of Ar bombardment and the role of radical species towards the formation of the orthocarborane network and in turn the boron carbide thin films.

KEYWORDS: orthocarbones, boron carbide, reactive molecular dynamics simulations, data mining, machine learning

MODELING OF ARGON BOMBARDMENT AND DENSIFICATION OF LOW TEMPERATURE ORGANIC PRECURSORS USING REACTIVE MD SIMULATIONS AND MACHINE LEARNING

By

Asante-Boahen Kwabena

A Master's Thesis
Submitted to the Graduate College
Of Missouri State University
In Partial Fulfillment of the Requirements
For the Degree of Master of Science, Material Science

July 2021

Approved:

Ridwan Sakidja, Ph.D., Thesis Committee Chair

Tiglet Besara, Ph.D., Committee Member

Fei Wang, Ph.D., Committee Member

Julie Masterson, Ph.D., Dean of the Graduate College

ACKNOWLEDGEMENTS

I dedicate this thesis to my daughter, Bailee Janae Asante.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

1.1 Boron Carbide

Boron carbide is a high strength ceramic which can be used in so many applications due to some of the unique mechanical properties it possesses. It currently falls third to Diamond and cubic boron nitride as the hardest known material in nature alongside having a very low density (Raucoules et al., 2011). This combined property makes it ideal in instances where strength to weight ratio comes into play. Also, it has a high melting temperature above 2500 K which makes it thermally stable (Thévenot, 1990). In addition to its high abrasion resistance, the thermal stability and mechanical strength of boron carbide gives it the unique physical properties well suited for body armor components and abrasives. Due to Boron-10 having a high cross-section for neutron absorption, boron carbide becomes an important material for nuclear reactors and neutron detection(Blevins & Yang, 2020) (Bigdeloo & Hadian, 2009). Boron carbide also has several applications in the semiconductor industry; thin-film boron carbide can be used as insulators for low-dielectric-constant intra/interlayer dielectrics (Nordell et al., 2015), patterning materials for semiconductor devices and optical films.

There a couple of growth techniques used to produce boron-carbide (Thévenot, 1990) (Suri et al., 2010). Some of these are hot pressing and pressureless sintering. These synthesis processes require high temperature sintering, which makes it an energy intensive process. A major limitation with this synthesis process is that it cannot be easily integrated with semiconductor device processing and oftentimes do not yield the desired electronic properties (Bandyopadhyay et al., 1984) (Wood & Emin, 1984). Also, these experimental studies are

conducted over a long time period for an extremely limited number of materials, and it imposes

high requirements in terms of resources and equipment.

1.2 Crystal Structure

Materials which are rich in boron usually have an icosahedron crystal structure. An

Icosahedron is a polyhedron that has 20 faces. An icosahedron is made up of 20 equilateral

triangles and 30 edges, with 5 faces coming to a point at each of the 12 vertices. Figure 1.1

shows the geometry of a regular icosahedron.
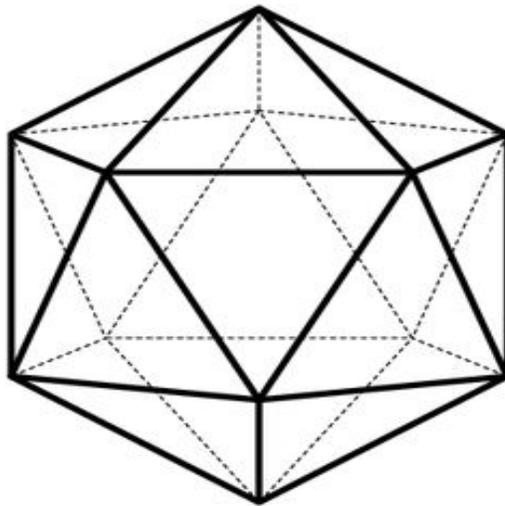
Figure 1.1 Geometry of a regular icosahedron

The ideal, most symmetric form of the structure of boron carbide is described as a

rhombohedral unit cell that is made of one icosahedral B_{12} unit and one C-B-C chain. The first

and third atoms of the chain are both connected to three different icosahedra. A typical

icosahedron typically has 11 boron atoms and 1 carbon atom ($B_{11}C$) or 12 boron atoms (B_{12}).

The structure shows the formation of icosahedral configuration because of the valence electron deficiency of boron in the B-B and B-C covalent bonds. There are 6 polar sites in every icosahedron, 3 in each opposite end with 6 equatorial sites between them. A two-center bond directly bonds the atoms in the polar sites to the neighboring icosahedra parallel to the cell edge. Figure 1.2 shows the crystal structure of Boron Carbide
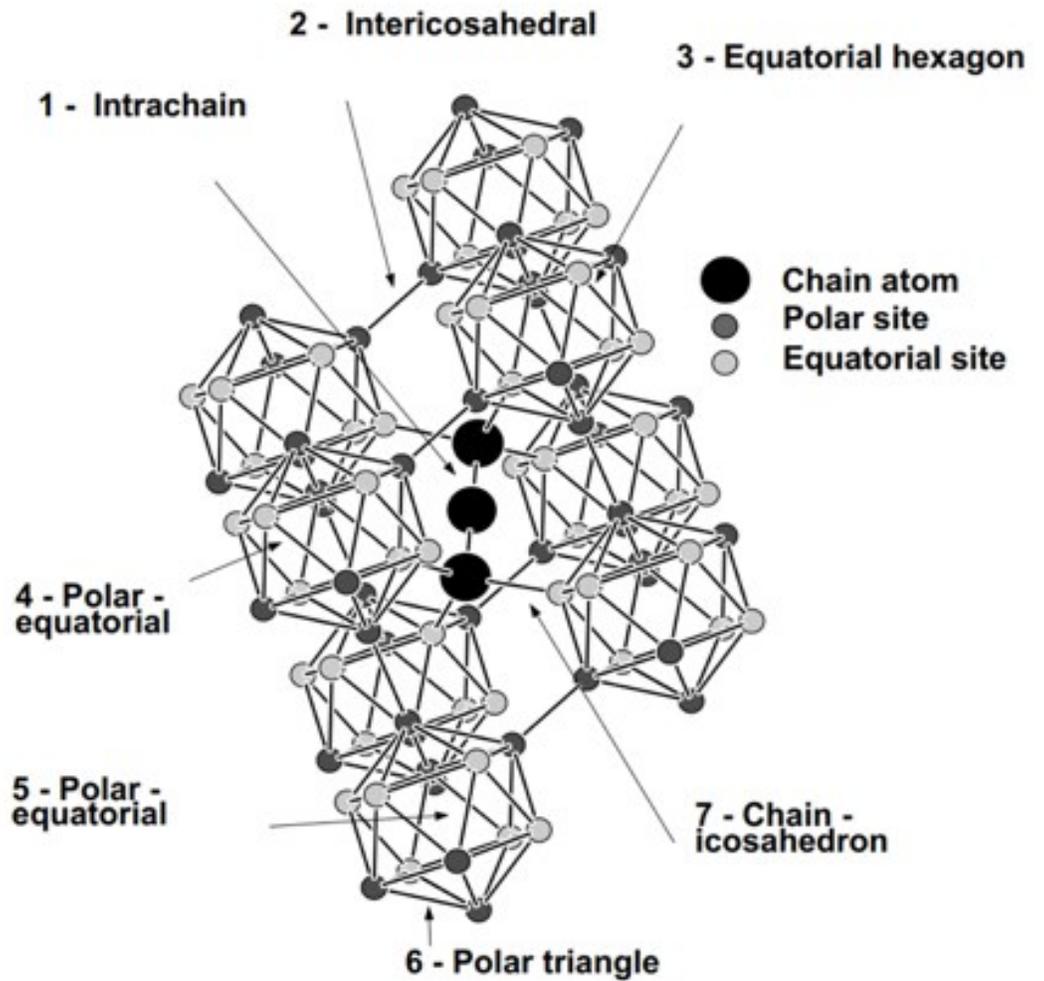
Figure 1.2 Crystal Structure of Boron Carbide (Raucoules et al., 2011)

There are two ways of connecting the equatorial atoms; either by either by connecting them directly to their neighboring icosahedra or to the chains. In $B_{11}C$, the carbon atoms are mainly in the polar sites, the chain is C-B-C for $B_{11}C_p$ and C-C-C for $B_{12}C_3$. Carboranes are sometimes used by experimentalists as organic precursor to produce hydrogenated boron carbide (Billa et al., 2009) (Zhang et al., 1998)-(Schulz et al., 2008). The basic chemical formula of carboranes is $B_{10}C_2H_{12}$ with the possible isomers being orthocarborane (o-carborane), paracarborane (p-carborane) and metacarborane (m-carborane). The icosahedral structures of these molecules are slightly distorted with ten boron atoms and two carbon atoms in their inner cage and twelve surrounding hydrogen atoms that points outwards and bonded with an atom of the inner cage (Gamba & Powell, 1996). The most dominant amongst these molecules is the orthocarborane.

All the three isomers of Boron Carbide are solids at room temperature and pressure with crystal FCC structures that have cell dimensions of 9.86 Å and an atomic number of 4 (Baughman, 1970). All the isomers have varying melting points and crystalline phases. Orthocarborane has the highest melting point of the three, with a melting point of 570K, followed by metacarborane at 546K and paracarborane at 534K. The formation of the crystal
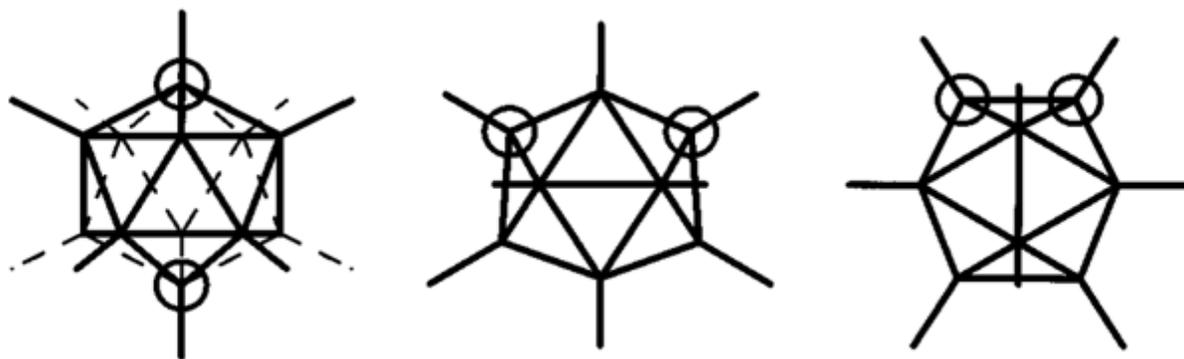
Figure 1.3 p-carborane, m-carborane and o-carborane molecules. The circles show the location of the C atoms[26].

structures comes as a result of the presence of weak van der Waal forces. The phase transitions of o-carborane occur at 274K and 167K, that of m-carborane occur at 277K and 165K and p-carborane at 303K and 240K (Leites, 1992). Phases I and II are orientally disordered for all the three compounds whereas phase III is orientally ordered (Baughman, 1970)

1.3 Machine Learning

One major limitation experimentalists in materials science face is conducting research over a long time frame with limited resources. These limitations have led to the need of introducing machine learning tools in the study and synthesis of amorphous boron carbide. Figure 1.3 shows the locations of the C atoms in carborane molecules

This introduction will cut down the time and the high cost involved in materials synthesis. Machine Learning, in recent years, has proved to have superhuman abilities in various disciplines like image classification, image and speech recognition, fraud detection, amongst many others. Its introduction into the field of solid-state systems has been widespread of late. Previously, computational methods like density functional theory (DFT), Monte Carlo simulations, and molecular dynamics were what brought about the computational revolution in materials science (Schmidt et al., 2019).

The continual increase in computing power and the development of more efficient codes have made way for computational high-throughput studies of large material groups in order to screen for ideal experimental candidates. The large-scale simulations alongside the experimental high through-put studies have been producing a large amount of data, making the use of machine learning methods in material science possible. Like the introduction of any novel approach in any field, machine learning has received a few criticisms in the field of material science. One of such main criticisms is the lack of knowledge and understanding arising from their use. This is

because machine-built models are sometimes too complex and alien for humans to understand. This notwithstanding, there has been numerous excellent reviews of machine learning in the fields of materials science and atomistic simulations. The Machine learning tool employed in this study is the Waikato Environment for Knowledge Analysis (Weka) which was developed by the University of Waikato in New Zealand. Weka is a collection of machine learning algorithms for solving real-world data mining problems.

1.4 Molecular Dynamics Simulation

Computer simulations in material science are mainly done to replicate the behavior of a complex system on a digital computer. It drastically cuts down on the cost, limitation in resources and time taken to undertake experiments. Simulations are classified into three main kinds according to the scale in which they are performed: atomic scale, the micro scale and the macro scale. The kind of simulation used in this study is the molecular dynamics which falls under the atomic scale. Simulations at this scale basically focuses on structure, mechanical properties, thermodynamic properties and kinetic properties of materials. In any classical molecular dynamics simulation, the trajectory of a system can be found by solving Newton's equation of motion by using information of the object's initial position, velocity and force.

$$m_i \frac{\delta r_i}{\delta t^2} = F_i$$

$$F_i = -\frac{\delta U(r_1, \ldots \ldots, r_n)}{\delta r_i}$$

m_i and r_i in the equations above represent the mass and position of the ith object respectively. The potential of the system is any given configuration is represented by U. When the correct potential is used in solving the equation, the trajectory of a system with atoms, grains, mesoscale particles,

6

or even huge planets can be accurately calculated. From the derivation of the potential, U, force information on every particle can be calculated. The number of differential equations that need to be solved will correspond to the number of objects. Practically, when integration is used instead of differential equations, it reduces the time taken to solve the equation. In figure 1.4 below is a pictorial representation of modeling times and their length scales.
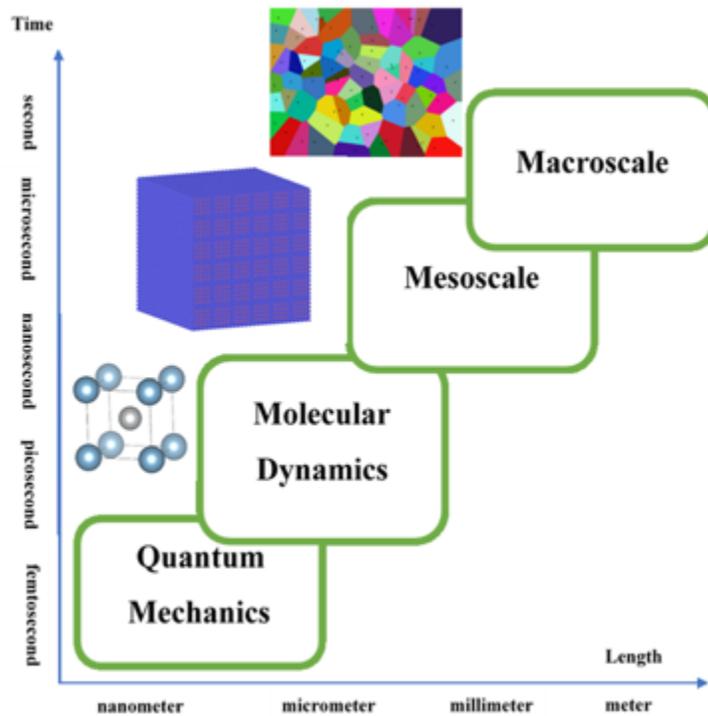
Figure 1.4 Modeling time and length scales (Baishnab et al., 2020)

$$-m_i \frac{dv_i}{dt} = \sum_j F_2 \left(r_i, r_j \right) + \sum_{j, k} F_3 \left(r_i, r_j, r_k \right) + \cdots \dots$$

$$\frac{dr_i}{dt} = v_i$$

In the equation above, F_2 represents the force between two atoms and F_3 is the force between three atoms. There is an adjustment on the force information as well as the positions of the atoms after each iteration.

The atoms keep vibrating and will not stay at rest at their energy minimum. After the simulation runs for a good amount of time, Boltzmann distribution is sampled. This sampling is an efficient way of exploring the surface's energy. Boltzmann distribution in this case refers to having the atoms arranged in a particular way which correlates with the system's energy. Alder and Wainwright (Alder & Wainwright, 1957), in 1957, published the first simulation work where they sought to solve the classical equations of motion of several hundred particles by using fast electronic computers. The 32-particle system had square potentials which showed the phase transition through a hard sphere model. The hard sphere model makes overlapping of atoms in the system impossible, just like macroscopic metal balls. After directly solving Newton's force equation, a system with three constant quantities is obtained. The three quantities are the number of atoms, volume, and energy. The energy of the system, its total and angular momentums do not change because it is an isolated system. This kind of system is termed the NVE or microcanonical ensemble. This system does not exchange energy or particles with its environment, therefore, does not have pressure and temperature control.

However, in experimental setting, the synthesis or characterization parameters are quantities like temperature, pressure, or volume. The pressure is mostly kept constant, such as synthesis in a high vacuum chamber, high vacuum, or atmospheric conditions. In taking measurements of bulk modulus, volume is kept constant and in annealing of samples, temperature is kept constant. Aside NVE, different statistical ensembles can be used to mimic experimental conditions. These ensembles are NVT, which is constant temperature and constant

8

volume, NPT, which is constant temperature and constant pressure, NST, which is constant temperature and constant stress and then NPH, which is constant pressure and constant enthalpy.

1.4.1 NVT Ensemble. Also known as the canonical ensemble, it has constant-temperature and constant-volume. One advantage this ensemble has is that it allows the pressure of a system to be fixed for materials regardless of their phase. It is a good way to perform conformational exploration of molecules when done without periodic boundary conditions. This mostly comes in handy when dealing with chemical reactions in gaseous materials as system pressure is irrelevant.

1.4.2 NPT Ensemble. This is the constant-temperature and constant pressure ensemble that enables a user to control both system pressure and temperature by allowing the volume to change and energy to flow in and out of the system. For the desired pressure to be obtained, the unit cell vectors have to be adjusted in every step. This is a nice way to fix the densities and equilibration volume of the system under a predetermined system pressure. Before undertaking any simulation with any material or interatomic potential, it is advisable to equilibrate the initial structure using NPT or NVT ensembles.

1.4.3 NST Ensemble. The constant-temperature, constant-stress ensemble is mainly applied in the study of the stress-strain relationship in polymeric or metallic materials. It can be said to be an extension of the constant-pressure (NPT) ensemble. It allows the user to tune the xx, yy, zz, xy, yz and zx components of the stress tensor.

1.4.4 NPH Ensemble. This is the constant-pressure, constant-enthalpy ensemble. Though this ensemble is equivalent to the NVE ensemble, it is a unique way of controlling the system's pressure without controlling the system's volume. Enthalpy H, (E + PV) remains constant when pressure is fixed without temperature control.

In performing simulations, it is very important to maintain a constant volume and pressure. A simple way that the pressure can be changed is using the piston-like mechanism. In controlling the temperature, the user must be extra careful and perform extensive calculations. Controlling the velocity may cause constraints on the volume and pressure since temperature depends on the velocity of all the particles. The theory behind the thermostat should be understood in order to comprehend the control of temperature. Thermostats are similar to heat baths. The temperature of materials is regulated by heat bath entity. Temperature in molecular dynamics can be defined by the equation (van Duin et al., 2001):

$$\frac{kT_{MD}}{2} = \frac{1}{6N} \sum_{i=1}^{N} \sum_{j=1}^{3} m_i v_{ij}^2$$

k = Boltzmann constant

T_{MD} = temperature

N = number of particles

m = mass of particles

v = velocity of particles

for three-dimensional velocity components j=3.

The purpose of the summation is to create a distribution. In simulation, the temperature can be controlled by restricting the average velocity within the proper distribution. Usually, the beginning temperature is set by assigning velocities at random under Maxwell-Boltzmann distribution. Another way to do this is by using the Anderson approach (Andersen, 1980). In this method, a particle is allowed to interact with the heat bath, in each time step, then a new velocity

is assigned under the Gaussian distribution. The time it takes to get to the targeted temperature, in this approach, depends on the number of collisions per unit time. This can be bypassed by introducing an additional force (Langevin thermostat).

Another approach that offers a more accurate solution is the Nosé-Hoover thermostat. In this approach, an additional degree of freedom for heat bath is introduced.

$$H_{nh} = \sum_i \frac{p_i^2}{2m_i s^2} + U(r, q) + \frac{p_s^2}{2Q} + LkT.\ln(s)$$

H_{nh} = heat bath

L = number of independent momentum degrees of freedom

Q = An imaginary mass

p_s = momentum of heat bath variable

When s(t) =1, the original Hamiltonian is associated with four interrelated equations.

$$\dot{r} = \frac{m_i v_i}{s}$$

$$\dot{v} = -\frac{1}{m_i} \frac{\delta U}{\delta r_i}$$

$$\dot{s} = \frac{p_s}{Q}$$

$$\dot{p}_s = \sum_i m_i s \dot{r}_i^2 - LKT$$

The last two equations refer to velocity control and they can be rewritten as:

$$\frac{ds}{dt} = \frac{Lk(T - T_{target})}{Q}$$

The Verlet algorithm is the widely used algorithm for time integration of the equation of motion in molecular dynamics. Taylor expansion up to the third order can be used to explain this algorithm. When the position (r) of any particle for a certain time (t) with timestep δt is known, the equations of next and previous steps according to Taylor expansion are (respectively):

$$r_i(t + \delta t) = r_i(t) + \dot{r}_i(t)\delta t + \frac{1}{2}\ddot{r}_i(t)\delta t^2 + O(\delta t)^3$$

$$r_i(t - \delta t) = r_i(t) - \dot{r}_i(t)\delta t + \frac{1}{2}\ddot{r}_i(t)\delta t^2 - O(\delta t)^3$$

Adding the two equations gives us the Verlet equation:

$$r_i(t + \delta t) \approx 2r_i(t) - 2r_i(t - \delta t) + \frac{F_i(t)}{m_i}\delta t^2 + O(\delta t)^3$$

It has been studied that the computation of force expression is the costliest calculation and most time-consuming (van Duin et al., 2001). Each particle in a group of particles, theoretically interacts with all other particles. For a system with n particles, the interaction equations would be n^2. But most of the interactions can become negligible after a number of nearest neighbors. The most used approach is by selecting a desired cutoff point r_c for each particle and putting into consideration only the forces of the particles that lie within the cutoff distance r_c. The cutoff distance usually specifies the number of interactions of nearest neighbors' that should be taken into consideration. Selecting a reasonable cutoff distance in classical molecular dynamics creates a balance between the cost of simulation and the expected accuracy. The particle-mesh algorithm is an approximation method that can be used to reduce the simulation load for long range forces like Coulomb interaction (Plimpton, 1995).

Out of the numerous tools used in modeling, Classical Molecular Dynamics (CMD) is the widely used. Just like all other approaches, this method has its limitations. To begin with, atoms in CMD are treated as a point mass. Without regarding the quantum effect, CMD usually has very few errors. De Broglie's hypothesis states that all particles can be treated as matter waves and the wavelength can be calculated from the equation $\lambda = mv/h$. Most atoms at room temperature have wavelengths of about 0.2 Å whilst their atomic distance usually ranges between 1 – 3 Å. When dealing with atoms like hydrogen or helium which are light in weight or when undertaking ultra-high temperature calculations, neglecting the quantum effect may lead to errors in results. Also, in order to achieve the desired accuracy with the Verlet algorithm, the maximum timestep used should be 1-2 femtoseconds. Typically, a fraction of a femtosecond is used. With the availability of complex computational facilities now, simulation can be continued for a few billion steps, which is going to calculate a system response for only a few microseconds. With CMD involving atomic scale modeling, although work can be done with millions of atoms, in reality, it is only going to be a cube with sides of a few hundred nanometers. For homogenous systems like crystalline materials, periodic boundary conditions can be used to suppress the issue whereas for non-homogenous systems like amorphous materials, it is very difficult to make (*Multiscale Modeling and Analysis for Materials Simulation*, 2011). Because of some these limitations, simulation results can most times be deceptive. In modeling, for example, heating or cooling of a material at a heating rate of 1K/ns is considered very slow (Sandoval & Urbassek) but in experimental setting, it is not possible to achieve a heating rate of 10^9 K/s. Likewise, for compression in simulation, a strain rate that is commonly used is 10^{-4}/ps which converts to a rate of 10^8/s (Wen et al., 2018),(Li et al., 2018) but in experimental setting, a very high strain rate is considered $10^4 – 10^6$ s^{-1} (Ramesh, 2008). CMD

works perfectly with Born-Oppenheimer approximation, where lighter electrons adjust with changes in the position of heavier nuclei. This is the reason why it is difficult for CMD to model formation and breaking of chemical bonds. The way of obtaining the interatomic potential is by using some fitting techniques (Duff et al., 2015)-(Brommer & Gähler, 2007) for parameters like elastic constants, thermal expansion from highly accurate density functional theory (DFT) calculations, or experimental data. One setback with these potentials is they are not transferrable. A typical example is how parameters for Ni in Ni-Al may not work for Ni-Al-Co. Potentials are mostly very element specific and are suitable for only some studies. Developing a potential for modeling purposes takes a long time and it comes with its limitations. The interatomic potential used in this study is the reactive force field (ReaxFF). This potential has a lot of advantages over the traditional force field.

The molecular dynamics simulations in this study were performed using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) (Plimpton, 1995). This simulator was developed by Sandia National Laboratories, which is a US Department of Energy laboratory. The reason for choosing LAMMPS over all the other credible codes like MDACP (Watanabe et al., 2013), (Watanabe et al., 2011), MDSPASS (Iskandarov et al., 2011)and IMD (Stadler et al., 1997) is because LAMMPS has all the methods required for this study since most codes are developed to fit particular studies. The NAMD (Phillips et al., 2005) code for instance, was developed specifically for biomolecular studies.

1.4.5 LAMMPS

LAMMPS is a molecular dynamic simulator that can model atoms or particles in all the three main states of matter. It can model particles up to a few hundred million in two-dimensional and three-dimensional systems. It can handle simulations on atomic, microscopic, mesoscopic and continuous scales. Due to its wide range of applicability, it has been used in modeling metals, semiconductors, biomolecular and granular systems.

LAMMPS has been programmed to work most efficiently on parallel computing in supercomputers or shared memory boxes that use message-passing techniques although it runs perfectly in serial with one core on personal computers. It is designed to function on systems that support the Message Passing Interface (MPI).

Previous versions of LAMMPS were written in Fortran (F77 and F90) but the recent version is written in C++ programming language. Since there mostly arises the need for upgrades in the code, it is developed in a manner where such upgrades and modifications can be easily executed. This compatibility mode makes it easy to keep up with the inclusion of newly developed force fields, pair styles and several other components of the analysis. This user-friendly interface also makes it possible for users to add relevant features to an existing code. In defining the interactions among a group of particles, LAMMPS makes use of Newton's equations of motion. These particle types are mainly electrons, charged/uncharged atoms, molecules, or radical grain.

1.4.6 PACKMOL

PACKMOL is a code that creates an initial point for simulations by packing molecules in defined regions of space. In this study, random structures with variable sizes of simulation cells were created by using PACKMOL.

PACKMOL can create a predefined simulation cell which has initial coordinate point of molecules. In order to achieve the expected structure, ranges of spatial restrictions can be put in place when producing the coordinates. An example is the minimum distance between molecules and their orientation in the three-dimensional space. The interaction of the large repulsive van der Waals forces can be limited at the beginning of the simulation due to the minimum distance limitation. By only specifying the coordinates of one molecule or unit, a user can make complex mixtures of different material phases by packing millions of molecules. There are a good number of output files that PACKMOL supports. Some of these are XYZ, TINKER, PDB and MOLDY.

1.4.7 ReaxFF

Prior to the introduction of atomic-scale modeling, the in-depth study of novel materials used to be a herculean task since it is very expensive and difficult to come by the experimental equipment. Simulation methods that involve quantum-mechanics have recently been wide-used for different kinds of materials due to their easy access, high level of results accuracy and then having user friendly interfaces. One limitation with this is the computational approach that comes with it. It makes it applicable to a smaller number of systems. These systems usually are not enough to fully comprehend the behavior of a material as compared to a system with larger molecules or amorphous structure. The development of empirical force fields (EFF) was in a

quest to overcome these limitations. Data sets that are obtained from either quantum mechanics or experimental are fitted against the parameters of EFFs. This makes the fitted data parameters less credible than the original data.

Methods involving EFF have been successful in condensed phase systems, but they could not describe reactive systems and bond energy to the dissociation limit. This is as a result of how EFFs are trained. They are trained where bonds remain close to their equilibrium. Reactive force field (ReaxFF) is a force field which permits the breakage and formation of bonds. It is a bond order-based force field. It makes use of the bond order obtained from the empirical distance between atoms. Unlike EFFs, ReaxFF is very transferrable and works best with all the three main phases of matter. The ReaxFF code was written to include (1) continuous forces and energy including reactions, (2) one particular element has only one type of force field atom, (3) no pre-defined reactive site is required (van Duin et al., 2001). The ability of ReaxFF to transfer its parameters to similar compounds has made possible an avenue to a different modeling scale which was impossible before. An example is how the parameters of oxygen atoms in solid metal oxide can be transferred and used in liquid water or gaseous O_2 molecules (Senftle et al., 2016). This makes it easy to model variable aspects of a system with more than one phase. Another practical example is the transition from the gaseous phase (Argon and orthocarborane) to the solid phase (amorphous hydrogenated boron carbide) in this study.

1.4.8 Methodology: ReaxFF is developed to include both reactive and non-reactive reactions in its descriptions whilst avoiding bond order formation and polarizable charge description. ReaxFF operates on the energy equation:

$$E_{system} = E_{bond} + E_{over} + E_{angle} + E_{tors} + E_{vdwaals} + E_{coulomb} + E_{specific} \text{ (Senftle et al., 2016)}$$

E_{bond} = the bond forming energy which is gotten from bond order and related to interatomic distance

E_{over} = Over coordination energy; stops over coordination of atoms

E_{angle} = Angle strain

E_{tors} = torsional energy

$E_{vdwaals}$ = van der Waal energy of every atom

$E_{coulomb}$ = Coulombic interaction

$E_{specific}$ = specific energy terms; unrelated to terms above

ReaxFF has components that are dependent on bond order and others which do not. The code calculates bond order directly from interatomic separation. The bond order calculation follows the following equation:

$$BO_{ij} = BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi}$$

Can also be written as $BO_{ij} = \exp\left[p_{bo1}(\frac{r_{ij}}{r_o^{\sigma}})^{p_{bo2}}\right] + \exp\left[p_{bo3}(\frac{r_{ij}}{r_o^{\pi}})^{p_{bo4}}\right] + \exp\left[p_{bo5}(\frac{r_{ij}}{r_o^{\pi\pi}})^{p_{bo6}}\right]$

BO in the equations above represents the bond order between the atoms i and j and r_{ij} is the interatomic distance. The P_{bo} terms are obtained empirically whereas r_o represents the bond length at equilibrium. There is a continuous transition between the pi and sigma bonds. Because of this continuity, forces that exist between the atoms can be calculated from taking the derivatives of the potential energy surface. The covalent interactions that the formula describes

makes the force field to correctly provide reaction barriers. Information on angle strain and bond energy is obtained from the bond order correction. Calculation of the coulombic interaction is done in each iteration from a charge equilibrium feature. Figure 1.5 below depicts an overview of the energy component of ReaxFF.
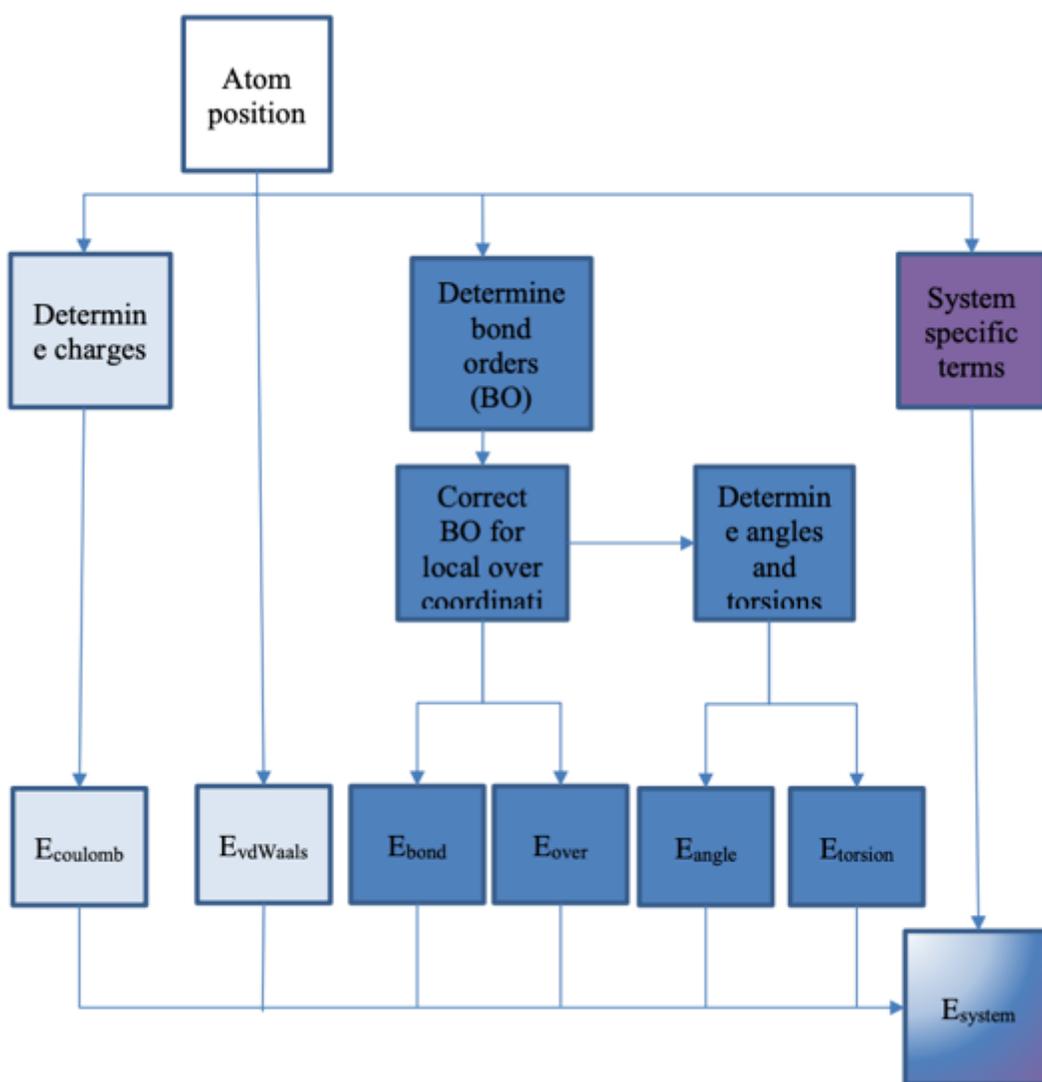
Figure 1.5 Energy components of ReaxFF (Car & Parrinello, 1985)

1.5 Quantum Calculation

Due to the already discussed limitation pertaining to the use of quantum mechanical calculations, most of the calculations in this study is done with classical molecular dynamics using ReaxFF. That regardless, quantum mechanical calculations were used to determine the extent where the force field can be applied to hydrogenated boron carbide. In calculating the substitution of core nuclei and valence electron exchange-correlation term, the density functional theory (DFT) which is a form of QM calculation was used. DFT is able to define multiple electron interactions in atoms by using spatially dependent electron density functionals. DFT since the 1970s was extensively used in many science disciplines although it did not live up to the level of expected accuracy. This, however, was no longer the case after some modifications were made to the theory in the 1900s.

DFT has a relatively higher accuracy than classical MD and has potential files available for each element on the periodic table. Also, it provides details of the electronic structure of elements, which is not the case with CMD. DFT operates on the following Hamiltonian equation:

$$H_{NVE} = \frac{1}{2} \sum_{i=1}^{N} m_i \, r_i^2 + E[\varphi(r_1, r_2, \dots, r_N)]$$

The main concept of the equation is that a system's potential energy can be derived from their electron structure at each step. The Kohn-Sham single-electron wavefunctions for the ground state is denoted by $\varphi(r_1, r_2, \dots, r_N)$. Information on force is given by the ground state

energy data. From this information, one iteration can be completed by moving the atoms to the next step. This DFT-based approach is called ab-initio molecular dynamics (AIMD).

$$H_{CP} = H_{NVE} + \frac{1}{2}\sum_j \mu \int |\psi_j(r)|^2 dr + L_{ortho}$$

The above equation is what the Car-Parrinello method, the first AIMD approach is based on. The first term, H_{NVE} is taken from the previous equation and the next term refers to the imaginary mass' kinetic energy, μ, which is also an expression of the electronic degrees of freedom. The last term, L_{ortho} refers to the orthogonality of the wave function of the single electron. The equation of motion is now written as:

$$\begin{cases} \mu\psi_i(r) = -H_{CP}\psi_i(r) + \sum_j \psi_j(r)\Lambda_{j,i} \\ m_i\ddot{r}_i = F_i \end{cases}$$

By bypassing the Kohn-Sham equations through self-consistent computation, the level of accuracy will not be altered although several codes in use currently utilize more efficient diagonalization algorithm of the Kohn-Sham matrix. There are quite a number of open-source and commercial codes which are similar to CMD. Some of the popular open-source codes are Quantum ESPRESSO (Giannozzi et al., 2009), Siesta (Soler et al., 2002) and ABINIT (Gonze et al., 2002) whereas some of the commercial codes are VASP (Kresse & Furthmüller, 1996b), Gaussian (*Expanding the limits of coputational chemistry*) and CPMD ((IAS), 2017). The code employed in this study is the Vienna Ab-initio Simulation Package (VASP). It has potentials

suited for all the elements on the periodic table and it also used in the study of a wide variety of materials. It also studies the surfaces and interfaces of their phases. The libraries VASP operates on is the fast Fourier transform (FFT) and linear algebra.

COMPUTATIONAL DETAILS

2.1 Energy Minimization and Bond Dissociation Energy

In a previous work done by Baishnab et al., (Baishnab et al., 2020), they performed the quantum mechanical calculations in their study using the Vienna Ab-initio Simulation Package (VASP) (Kearley et al., 2006). VASP was used in setting the plane wave basis and projector-augmented wave (PAW) based pseudo-potentials. (Kearley et al., 2006), (Kresse & Joubert, 1999)-(Kresse & Furthmüller, 1996a). The ultra-soft pseudopotential was used in the calculation involving the valence electron exchange-correlation term and substitution of core nuclei. To improve on the local spin density (LSD) description of atoms and molecules, the version of generalized gradient approximation (GGA) exchange-correlation function of DFT that was used is the Perdew Burke Erzerhof (PBE). The k-space matrix was created with the Monkhorst-Pack scheme.

Quantum mechanical calculations were carried out on several components like the energies of bond dissociation and energies of ground state of important compounds like orthocarborane, BH_3, CH_4, B_2H_4, and others. In calculating the ground state energies, a cell that was as big to fit a single molecule of the mentioned compounds was selected. The cell also had to have very little interaction between the molecule and its periodic image. For the molecule of orthocarborane, a cell with 12 Å at the edges was selected whereas a cell having 6 Å at the edges was selected for the BH_3 molecule. With the energy minimization calculations, the K-points they used were 4 4 4.

In generating a bond dissociation energy curve, the structurally minimized molecule was taken and an atom of a relevant bond was slowly distanced till a point where the bond can be said to be fully broken. The system's total energy was calculated at different distances of an atom and then generate points of distance vs energy. Whiles carrying out this process, all the other atoms of the molecule are fixed at their initial positions. When calculating the B-H bond dissociation energy for orthocarborane, the structurally minimized orthocarborane molecule was obtained first. Whiles all other atoms were kept fixed at their initial positions, the distance between one boron atom connected with a hydrogen atom was varied and the energy at all distances were calculated. The ReaxFF potential was also used to undertake similar calculations.

2.2 Argon bombardment

The main aim of this phase of the study is to comprehend the results of ion bombardment on orthocarborane molecules. The bombardment in this study takes place in a gaseous phase which is the mixture of bombarding ion and Orthocarborane (Schulz et al., 2008), (Nordell et al., 2015). Before the simulation could take place, an initial structure had to be created and a molecule of orthocarborane had to be made. In making the orthocarborane, two boron atoms in a uniform B_{12} icosahedral structure were replaced with two carbon atoms. I then added 12 outward pointing hydrogen atoms to form an icosahedra. The structure of the orthocarborane comes out slightly distorted due to the presence of carbon and also unidentical lengths of BH and CH bonds. This structure was then minimized with LAMMPS to generate the second structure at 0K. The LAMMPS command used in minimizing the structure is

minimize 0.0 1.0e-8 1000 100000

minimize etol ftol maxiter maxeval

etol = stopping tolerance for energy (unitless)

ftol = stopping tolerance for force (force units)

maxiter = max iterations of minimizer

maxeval = max number of force/energy evaluations (LAMMPS documentation)

Figure 2.1 shows the structures of the orthocarborane molecule before and after relaxation.

| a. | b. |
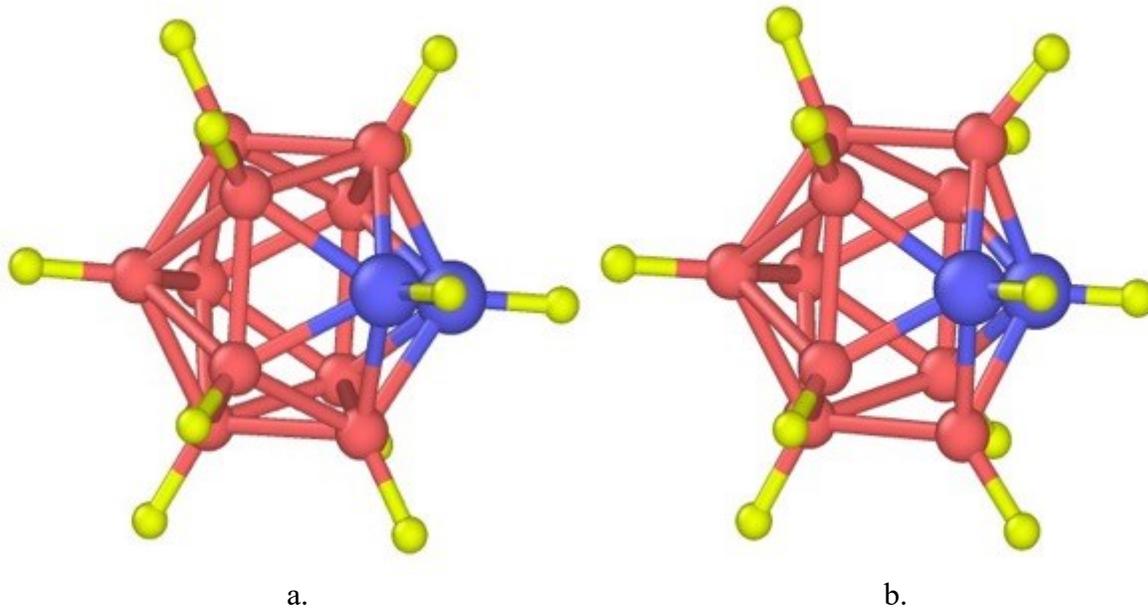
Figure 2.1 Orthocarborane molecule: (a) before relaxation and (b) after relaxation (Baishnab et al., 2020)

Minimizing the initial structure is by changing the coordinates of the atoms to find the minimum potential energy. The minimizer stops the movement of atoms in a particular iteration and allows relaxation of overlapping atoms in instances of very large energy and forces. When any of the stopping conditions are met, minimization comes to an end.

The next step was creating a bigger gaseous mixture structure of argon atoms and orthocarborane molecules. The initial simulation cell had a 100 Å edge length. A thousand relaxed orthocarborane units alongside varied number of argons (highest being 25) were

randomly packed inside the cell. The molecules were kept at a minimum distance of 2 Å apart

from one another. PACKMOL was used in building the structure. Figure 2.2 shows the initial

structure for argon bombardment.
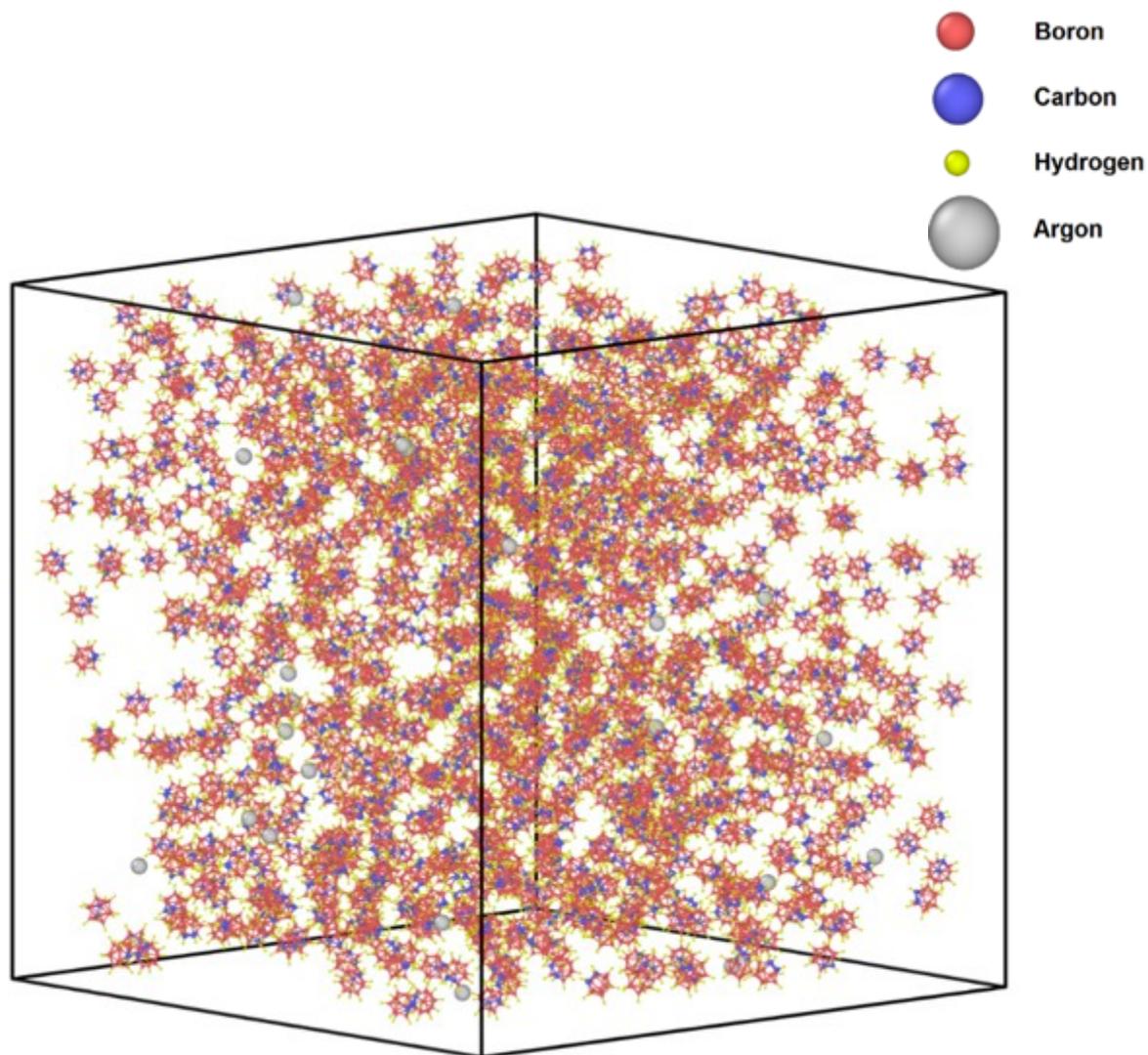
Figure 2.2 Initial structure for argon bombardment (atoms are not represented with their
relative sizes)

The cell size was extended by 1 angstrom in all six directions. This was done to prevent the atoms from overlapping at the edge of the simulation cell. Minimization was performed one more time before the argon bombardment. In figure 2.3, there is a pictorial representation of the 2D view of the initial simulation cell.
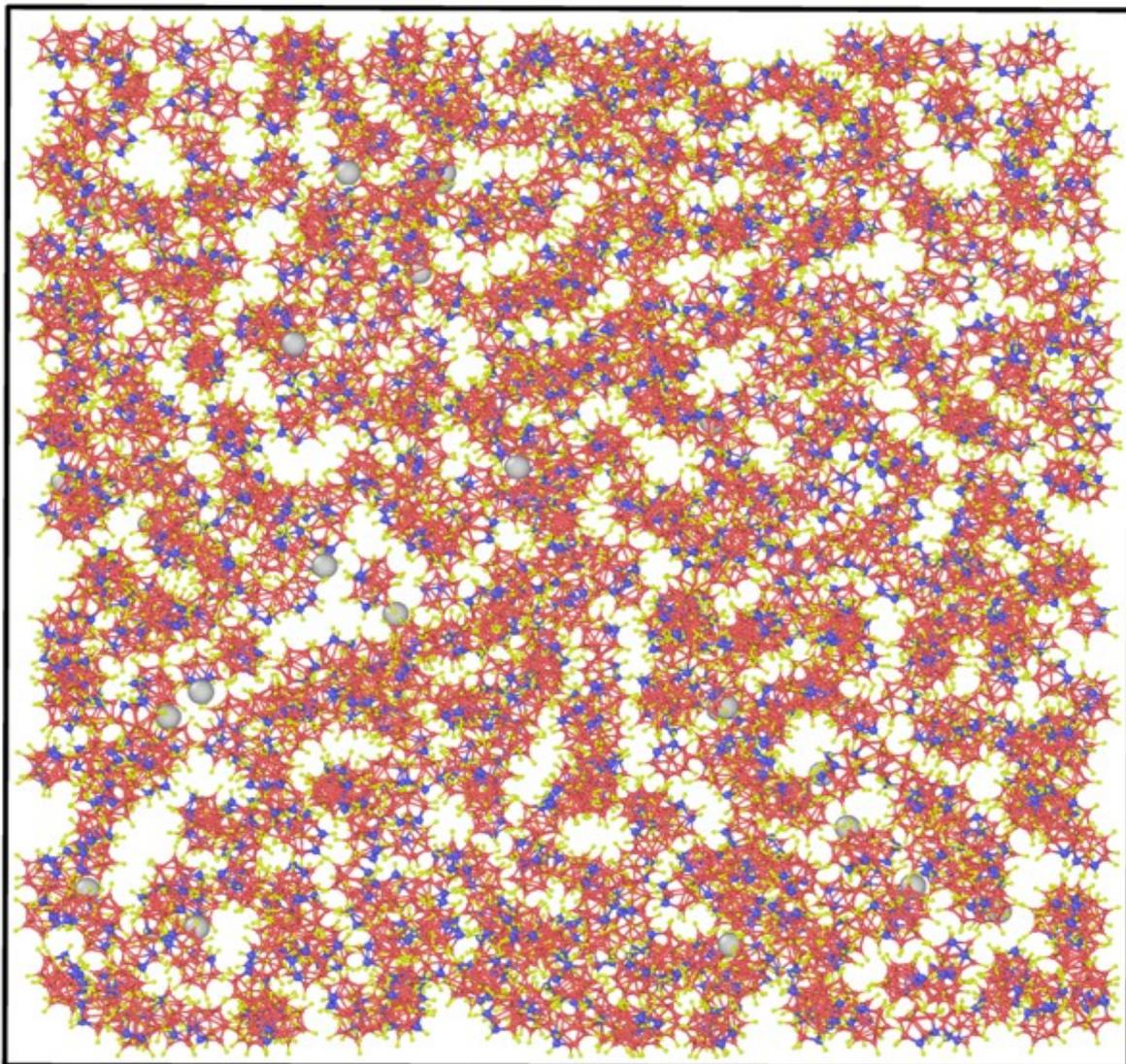
Figure 2.3 Initial simulation cell 2D view with after the edge correction (atoms are not represented with their relative size)

Another important aspect of the simulation was the kinetic energies of the atoms. To begin with, a single argon atom was modeled unto a single orthocarborane and the outcome was observed. The kinetic energy ranged from a few electron volts to up to a few hundred electron volts. At lower kinetic energies (typically 30 eV and below), there were no significant effect of the bombardment observed (fig. 2.1a). There was no ejection of species at this kinetic energy. On the contrary, the orthocarborane molecules saw a significant damage at kinetic energies 100 eV and beyond Higher kinetic energies detached different species from the orthocarborane. (fig. 2.1b) When the energy is tuned to few hundreds of electron volts, a single argon atom is able to break several molecules through severe collisions and the ejected species as well, having high energy are also able to damage neighboring molecules. The highest kinetic energy that was used in this study was 600 eV (Kinetic energy of argon in experimental setup) and the energies were varied in increments of 50 eV. The time step used in modeling high kinetic interactions of such nature should not be above 0.1 fs (Jensen et al., 2012). The time step used in our study was 0.05 fs and it relatively gave better results than 0.1 fs. The simulations were performed for 5 ps and it was enough time for the orthocarborane molecules to absorb kinetic energy of the argon atoms. The simulation was an NVE (constant number of particles, volume, and energy) process. After each NVE simulation, An NVT (constant number of particles, volume, and temperature) process is carried out to cool the structure to room temperature (300K). The Noose-Hoover thermostat is used to control the process of cooling in LAMMPS.

Figure 2.4 shows the collision of various pairs after 150 fs at 30 eV and 190 eV.

a   b
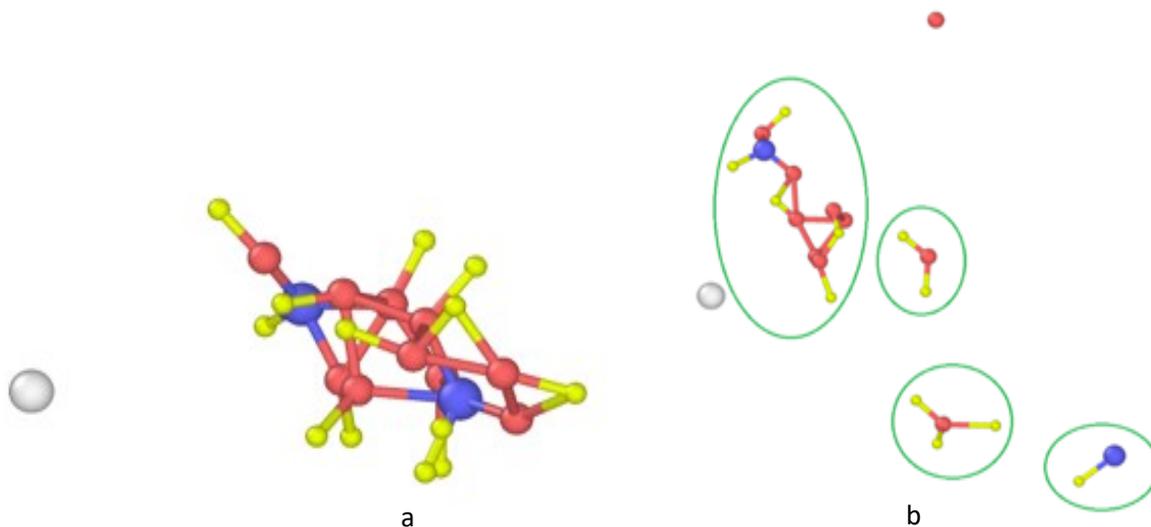
Figure 2.4 Snapshot after 150 fs of collision (a) 30 eV, (b) 190 eV

After the bombardment, I started analyzing the pair distribution functions of the various pairs at different times. The time evolution for the species that were created after the bombardment. At different points in the simulation, the distances of the various species changed, hence it required using different cut-off distances in counting the number of species.

2.3 Free radicals and Orthocarborane interaction

The bombardment of the argons gave a head way to further studies in this work. Now the free radicals from the bombardment interacted with the undamaged orthocarborane and this interaction was analyzed. The species formed were as a result of the energy the argons possessed and also the place of impact with the molecules. Some of the formed species were BH, CH, BH, H_2, BH_3, CH_2 and CH_3. Free boron, hydrogen and carbon atoms became available after few hundred femtoseconds into the bombardment. All the other species, but BH_3 and H_2 did not have a neutral charge and were active at different levels. I packed 1000 random units of

orthocarborane and varied the number of free radicals whiles equilibrating them at different temperatures. To. Begin with, a cubic cell with 92 angstroms at all sides was created and isolated carbon and boron atoms were put in it. A distance of 2 angstroms was maintained between the atoms and molecules. For every carbon that was that was chosen, 5 as much boron were also taken, since this is their ratio in the orthocarborane molecule. The number of boron was varied from 50 to 200 and that of carbon was varied from 10 to 400. The simulations were performed at
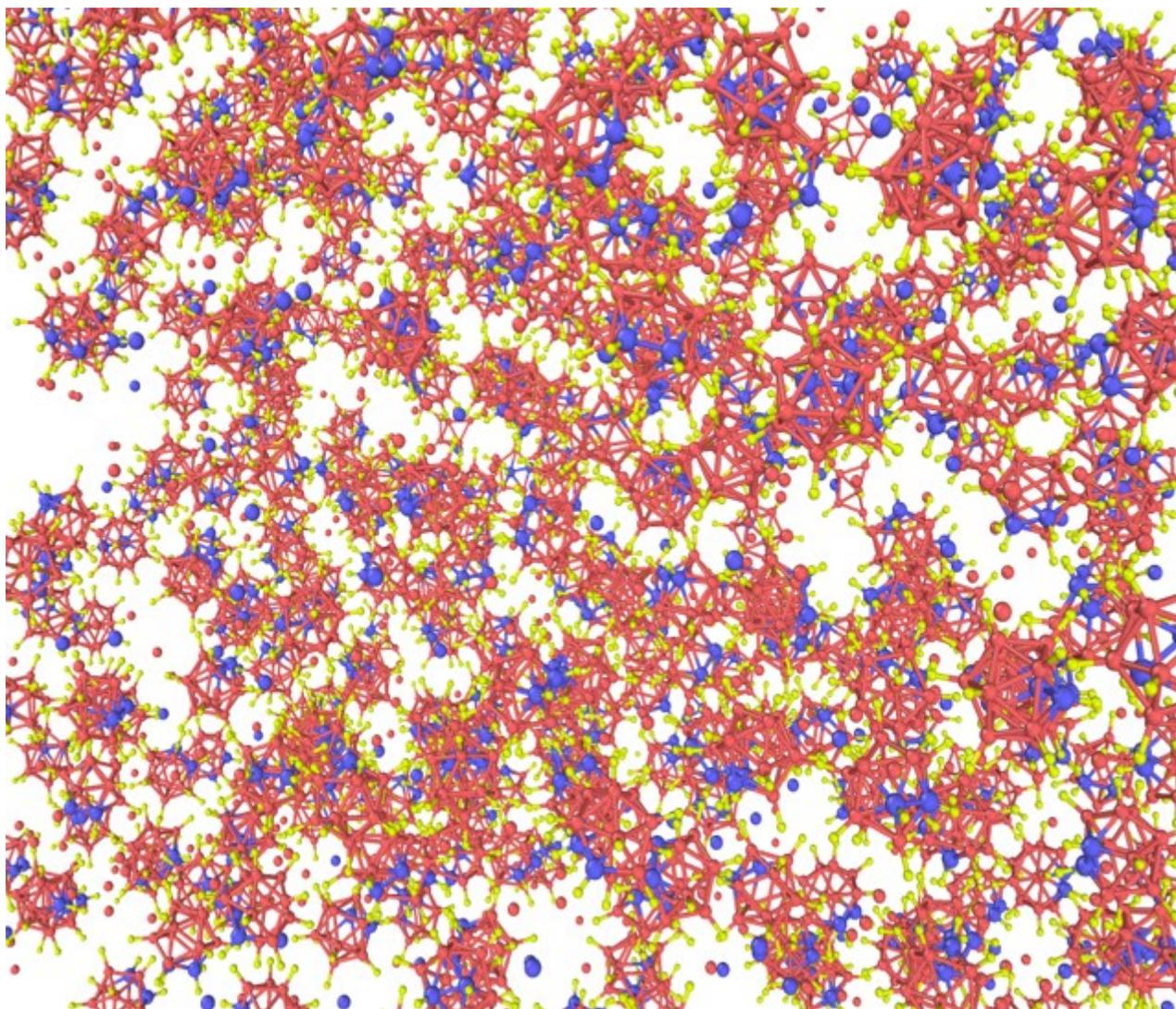
Figure 2.5 Initial structure for orthocarborane and free radicals (free boron and carbon atoms)

300K, 500K, and 800K. Figure 2.5 shows the initial structure for o-carborane and free radicals.

In our quest to keep the simulations organized, the radicals were modeled in ordered pairs. For example, BH and CH, BH_2 and CH_2, BH_3 and CH_3. What this means is that, in place of 1000 molecules of orthocarborane of isolated B and C, I put in BH and CH, BH_2 and CH_2, or BH_3 and CH_3. The ratio of boron to carbon was kept constant in all cases and the number of varied radicals also remained just like before. I used ReaxFF to relax all the free radicals individually at 0K after which they were equilibrated at 300K, 500K and 800K. This simulation was done under the NVT ensemble with a time step of 0.25 fs with 10^6 steps, spanning through a total time of 250 ps.

RESULTS & DISCUSSION

3.1 Radical Species Analysis

The overall goal of the simulation study is to understand the dynamics taken place during the Ar bombardment onto the ortho-carborane precursor to generate a range of free radical species which will react to one another or with the remaining precursors to form aggregates. The chemical reactions during the Ar bombardment generally involve complex and multi-step processes, and thus in this study the focus of the simulation study mainly looks at on the major chemical reaction event. This is defined by examining the number of gas species produced during the reactive molecular dynamics (MD) simulations.

In addition, the study concentrates on examining two main variables:

1) The effect of concentration of ortho-carbonate gases i.e., the ratio of the number of ortho-carbonane relative to the number of Ar gases.

2) The energy level of Argon bombardment that initiates the chemical reactions.

The main reason why these variables were chosen for MD simulations is because they can be controlled experimentally(Nordell et al., 2015). The concentration ratio of ortho-carborane/Ar gas mixture can be adjusted from the flow rate of, and solution concentration of ortho-carbonate/Ar mixture introduced into the chamber as gas molecules. Similarly, the energy level the ionic Ar bombardment for this plasma-enhanced chemical vapor deposition can be experimentally controlled through radio frequency (RF) power and the pressure concentration(Nordell et al., 2015).

The previous modeling study has also evaluated the role of energy level(Baishnab et al., 2020), but the concentration ratio used was relatively high e.g. 1000 ortho-carborane for 25 Ar gas (40

to 1) ratio. In addition, the MD simulations used to model the Ar bombardment processes were only up to 100,000 time steps (each step is ¼ fs) resulting observations up to about 25 ps(Baishnab et al., 2020). In the case of a high ratio of the o-carborane/Ar concentration, the previous study showed that there is a relatively fast transfer of the kinetic energy from Argon gas to the o-carborane. In fact, after only a few ps of simulations (1ps = 10,000 steps), almost all reactions have completed(Baishnab et al., 2020). This is mainly because the Ar gas no longer retains its energy due to the collisions with many o-carbonates.

In this study, the concentration ratio is lowered to 800-25 (36 to 1) and 100-25 (4-1) in comparison to the previous study(Baishnab et al., 2020). The first ratio of 36-1 is quite close to the previous study (Baishnab et al., 2020) so that I can compare the results. The second ratio is much closer to the experimentally used ratio. The goal here is to see if there is any trend caused by changing the concentration ratio. For each ratio, I varied the initial kinetic energy of Argon gas by assigning the initial velocity. For simplicity, similar to the previous study(Baishnab et al., 2020), the values of some of the axial velocities were assigned randomly, then the total square root of the whole velocity is set constant to maintain the same kinetic energy. The simulation was done under NVE ensemble and thus the total energy (kinetic + potential energy) was kept constant. Figures 3.1 and 3.2 show plots of timestep vs species count for 100 orthocarboranes at 50eV and 350 eV
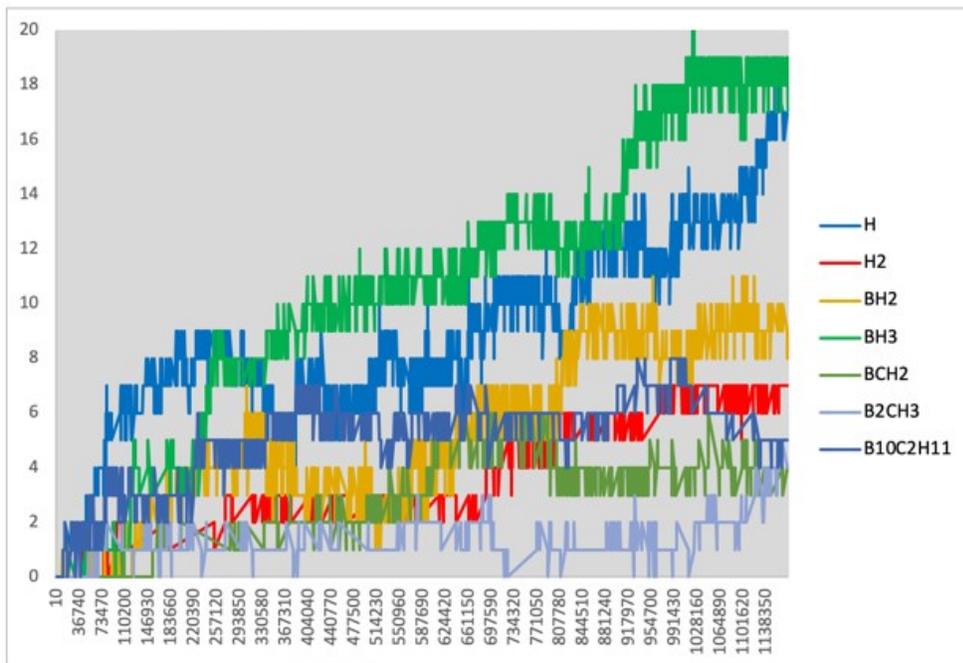
Figure 3.1 A plot of timestep vs species count for 100 orthocarboranes at 50eV
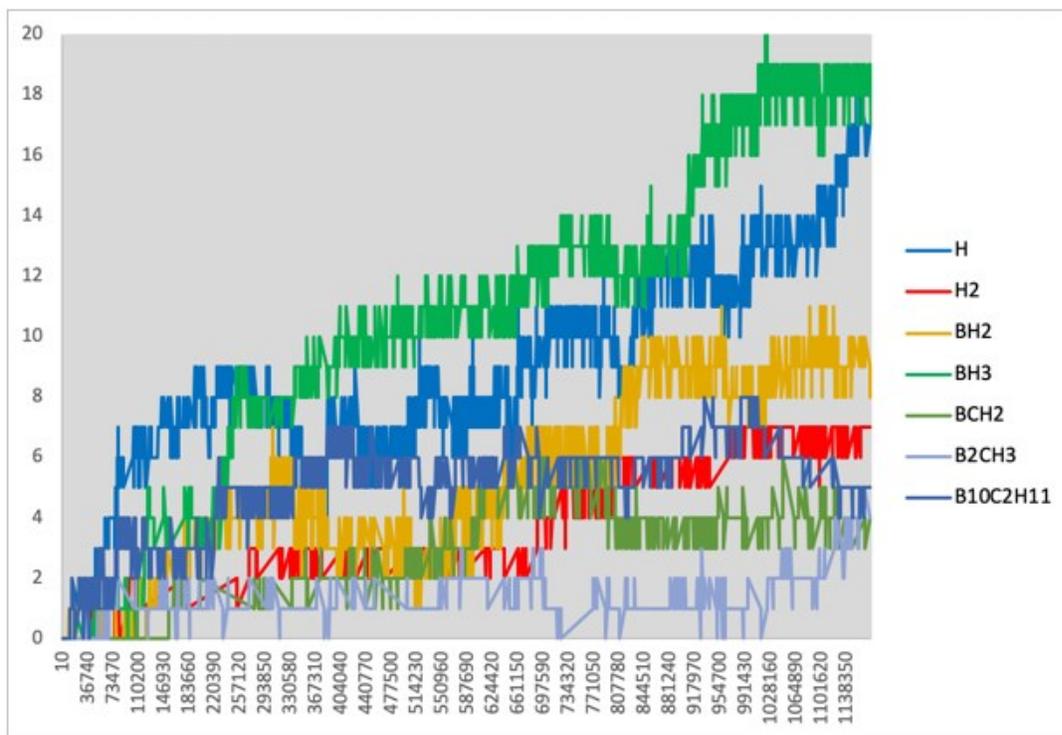
Figure 3.2 A plot of timestep vs species count for 100 orthocarboranes at 350eV

Figures 3.3 and 3.4 show plots of timestep vs species count for 100 orthocarboranes at 500 eV and 600 eV
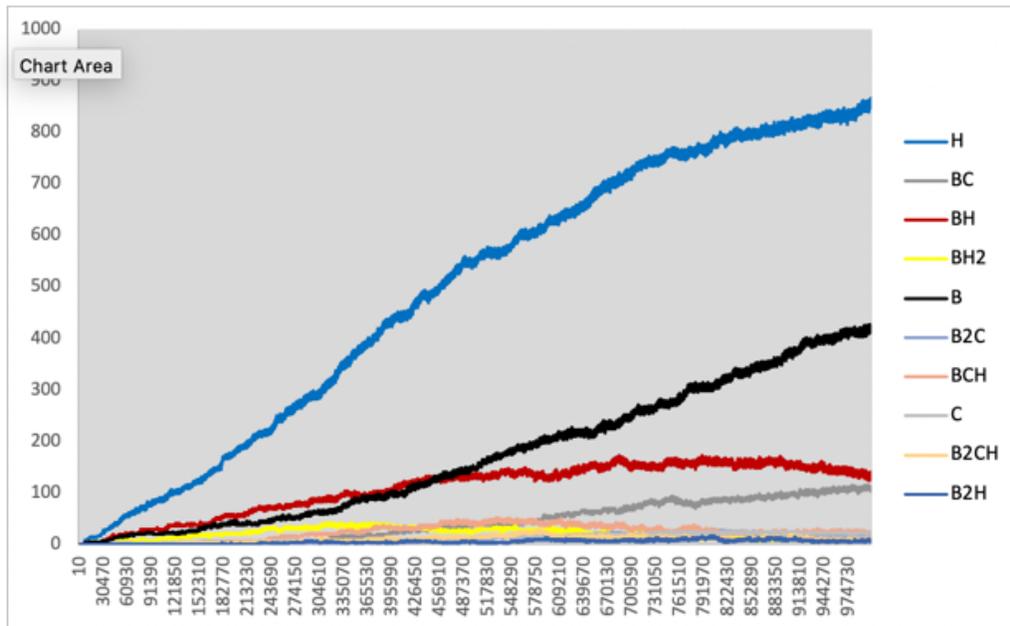
Figure 3.3 A plot of timestep vs species count for 100 orthocarboranes at 500eV
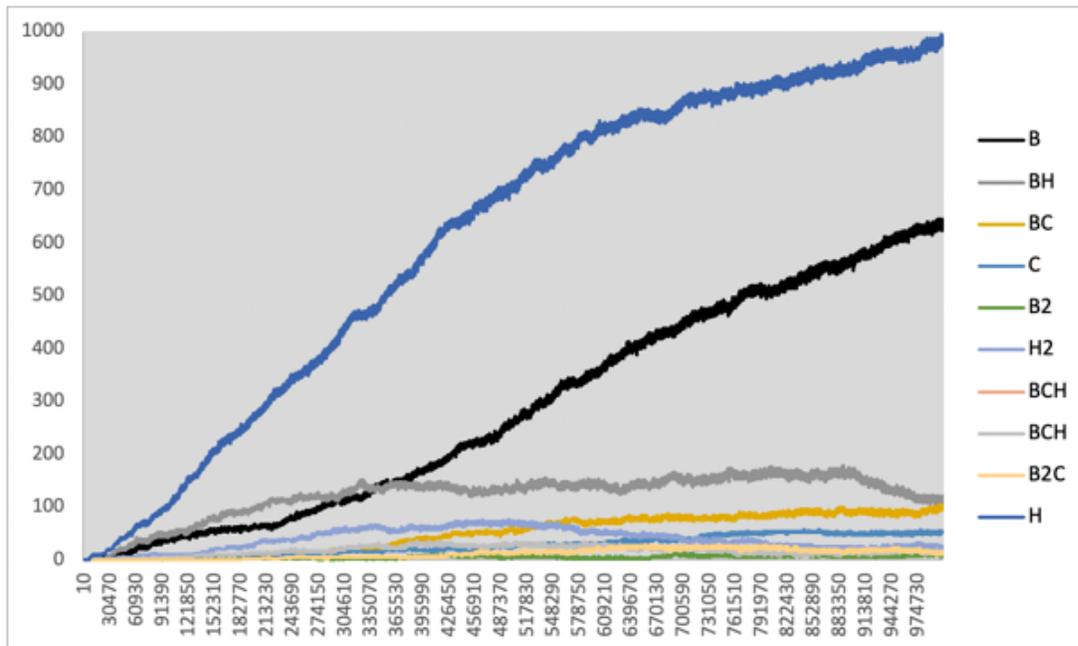
Figure 3.4 A plot of timestep vs species count for 100 orthocarboranes at 600eV

Figures 3.5 and 3.6 show plots of timestep vs species count for 800 orthocarboranes at 50 eV and 350 eV
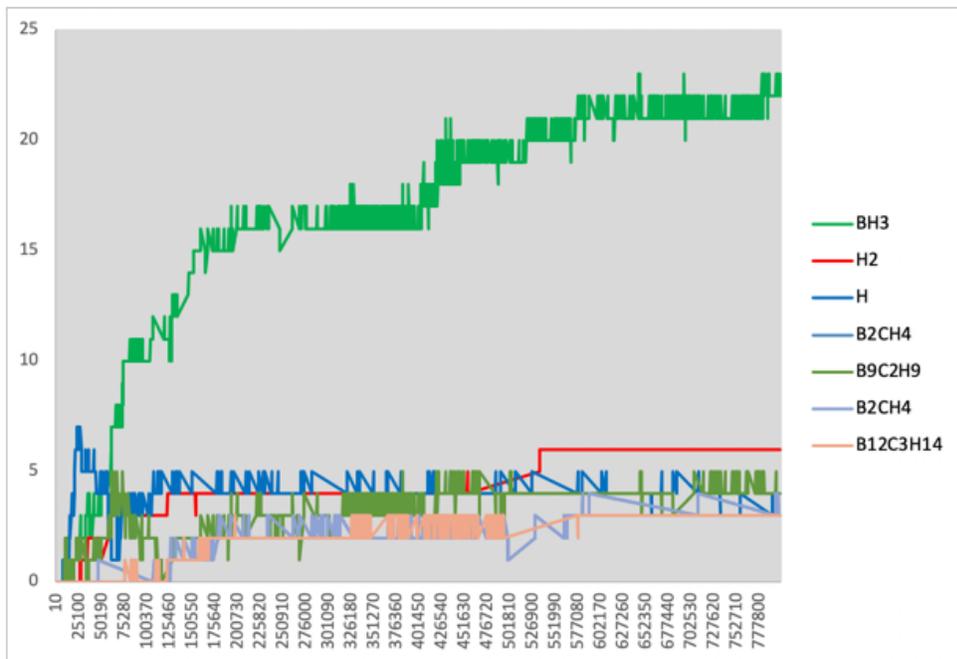
Figure 3.5 A plot of timestep vs species count for 800 orthocarboranes at 50eV
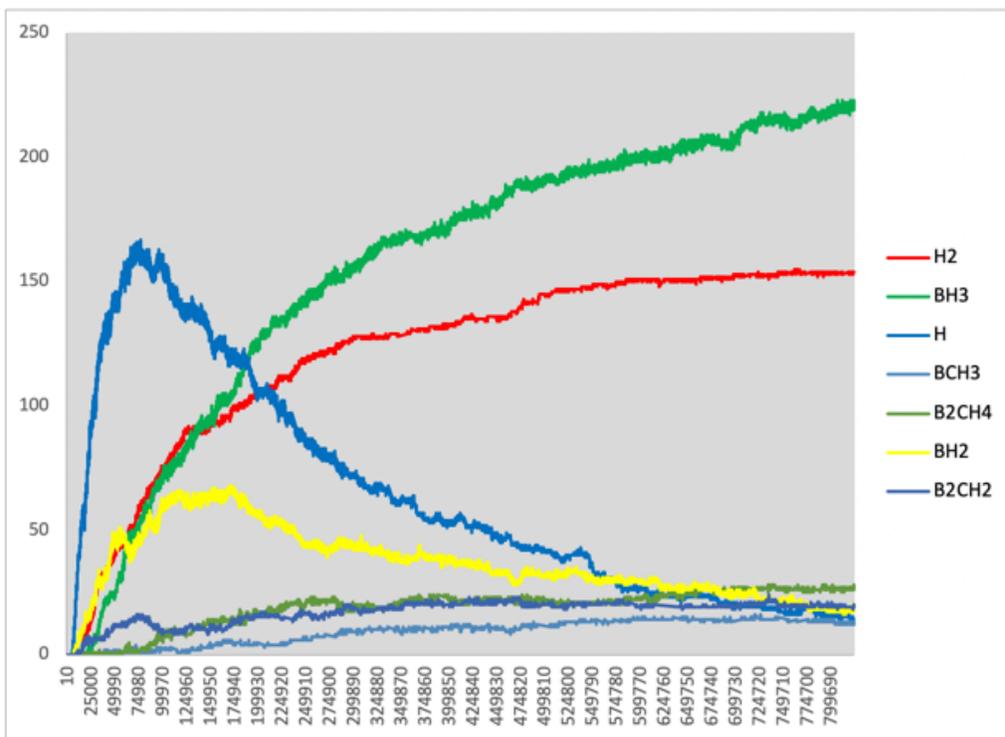
Figure 3.6 A plot of timestep vs species count for 800 orthocarboranes at 350 eV

Figures 3.7 and 3.8 show plots of timestep vs species count for 800 orthocarboranes at 500 eV and 600eV
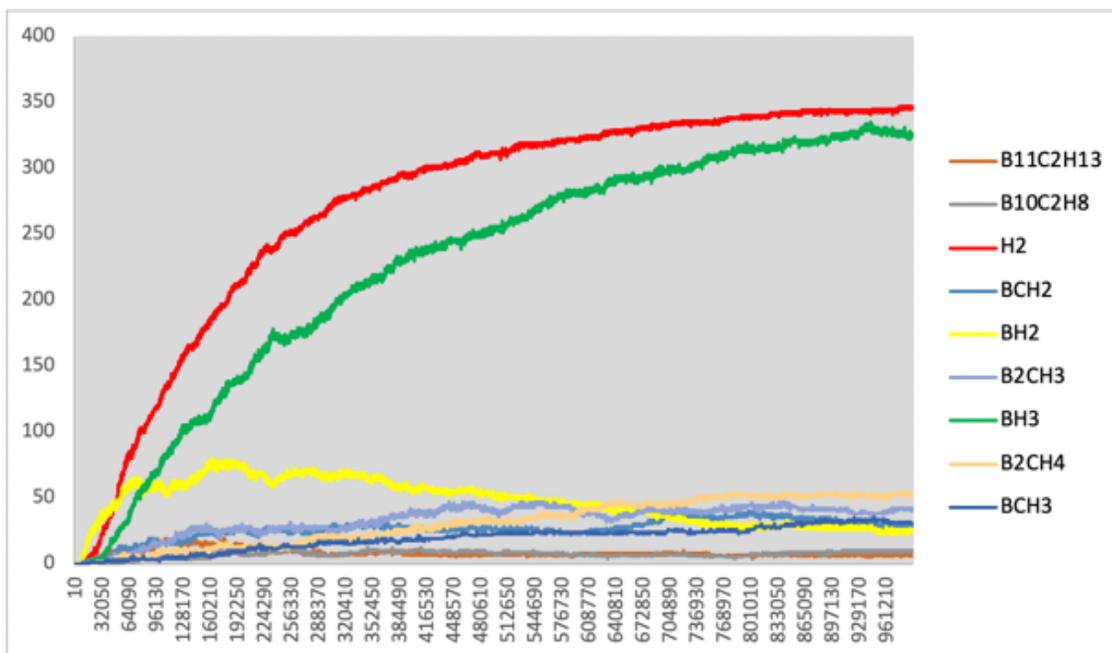
Figure 3.7 A plot of timestep vs species count for 800 orthocarboranes at 500eV
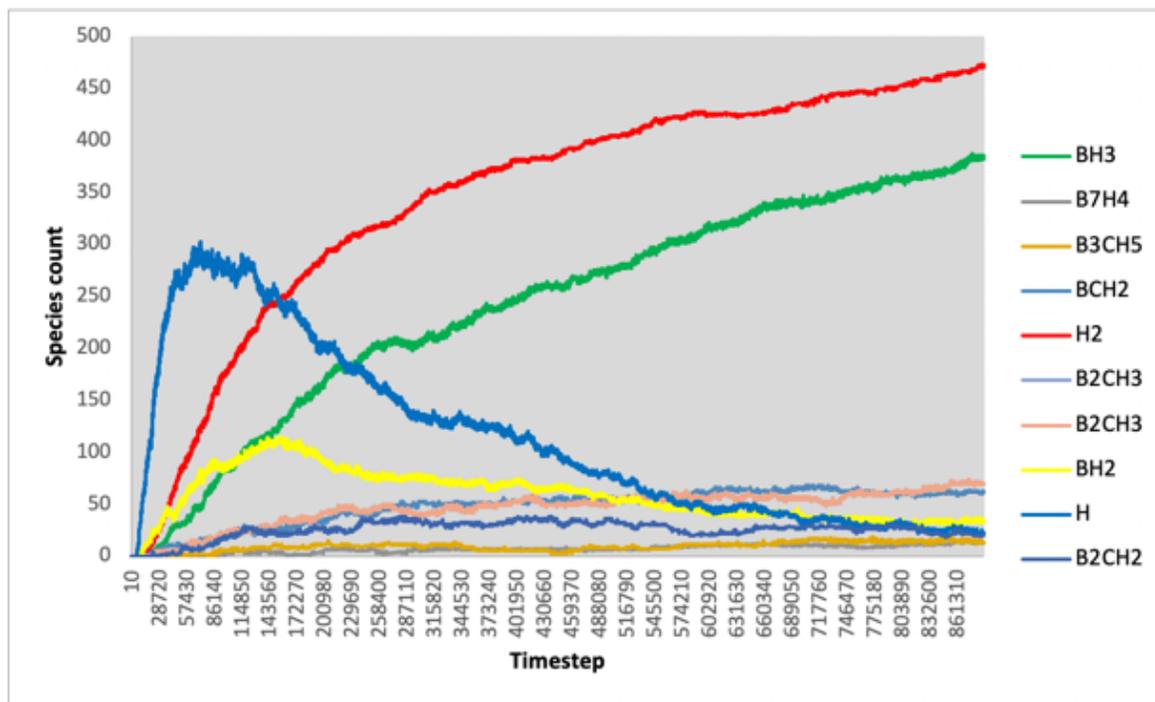
Figure 3.8 A plot of timestep vs species count for 800 orthocarboranes at 600eV

Figures 3.1-3.4 are plots of the major species produced during the bombardment

of Ar onto 100 o-carborane with the initial kinetic energy level of 50, 350, 500 and

600 eV respectively. Figures 3.5-3.8 show the number of species generated

from Ar bombardment also with the same range of energy onto 800 o-carbonare.   A number

of observations can be made based on these figures:

In general, the species with the large counts produced by 100 o-carborane are made of

free radicals especially at the high energy levels. Figures 3.1-3.4, all for 100 o-carborane cases,

show the larger counts belong to H, B and BH radicals. Even for the case of 50 eV, I can still see

the H radicals form in competition with BH_3 neutral gas. This can be understood by the fact that

at low concentration ratio, the Ar gas can retain its kinetic energy for a longer time period as it

only encounters smaller sizes of o-carborane molecules or agglomerates. Thus, there are more

opportunities to break up bonds and this results in more radicals generated. In each case, with the

exception of the 50 eV, a steady rise of counts for B, H and additional radicals like BH is seen.

In the case of 50 eV, as the initial energy is already low, parts of the free radicals generated

started to react to one another and formed more stable molecules like BH_3. Overall, the case of

100 o-carborane shows the dominant species of free radicals produced during

the Ar bombardment.

This trend contrasts with the case of 800 o-carborane as shown in Figures 3.5-3.8 where

at longer simulation periods, the steady growth of stable molecules such as H_2 and BH_3 instead

can be seen. The formation of these stable molecules is consistent with the previous finding

by Baishnab, et.al. (Baishnab et al., 2020). Here, the large amount of o-carbonare within the

system effectively slows down the energetic Ar gases and thus the free radicals can now interact with one another forming more chemically stable molecules.

In the case of 800 o-carborane for effectively all energy levels, the rise of H free radicals initially can be seen similar to the case of 100 o-carborane. But the difference is that it only lasts for a short amount of time. Figure 3.6 gives a good example. This is the case of 800 o-carborane with the energy level of 350 eV. Initially, the gas species are dominated by the formation of H gases. However, after around 80,000 steps (20 ps), there is a drastic decline. This is due to the rapid rise in the relatively more stable gas molecules including H_2 and BH_3 (and also BH_2). Beyond this point forward, the more stable gases dominate the population. This initial increase in H was also observed by the previous study (Baishnab et al., 2020), but in the current study, the MD simulations were extended up to 800,000 – 1,000,000 steps allowing the rest of the radicals to interact and more reactions to occur.

Overall, the results of 800 o-carborane confirmed the previous finding about the major reactions producing the stable gases. This means for the more concentrated system like that of 800 o-carborane system or 1000 o-carborane system, it will be observed that the majority of the species are comprised of these steady molecules.

Unlike the previous study, this study however also found out that when the concentration ratio is lowered on the other hand, more free radical population can be observed in the population. In fact, stable gases like H_2 or BH_3 are much less than in population. This was not previously observed in the prior study. Figures 3.9 and 3.10 show plots of timestep vs species count for H at 100 orthocarboranes and BH_3 at 800 orthocarboranes at variable energies.

39
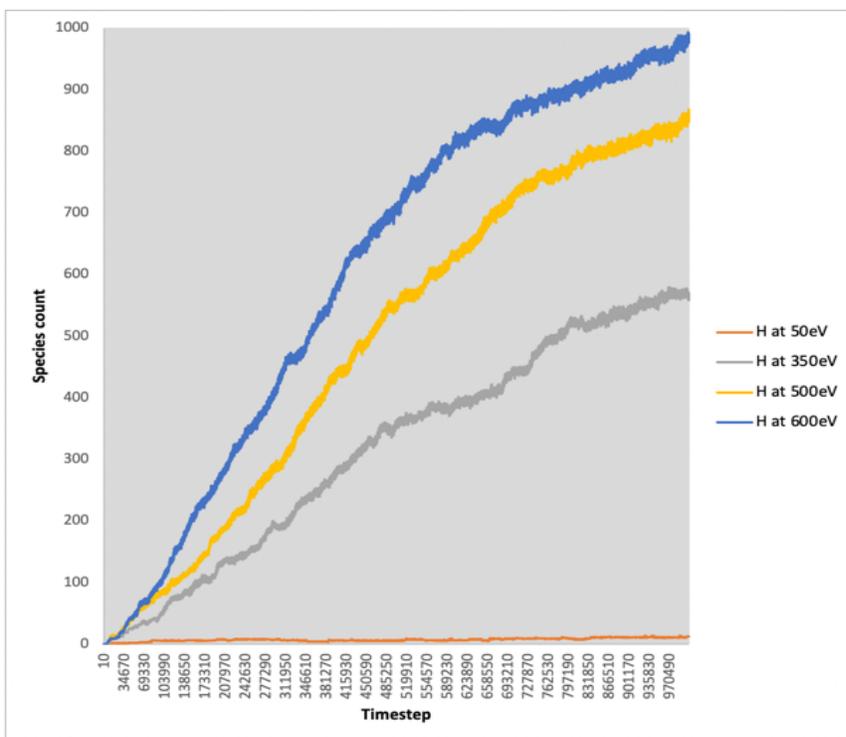
Figure 3.9 A plot of timestep vs species count for H at 100 orthocarboranes and variable energies.
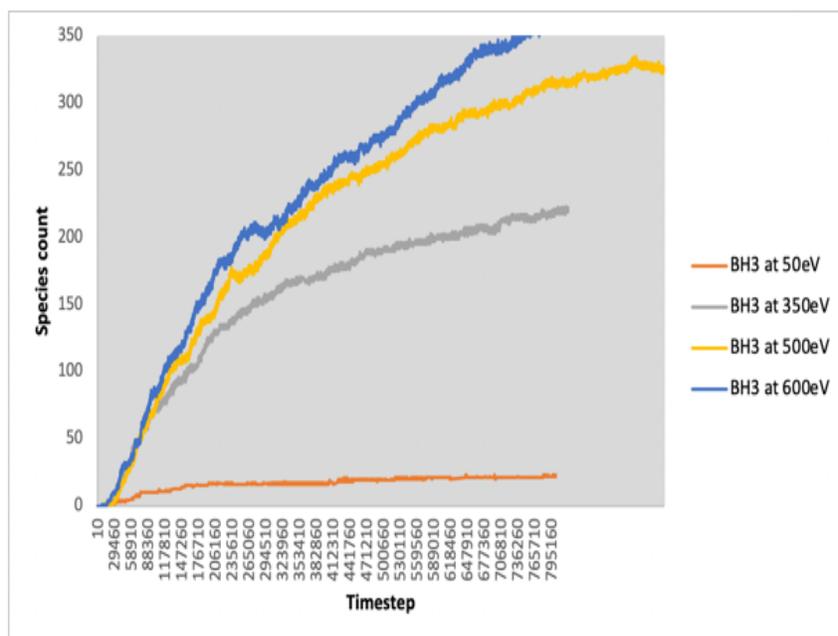
Figure 3.10 A plot of timestep vs species count for BH$_3$ at 800 orthocarboranes and variable energies.

Figures 3.11 and 3.12 show plots of timestep vs species count for $B_{10}C_2H_{12}$ at 100 and 800 orthocarboranes with variable energies
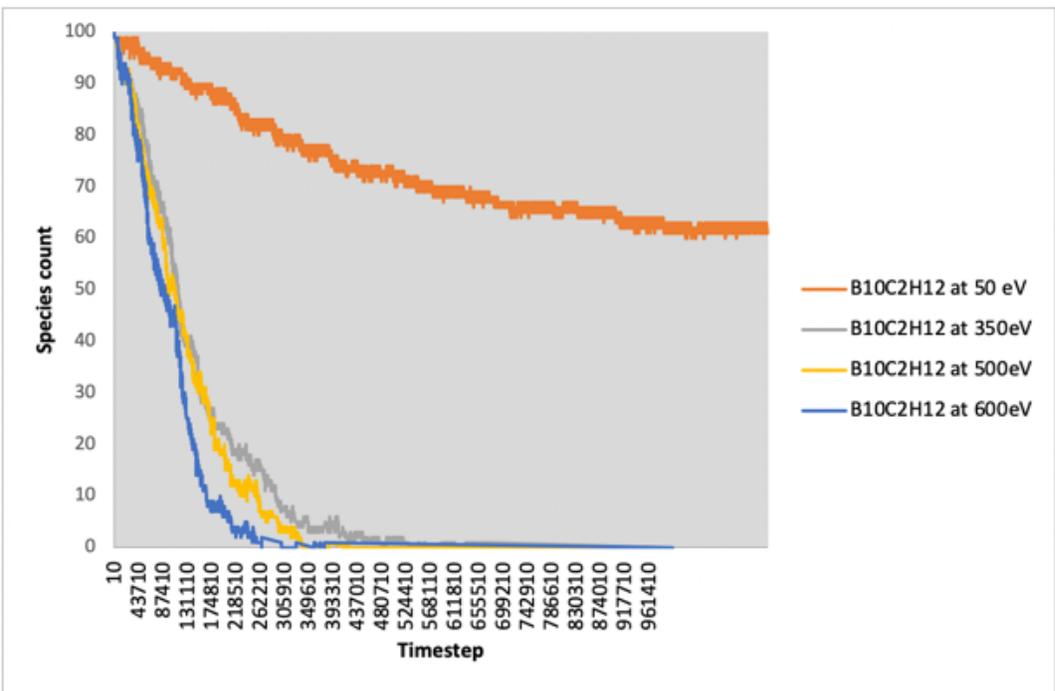
Figure 3.11 A plot of timestep vs species count for $B_{10}C_2H_{12}$ at 100 orthocarboranes and variable energies.
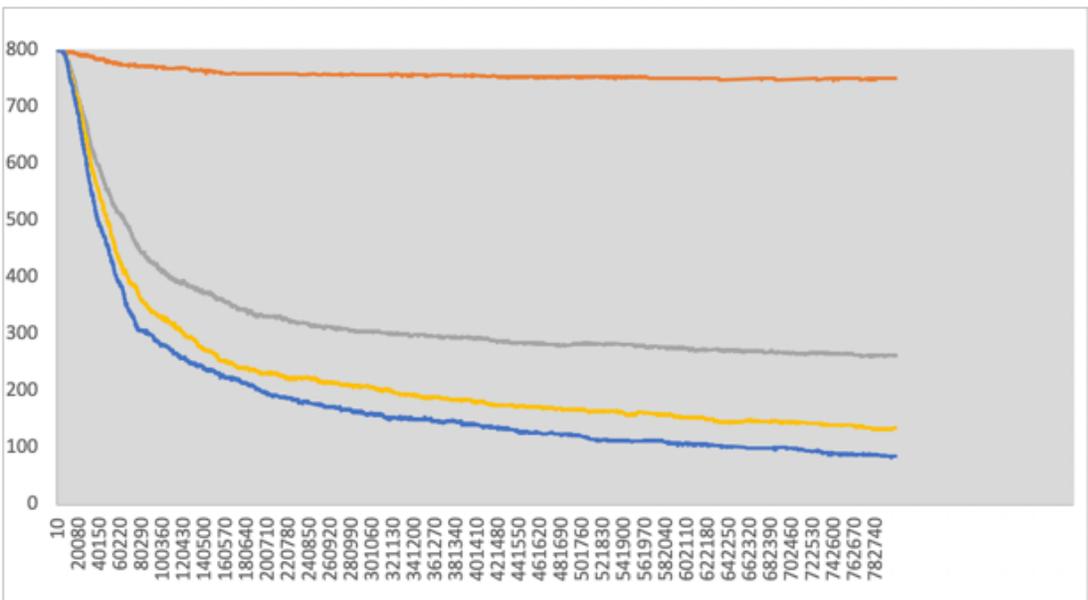
Figure 3.12 A plot of timestep vs species count for $B_{10}C_2H_{12}$ at 800 orthocarboranes and variable energies

Figures 3.9 - 3.12 are plots of the dominant species for each concentration of orthocarborane at various temperatures. Figures 3.9 and 3.10 are plots to compare the species counts of H and BH_3 with respect to time at various energy levels. From the plots, it can be concluded that there is generally a significant drop in species count as the energies increase. This is mainly due to the vigorous activity of the Ar since it becomes more active at higher energy levels and therefore knocks off more atoms which in turn produces more radicals. This is the same case for the o-carborane count as well (Fig 3.11 and 3.12). At 50eV, the Ar atoms are not as active since they have low energy and therefore the number of o-carborane knocked off are not as many as it is in the case of 600eV.

3.2 Machine Learning Analysis on Species

To gain more insights into the dynamics of the gas population, I performed Machine Learning analysis as implemented in the Weka code. I implemented the M5P algorithm for the classification. In short, the algorithm combines the best decision tree model and from which specific linear regression models can be assigned. This approach is suitable to describe the chemical reactions as the counts of species generated can be highly dependent of one another. For simplicity, I only focused on applying ML algorithm to predict the amount of the major products produced in both cases, namely the free radicals of H and/or B for the 100 o-carborane case and the stable gases of H_2 and/or BH_3. From the previous figures, I can see that the free radicals of H and B grow quite concurrently. Similarly, the number of stable gases of H_2 and BH_3 also increases in the same manner. Thus, there is a strong interdependency between these pairs of gases. Thus, in the ML models, I made two separate sub models; one with both gases present and another when only one of them (e.g., only H_2 without the BH_3) is present. The

main idea here is to test the interdependency of the quantity of these gases with the rest of the

radical species. All models, unless mentioned otherwise, use 10-fold cross validation. In figure

3.13 below, I have a plot of Predicted H with B (100 o-carborane, 600eV)
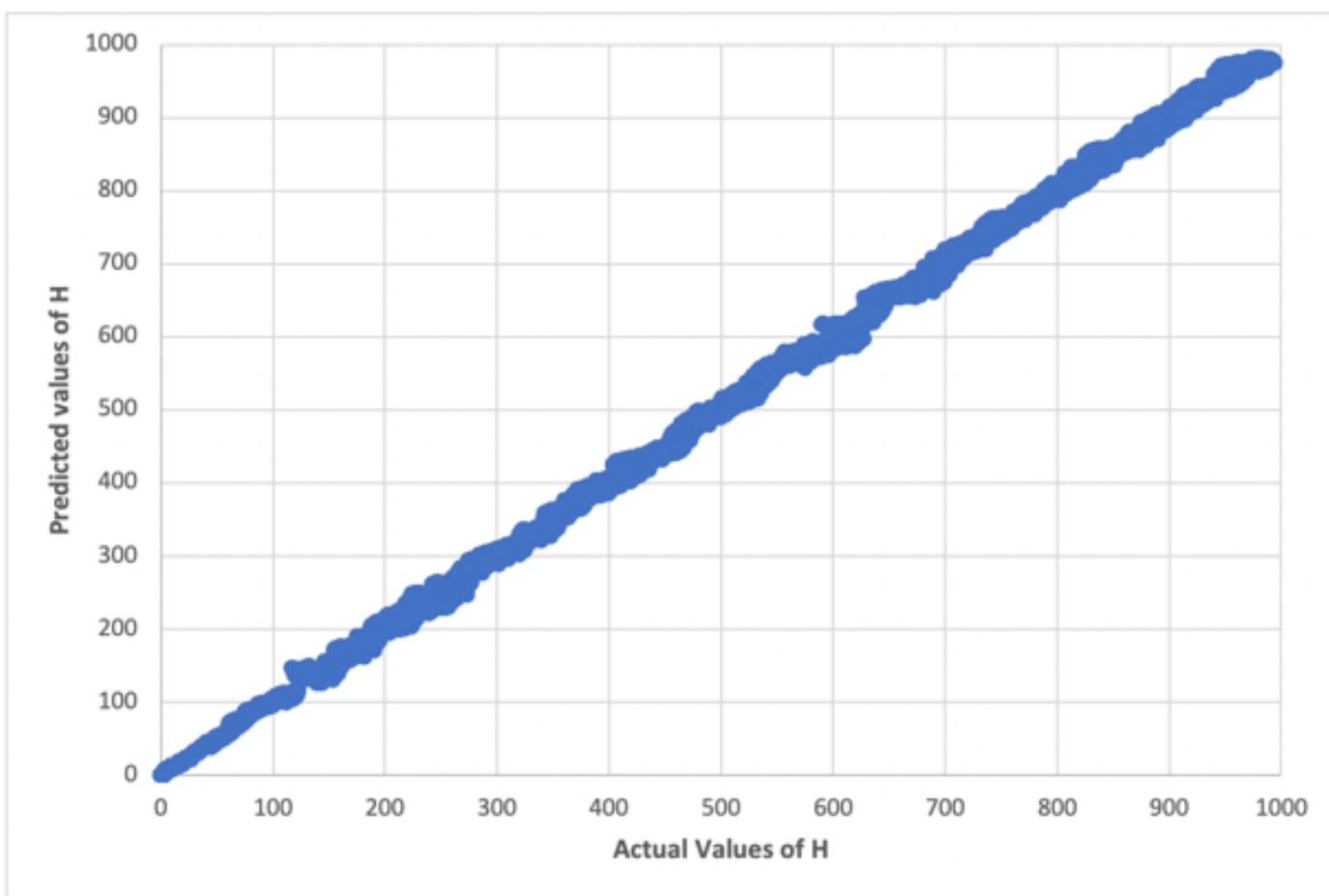
Figure 3.13 A plot of Predicted H with B (100 o-carborane, 600eV)

In figure 3.14 I show a plot of Predicted H without B (100 o-carborane, 600eV)
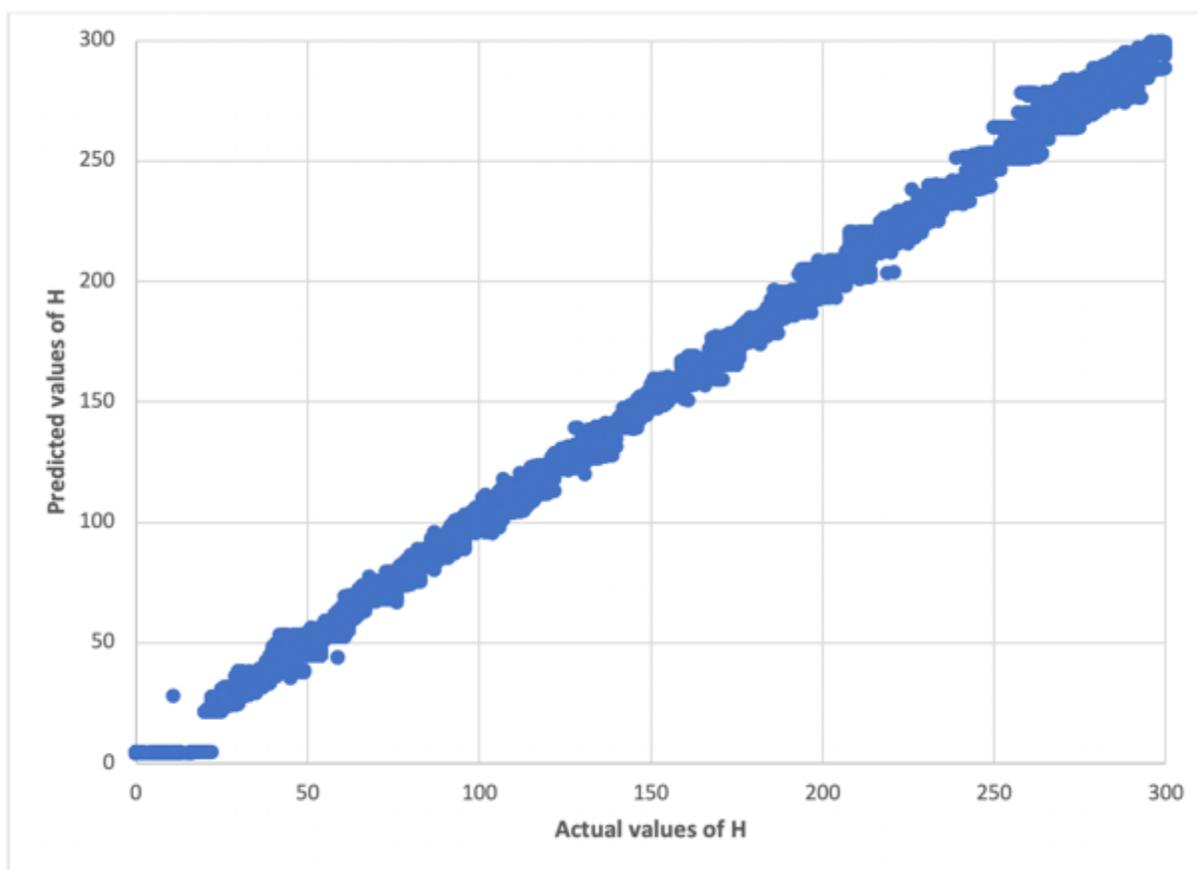
Figure 3.14 A plot of actual vs predicted values of H without B (100 o-carborane, 600eV)

Figures 3.13 and 3.14 show the comparison between the actual versus predicted for the case of free radical H from the 100 o-carborane case with 600 eV. In both cases, the results are very good with a correlation factor of close 1.0. The appendix lists the M5P model developed for H radicals. Here, it can be seen that indeed the H content is highly dependent of the amount of B when B is included as a free variable. This is expected due to the apparent concurrent growth

observed. When the B variable is removed, it can be seen that the M5P model can still predict the amount of H radical quite accurately, revealing the dependency of H radical amount to firstly the BC radical and secondly to either H_2 or C. This finding supports the fact that in the case of the 100 o-carborane, the formation of the radicals like H is highly dependent on the rest of radicals including those that combine B and C such as BC or BCH. In figures 3.15 and 3.16, I show plots of actual vs predicted values of H_2 with and without BH_3 (800 o-carborane, 600eV)
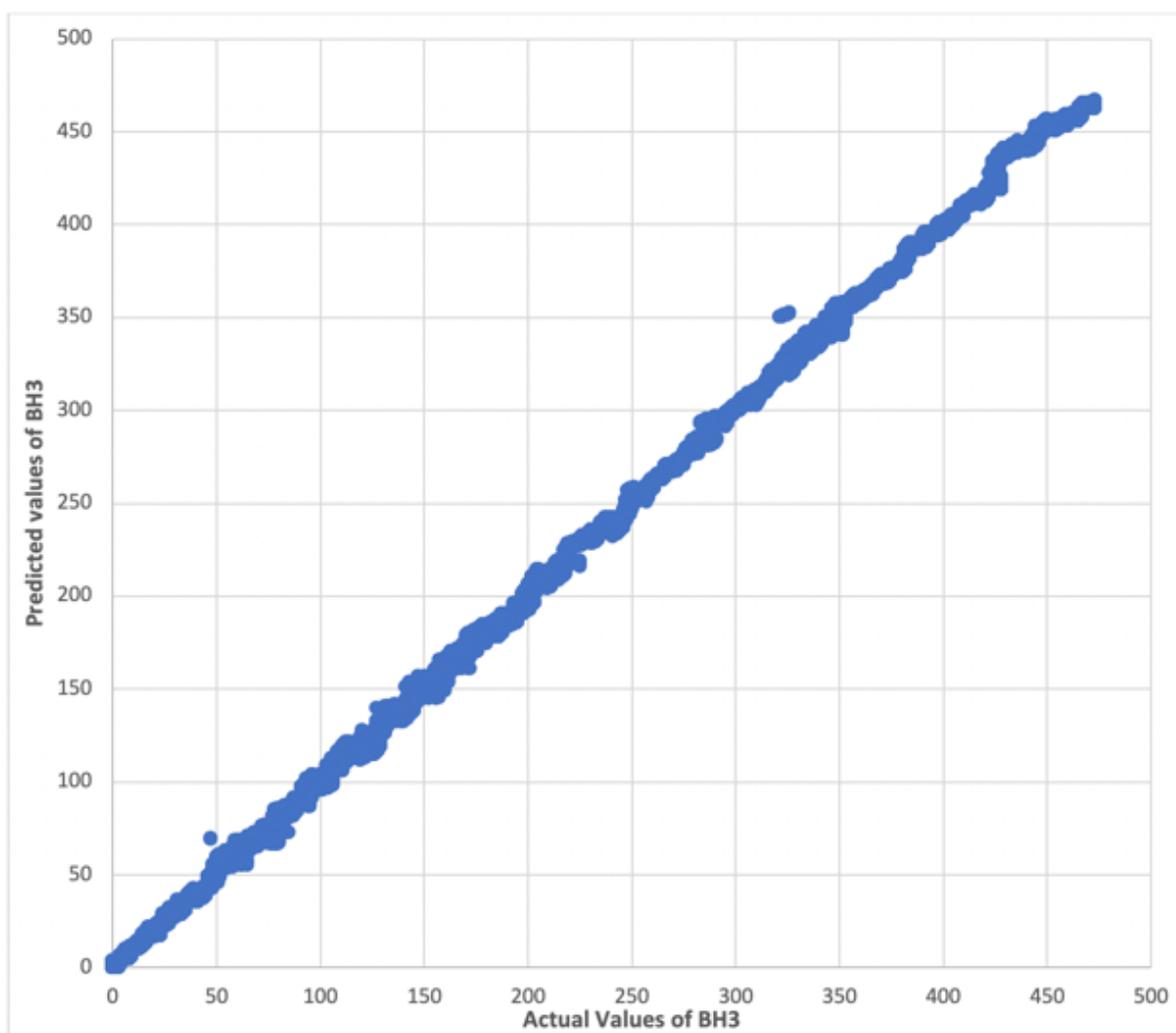
Figure 3.15 A plot of actual vs predicted values of H_2 with BH_3 (800 o-carborane

45
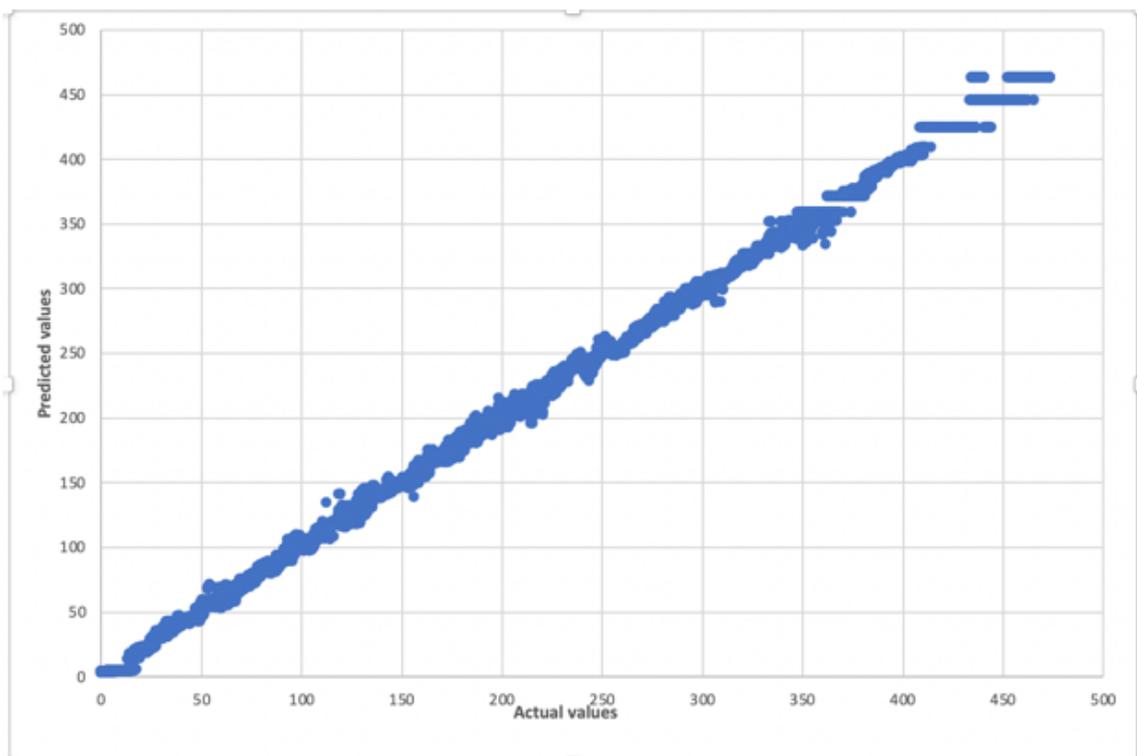
Figure 3.16 A plot of actual vs predicted values of H_2 without BH_3 (800 0-carborane, 600eV)

Figures 3.15 and 3.16 show the comparison between the actual versus predicted for the case of stable molecules of H_2 with or without the inclusion of BH_3 as a variable. Not surprisingly, when BH_3 is added, as also detailed in Appendix, it can be seen the M5P model is consisted of BH_3 as a major determining factor. This is consistent with what is expected based on the similar type of growth of the two species. When the BH_3 is removed on the other hand, a more complex M5P model is seen. The model shows the decision tree is determined now by the amount of the mixed B and C radicals such as BCH_2, H and a series of BHx radicals. While the role of H and BHx radicals are quite expected, the major role of mixed radicals was not initially expected. This suggests that even with the highly concentrated system like 800 o-carborane, there is still dependency toward mixed radicals. Similar to the case of 100 o-carborane, they are

quite hidden/masked because of the dominant effect of either B or BH_3. ML algorithms allow us to see more insights into the dynamics of the gaseous formations.

I also tested the M5P models toward the H radicals for the case of 800 o-carborane. As shown in the previous figures, the initial increase of H radicals followed by a drop. This was suspected due to the rise in the BH_3/H_2 content in the gas population. Figure 3.15 shows the M5P model for H radicals. As expected, the decision trees are dominated by either H_2 or BH_3 contents, followed by BH_2 or BCH_2. In figures 3.17 and 3.18, I show plots of H predicted with H_2 and BH_3 included and not included using Machine Learning.
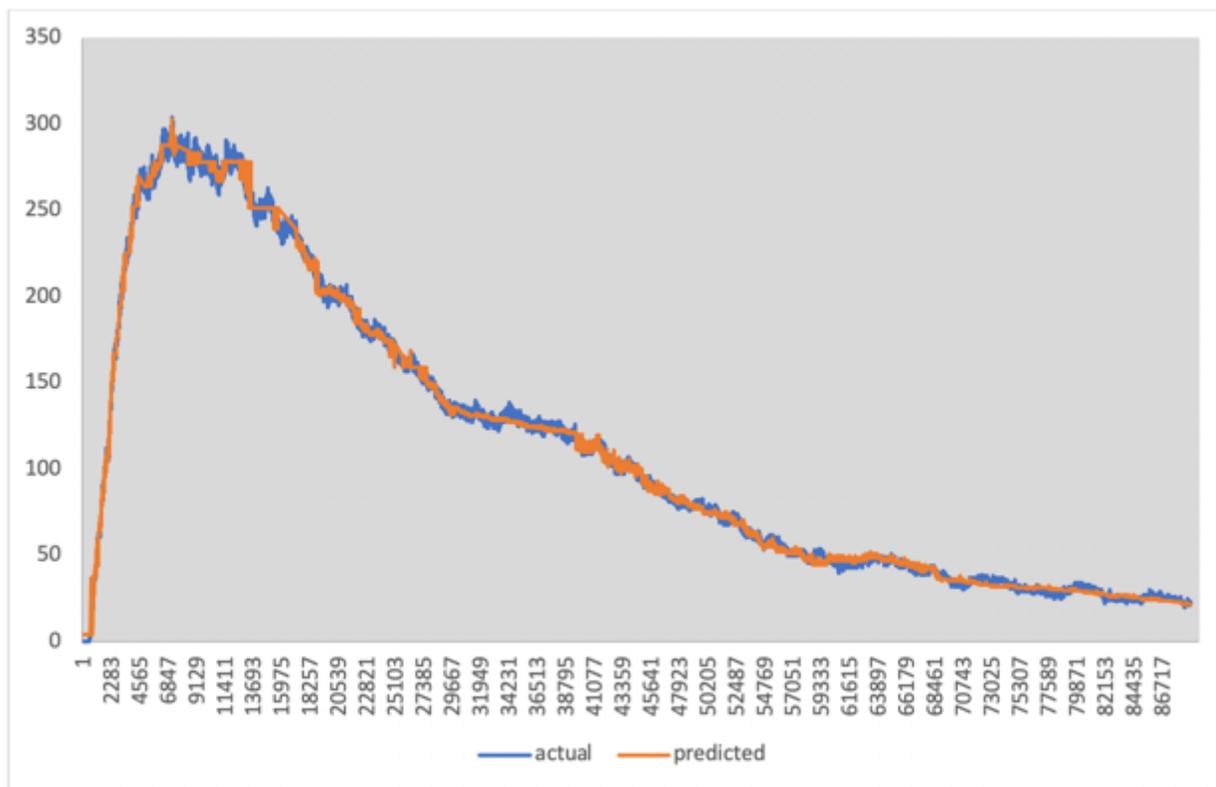
Figure 3.17 A plot of H predicted with H_2 and BH_3 included using Machine Learning
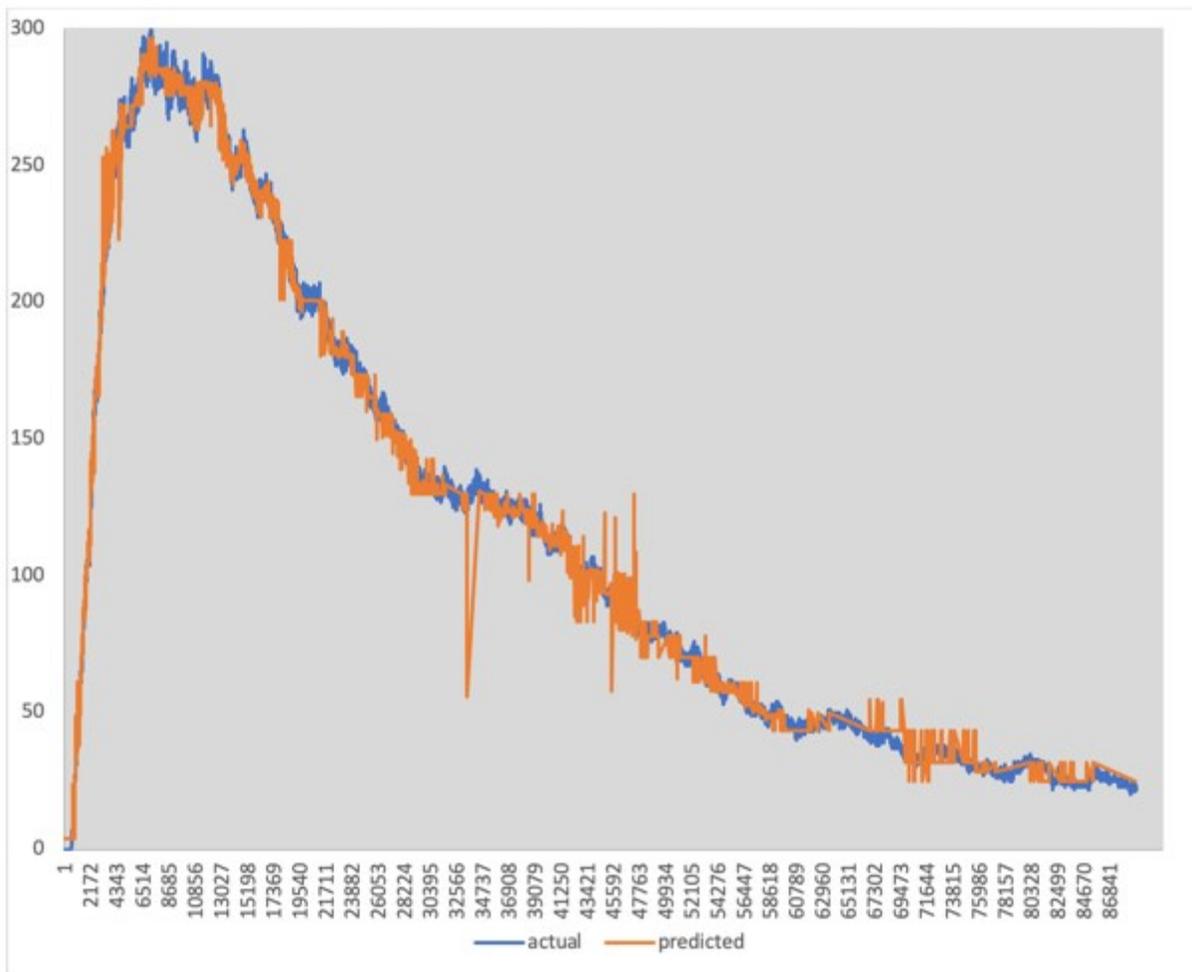
Figure 3.18 A plot of H predicted without H_2 and BH_3 using Machine Learning.

Lastly, I used the whole training set to test the M5P algorithm to model the amount of H radicals with the simulation time. Figure 3.17 is with the BH_3 and H_2 included. Figure 3.18 is when both gases are removed as variables. The first model (figure 3.17) gives the best model as expected. But the second model shows a reasonably accurate prediction as well. As shown in Appendix, a model complex model that involves mixed B and C radicals can be used to replace the more obvious variables of BH_3 and H_2 to help predict the H content.

In summary, using machine learning algorithm, quantitative models that show the interdependence of the amount of these gases can be developed. In the future, these models can

be converted into kinetic models to predict the overall chemical reactions during a complex

process such as ionic bombardments.

CONCLUSION

I was successful in modeling the argon bombardment process using ReaxFF as the potential force field using MD simulations. Enhancing on previous works done on this project, I run the simulation relatively longer (almost ten times). This enabled us to see and successfully track radical species that mainly causes the densification process and also those that showed up after a longer time. Again, modeling the simulation at lower energies that represent experimental setting has given us a fair view of how the interaction will happen experimentally. The simulation gives a fair understanding on the role hydrogen content plays in the densification process which in turn affects the mechanical properties of thin films. Since the energy of argon has an effect on the quantity of hydrogen gas formed and the entire densification process at large, one can control the hydrogen content in thin films by choosing the right initial power input. Machine Learning tools were also used in tracking and predicting the codependency of the radical species on one another. It was used in predicting the radical species which would have been difficult to just eyeball. It had a very accurate prediction accuracy.

REFERENCES

(IAS), I. f. A. S. (2017). JUQUEEN: 3.15.3. https://www.fz-juelich.de/ias/jsc/EN/Expertise/Support/Software/Chemistry/CPMD.html?nn=1025784

Alder, B. J., & Wainwright, T. E. (1957). Phase Transition for a Hard Sphere System. Journal of Chemical Physics, 27, 1208-1209. https://doi.org/10.1063/1.1743957

Andersen, H. C. (1980). Molecular dynamics simulations at constant pressure and/or temperature. The Journal of Chemical Physics, 72(4), 2384-2393. https://doi.org/10.1063/1.439486

Baishnab, N., Khadka, R., Paquette, M. M., Rulis, P., Oyler, N. A., Hwang, J., & Sakidja, R. (2020). Role of generated free radicals in synthesis of amorphous hydrogenated boron carbide from orthocarborane using argon bombardment: a ReaxFF molecular dynamics study. Materials Research Express, 6(12), 126461. https://doi.org/10.1088/2053-1591/ab7b2c

Bandyopadhyay, A. K., Beuneu, F., Zuppiroli, L., & Beauvy, M. (1984). The role of free carbon in the transport and magnetic properties of boron carbide. Journal of Physics and Chemistry of Solids, 45(2), 207-214. https://doi.org/https://doi.org/10.1016/0022-3697(84)90120-3

Baughman, R. H. (1970). NMR, Calorimetric, and Diffraction Study of Molecular Motion in Crystalline Carboranes. The Journal of Chemical Physics, 53(10), 3781-3789. https://doi.org/10.1063/1.1673841

Bigdeloo, J., & Hadian, A. (2009). Synthesis of High Purity Micron Size Boron Carbide Powder from B2O3/C Precursor. POSTER PAPER International Journal of Recent Trends in Engineering, 1.

Billa, R. B., Hofmann, T., Schubert, M., & Robertson, B. W. (2009). Annealing effects on the optical properties of semiconducting boron carbide. Journal of Applied Physics, 106(3), 033515. https://doi.org/10.1063/1.3190679

51

Blevins, J., & Yang, G. (2020). Enabling Ga2O3's neutron detection capability with boron doping and conversion layer. Journal of Applied Physics, 128(15), 155706. https://doi.org/10.1063/5.0015522

Brommer, P., & Gähler, F. (2007). Potfit: effective potentials fromab initiodata. Modelling and Simulation in Materials Science and Engineering, 15(3), 295-304. https://doi.org/10.1088/0965-0393/15/3/008

Car, R., & Parrinello, M. (1985). Unified Approach for Molecular Dynamics and Density-Functional Theory. Physical Review Letters, 55(22), 2471-2474. https://doi.org/10.1103/PhysRevLett.55.2471

Duff, A. I., Finnis, M. W., Maugis, P., Thijsse, B. J., & Sluiter, M. H. F. (2015). MEAMfit: A reference-free modified embedded atom method (RF-MEAM) energy and force-fitting code. Computer Physics Communications, 196, 439-445. https://doi.org/https://doi.org/10.1016/j.cpc.2015.05.016

Expanding the limits of coputational chemistry. Retrieved July 18, 2021 from https://gaussian.com/

Gamba, Z., & Powell, B. M. (1996). The condensed phases of carboranes. The Journal of Chemical Physics, 105(6), 2436-2440. https://doi.org/10.1063/1.472111

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G. L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P., & Wentzcovitch, R. M. (2009). QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. Journal of Physics: Condensed Matter, 21(39), 395502. https://doi.org/10.1088/0953-8984/21/39/395502

Gonze, X., Beuken, J.-M., Caracas, R., Detraux, F., Fuchs, M., Rignanese, G. M., Sindic, L., Verstraete, M., Zerah, G., Jollet, F., Torrent, M., Roy, A., Mikami, M., Ghosez, P., Raty, J.-Y., & Allan, D. C. (2002). First-principles computation of material properties: The ABINIT software project. Computational Materials Science, 25, 478-492. https://doi.org/10.1016/S0927-0256(02)00325-7

Iskandarov, A. M., Dmitriev, S. V., & Umeno, Y. (2011). Temperature effect on ideal shear strength of Al and Cu. Physical Review B, 84(22), 224118. https://doi.org/10.1103/PhysRevB.84.224118

Jensen, B. D., Bandyopadhyay, A., Wise, K. E., & Odegard, G. M. (2012). Parametric Study of ReaxFF Simulation Parameters for Molecular Dynamics Modeling of Reactive Carbon Gases. Journal of Chemical Theory and Computation, 8(9), 3003-3008. https://doi.org/10.1021/ct300491d

Kearley, G. J., Johnson, M. R., & Tomkinson, J. (2006). Intermolecular interactions in solid benzene. The Journal of Chemical Physics, 124(4), 044514. https://doi.org/10.1063/1.2145926

Kresse, G., & Furthmüller, J. (1996a). Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. Computational Materials Science, 6(1), 15-50. https://doi.org/https://doi.org/10.1016/0927-0256(96)00008-0

Kresse, G., & Furthmüller, J. (1996b). Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. Physical Review B, 54(16), 11169-11186. https://doi.org/10.1103/PhysRevB.54.11169

Kresse, G., & Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. Physical Review B, 59(3), 1758-1775. https://doi.org/10.1103/PhysRevB.59.1758

LAMMPS documentation. Retrieved July 18, 2021 from https://docs.lammps.org/minimize.html

Leites, L. A. (1992). Vibrational spectroscopy of carboranes and parent boranes and its capabilities in carborane chemistry. Chemical Reviews, 92(2), 279-323. https://doi.org/10.1021/cr00010a006

Li, N.-L., Wu, W.-P., & Nie, K. (2018). Molecular dynamics study on the evolution of interfacial dislocation network and mechanical properties of Ni-based single crystal superalloys. Physics Letters A, 382(20), 1361-1367. https://doi.org/https://doi.org/10.1016/j.physleta.2018.03.031

Multiscale Modeling and Analysis for Materials Simulation. (2011). (Vol. Volume 22) [doi:10.1142/8212]. WORLD SCIENTIFIC. https://doi.org/doi:10.1142/8212

Nordell, B. J., Karki, S., Nguyen, T. D., Rulis, P., Caruso, A. N., Purohit, S. S., Li, H., King, S. W., Dutta, D., Gidley, D., Lanford, W. A., & Paquette, M. M. (2015). The influence of hydrogen on the chemical, mechanical, optical/electronic, and electrical transport properties of amorphous hydrogenated boron carbide. Journal of Applied Physics, 118(3), 035703. https://doi.org/10.1063/1.4927037

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., & Schulten, K. (2005). Scalable molecular dynamics with NAMD [https://doi.org/10.1002/jcc.20289]. Journal of Computational Chemistry, 26(16), 1781-1802. https://doi.org/https://doi.org/10.1002/jcc.20289

Plimpton, S. (1995). Fast Parallel Algorithms for Short-Range Molecular Dynamics. Journal of Computational Physics, 117, 1-19. https://doi.org/10.1006/jcph.1995.1039

Ramesh, K. T. (2008). High Rates and Impact Experiments. In W. N. Sharpe (Ed.), Springer Handbook of Experimental Solid Mechanics (pp. 929-960). Springer US. https://doi.org/10.1007/978-0-387-30877-7_33

Raucoules, R., Vast, N., Betranhandy, E., & Sjakste, J. (2011). Mechanical Properties of Icosahedral Boron Carbide Explained from First Principles. Physical Review B, 84, 31012. https://doi.org/10.1103/PhysRevB.84.014112

Sandoval, L., & Urbassek, H. M. Finite-size effects in Fe-nanowire solid-solid phase transitions: a molecular dynamics approach. (1530-6992 (Electronic)).

Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials, 5(1), 83. https://doi.org/10.1038/s41524-019-0221-0

Schulz, D., Lutfurakhmanov, A., Mayo, B., Sandstrom, J., Bunzow, D., Qadri, S., Ruqiang, B., Chrisey, D., & Caruso, A. (2008). Characterization of a-B5C:H prepared by PECVD of orthocarborane: Results of preliminary FTIR and nuclear reaction analysis studies. Journal of Non-Crystalline Solids, 354, 2369-2371.

Senftle, T. P., Hong, S., Islam, M. M., Kylasa, S. B., Zheng, Y., Shin, Y. K., Junkermeier, C., Engel-Herbert, R., Janik, M. J., Aktulga, H. M., Verstraelen, T., Grama, A., & van Duin, A. C. T. (2016). The ReaxFF reactive force-field: development, applications and future directions. npj Computational Materials, 2(1), 15011. https://doi.org/10.1038/npjcompumats.2015.11

Soler, J. M., Artacho, E., Gale, J. D., García, A., Junquera, J., Ordejón, P., & Sánchez-Portal, D. (2002). The SIESTA method forab initioorder-Nmaterials simulation. Journal of Physics: Condensed Matter, 14(11), 2745-2779. https://doi.org/10.1088/0953-8984/14/11/302

Stadler, J., Mikulla, R., & Trebin, H. R. (1997). IMD: A Software Package for Molecular Dynamics Studies on Parallel Computers. International Journal of Modern Physics C, 08(05), 1131-1140. https://doi.org/10.1142/S0129183197000990

Suri, A. K., Subramanian, C., Sonber, J. K., & Murthy, T. S. R. C. (2010). Synthesis and consolidation of boron carbide: a review. International Materials Reviews, 55(1), 4-40. https://doi.org/10.1179/095066009X12506721665211

Thévenot, F. (1990). Boron carbide—A comprehensive review. Journal of the European Ceramic Society, 6(4), 205-225. https://doi.org/https://doi.org/10.1016/0955-2219(90)90048-K

van Duin, A. C. T., Dasgupta, S., Lorant, F., & Goddard, W. A. (2001). ReaxFF: A Reactive Force Field for Hydrocarbons. The Journal of Physical Chemistry A, 105(41), 9396-9409. https://doi.org/10.1021/jp004368u

Watanabe, H., Suzuki, M., & Ito, N. (2011). Efficient Implementations of Molecular Dynamics Simulations for Lennard-Jones Systems. Progress of Theoretical Physics, 126(2), 203-235. https://doi.org/10.1143/PTP.126.203

Watanabe, H., Suzuki, M., & Ito, N. (2013). Huge-scale molecular dynamics simulation of multibubble nuclei. Computer Physics Communications, 184(12), 2775-2784. https://doi.org/https://doi.org/10.1016/j.cpc.2013.07.023

Wen, Z. X., Wang, J. P., Wu, Y. W., Zhou, K. J., & Yue, Z. F. (2018). Atomistic simulation analysis of the effects of void interaction on void growth and coalescence in a metallic system. Current Applied Physics, 18(6), 744-751. https://doi.org/https://doi.org/10.1016/j.cap.2018.03.009

Wood, C., & Emin, D. (1984). Conduction mechanism in boron carbide. Physical Review B, 29(8), 4582-4587. https://doi.org/10.1103/PhysRevB.29.4582

Zhang, D., McIlroy, D. N., O'Brien, W. L., & De Stasio, G. (1998). The Chemical and Morphological Properties of Boron–Carbon Alloys Grown by Plasma-Enhanced Chemical Vapour Deposition. Journal of Materials Science, 33(20), 4911-4915. https://doi.org/10.1023/A:1004422016254

LAMMPS input scripts

Structural minimization scripts

units real

boundary   p p p

atom_style charge

read_data 100_ortho_25_argon.BCH

##hybrid potentials

##ReaxFF

Potential

pair_style hybrid reax/c lmp_control lj/cut 11.0

pair_coeff * * reax/c ffield.reax.hcb B C H NULL

##lj

pair_coeff 1 4 lj/cut 0.150885 3.4265

pair_coeff 2 4 lj/cut 0.115303 3.385

pair_coeff 3 4 lj/cut 0.13283 3.054

pair_coeff 4 4 lj/cut 0.23983 3.4

mass 1 10.811

mass 2 12.0107

mass 3 1.00794

mass 4 39.948

##

#########

```
group PKA type 4


velocity  all create 2.0 893267 rot yes dist gaussian


neighbor  0.3 bin

neigh_modify    every 1 delay 0 check yes

restart 5000 restart.*.dens17


#Outputs and Run

thermo   100

thermo_style custom step temp pe etotal press vol pxx pyy pzz lx ly lz xy xz yz


dump      1 all custom 100 dump_total_600_eV.reax.amorphous_b4cH id type q x y z


fix       1 all qeq/reax 1 0.0 10.0 1e-6 param_bch.qeq
```

```
fix          4 all reax/c/species 1 1 10 species_600_eV.out element B C H Ar cutoff 1 4 1 cutoff 2

4 1 cutoff 3 4 1 cutoff 4 4 1

fix          2 all nvt temp 2 2 100.0

timestep     0.05


run  5000


#Starting PKA run

#First unfix

unfix        2



############################################################

group PKA1 id == 2401

group PKA2 id == 2402

group PKA3 id == 2403
```

group PKA4 id == 2404

group PKA5 id == 2405

group PKA6 id == 2406

group PKA7 id == 2407

group PKA8 id == 2408

group PKA9 id == 2409

group PKA10 id == 2410

group PKA11 id == 2411

group PKA12 id == 2412

group PKA13 id == 2413

group PKA14 id == 2414

group PKA15 id == 2415

group PKA16 id == 2416

group PKA17 id == 2417

group PKA18 id == 2418

group PKA19 id == 2419

group PKA20 id == 2420

group PKA21 id == 2421

group PKA22 id == 2422

group PKA23 id == 2423

group PKA24 id == 2424

group PKA25 id == 2425

##

####

velocity PKA1 set -0.093421963 0.4946317 -0.189871321 units box

velocity PKA2 set -0.139243516 0.347723005 -0.386185521 units box

velocity PKA3 set -0.399808671 0.231396691 -0.275767835 units box

velocity PKA4 set -0.514487782 -0.15490504 -0.027313939 units box

velocity PKA5 set -0.04688941 0.291898419 0.449484148 units box

velocity PKA6 set 0.235841762 -0.241773637 -0.418764221 units box

velocity PKA7 set 0.148291693 -0.28918918 0.428740591 units box

velocity PKA8 set 0.331223185 0.422913633 0.029573001 units box

velocity PKA9 set -0.201536712 -0.256911739 0.427572933 units box

velocity PKA10 set -0.331223185 0.370791828 0.205533267 units box

velocity PKA11 set -0.451201104 -0.227714197 -0.184399321 units box

velocity PKA12 set 0.37372317 0.04716364 -0.384117019 units box

velocity PKA13 set 0.465917886 -0.243794817 0.113683389 units box

velocity PKA14 set -0.47935762 0.240534278 -0.042412684 units box

velocity PKA15 set 0.435247611 0.306832639 -0.076501968 units box

velocity PKA16 set 0.456246172 0.285050829 0.00497558 units box

velocity PKA17 set 0.5376679 0.010226022 0.015746691 units box

velocity PKA18 set -0.07487452 0.027882498 0.532029767 units box

velocity PKA19 set 0.393465097 0.289130711 0.225893668 units box

velocity PKA20 set 0.338571622 -0.354569871 -0.221559845 units box

velocity PKA21 set 0.446018595 0.197370991 0.22704935 units box

velocity PKA22 set -0.111855581 0.46464158 0.247054312 units box

velocity PKA23 set -0.359989344 0.356232135 -0.181509338 units box

velocity PKA24 set -0.345816927 -0.224461303 0.345640099 units box

velocity PKA25 set -0.451201104 -0.227714197 -0.184399321 units box

##
#####

compute mype all pe/atom

##
#####

fix   2 all nve

run   100000

unfix 2

ReaxFF potential file

Reactive MD-force field

 39   ! Number of general parameters

   50.0000 !Overcoordination parameter

4.3822 !Overcoordination parameter

21.2839 !Valency angle conjugation parameter

3.0000 !Triple bond stabilisation parameter

6.5000 !Triple bond stabilisation parameter

53.9706 !C2-correction

1.0053 !Undercoordination parameter

9.0000 !Triple bond stabilisation parameter

7.6280 !Undercoordination parameter

14.5067 !Undercoordination parameter

-10.0198 !Triple bond stabilization energy

0.0000 !Lower Taper-radius

10.0000 !Upper Taper-radius

2.8793 !Not used

33.8667 !Valency undercoordination

25.6125 !Valency angle/lone pair parameter

1.1177 !Valency angle

1.9645 !Valency angle parameter

6.1431 !Not used

6.6623 !Double bond/angle parameter

0.1809 !Double bond/angle parameter: overcoord

3.9954 !Double bond/angle parameter: overcoord

-2.4837 !Not used

4.8815 !Torsion/BO parameter

10.0000 !Torsion overcoordination

2.3276 !Torsion overcoordination

-1.2327 !Conjugation 0 (not used)

1.7905 !Conjugation

1.5591 !vdWaals shielding

0.1000 !Cutoff for bond order (*100)

2.8921 !Valency angle conjugation parameter

1.6356 !Overcoordination parameter

5.6937 !Overcoordination parameter

2.5067 !Valency/lone pair parameter

0.5000 !Not used

20.0000 !Not used

5.0000 !Molecular energy (not used)

0.0000 !Molecular energy (not used)

1.6052 !Valency angle conjugation parameter

4   ! Nr of atoms; cov.r; valency;a.m;Rvdw;Evdw;gammaEEM;cov.r2;#

alfa;gammavdW;valency;Eunder;Eover;chiEEM;etaEEM;n.u.

cov r3;Elp;Heat inc.;n.u.;n.u.;n.u.;n.u.

ov/un;val1;n.u.;val3,vval4

C   1.3647 4.0000 12.0000 1.9091 0.1597 0.8712 1.2018 4.0000

   9.5729 2.7769 4.0000 35.6314 79.5548 5.7254 6.9235 0.0000

   1.2661 0.0000 -0.0526 5.0514 29.6014 11.9957 0.8563 0.0000

  -17.6107 2.9280 1.0564 4.0000 2.9663 1.6737 0.1421 14.0707

H   0.6650 1.0000 1.0080 1.6054 0.0601 0.7625 -0.1000 1.0000

   9.3943 4.3633 1.0000 0.0000 121.1250 3.8196 9.8832 1.0000

-0.1000 0.0000 -0.1611 3.5942 2.8307 1.0000 1.0698 0.0000

-18.1553 3.0626 1.0338 1.0000 2.8793 1.2669 0.0139 12.4538

B   1.2568 3.0000 10.8110 1.2701 0.1586 0.5400 0.9900 3.0000

10.0865 2.0000 3.0000 31.9142 80.0000 2.3294 5.0957 0.0000

-1.3000 0.0000 -2.3700 13.1142 3.5132 0.4765 0.0000 0.0000

-7.3948 3.5000 1.0564 3.0000 2.8413 1.8463 0.2767 16.2666

X -0.1000 2.0000 1.0080 2.0000 0.0000 1.0000 -0.1000 6.0000

10.0000 2.5000 4.0000 0.0000 0.0000 8.5000 1.5000 0.0000

-0.1000 0.0000 -2.3700 8.7410 13.3640 0.6690 0.9745 0.0000

-11.0000 2.7466 1.0338 6.2998 2.8793 0.0000 0.0000 0.0000

6   ! Nr of bonds; Edis1;LPpen;n.u.;pbe1;pbo5;13corr;pbo6

pbe2;pbo3;pbo4;n.u.;pbo1;pbo2;ovcorr

1 1 142.9877 117.7932 70.0184 0.2152 -1.0820 1.0092 50.0568 0.1436

0.1120 -0.1904 8.5003 1.0000 -0.0966 5.9567 1.0000 0.0000

1 2 167.5082 0.0000 0.0000 -0.4457 0.0000 1.0000 6.0000 0.6385

18.9826 1.0000 0.0000 1.0000 -0.0093 8.5218 0.0000 0.0000

2 2 187.8546 0.0000 0.0000 -0.3116 0.0000 1.0000 6.0000 0.6810

   8.6056 1.0000 0.0000 1.0000 -0.0186 5.7036 0.0000 0.0000

1 3 87.2046 87.2719 0.0000 0.6389 -0.5211 1.0000 18.9617 0.3999

   3.3394 -0.0684 14.3527 1.0000 -0.2034 7.1462 1.0000 0.0000

2 3 156.4418 0.0000 0.0000 -0.4687 -0.3000 1.0000 25.0000 0.5711

   4.9151 0.0000 0.0000 1.0000 -0.0448 5.0368 1.0000 0.0000

3 3 95.8586 0.0000 0.0000 0.4809 -0.2500 1.0000 25.0000 0.4061

   0.4195 -0.2000 15.0000 1.0000 -0.1125 5.2754 1.0000 0.0000

3   ! Nr of off-diagonal terms; Ediss;Ro;gamma;rsigma;rpi;rpi2

1 2 0.0456 1.7209 10.4263 1.0384 -1.0000 -1.0000

1 3 0.1479 1.6161 10.0027 1.5561 1.2444 -1.0000

2 3 0.0493 1.5271 10.7353 1.1092 -1.0000 -1.0000

11   ! Nr of angles;at1;at2;at3;Thetao,o;ka;kb;pv1;pv2;val(bo)

1 1 1 77.0860 49.1556 0.7273 0.0000 0.0933 15.5317 1.0400

1 1 2 70.3831 11.5011 7.4039 0.0000 0.0337 0.0000 1.0400

2 1 2 69.7271 12.9817 2.0676 0.0000 0.0839 0.0000 1.0400

1 2 2   0.0000   0.0000   6.0000   0.0000   0.0000   0.0000   1.0400

1 2 1   0.0000   3.2164   7.2937   0.0000   0.0000   0.0000   1.0400

2 2 2   0.0000   27.9213   5.8635   0.0000   0.0000   0.0000   1.0400

2 3 2   65.0000   25.0000   2.0000   0.0000   2.0000   0.0000   1.0400

1 1 3   61.5194   22.1710   1.6392   0.0000   2.0969   0.0000   1.0400

1 3 3   68.9652   20.0561   0.9409   0.0000   0.4993   0.0000   1.0400

3 1 3   76.7750   21.9876   0.9829   0.0000   0.1046   0.0000   1.0400

1 3 1   0.0000   23.2250   1.0000   0.0000   1.0000   0.0000   1.0400

6   ! Nr of torsions;at1;at2;at3;at4;;V1;V2;V3;V2(BO);vconj;n.u;n

1 1 1 1   0.0000   35.6556   0.2614   -6.3913   -1.7021   0.0000   0.0000

1 1 1 2   0.0000   61.9992   0.3134   -6.6967   -1.8570   0.0000   0.0000

2 1 1 2   0.0000   46.5929   0.3206   -5.5976   -1.0363   0.0000   0.0000

0 1 2 0   0.0000   0.0000   0.0000   0.0000   0.0000   0.0000   0.0000

0 2 2 0   0.0000   0.0000   0.0000   0.0000   0.0000   0.0000   0.0000

0 1 1 0   0.0000   12.4562   0.0000   -3.6133   -1.2327   0.0000   0.0000

Machine Learning decision tree for 100 o-carborane at 600eV

=== Run information ===

Scheme: weka.classifiers.trees.M5P -M 4.0 -num-decimal-places 4
Relation: bk16-weka.filters.unsupervised.attribute.Remove-R1
Instances: 100500
Attributes: 9
        B
        BH
        BC
        C
        B_2
        H_2
        BCH
        B_2C
        H
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

B <= 183.5 :
| H2 <= 32.5 :
| | BCH <= 6.5 : LM1 (10930/1.101%)
| | BCH > 6.5 : LM2 (9429/2.514%)
| H2 > 32.5 :
| | C <= 14.5 : LM3 (9334/2.042%)
| | C > 14.5 : LM4 (12045/2.575%)
B > 183.5 :
| B <= 385.5 :
| | B <= 270.5 : LM5 (9425/2.659%)
| | B > 270.5 : LM6 (11496/1.709%)
| B > 385.5 :
| | B <= 527.5 : LM7 (20692/2.106%)
| | B > 527.5 : LM8 (17149/2.068%)

LM num: 1
H =
        1.5426 * B
        + 0.8425 * BH
        + 0.0021 * BC
        + 0.001 * C
        + 0.0061 * H_2

+ 0.0133 * BCH
+ 0.0012 * B2C
+ 0.6335

LM num: 2
H =
0.0017 * B
+ 1.7998 * BH
+ 0.0021 * BC
+ 0.001 * C
+ 2.2504 * H_2
+ 0.015 * BCH
+ 4.9219 * B2C
+ 13.3895

LM num: 3
H =
1.644 * B
+ 1.5168 * BH
+ 1.2578 * BC
+ 0.0004 * C
+ 0.0012 * H2
+ 0.002 * BCH
+ 0.0022 * B2C
+ 19.4005

LM num: 4
H =
1.4436 * B
+ 0.7572 * BH
+ 2.2683 * BC
+ 0.0005 * C
+ 0.0012 * H_2
+ 0.002 * BCH
+ 0.0022 * B2C
+ 130.586

LM num: 5
H =
1.1937 * B
+ 0.0002 * BH
+ 0.0004 * BC
- 0 * C
+ 0.0006 * H2
+ 0.0007 * BCH
+ 0.0003 * B2C

+ 397.3597

LM num: 6
H =
      0.5827 * B
      + 0.0002 * BH
      + 0.0004 * BC
      - 0 * C
      - 1.9897 * H2
      + 0.0007 * BCH
      + 0.0003 * B2C
      + 706.0909

LM num: 7
H =
      0.5943 * B
      + 0.0002 * BH
      + 0.0004 * BC
      - 0 * C
      + 0.0005 * H2
      + 0.0007 * BCH
      + 0.0003 * B2C
      + 595.0817

LM num: 8
H =
      0.5994 * B
      + 0.0002 * BH
      + 0.0004 * BC
      - 0 * C
      + 0.0005 * H2
      + 0.0007 * BCH
      + 0.0003 * B2C
      + 593.7369

Number of Rules : 8

Time taken to build model: 1.18 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient   0.9998
Mean absolute error   5.1581
Root mean squared error   6.6076
Relative absolute error   1.9044 %

Root relative squared error   2.145 %
Total Number of Instances   100500
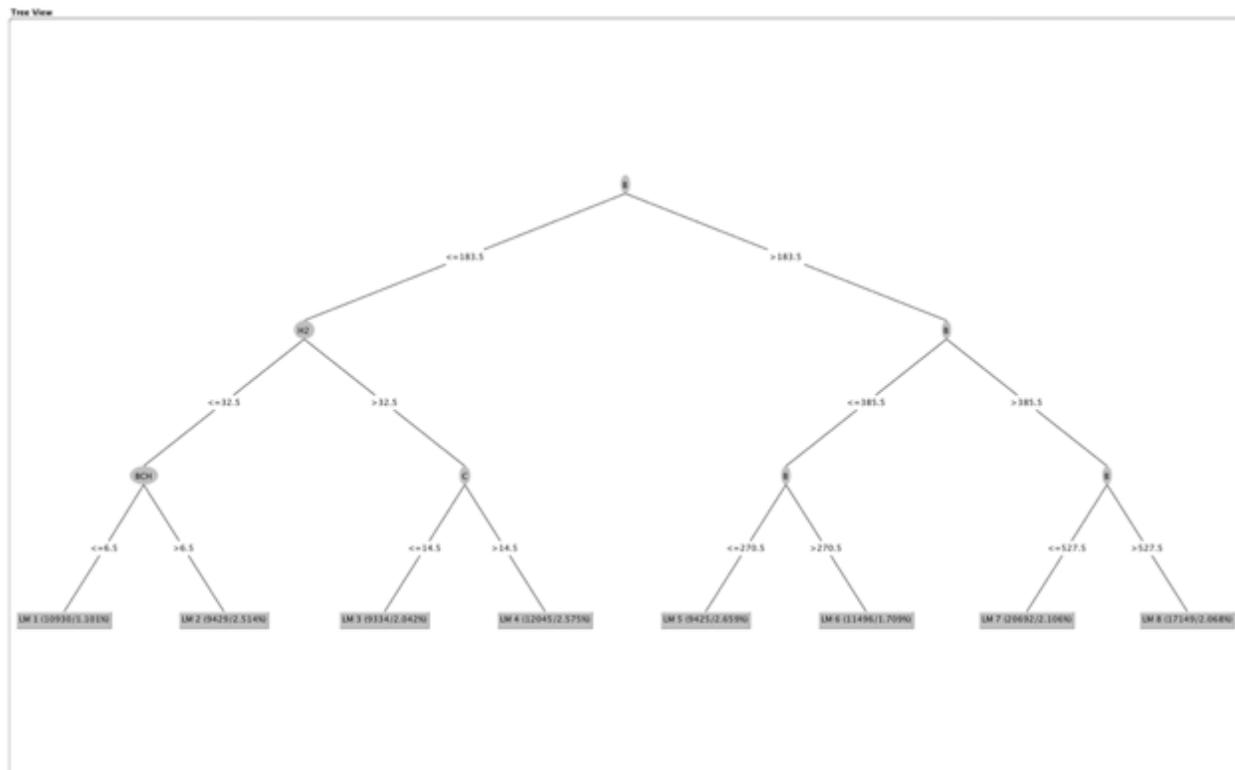
Machine Learning decision tree for 800 o-carborane at 600eV

=== Run information ===

Scheme: weka.classifiers.trees.M5P -M 4.0 -num-decimal-places 4
Relation: Book211-weka.filters.unsupervised.attribute.Remove-R1
Instances: 88974
Attributes: 9
        BH3
        B7H4
        B3CH5
        BCH2
        H2
        B2CH3
        BH$_2$
        H
        B2CH2
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

BH3 <= 209.5 :
| BH3 <= 105.5 :
| | BH_2 <= 48.5 :
| | | BH3 <= 8.5 : LM1 (2716/1.498%)
| | | BH3 > 8.5 :
| | | | BH3 <= 16.5 : LM2 (1015/2.047%)
| | | | BH3 > 16.5 : LM3 (1243/2.564%)
| | BH_2 > 48.5 :
| | | BH3 <= 65.5 : LM4 (3368/2.724%)
| | | BH3 > 65.5 : LM5 (3983/3.059%)
| BH3 > 105.5 :
| | BH3 <= 166.5 :
| | | BH3 <= 141.5 : LM6 (4580/3.134%)
| | | BH3 > 141.5 : LM7 (3195/1.529%)
| | BH3 > 166.5 :
| | | H <= 151.5 : LM8 (3662/2.602%)
| | | H > 151.5 : LM9 (7458/1.66%)
BH3 > 209.5 : LM10 (57754/3.433%)

LM num: 1
H2 =
        0.0164 * BH3
        + 0.001 * B2CH3
        + 0.0057 * BH_2
        + 0.1105 * H
        + 0.0013 * B2CH2
        - 1.6721

LM num: 2
H2 =
        0.0426 * BH3
        + 0.001 * B2CH3
        + 0.0112 * BH_2
        + 0.3092 * H
        + 0.0013 * B2CH2
        - 32.2573

LM num: 3
H2 =
        2.3019 * BH3
        + 0.001 * B2CH3
        + 0.9756 * BH_2
        + 0.0006 * H

+ 0.0013 * B2CH2
- 40.003

LM num: 4
H2 =
     0.8958 * BH3
     - 0.0984 * B2CH3
     + 0.894 * BH_2
     - 0.0001 * H
     + 0.0043 * B2CH2
     + 10.9239

LM num: 5
H2 =
     1.0792 * BH3
     + 0.0028 * B2CH3
     + 0.0048 * BH_2
     - 0.0001 * H
     + 1.2349 * B2CH2
     + 66.0366

LM num: 6
H2 =
     0.5317 * BH3
     - 0.0002 * B2CH3
     + 0.0011 * BH_2
     - 0.614 * H
     - 0.0008 * B2CH2
     + 330.1513

LM num: 7
H2 =
     0.0049 * BH3
     - 0.0002 * B2CH3
     + 0.0011 * BH_2
     - 0.6031 * H
     - 0.0008 * B2CH2
     + 404.2896

LM num: 8
H2 =
     1.3494 * BH3
     - 0.0002 * B2CH3
     + 0.0011 * BH_2
     - 1.1575 * H
     - 0.0024 * B2CH2

+ 221.0257

LM num: 9
$H_2 =$

    0.0024 * BH3
    - 0.0002 * B2CH3
    + 0.0011 * BH_2
    - 0.532 * H
    + 0.3576 * $B2CH_2$
    + 392.5987

LM num: 10
$H_2 =$

    0.5007 * BH3
    - 0.6895 * B7H4
    + 0.0002 * B2CH3
    + 0.0002 * BH_2
    - 0.2527 * H
    + 0.0004 * B2CH2
    + 287.1497

Number of Rules : 10

Time taken to build model: 0.94 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient	0.9995
Mean absolute error	2.9605
Root mean squared error	3.7059
Relative absolute error	3.2097 %
Root relative squared error	3.1227 %
Total Number of Instances	88974