Enumeration Method for Search of Low-energy Structures on Graphene

Joseph Lawrence

A senior thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Bachelor of Science

Bret Hess, Advisor

Department of Physics and Astronomy

Brigham Young University

August 2016

ABSTRACT

Enumeration Method for Search of Low-energy Structures on Graphene

Joseph Lawrence
Department of Physics and Astronomy, BYU
Bachelor of Science

Structures of metal on graphene play important roles in batteries and hydrogen storage, and discovery of new structures could improve these applications. Through computational methods, the formation energy of a proposed structure can be calculated. Enumeration is the process of creating lists of possible structures to search for low formation energy. We present an enumeration method that allows for efficient treatment of vacancies and apply it to the case of six sites per two carbon atoms. If the separation between two sites in the lattice is less than a specified distance, the pair is referred to as forbidden. The enumeration skips structures where at least one forbidden pair has both sites non-vacant, speeding up the process exponentially by volume. This allows for enumeration of structures of high volumes that would be impossible using standard methods. The search method is applied to the case of scandium, titanium, yttrium, and zirconium on graphene. We report results of several low-energy structures, many of which do not currently appear in the literature. These structures may be useful in future applications of batteries and hydrogen storage.

Keywords: graphene, enumeration, density functional theory, cluster expansion, scandium, titanium, yttrium, zirconium

ACKNOWLEDGMENTS

I would like to thank Dr. Bret Hess for his advisement on this project. I would also like to thank the Physics and Astronomy department at Brigham Young University for the funding that made this work possible.

PHSCS 416, Writing in Physics, provided considerable help in the steps of writing and revising this thesis.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

Graphene has many applications in both nanotechnology and batteries. When graphene is doped with a metal, it acts as a conductor and can be used in electronics. By doping graphene with a transition metal at the right concentration and configuration, a very effective cathode or anode can be created. Additionally, metal on a graphene layer can be used to facilitate hydrogen storage. The metal atoms allow hydrogen atoms to bind to the structure, storing them for later use such as in a fuel cell. This is an alternative to storing hydrogen in liquid or gas form, and is a preferable method for its efficiency and safety [1, 2].

In fabrication of these nanostructures, the bonds between graphene and metals must be understood. Due to the large number of possible metals and their configurations, studies of graphene-metal interaction are varied and far from complete. This thesis outlines computational methods used in searching for stable, low-energy structures. By calculating energies computationally rather than experimentally, thousands or even millions of structures can be searched.

An endless number of possible structures on graphene exist—any combination of atoms in any

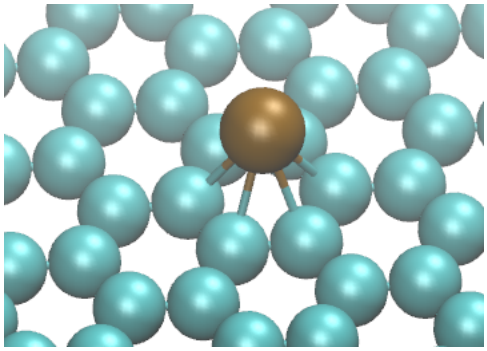

Figure 1.1 An adatom on graphene.

configuration could be considered. To limit the search to a reasonable number of structures, I narrowed the search to structures consisting of a single metal element on graphene. (see Fig. 1.1) The search was further narrowed to consider only scandium, titanium, yttrium, or zirconium, which are the lightest of the transition metals and most useful for batteries. In determining how favorable a structure is, many characteristics could be considered. Formation energy was chosen as the primary characteristic to compare because it indicates stability of the structure. A lower formation energy indicates greater stability and thus a greater likelihood that the structure would form in the lab. Other characteristics will be referred to later to deepen the analysis of specific structures.

Formation energy is often calculated using density functional theory (DFT), a computational method that is performed on one structure at a time. Depending on the complexity of the structure, this can take a few minutes to several hours per structure. In searching a large set for low-energy structures, this process can be time consuming. The search is made more efficient by the use of a predictive model called cluster expansion. Once a few hundred formation energies have been calculated, cluster expansion is used to predict which uncalculated structures will have the lowest formation energies. This makes it so only a fraction of the structures in the set need to be calculated, allowing for searches of sets including millions of structures.

Previous work has been done to systematically generate all possible structures given a particular configuration. [3, 4] This process is called enumeration. The computer code developed by Hart

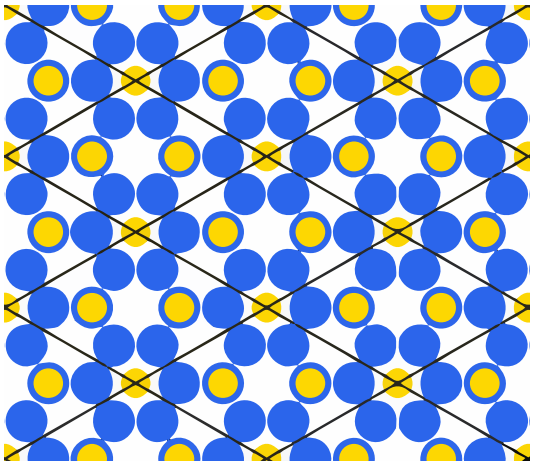

Figure 1.2 A low-energy configuration of titanium on graphene as reported by Sivek [2]. Blue circles represent carbon atoms and yellow circles represent titanium atoms.

and Forcade takes an input lattice of possible atom positions, and generates a complete list of possible structures, excluding any that are rotationally or translationally equivalent. This method was modified for the case of metal on graphene, and the resulting set was searched for low-energy structures.

1.2 Past Results

Past work has identified a few low-energy structures involving transition metals on graphene. This work presents a comprehensive search of structures that is much more complete than the guess-and-check methods previously employed. To give context, a particular low-energy structure of titanium reported by Sivek is shown in Figure 1.2 [2].

This structure is reported to have a formation energy of -0.93 eV per Ti atom, which was the lowest energy of the structures investigated by the paper. In the results section of this thesis, several other low-energy structures are shown, giving a more complete picture of the behavior of these structures.

1.3 Purpose

The purpose of this thesis is to present a modified version of the enumeration method , applied to metal on graphene. The modifications allow the inclusion of six possible sites for metal atoms per cell. A cell is a unit of the lattice that includes two carbon atoms and any atoms added above the graphene layer. Each of the six sites in a cell may be occupied by a metal atom or vacant, which allows for a wide variety of concentrations and configurations.

Normally, enumeration with that many possible sites would be lengthy. Enumeration of six or more cells with six sites each can take days and involve billions of structures. The proposed method skips a large portion of the structures by including only structures that have sufficient space between the metal atoms, using a specified distance dependent on the atom size. This allows for efficient enumeration of structures that are both large and complex.

This thesis presents a search which was used to find low-energy structures of scandium, titanium, yttrium, and zirconium at varying concentrations. The following chapter of this thesis outlines the methods used in this search. The final chapter reports the results of the search, analyzing the structures of greatest interest and comparing them to the literature.

Chapter 2

Methods

2.1 Program Outline

Our group has developed a program that performs all the steps of the low-energy search. [5] The user gives some input parameters, then the program performs the steps of enumeration, VASP calculations and cluster expansion to find the structures with the lowest formation energy.

The input parameters define the positions of sites in each cell. Also defined are the atoms that can occupy these sites, including an option that allows sites to be vacant. The maximum volume is given, where volume is defined to be the number of cells in a structure. This is also where additional parameters are given for calculation accuracy and program flow.

The following subsections give some detail about the key elements of the program—VASP calculations, cluster expansion and enumeration.

2.1.1 VASP

The formation energy of a given structure can be calculated using VASP, a computational package that implements density functional theory (DFT). VASP starts with initial atom positions and finds

the lowest possible energy state through iterative relaxation. First, the electrons in the structure are relaxed, finding the most favorable electron configuration. Then VASP relaxes the positions of the ions to find a lower energy configuration. These two types of relaxation are repeated until the process converges and the lowest possible energy state is found.

VASP reports a total energy for the structure. To calculate the formation energy, the total energy is compared to reference energies for each element. For metals, this reference energy is the energy per atom of an isolated hexagonal monolayer. For carbon, the reference energy is the energy per atom of graphene. The formation energy per adatom (atom added to the graphene layer) can be calculated by

$$FE = \left(\frac{E_{struct}}{N_C} - \sum_i \left[x_i E_{ref,i}\right] - E_{ref,C}\right) * \frac{N_C}{N_{adatoms}},$$

where E_{struct} is the energy reported by VASP, N_C is the number of carbon atoms, x_i is the concentration per carbon of the ith adatom, $E_{ref,i}$ is the reference energy for the ith adatom, $E_{ref,C}$ is the reference energy for carbon, and $N_{adatoms}$ is the total number of adatoms.

2.1.2 Cluster Expansion

Rather than using VASP to calculate the energy of each enumerated structure, cluster expansion takes a few VASP energies and predicts the energies of the remaining structures. The details of cluster expansion are explained by Rosenbrock et al. [6] and is implemented in the UNCLE code by Nelson et al. [7]

Cluster expansion is analogous to the simpler case of approximation of functions. For example, imagine you have n points in xy coordinates. Any n points can be fit exactly with an nth order polynomial. If not all the points are known, a lower-order polynomial can be fit to the known points, and the polynomial can often be a good approximation for the unknown points.

Cluster expansion works in a similar manner. For sites on a lattice, clusters of these sites are

chosen. Two sites make a 2-body cluster, three sites make a 3-body cluster, and so on. For a given structure, the occupation of these clusters is considered. Fitting coefficients are multiplied by the occupation of each of these clusters, and the sum of all the terms is the predicted formation energy. If one knows the formation energy of every structure, coefficients can be assigned to the clusters that fit the formation energies exactly. This is often a good approximation before all the structures have been calculated, and can be used to predict formation energies of uncalculated structures. Previously calculated energies are used to fit the coefficients, and the formula is then used to predict the energies of the remaining structures.

After formation energies are predicted for each structure, the lowest-energy structures of each concentration are sent to VASP for accurate energy calculation. Those new energies are added to the cluster expansion fit, and the program alternates iteratively between VASP and cluster expansion until the program stops finding new low-energy structures.

2.1.3 Enumeration

Before the search can begin, a list of structures to search must be generated. This process is called enumeration, and is done systematically following the approach described by Hart and Forcade [3]. A summary of the method is given here.

The cell is the basic unit of a structure (see Fig. 2.1). Each cell can have one or more sites for possible atom locations. The cell has boundaries defined by three vectors coming from the origin, called parent lattice vectors. Cells are placed next to each other to repeat periodically in all directions. For the case of graphene, the layers are kept 15 ÃĚ apart so the interaction between them is negligible.

The labeling of a particular cell is a string that represents the occupation of each site in the cell. A volume-1 structure is a structure with only one cell repeated periodically. The majority of structures are of higher volume. The volume indicates how many cells combine together to create

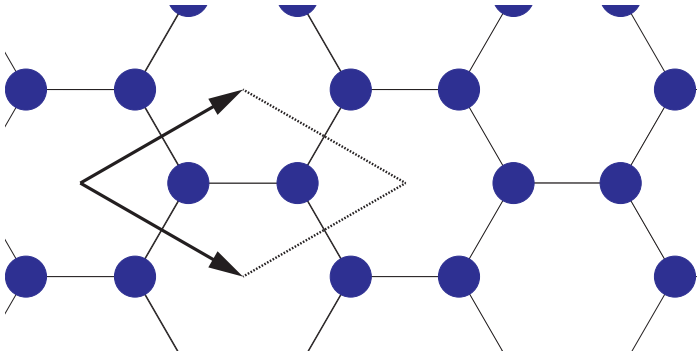

Figure 2.1 An example cell on a graphene lattice. Blue circles represent carbon atoms. The two parent lattice vectors start at the origin of the cell and define the cell boundaries. A third parent lattice vector (not shown) points out of the page in the z direction.

a supercell , and the supercell is repeated periodically to create a structure. New lattice vectors describe the size of the supercell, which is a linear combination of the parent lattice vectors. Different linear combinations of the parent lattice vectors create different sizes and shapes of supercells.

Once the size of the cell, the coordinates of its sites, and the maximum volume is determined, the enumeration code generates a list of all possible supercells for each volume. This list is then reduced by symmetry of the parent lattice vectors to remove rotationally equivalent supercells. For each remaining supercell, a list of all possible structures is created. This has a length of the number of possible atoms at each site to the power of the total number of sites. This list is then reduced for translational equivalence along the parent lattice vectors, as well as rotational equivalence about the parent lattice vectors.

2.2 New Enumeration Method

For enumeration of structures on graphene, we can define a cell to contain two carbon atoms in the graphene layer. To make the search as complete as possible, we choose to include six sites for adatoms above the layer. This includes one hollow site, two top sites and three bridge sites (see Fig. 2.2). Since the two carbon atoms are constant for each cell, they are not considered in

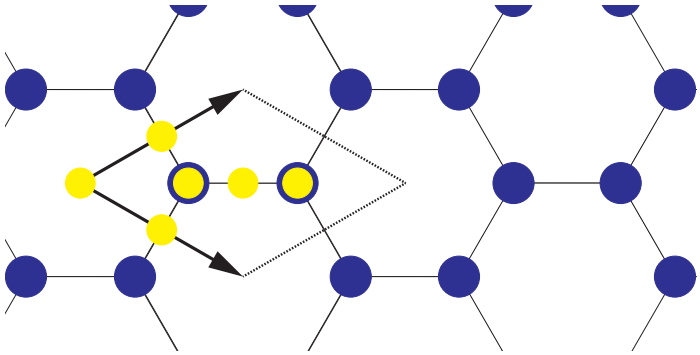

Figure 2.2 The six sites in a unit cell. Blue circles represent carbon atoms and yellow circles represent sites for metal atoms above the graphene layer. There are three different kind of sites: One hollow site is at the cell's origin, two top sites are above the carbon atoms, and three bridge sites are between pairs of carbon atoms.

the enumeration and are simply added in for VASP calculations. The cell can be represented as a six-digit label, with each digit in the label describing the content of each of the six adatom sites.

Only a small number of the possible structures need to be considered. Each site can be occupied by a metal atom or be vacant. These sites are close enough together that if all of them cannot be filled with metal atoms. If two atoms are too close together, the VASP energy cannot converge. We can define a minimum distance between atoms to be 0.75 times the sum of their atomic radii. Any structures containing atoms closer than this distance is said to contain a forbidden pair and is an invalid structure.

It is possible to run the standard enumeration code and remove structures with forbidden pairs afterwards. However, this is extremely time consuming since huge numbers of structures would have to generated and most of them would contain forbidden pairs. Instead, the proposed method runs the enumeration from scratch, skipping entirely any cells that contain forbidden pairs.

To begin, every possible volume-1 structure is generated, including rotationally and translationally equivalent structures. For the binary case with six sites that makes $2^6 = 64$ possible volume-1 structures, which are referred to as tiles. Any structure can be made up of these tiles. However, many of these tiles contain forbidden pairs, and can be dropped from the enumeration. For the tran-

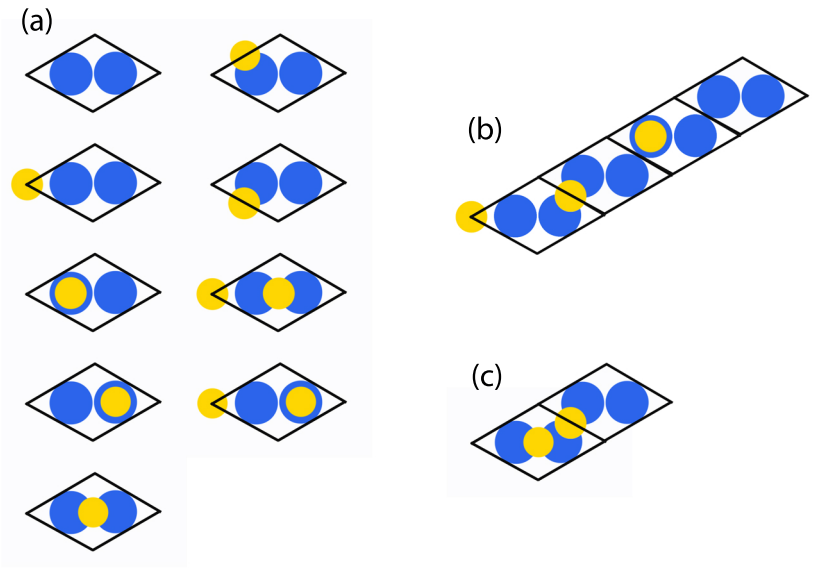

Figure 2.3 Tiles that are used to create structures. (a) shows the nine tiles that are possible with metal on graphene. (b) shows some tiles combined to form a valid supercell. (c) shows a combination of tiles that creates a forbidden pair.

sition metals we are considering, only nine of the sixty-four possible tiles do not contain forbidden pairs (see Fig 2.3).

For each supercell, tiles are combined systematically to create structures. Since the number of valid tiles has been significantly reduced, this creates far fewer structures than standard enumeration. After each structure is created, it is tested to see if any forbidden pairs have been introduced across tile boundaries. Additionally, since the supercell is repeated periodically, tests are made for forbidden pairs across the supercell boundary.

For each remaining structure, a list of translationally equivalent structures is created. The structure is only retained if its label is the lowest numerically on the list. Similarly, rotationally equivalent structures are removed, using symmetries of the lattice vectors. This ensures that only one unique structure is retained from each group of translationally or rotationally equivalent structures.

Chapter 3

Results

3.1 Enumeration Results

Table 3.1 compares the number of structures considered by standard enumeration and forbidden pairs enumeration. The number of structures considered is a good indicator of both time and memory required for enumeration. Volume 6 could not finish standard enumeration in a few days, whereas forbidden pair enumeration could finish the same set in about an hour.

The data in Table 3.1 can be approximated by a power law. The increase in speed of the forbidden pairs method is about 3.8^V, where V is the maximum volume enumerated. According to this formula, volume 6 is over 3000 times faster to enumerate using the forbidden pairs method. This exponential increase in speed by volume allows for enumerations of volume and complexity that would be impossible using standard enumeration.

3.2 Newly Discovered Structures

The new enumeration method was used to perform a search of scandium, titanium, yttrium, and zirconium each on graphene. This search was complete up to volume six. The search resulted in

Table 3.1 A comparison of number of structures considered by standard enumeration versus forbidden pair enumeration for the case of six sites per unit cell. For each volume, the total number of enumerated structures is given for three cases: First, the number of structures output by standard enumeration after all calculations; Second, the total number of structures considered by the forbidden pairs method; Third, the number of unique structures after all forbidden pairs are removed in either case.

Volume	Standard Enumeration	Forbidden Pairs	Final Usable Structures
1	24	9	4
2	748	90	17
3	34408	1548	72
4	2581020	21231	340
5	111800004	139329	1152
6	?	1733652	6929

hundreds of low-energy structures for each element.

Figures 3.1, 3.2, 3.3, and 3.4 show the results of the low-energy search for each element. The graph in each figure shows the calculated and predicted formation energies for structures of varying concentration. The reported concentration is the site concentration, meaning the fraction of the sites in the supercell are filled with a metal. Of greatest interest are the structures with the lowest VASP energies at each concentration. Though the cluster expansion energies are shifted somewhat, the ordering of the structures is usually quite accurate, leading to useful predictions of what the lowest-energy structures are. The search stopped when new low-energy structures stopped being discovered, so only about a third of the structures have VASP energies.

The red line on each figure indicates the convex hull. The structures on the convex hull are the lowest-energy structures of their respective concentrations. Structures that would make the line not convex are skipped, meaning only a few structures are included. In experiment, these are the structures that are most likely to form. Images of these structures are included in each figure below

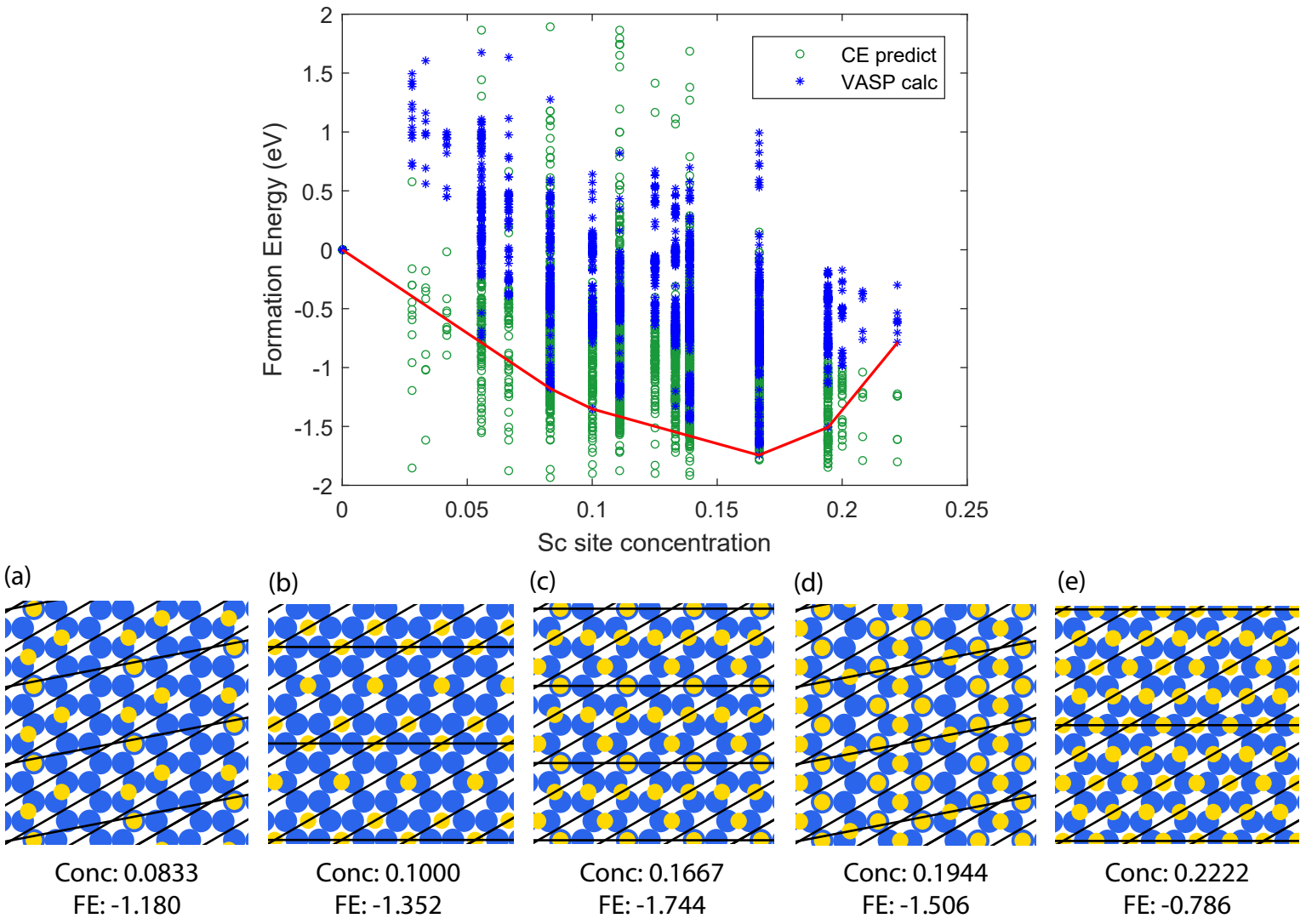

Figure 3.1 Formation energy vs. Sc site concentration for structures in the search set. The green circles indicate energies predicted by cluster expansion. Blue stars indicate energies calculated by VASP. The red line indicates the convex hull. A representation of each structure that lies on the convex hull is shown, as well as the structure's site concentration and formation energy in eV.

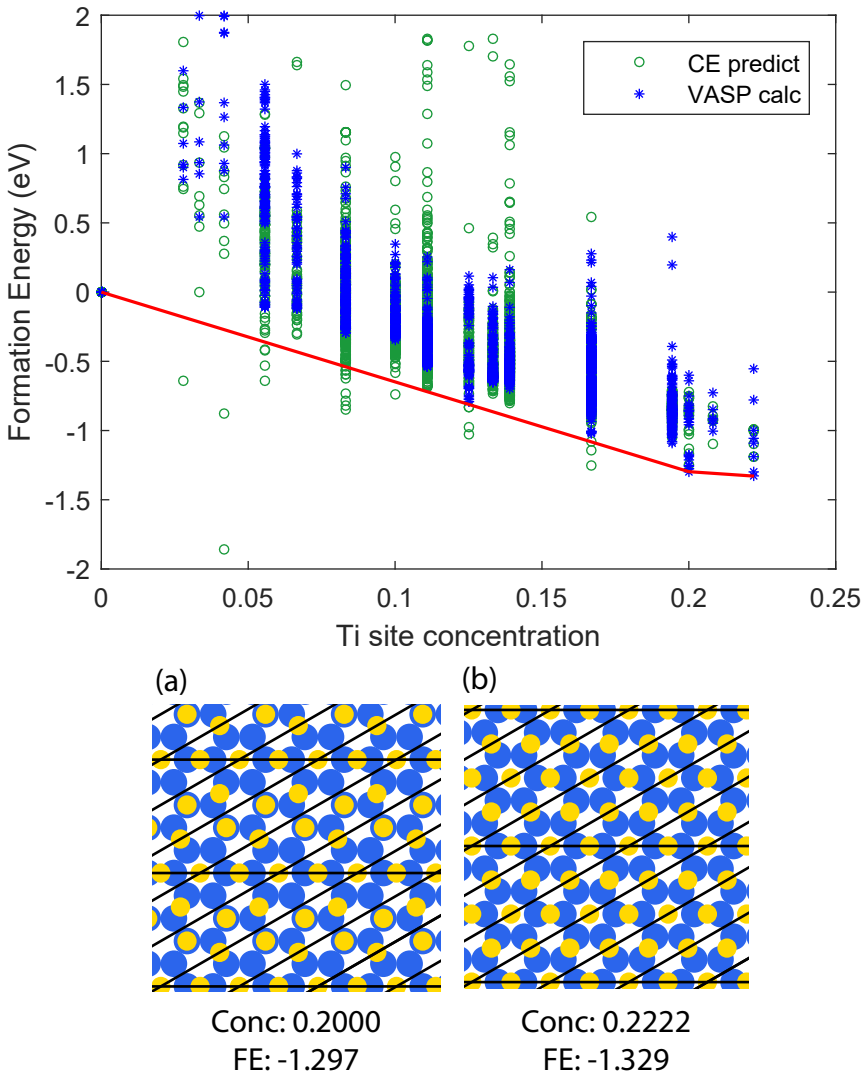

Figure 3.2 Formation energy vs. Ti site concentration for structures in the search set. The green circles indicate energies predicted by cluster expansion. Blue stars indicate energies calculated by VASP. The red line indicates the convex hull. A representation of each structure that lies on the convex hull is shown, as well as the structure's site concentration and formation energy in eV.

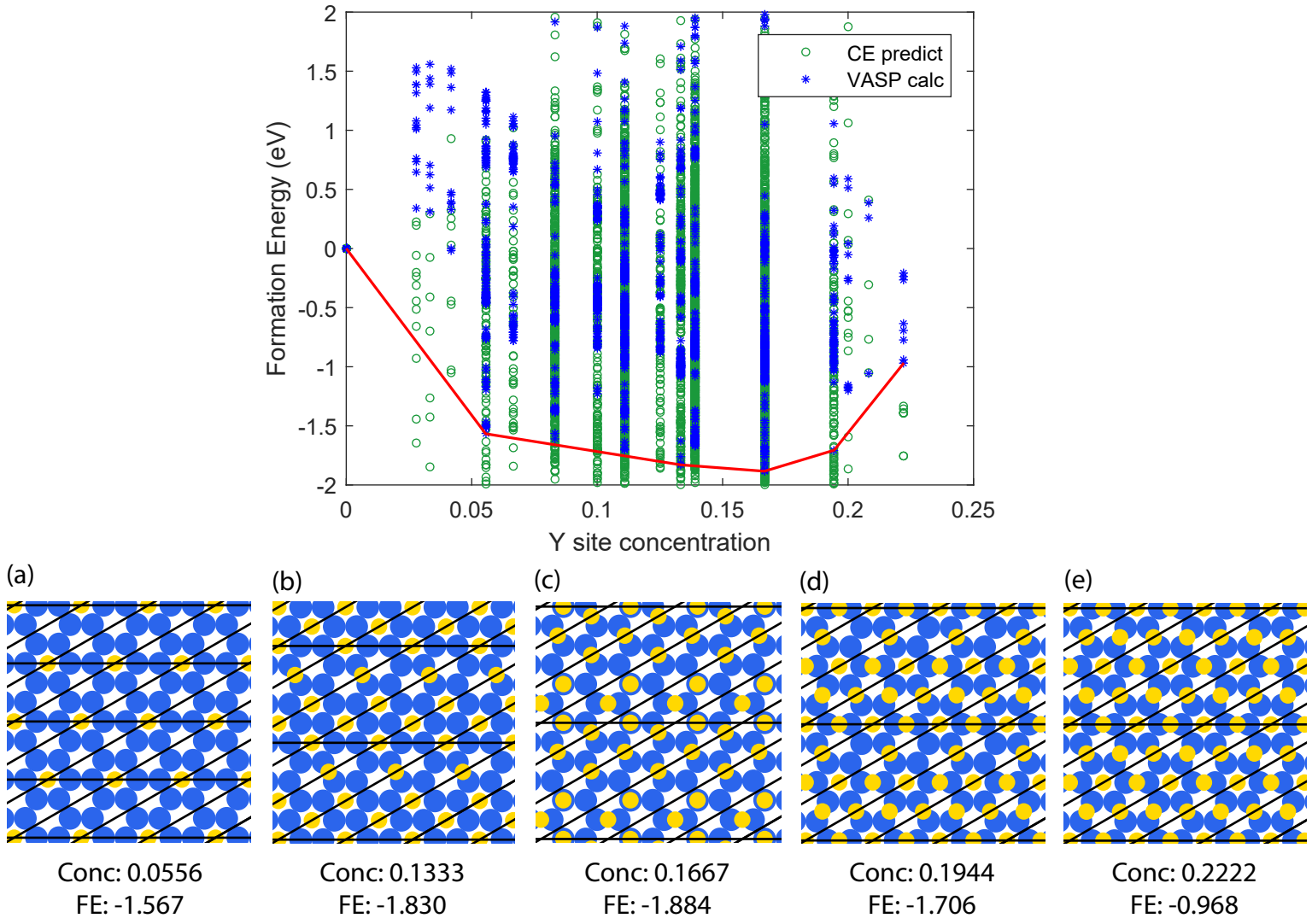

Figure 3.3 Formation energy vs. Y site concentration for structures in the search set. The green circles indicate energies predicted by cluster expansion. Blue stars indicate energies calculated by VASP. The red line indicates the convex hull. A representation of each structure that lies on the convex hull is shown, as well as the structure's site concentration and formation energy in eV.

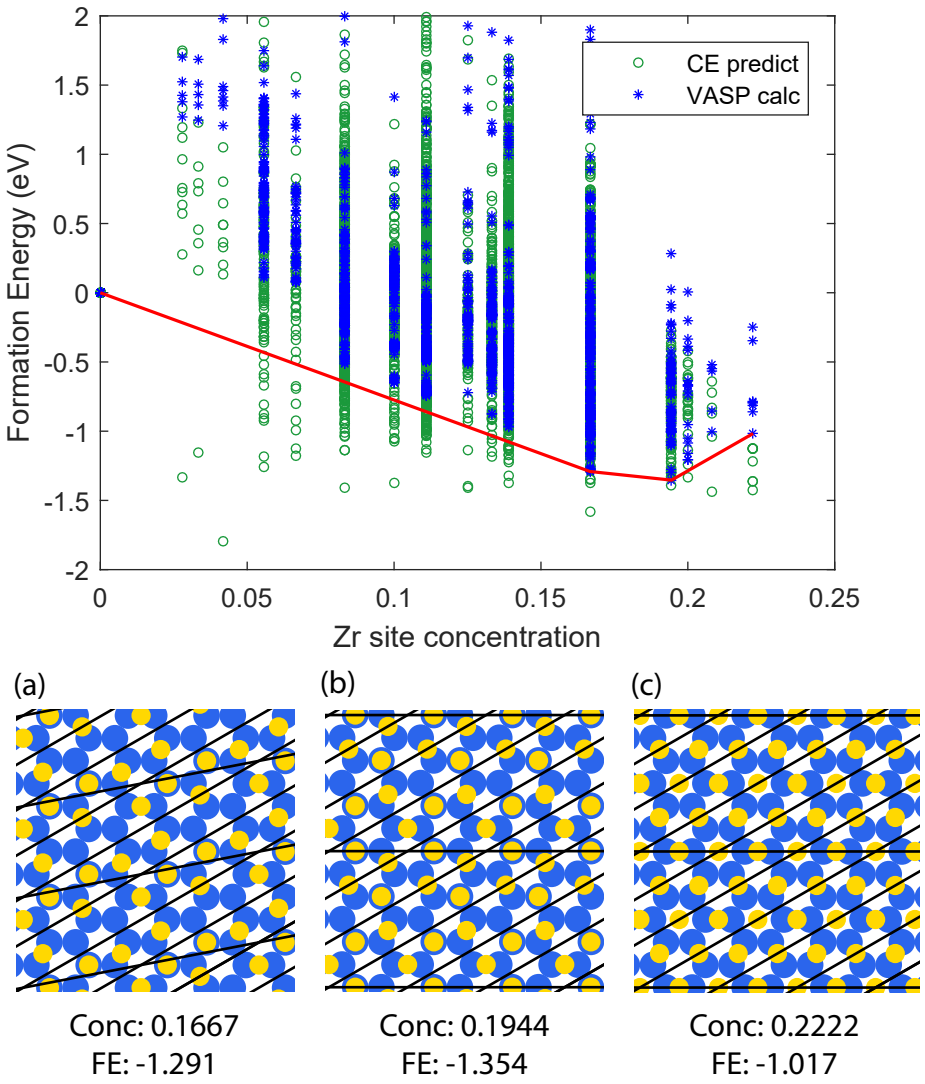

Figure 3.4 Formation energy vs. Zr site concentration for structures in the search set. The green circles indicate energies predicted by cluster expansion. Blue stars indicate energies calculated by VASP. The red line indicates the convex hull. A representation of each structure that lies on the convex hull is shown, as well as the structure's site concentration and formation energy in eV.

the graph of formation energies. Only the initial positions are shown, so it can be expected that these structures will relax in their positions somewhat.

Each metal had different low-energy structures on the convex hull, though it is interesting to note that all four metals formed the same lowest-energy structure at concentration 0.222. This structure consisted of two metal atoms for every three carbon, spread evenly in a lattice of equilateral triangles. This is the same structure reported by Sivek for titanium in Section 1.2. Sivek's reported formation energy for this structure (-0.93 eV) is considerably different than the energy calculated in this work (-1.392 eV). This can be attributed to the low-precision setting used for the search. Despite the bias error, we agree that the reported structure has the lowest formation energy for titanium. Further work could involve recalculating these low-energy structures at high-precision to get their absolute formation energies.

Other work has shown that in addition to bias error, the formation energies calculated in our search may have an ordering error of about 20 meV. This is due to using supercells of different sizes and shapes in the calculations. In some cases, there are multiple structures in a concentration within 20 eV of the lowest energy. This means that any one of those structures may be the one that actually forms. However, these structures are usually very similar and often relax to the same end positions, so only one structure is shown at each concentration.

These figures show some general trends. The group 3 elements, scandium and yttrium, have a minimum in the convex hull at a lower concentration than the group 4 elements, titanium and zirconium. This allows for more structures in the convex hull. Additionally, both of the group 3 elements have lower minimum formation energies, yttrium having the lowest at -1.884 eV.

3.3 Future Work

The search presented in this thesis dealt with a very limited subspace of the structures that could be considered. The search could be expanded from the binary case to include structures with three or four elements. These structures could be made up of multiple metals as well as hydrogen or oxygen. This could lead to understanding of the characteristics of the different elements in these structures.

The new enumeration method quickly generates a complete list of structures, but this method could be sped up even further using another tile strategy that has yet to be implemented. Certain combinations of tiles will never work next to each other due to atoms being too close together across the boundary. The current method still generates these structures and removes them individually in testing for forbidden pairs. However, a search could be made beforehand that identifies pairs of tiles that are forbidden, and skips generation of those structures. This would create the same number of structures but drastically reduce the time required, allowing enumeration of higher volume structures.

3.4 Conclusion

When used effectively, cluster expansion of adatoms on graphene can expedite the process of searching for low-energy structures, and lead to the discovery of new interesting structures. This allows for a complete search of the set of possible structures, rather than having to guess and check. As this method is utilized, unexpected low-energy structures may be discovered that would not be found without doing a complete search. This was shown to be the case with titanium.

The new enumeration method allows for the creation of more complicated structures much faster than standard enumeration would require. Before the new enumeration method was implemented, the most complicated structures considered had two top sites per cell. Implementing the

method allowed for the addition of three bridge sites and a hollow site per cell, leading to structures that were much more varied. With this search method, low-energy structures on graphene can be found which may lead to the creation of better batteries, as well as facilitate hydrogen storage.

Bibliography

[1] Y. Sanchez-Paisal, D. Sanchez-Portal, N. Garmendia, R. Munoz, I. Obieta, J. Arbiol, L. Calvo-Barrio, and A. Ayuela, "Zr-metal adhesion on graphenic nanostructures," Applied Physics Letters **93,** 053101 (2008).

[2] J. Sivek, H. Sahin, B. Partoens, and F. M. Peeters, "Adsorption and absorption of boron, nitrogen, aluminum, and phosphorus on silicene: Stability and electronic and phonon properties," Physical Review B **87,** 085444 (2013).

[3] G. L. W. Hart and R. W. Forcade, "Algorithm for generating derivative structures," Physical Review B **77,** 224115 (2008).

[4] G. L. W. Hart and R. W. Forcade, "Generating derivative structures from multilattices: Algorithm and application to hcp alloys," Physical Review B **80,** 014120 (2009).

[5] E. Swenson, "Nanostructures formed by calcium substitution in graphane,", Senior Thesis, Brigham Young University, (2016).

[6] C. W. Rosenbrock, B. Bieniek, and V. Blum, "Hands-On Tutorial on Cluster Expansion IPAM Los Angeles, California, July 2014," (2014).

[7] L. J. Nelson, V. Ozolins, C. S. Reese, F. Zhou, and G. L. W. Hart, "Cluster expansion made easy with Bayesian compressive sensing," Physical Review B **88,** 155105 (2013).

Index