Summer 2024

Crystal Structure Prediction of Metal Chalcogenides

Qi Zhang
Missouri University of Science and Technology

Department: Physics

Recommended Citation

Zhang, Qi, "Crystal Structure Prediction of Metal Chalcogenides" (2024). *Doctoral Dissertations*. 3356.
https://scholarsmine.mst.edu/doctoral_dissertations/3356

CRYSTAL STRUCTURE PREDICTION OF METAL CHALCOGENIDES

by

QI ZHANG

A DISSERTATION

Presented to the Graduate Faculty of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

PHYSICS

2024

Approved by:

Dr. Alexksandr Chernatynskiy, Advisor
Dr. Julia E. Medvedeva
Dr. Thomas Vojta
Dr. Yew San Hor
Dr. Amitava Choudhury

ABSTRACT

A novel crystal structure prediction (CSP) method has been developed to predict energetically favorable (stable) structures based on targeted chemical compositions. It leverages the structural characteristics of recurring motifs featured in many crystals and symmetry restrictions from space groups to effectively lower the degrees of freedom of a system when conducting CSP simulations.

The proposed method is applied to predicting low-energy structures of two metal chalcogenide systems: Li_3PS_4 and $Na_6Ge_2Se_6$. Both systems feature rigid bodies in their structures as building blocks, making them particularly suited to the proposed method. The validity and effectiveness of this method are demonstrated by not only identifying the experimentally observed phases for both chemical compositions but also predicting a novel $Na_6Ge_2Se_6$ phase that possesses much lower energy than the currently known lowest energy phase, as evaluated through first-principles calculations.

This predicted $Na_6Ge_2Se_6$ phase is then directly compared to the observed one in terms of structural, elastic, electronic, phonon, thermal, and optical properties. The results display differences and similarities across every aspect as well as showing the mechanical and dynamic stability of the predicted $Na_6Ge_2Se_6$ phase and a lower free energy at elevated temperatures than its observed counterpart.

The method is then extended to quaternary systems: $LiZnPS_4$ and $LiMnPS_4$. The lowest energy structure identified through CSP simulations for $LiZnPS_4$ coincides with the synthesized phase. For $LiMnPS_4$, an even lower energy phase is predicted that has never been reported before. This phase is dynamically stable and possesses lower free energy at elevated temperatures. Additionally, a few metastable candidates are also identified.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Dr. Alexksandr Chernatynskiy, for his unwavering support, guidance, and encouragement throughout my research. His insightful advice and extensive knowledge of computational physics have been invaluable to the completion of this dissertation. I am incredibly fortunate to have had the opportunity to work under his mentorship.

I would also like to extend my sincere thanks to the members of my thesis committee, Dr. Julia E. Medvedeva, Dr. Thomas Vojta, Dr. Yew San Hor, and Dr. Amitava Choudhury, for their time, constructive feedback, and valuable suggestions, which have greatly enhanced the quality of my work. Special thanks to Dr. Amitava Choudhury, as these projects would not have been possible without his expertise in chemistry.

Finally, I am profoundly grateful to my parents for their unconditional love, constant support, and understanding. I also extend my heartfelt thanks to my friends for their encouragement, which has been a source of strength throughout this journey, especially during the difficult times of the pandemic.

TABLE OF CONTENTS

Page

SECTION

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

Table Page

1. INTRODUCTION

Crystal structure is fundamentally crucial for understanding a material's properties, as it directly or indirectly influences almost every characteristic. The ability to determine the structure allows scientists to predict a material's properties even before it is synthesized, making crystal structure prediction (CSP) essential in computational materials design. However, predicting crystal structures remains a complex challenge. In 1988, John Maddox famously remarked in a Nature article, "One of the continuing scandals in the physical sciences is that it remains impossible to predict the structure of even the simplest crystalline solids from a knowledge of their composition."[77] While its accuracy is debatable nearly 40 years later, Maddox's statement still underscores the formidable nature of crystal structure prediction.

1.1. CHALLENGES OF CRYSTAL STRUCTURE PREDICTION

1.1.1. What is Crystal Structure Prediction. In the early days of structure prediction, scientists relied heavily on imagination and intuition. Despite these limitations, there were some notable discoveries. In 1611, mathematician Johannes Kepler predicted the structure of ice in his treatise "On the Six-Cornered Snowflake,"[10, 56] suggesting what we now know as hexagonal close packing. Although this was not the correct structure for ice, it accurately describes elements like Be, Mg, and Cd under normal conditions. Another significant example is August Kekulé's famous vision of a snake biting its tail, which inspired him to propose the cyclic structure of benzene[119]. In 1897, William Barlow modeled the rocksalt structure. However, without the grounding in physical laws that modern crystal structure prediction relies on, these early efforts, while visionary, fall short of today's standards in the field. When we talk about CSP today, it involves determining the stable crystal structure at given pressure and temperature conditions, based solely on the chemical formula. However, for any given chemical composition, there are

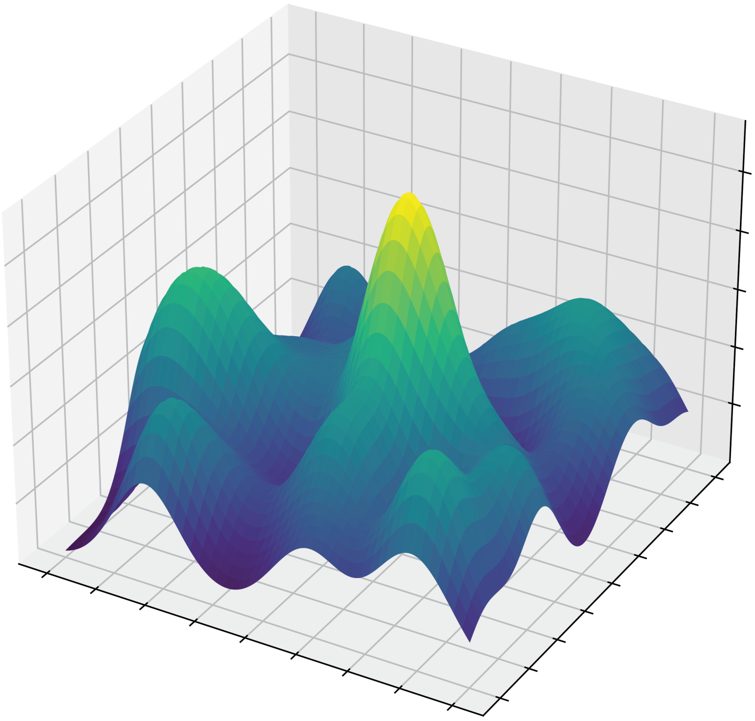


Figure 1.1. A schematic illustration of the potential energy surface (PES).

countless possible atomic arrangements that can theoretically be formed in a laboratory. These arrangements represent all potential locations on a complex potential energy surface (PES) as depicted in Figure 1.1. Among these configurations, a finite number of special structures stand out under specific thermodynamic conditions (pressure, temperature, chemical potential). These structures may represent the lowest energy (most stable) or possess extreme values of certain properties (such as hardness, density, band gap, superconducting transition temperature). Thus, from the perspective of scientific computing, CSP can be regarded as a global optimization problem on a complex PES to find the most favorable candidate structures. This problem can be split into two sub-questions: how to sample the configuration space and how to estimate the stabilities of the sampled structures.

1.1.2. Potential Energy Surface Sampling. The number of distinct points on the PES can be estimated using the following equation[82]:

$$C = \binom{V/\delta^3}{N} \prod_i \binom{N}{n_i} \tag{1.1}$$

Where N is the number of atoms in the unit cell of volume V, δ is a relevant discretization parameter and n_i is the number of atoms of i-th type in the unit cell. When using $\delta = 1\mathring{A}$ and assuming an atomic volume of 10 $\mathring{A}^3$, C approximates to 10^N, which is an astronomically large number. The dimensionality of the PES d, can be expressed as $d = 3N + 3$, where $3N-3$ degrees of freedom correspond to atomic positions, and the remaining six dimensions represent lattice parameters. For a system with 50 atoms in the unit cell, the PES would thus be 153-dimensional. Thus, the complexity of CSP increases exponentially with system size, making it a non-deterministic polynomial-time hard (NP-hard) problem. It means the problem's scaling with system size is faster than any polynomial. Given the vast number of potential solutions, simple exhaustive search strategies become impractical.

1.1.3. Structure Stability Evaluation. Based on density functional theory (DFT)[26, 41, 87, 90], sophisticated ab initio methods have been developed to accurately evaluate various properties of materials, which are implemented in software packages such as the Vienna Ab-initio Simulation Package (VASP)[62, 63], Quantum ESPRESSO[31], ABINIT[35], CP2K[43], and WIEN2k[14]. These first-principles methods are commonly used to calculate the energy or enthalpy and determine the thermodynamic stability of candidate structures. Molecular dynamics with classical interatomic potentials is another option which is widely used for large-scale materials studies[34]. However, these empirical potentials are less accurate than first-principles methods, especially when complex electron interactions are involved. Furthermore, a complex PES features countless local energy minima with very small energy differences. Classical potentials fall short in accurately predicting the energy ranking of structures, which is crucial for CSP. Unfortunally, compared to less accurate approaches such as classical molecular dynamics, ab initio methods require significantly more computational time. This is particularly challenging for CSP, which often necessitates evaluating the stabilities of thousands of structures due to the vastness of the PES. Additionally, the computational cost of ab initio methods for local optimizations grows rapidly with the number of atoms, making comprehensive CSP for complex crystal structures often unfeasible.

1.2. ADDRESSING THE CHALLENGES

1.2.1. Promising Region of PES. For the first challenge, various algorithms have been developed for exploring the PES, including simulated annealing [57, 86], minima hopping [5], basin hopping [118], metadynamics [65], and random sampling [92]. However, all of these methods suffer from the so-called “curse of dimensionality” due to the typically very high-dimensional nature of the PES. This high dimensionality is directly related to the number of degrees of freedom required to model a material system, as explained above. Hence, the accessible size of the system is still limited and the effectiveness of a CSP method hinges upon its ability to identify and explore the most promising region of PES efficiently.

1.2.1.1. Rigid body restriction. To accomplish this objective, some CSP approaches were developed on the basis of a ”self-improving” strategy, which assumes that the chemically desired structures often share similar geometric motifs such as bond lengths and the coordination of atoms are located on the low-lying regions of the PES. A typical example of CSP using the self-improving strategy is the genetic algorithm[1, 21, 33, 72, 81, 111, 123, 125], where low-energy structures serve as parents to procreate new candidates by a ”crossover” operation. The basic tenet here is that offsprings resemble their parents, and procreation is rewarded for success and an in-depth exploitation is performed in the most promising region on the PES. Another method is the automated assembly of secondary building units (AASBU)[32]. This technique utilizes prior information to enforce connections between individual atoms or rigid bodies, treating them as building units. By doing so, it effectively reduces the degrees of freedom when performing structure searches which is based on a combination of a simulated annealing procedure and ”cost function” minimizations.

1.2.1.2. Space group restriction. On the other hand, statistical analysis of modern structural databases on inorganic crystals reveals a striking preference for certain space groups over others. Previous analysis using data from ICSD-2006 indicates that 67% of all crystal structures are found in only 10% of all space groups, while 31% of the space groups are empty or rare [114]. Furthermore, the preference for occupation of Wyckoff positions (WPs) within space groups has also been summarized [113]. For example, in most structures of inorganic crystals, anions occupy less symmetric positions than cations. This information can be highly beneficial in locating the most promising regions of the PES for CSP if needed.

1.2.2. Machine Learning Interatomic Potentials. The second challenge is less prominent but still significant despite major developments in computer science. Machine-learning interatomic potentials (MLIPs) have gained significant attention in computational physics and materials simulations recently[100]. Because A well-trained MLIP can bridge the gap between the accuracy of ab initio methods and the computational efficiency of traditional force fields. Recent advancements have led to the development of universal MLIPs, which are general models capable of simulating a wide range of systems across most of the periodic table. Based on graph neural networks (GNN)[98], universal MLIPs such as M3GNet[19], CHGNet[22], MatGL[19], and MACE[11] have been developed. Assessments of their accuracy suggest that, although further optimization and training are needed for a broader range of applications, universal MLIPs have shown great precision when the right MLIP is chosen for the appropriate application[27, 126].

1.2.3. Dissertation Outline. Aiming to address the intricate difficulties imposed by CSP, this dissertation reports a new CSP method that incorporates recurring motifs found in many crystal structures along with the symmetry of space groups to effectively reduce the system's degrees of freedom. This method is implemented in an open-source Python package called Space Group Restricted Crystal Structure Prediction (SGRCSP), developed by the author of this dissertation. The package is available at the public GitHub repository: `https://github.com/ColdSnaap/sgrcsp.git`, which enables automated simulations with a simple command. The machine-learning interatomic potential CHGnet is also integrated into the code as an option for energy calculations of sampled structures.

The algorithm of the proposed method and the tutorial for SGRCSP are detailed in Section 2.2. Its validity and effectiveness are demonstrated in Section 3 by running simulations on two metal chalcogenide systems: Li_3PS_4 and $Na_6Ge_2Se_6$, using both the proposed method and a popular CSP package, USPEX[33, 74], which leverages a evolutionary algorithm. The results report multiple candidate metastable phases for Li_3PS_4 and $Na_6Ge_2Se_6$, notably identifying a $Na_6Ge_2Se_6$ phase with much lower energy than the only experimentally observed $Na_6Ge_2Se_6$ phase[23].

In Section 4, the new $Na_6Ge_2Se_6$ phase is examined in greater detail, with a comprehensive analysis from structural, elastic, electronic, phonon, thermal, and optical perspectives. A direct comparison is also made with the observed phase from these perspectives.

In Section 5, the method's applications are further extended from predicting ternary metal chalcogenides to quaternary metal chalcogenides, namely $LiMnPS_4$ and $LiZnPS_4$. The ground state of $LiZnPS_4$ is successfully predicted, and a lower energy phase is identified for $LiMnPS_4$ than the experimentally observed one, further affirming the method's effectiveness.

The conclusion, along with the limitations of the method and further work needed, is detailed in Section 6.

2. SPACE GROUP RESTRICTED CRYSTAL STRUCTURE PREDICTION

One major challenge for CSP is achieving effective searching on the PES, as introduced in Section 1.2.1. One way to address this challenge is by leveraging structural characteristics of targeted compounds from experimental observations and empirical evidence. Recurring motifs or rigid bodies are often seen in many structures. Among these, metal chalcogenides have gained considerable interest recently for their potential applications in energy conversion/storage, optoelectronics, and thermoelectrics. These areas are vital for supporting the advancement of current technologies and industries, especially in addressing the global energy crisis as countries aim to transition from fossil fuels to renewable, efficient, and eco-friendly energy solutions.

2.1. METAL CHALCOGENIDES

The general formula for metal chalcogenides is ABX, where A represents a metal (or a combination of 2-3 different metals), B is a main group element, and X denotes a chalcogen. These compounds have emerged as leading candidates for electrolytes in lithium/sodium solid-state batteries[46], owing to their abundant raw material availability, decent room-temperature ionic conductivity, and lower mechanical stiffness. Complex metal chalcogenides also stand out as some of the most efficient photovoltaics materials[2], favored for their tunable bandgap, non-toxic and stable nature. In nonlinear optics applications[20], particularly in the IR range where many materials face fundamental limitations, chalcogenides are increasingly utilized. Their application as cathode materials in rechargeable magnesium batteries, seen as a next-generation alternative to lithium-ion batteries due to high energy density, addresses the challenge of slow diffusion kinetics in most cathode materials by offering larger surface areas and shorter migration paths[96]. Additionally, thermoelectrics[103] and photocatalysis[37], are other possible areas of applications for chalcogenides. Despite significant interest, much of the chemical space of complex metal chalcogenides remains uncharted, both experimentally and computationally, due to the vast array of possible atomic combinations. This unexplored territory presents substantial opportunities for discovering and developing even more effective materials for a variety of applications.

Many complex chalcogenides are characterized by recurring structural motifs. For instance, the BX_4 tetrahedron moiety is extensively studied, with compounds containing this motif being actively developed across various fields[4, 18, 54, 68, 69, 71, 80, 85, 101, 112]. Another example is the B_2X_6 dimer unit, foundational in recently synthesized non-linear optical materials[8]. These blocks are notably stable, as demonstrated by one of the synthesis routes called the metathesis reaction. Experimentally, this feature of complex chalcogenides enables researchers to synthesize structures with desired properties in a rational manner. Computationally, it allows us to leverage information on bond distances, angles, dihedrals, and connectivity constraints of the rigid blocks within the structure to simplify the structure search.

2.2. ALGORITHM

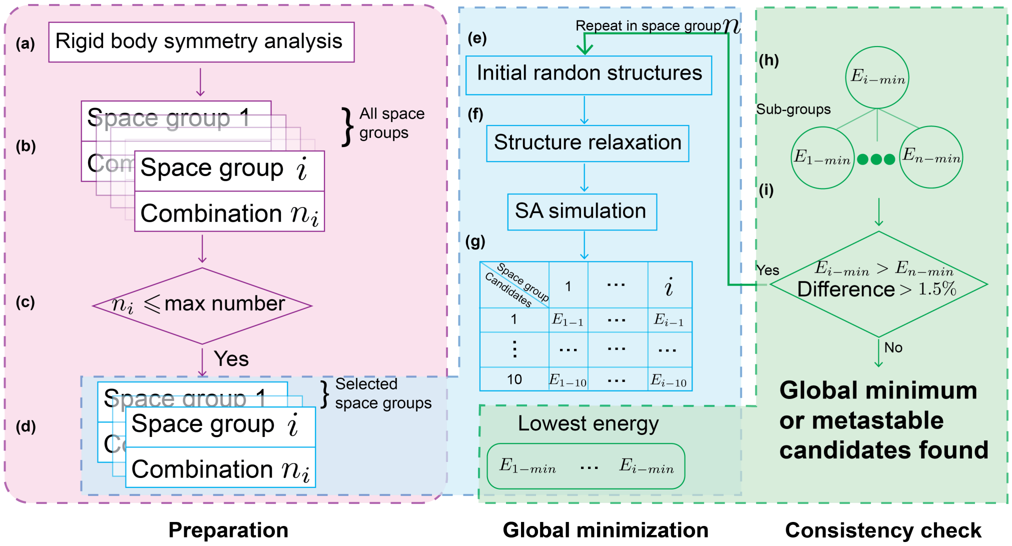


Figure 2.1. Schematic illustration of the simulation procedure.

2.2.1. Symmetry of Rigid Bodies. Determining the ground state structure in CSP represents a global optimization challenge within “phase space” of all possible structures, characterized by a dimensionality of $3N - 6$, where N signifies the number of atoms in the system. The subtraction of 6 accounts for the structure’s translational and rotational degrees of freedom. As the number of atoms in the crystal unit cell increases, this dimensionality rises linearly. Consequently, the volume of the phase space increases exponentially, a phenomenon often referred to as the ”curse of dimensionality.” However, the inherent symmetry within the crystal structure offers an advantage, allowing for a reduction in the number of independent atomic coordinates and, consequently, the dimensionality of the phase space wherein the structure resides. This dimensionality reduction directly translates to fewer exploratory attempts required to find the minimum energy configuration, thereby making identifying the ground state structure more efficient[74, 92].

Table 2.1. Examples of the WPs in space group $P2_1/c(14)$ and $Pmm2_1(31)$.

Space group	Multiplicity	Wyckoff letter	Site symmetry	Coordinates	
$P2_1/c$ (No.14)	4	e	1	(1)x, y, z	(2)$-x, -y + 1/2, -z - 1/2$
				(3)$-x, -y, -z$	(4) $x, -y + 1/2, z + 1/2$
	2	d	-1	(1)$1/2, 0, 1/2$	(2)$1/2, 1/2, 0$
	2	c	-1	(1)$0, 0, 1/2$	(2)$0, 1/2, 0$
	2	b	-1	(1)$1/2, 0, 0$	(2)$1/2, 1/2, 1/2$
	2	a	-1	(1)$0, 0, 0$	(2)$0, 1/2, 1/2$
$Pmm2_1$ (No.31)	4	b	1	(1)x, y, z	(2)$-x + 1/2, -y, z + 1/2$
				(3)$x + 1/2, -y, z + 1/2$	(4) $-x, y, z$
	2	a	m	(1)$0, y, z$	(2)$1/2, -y, z + 1/2$

For a crystal with a given symmetry, the atomic positions are classified by Wyckoff positions (WPs). The most general WP has the highest multiplicity, and the atoms found on it do not lie on any symmetry elements. The remaining sites are the so-called special positions, and the atoms that occupy these sites reside on symmetry elements of the cell. A WP is unique if all its coordinates are fixed. meaning they do not contain variables. A unique position can host only a single atom, whereas a non-unique WP can accommodate multiple atoms. For instance, consider space group $P2_1/c(14)$, The coordinates for its WPs are delineated in Table 2.1. Here, the most general WP is $4e$, which has a multiplicity of 4 and is not unique. In contrast, other WPs ($2a$, $2b$, $2c$, and $2d$) are unique, each with multiplicity of 2. However, the most crucial aspect of WPs for our applications is the site symmetry. This is because an arbitrary object can occupy specific WP in periodic crystals only if the symmetry group of that object contains the site symmetries as a subgroup.

In crystallography, atoms are typically considered as spherically symmetric point particles. As a result, the full rotation group of any atom inherently includes the site symmetry of a WP as a subgroup. This allows atoms to be located anywhere within the unit cell. However, for molecules or rigid bodies, their own symmetry groups impose restrictions on WP compatibility, determining if a rigid block can be situated at a specific WP. Additionally, the alignment between the symmetry group of the rigid block and the site symmetries frequently limits the block's orientation, effectively reducing the dimensionality of the phase space.

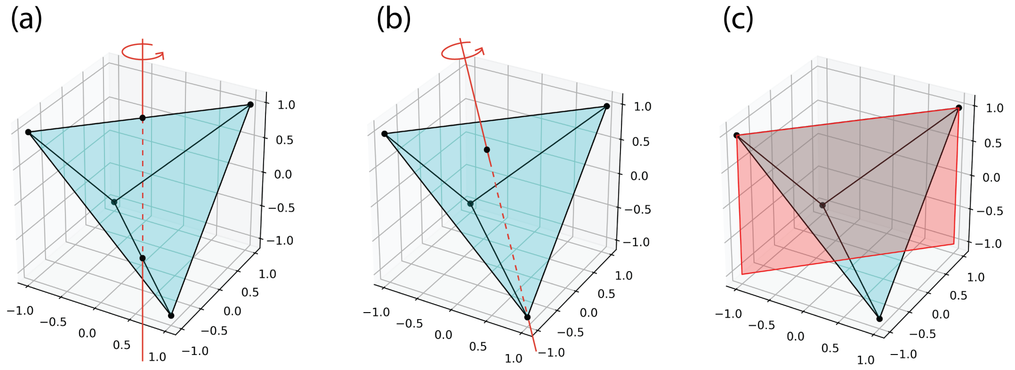


Figure 2.2. The symmetry elements of a perfect tetrahedron.

Consider a perfect tetrahedron rigid block within a crystal structure of space group $Pmm2_1$(31). The symmetry operations of a perfect tetrahedron form the symmetry group known as full tetrahedral symmetry, T_d. This group includes symmetry elements such as 2-fold rotational axes, 3-fold rotational axes, mirror planes, and 4-fold rotoreflections (improper rotations) as illustrated in Figure 2.2. We reference the site's symmetries for space group $Pmm2_1$(31) in Table 2.1. The general position $4b$ has only the identity operation,

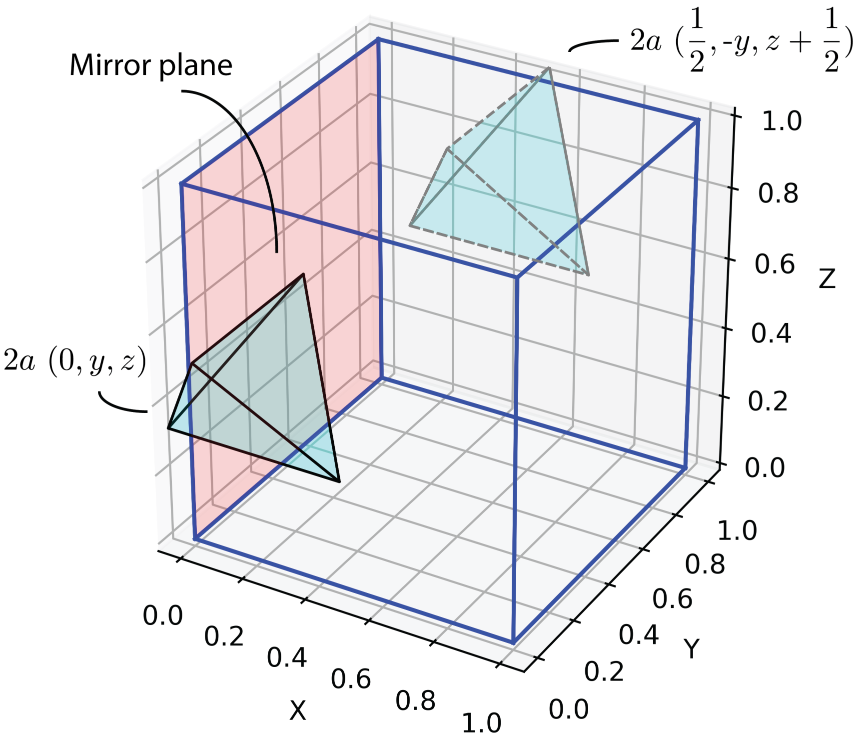


Figure 2.3. A tetrahedral rigid body is placed at WP 2a (0, y, z) with a symmetry element of a mirror plane in space group $Pmm2_1$(31).

which is a subgroup of any group, thus allowing the tetrahedron to occupy this WP. However, with a multiplicity of four, the crystal must contain four tetrahedra in its unit cell. The identity operation does not restrict the orientation of the tetrahedron, allowing it to rotate freely around three general axes. Besides the general WP, this group also includes special WP $2a$, which contains a mirror plane as shown in Figure 2.3. The combination of mirror and identity operations forms a subgroup of the tetrahedral symmetry group, making this WP compatible with the tetrahedron. With a multiplicity of two, two tetrahedra must be present in the unit cell. The mirror plane's specific orientation necessitates that the tetrahedron's mirror plane coincides with it, limiting its rotation to a single axis perpendicular to the mirror plane. Since neither WPs are unique, multiple tetrahedra can be placed at each of those positions. In practice, the total number of atoms or rigid bodies in the unit cell is capped due to limited computational resources, limiting possible combinations. If we restrict the crystal to contain only two tetrahedra per unit cell, the sole option is placing them at $2a$ WP, with only three independent variables describing their positions: y and z coordinates of the WP, and the rotation angle in the mirror plane. For four tetrahedra, two configurations are possible: i) all four at the general position $4b$ with arbitrary orientation or ii) two sets of tetrahedra at WP $2a$. Both cases require six independent coordinates to specify the unit cell. On the other hand, space group $P2_1/c(14)$ has special WPs whose symmetry elements consist solely of inversion, as shown in Table 2.1. This operation is not a symmetry of a tetrahedron. Therefore, the only option is to place all tetrahedra at general positions when searching for compounds with tetrahedra in this space group. This sets the minimum number of tetrahedra at four in the unit cell. Consequently, space group $P2_1/c(14)$ will be disregarded if we limit the total number of tetrahedra to two. Analyzing the symmetry of the rigid body and determining its compatibility WPs in a space group is the first step of the simulation as shown in Figure 2.1(a).

2.2.2. Enumerating Possible Configurations. With the idea of the symmetry, it is feasible to enumerate all possible atom arrangements for a certain compound composition as in Figure 2.1(b). For example, let us list all possible combinations of atomic arrangements to WPs for Li_3PS_4, where PS_4 is a tetrahedron rigid body, in the space group $Pmm2_1(31)$. From the previous discussion, there must be an even number of the tetrahedral units in the primitive cell, and for this example we will consider the smallest number of the stoichiometric units, i.e. two. In general, the size of the unit cell is limited by the computing power available for local minimization through DFT. In practice, we set the maximum number of atoms in the cell, and progressively consider possible number of stoichiometric units in the cell up to this limit. PS_4 rigid bodies can be at both $2a$ and $4b$ WPs based on site symmetries. However, since $4b$ WP has multiplicity of four, and we are considering only 2 stoichiometric units per cell we are left with the only possibility of PS_4 units at $2a$. The number of tetrahedral units sets the number of lithium atoms in the unit cell to six. Taking into account the multiplicities

Table 2.2. Numbers of combinations for Li_3PS_4 in different space groups with a maximum of three PS_4 rigid bodies in the unit cell.

Space group	Combinations	Space group	Combinations	Space group	Combinations
1	3	2	127	3	>1000
4	1	5	12	6	256
7	1	8	2	10	>1000
11	20	13	76	16	>1000
17	96	18	12	21	88
23	76	25	>1000	26	12
27	96	28	39	30	12
31	2	32	12	34	12
35	20	38	20	44	16
47	>1000	51	124	59	24
75	20	77	39	81	>1000
82	28	84	60	85	16
86	8	89	206	90	16
93	330	94	8	99	28
100	6	101	16	102	3
105	48	111	>1000	112	270
113	22	114	4	115	>1000
116	64	117	52	118	52
119	28	121	6	125	26
129	16	131	90	132	38
134	10	137	4	143	>1000
144	1	145	1	146	2
147	34	149	>1000	150	130
151	12	152	12	153	12
154	12	156	>1000	157	64
158	33	159	10	160	2
168	23	171	12	172	12
173	10	174	>1000	177	130
180	76	181	76	183	27
185	3	186	10	187	>1000
189	206	195	23	197	1
201	1	208	3	215	17
217	1	218	3	224	1

Note: Highlighted space groups have 20 or fewer combinations.Combinations exceeding 1000 are marked as ’> 1000’. Space groups not included in the table have no viable cases.

of the WPs, these atoms can be distributed only in the following two combinations: $\{4b,2a\}$ and $\{2a,2a,2a\}$. We list the numbers of possible combinations for every space group in

Table 2.2 with the restriction that the crystal contains not more than three tetrahedra per primitive cell. For all the missing space groups, it is not possible to arrange up to three stoichiometric units of Li_3PS_4 in that space group.

While the symmetry restrictions from the rigid body and the crystal space group reduce the dimensionality of the phase space for the global optimization, the number of the possible arrangements of rigid block and atoms among WPs is likely to be very large in some space groups. Examining the Table 2.2 we see that Li_3PS_4 in the space group $Pmma(51)$ has 124 possible arrangements of the rigid blocks and lithium atoms. Considering the structure prediction method involves many structure relaxations through DFT for each arrangement. It renders the search within this space group very time consuming. We circumvent this difficulty by taking the subgroups of a space group into consideration.

A subgroup of a particular space group contains only some of the symmetry operations of that space group. It means that there are fewer restrictions for the atomic positions from the symmetry considerations. Effectively, setting the structure within the subgroups will replace (large) enumerated number of possible combinations of atomic positions with random sampling in higher-dimensional PES. If the global minimum structure possesses the symmetry of the original group, then this will appear as "accidental" symmetry in the best structure in its subgroups, if the algorithm is successful. In the example of the tetrahedron rigid block above, when we are trying to find the lowest energy structure in the space group $Pmma(51)$, we can look at its subgroup $P2_1(4)$ instead, which only has 1 possible arrangements of atoms/rigid blocks among WPs. Of course, that process can be continued all the way to the space group $P1$ which contains only the identity operation. This will be equivalent to ignoring the symmetries of the crystal and corresponding reduction in the dimensionality of the phase space.

Figure 2.4 schematically shows how CSP in space group $Pc(7)$ can "accidentally" result in the same structure as in its supergroup $Pmn2_1(31)$. In $Pc(7)$, there is only one general WP $2a$, allowing atoms/rigid bodies to move freely within the entire cell. When the

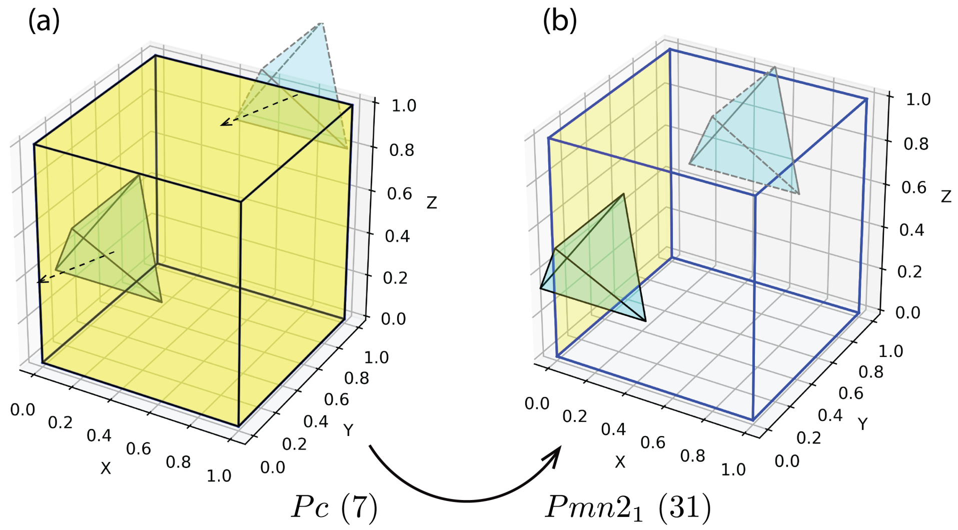


Figure 2.4. A schematic illustration shows how CSP in a subgroup can result in the same structure as its supergroup.

random sampling in $Pc(7)$ visits a configuration where the rigid bodies are at special WP $2a$ in space group $Pmn2_1(31)$ with an orientation that aligns its mirror plane with the mirror plane of the WP (or close enough that structure optimization can locate this configuration), it effectively samples the shared part of the PES between $Pc(7)$ and $Pmn2_1(31)$. This plays an important role in providing an consistency check to the CSP results as well as determining metastable structures.

2.2.3. Generation of The Initial Structure. The initial structure for global minimization is crucial for the efficient convergence of the optimization process. Besides atomic positions, the structure is also defined by crystal basis vectors. For convenience, it is more practical to use three lattice constants and three angles rather than basis vectors, and the space group symmetry imposes additional constraints on these parameters. A classic example is a cubic lattice where $a = b = c$ and $\alpha = \beta = \gamma = 90.0°$, leaving only one unique parameter to define. The initial values of lattice constants and angles are set randomly within the constraints that the structure should adhere to a specified density and desired lattice system.

At this stage, the allocation of atoms and rigid blocks to the WPs is already established. Therefore, we generate initial structures by randomly assigning atomic coordinates to all WPs and randomly orienting the rigid blocks as in Figure 2.1(e), if permissible.

2.2.4. Global Optimization. Once we specify the configuration of atoms and rigid bodies among the Wyckoff positions, we carry out a global optimization of the energy via a modified simulated annealing (SA) method.

The SA closely mirrors the natural process where a molten system's energy progressively reduces as the temperature decreases, ultimately arriving at the global minimum configuration - the most stable crystalline phase. The SA is initiated by simulating the system's dynamic evolution at some initial temperature using either molecular dynamics or a Monte-Carlo scheme. The initial temperature has to be chosen so that there are sufficient kinetic energy to overcome energy barriers between local minima. The temperature is then progressively lowered as the system continues its evolution. The temperature lowering gradually restricts transitions between the minima that have barriers greater than the current temperature, and eventually the system arrives and stays in the global minimum, if certain conditions are met[30]. In this work we employ a Monte-Carlo scheme to simulate the dynamics, which contains three ingredients: The probability distribution $g(\Delta x_i, T)$ to select a new random trial structure, the probability distribution $p(T, \Delta E)$ to accept the new structure and the cooling schedule $T(k)$. Here, Δx_i is a change of the parameter x_i to be optimized, i.e. symmetry-restricted positions of the atoms and rigid bodies, $T(k)$ is a temperature T at the step k of the simulation, ΔE is the energy difference between the new and the current structures. If one chooses Boltzmann SA, i.e. for each parameter the function $g(\Delta x_i, T)$ is

$$g(\Delta x_i, T) = (2\pi T^{-\frac{1}{2}}) exp[-\Delta x_i^2/(2T^2)] \tag{2.1}$$

and the acceptance probability as

$$p(T, \Delta E) = \frac{1}{1 + exp(\Delta E/T)} \tag{2.2}$$

Then it can be shown[30], that logarithmic reduction of the temperature, i.e. $T(k) = T_0/\ln k$ guarantees the convergence of the method. In practice, however, logarithmic reduction is too slow, i.e. it requires a prohibitive number of simulation steps, especially if the energy is evaluated via the first principles calculations. Thus, the convergence of the method is not assured. We note, that there are other choices possible for the distributions $g(\Delta x_i, T)$ and $p(T, \Delta E)$ which permit faster cooling schedule[57, 106], however applicability of these approaches for our optimization problem is beyond the scope of this work.

To improve the chances of finding the global minimum we introduce the following modifications to the SA method. SA methods typically record simulation results at predetermined intervals, followed by a local energy minimization, with the final post-minimization structures serving as candidate structures for the global minimum. Our approach extends beyond merely storing these candidates. We leverage the configurations of these candidates, including lattice parameters and atomic positions, as the starting point for the subsequent SA round. This strategy capitalizes on the fact that the first principles softwares, such as VASP[62, 63], conveniently implements symmetry-restricted structure optimization. Consequently, we can update the lattice parameters post-minimization instead of constantly perturbing the lattice randomly, allowing for a further reduction in dimensionality. During the SA process, only the atoms undergo random movement within their designated WPs, while the six lattice parameters remain constant.

Additionally, the bond lengths within the rigid bodies are adjusted based on the DFT optimization. This adjustment is particularly advantageous when only the general geometry of the rigid body is known, and an approximate bond length is set prior to the simulation. However, it is possible for the rigid body to disintegrate following a minimization step. To address this, we implement a rigid body integrity check after each energy minimization. If the rigid body is found to have broken apart, we store the relaxed structure in case the structure with broken rigid bodies results in a lower energy phase. However, the subsequent simulation will continue with the last relaxed structure that has intact rigid bodies.

Secondly, the term ΔE in the acceptance probability equation exhibits varying sensitivities to the optimization parameters. It is thus advantageous to scale Δx_i by the inverse of the energy change due to the variation of the parameter x_i alone. These sensitivities are periodically evaluated throughout the simulation run. Lastly, our findings suggest that gradually cooling the system is beneficial. However, since our methodology does not depend on a specific cooling schedule to ensure convergence, the exact nature of this dependence is not critically important. We implement a cooling schedule as $T(k) = T_0 \times 0.99^{k/N}$, where N represents the number of steps between subsequent local optimizations. In typical simulation runs for the applications discussed in the Section 3 and 5, we use the following parameter values: the initial temperature is set to $T_0 = 2$ K, the total number of steps is $k_{\text{max}} = 10^4$, and local optimization occurs every $N = 50$ steps.

After finishing the simulations for all space groups with feasible distributions of atoms and rigid blocks among WPs, we store the ten lowest energy structures for each combination, as depicted in Figure 2.1(g), for potential metastable structures. Subsequently, we conduct a consistency check in chosen space groups that exhibit low-energy structures, using the lowest energy structure from other space groups to determine whether the global minimum or metastable candidates have been identified. This step is also illustrated in Figure 2.1. For a specific space group, simulations within its subgroups should yield energies equal to or lower than those obtained within the group itself, provided the optimization in that subgroup was successful. Thus, we compare the lowest energies found in available subgroups of a given group. If they are consistent and the lowest among all discovered energies, it strongly suggests that the ground state has been identified. However, if energies within subgroups are higher than in the parent group by more than 1.5% from our experience, this indicates incomplete optimization in the subgroup, prompting further calculations with new random initial structures. The number of loops run for this process can be determined as needed; in this work, we set it to three. If, after three attempts, the energy within a subgroup remains lower than that within the group, the optimization success remains uncertain. Such a consistency check is vital because, in our implementation, SA does not guarantee the discovery of the global minimum, as previously discussed.

2.3. SGRCSP PYTHON PACKAGE

There are many CSP packages available today utilizing various simulation algorithms. AIRSS[92] uses the random structure searching method and determined the structure of SiH_4 under high pressure[91]. CALYPSO[122] uses particle swarm optimization (PSO) and predicted the high-pressure Li phase[73], as well as a chemical reaction of Xe with Fe/Ni at the temperatures and pressures found in the Earth's core, forming energetically favorable $XeFe_3$ and $XeNi_3$[128]. XtalOpt[72] employs an evolutionary algorithm to identify high-pressure ices[38] and condensed astatine[39]. USPEX[33, 74], also using an evolutionary algorithm, successfully predicted novel phases of Al with Sc or Ta[12] and high energy density materials such as sodium pentazolates NaN_5 and Na_2N_5[104].

Inspired by these success stories and with the intention of uncovering more energetically favorable and dynamically stable metal chalcogenide phases, this dissertation presents a novel CSP method. This method enables CSP simulations with rigid bodies and atoms restricted to designated WPs in a specific space group, making it particularly suitable for metal chalcogenides featuring different types of rigid bodies as constituents of their structures.

2.3.1. Introduction. Space Group Restricted Crystal Structure Prediction (SGRCSP) is an open-source package available at `https://github.com/ColdSnaap/sgrcsp.git`. The entire code is implemented in Python 3 and built with a few important external packages. Structure-related properties, such as symmetry analysis, are calculated with Pymatgen[84] and ASE[66]. Structure generation is carried out with PyXtal[28]. To support scientific computing and data processing tasks, the package also requires NumPy[83], SciPy[116], and Spglib[110]. The implementation details are covered in the following subsections. Please note that the package structure and implementation details may change over time. For the latest updates and version information, please refer to the public GitHub repository.

2.3.2. Input.txt File. The input.txt file contains all the user-defined parameters to control the initialization and simulation procedures during CSP. The parameters take the form of “Parameter = value,” similar to the format used in the INCAR file for VASP. Below is an example of an input.txt file:

```
# Molecules in the structures are MOL_1.xyz and MOL_2.xyz
MolFiles = MOL_1.xyz MOL_2.xyz
# Ratio between these two molecules is 1:3
MolNumber = 1 3
# Maximum number of molecules can be 3:9
RatioMaxMultiplier = 3
# Space groups to search for possible combinations
SpaceGroup = 1 4 7 28-31
# Temperature lowering schedule
TempStep = fast
# Energy calculator for random structures
Caltool = CHGnet
# Initial temperature
TempInitial = 2.0
# Initial volume scaler
Volume = 0.75
# CSP routine
Routine = routine1
# Perturbation type to sample structures
Perturbation = uniform
# Parameters for perturbation
```

```
Lattice = 0
Coordinate = 0.2
Rotation = 2.0
# Whether or not to relax the initial structure for CSP
InitialRelax = True
# Subcommand for first-principles calculations
SubCommand = mpirun vasp_std
# Number of steps for random structure update before cooling
InternalLoop = 50
# Terminate signal for CSP
StepMax = 10000
# Continue the current run
Continue = False
```

This is a typical input file for conducting CSP in certain space groups. The detailed explanations are as follows:

```
# Molecules in the structures are MOL_1.xyz and MOL_2.xyz
MolFiles = MOL_1.xyz MOL_2.xyz
```

One of the key features of SGRCSP is the ability to place a molecule/rigid body at a specific WP based on the symmetry of the rigid body itself. The setup for rigid bodies is done by creating rigid body files. The file format is as follows:

```
P S S S S

-6.43976        3.71576       -0.00000
-4.43976        3.71576       -0.00000
-7.10642        3.16038        1.80198
-7.10643        2.43289       -1.38196
-7.10643        5.55400       -0.42002
```

The first line consists of the atom types within the rigid body, and the Cartesian coordinates of the atoms are listed in the exact order after a blank line. The code will automatically calculate the geometric center and the symmetry of the rigid body, enabling precise placement of the rigid body at its designated WP. Therefore, one doesn't need to worry about the exact coordinates of the atoms but only their relative positions. This is particularly convenient when extracting rigid bodies from existing structures. If a user wants to create a rigid body from scratch with specific bond lengths, it is recommended to use free tools such as Avogadro[36] to intuitively draw the desired rigid bodies. Please note that a single atom is defined in the same way as rigid bodies, with only one atom type. You can simply place it at the origin as follows:

```
Li

0.00000        0.00000        0.00000
```

```
# Ratio between these two molecules is 1:3
MolNumber = 1 3
# Maximum number of molecules can be 3:9
RatioMaxMultiplier = 3
```

The `MolNumber` tag sets the ratio between the rigid bodies defined previously with the `MolFiles` tag. The order of the ratio numbers is important because it corresponds to the order of the molecule files. For example, if `MolFiles = Li.xyz PS4.xyz` and `MolNumber = 3 1`, the targeted chemical composition is Li_3PS_4. The `RatioMaxMultiplier` tag sets the maximum stoichiometric unit of the targeted chemical composition. For instance, `RatioMaxMultiplier = 3` means there can be nine Li and three PS_4 rigid bodies in the simulation box at maximum.

```
# Space groups to search for possible combinations
SpaceGroup = 1 4 7 28-31
```

The `SpaceGroup` tag specifies the space group to be used for CSP. The code will search for viable combinations of WPs based on the types of rigid bodies, the targeted chemical composition, and the stoichiometric unit number. If viable combinations are found, initial random structures will be generated within that space group for CSP. The `SpaceGroup` tag can be set using simple integers (e.g., 1 4 7) or as a range (e.g., 28-31), separated by spaces.

```
# Temperature lowering schedule
TempStep = fast
# Initial temperature
TempInitial = 2.0
```

The cooling schedule is a highly customizable parameter. The `TempStep` tag is used to select a cooling function that the user can define. For example, `TempStep = fast` indicates that the "fast" cooling function is used which is $T(k) = T_0 \times 0.99^{k/N}$. Users can add new functions and assign a name to each function. The `TempInitial` tag sets the initial temperature for CSP. Some cooling schedules might also require a `TempFinal` tag to set a final temperature.

```
# Energy calculator for random structures
Caltool = CHGnet
# Subcommand for first - principles calculations
SubCommand = mpirun vasp_std
```

The CSP process involves two types of structure energy evaluations. The first type is the calculation of energy for unrelaxed structures, which is required for the structure update procedure of SA. In addition to first-principles methods, we offer the option of using the MLIP CHGnet for faster energy calculations (set `Caltool = CHGnet` to switch to this method, otherwise use `Caltool = VASP` for the first-principles option). The second type is the energy calculation with structure relaxation. For this part, we prioritize accuracy and therefore use the first-principles method. The `SubCommand` tag is used for submitting first-principles calculation jobs and may need to be adjusted according to one's computing environment. Please note that the `Caltool` tag only controls which energy calculator to use during the random sampling phase. Even with `Caltool = CHGnet`, the final structure local optimization will be done using the first-principles method based on your sub command.

```
# Initial volume scaler
Volume = 0.75
# Whether or not to relax the initial structure for CSP
InitialRelax = True
```

The `Volume` tag controls the lattice parameters for the initial randomly generated structure. The initial structure must satisfy the following conditions: the lattice parameters should correspond to the correct crystal system according to the space group, the atoms must be in their designated WPs, and the minimum distance between atoms should be greater than a certain value (adjustable via the `MinDistance` tag, with a default value of `MinDistance = 1.4`). The smaller the volume scaler, the harder it is to generate an initial structure that fulfills all the conditions. If the code cannot find a viable structure after a certain amount of time, the volume scaler will be increased until an initial structure is successfully generated. Once a viable initial structure is found, it will be relaxed with symmetry restrictions through a first-principles method (if `InitialRelax = True`) to optimize the geometry before entering the random sampling phase.

```
# Perturbation type to sample structures
Perturbation = uniform
# Parameters for perturbation
Lattice = 0
Coordinate = 0.2
Rotation = 2.0
```

The Perturbation tag determines how random perturbation values are chosen for random structure sampling. Setting Perturbation = uniform suggests that the random perturbation values are drawn from a uniform distribution. For a uniform distribution:

- `Lattice = 0` means that the lattice parameters are not changed during the random sampling phase.
- `Coordinate = 0.2` indicates that the perturbations to atomic positions are within the range of 0 to 0.2.
- `Rotation = 2.0` suggests that the perturbations to the rotation angle for rigid bodies are within the range of 0 to 2.0.

The Perturbation tag is highly customizable like the cooling schedule, allowing users to add new distribution functions suited for different tasks. This flexibility enables tailored perturbations to better fit specific requirements in random structure sampling.

```
# CSP routine
Routine = routine1
# Number of steps for random structure update before cooling
InternalLoop = 50
# Terminate signal for CSP
StepMax = 10000
# Continue the current run
Continue = False
```

The `InternalLoop` tag determines how many steps of random sampling you want to perform before lowering the temperature. Including a few steps with a constant temperature in the SA process relatively increases the chance for the random structure to settle into a better configuration in terms of atomic positions with the particular temperature setting. This is especially beneficial when the cooling schedule is too fast to meet the convergence criteria for SA, which is exactly our case. The `Routine` tag determines the procedure of the simulation. This includes whether to perform a structure relaxation every certain number of random sampling steps and store it as a metastable candidate, or to enable perturbation on lattice parameters and skip structure relaxation altogether. These can be configured through customized routines. The `StepMax` tag sets the maximum number of random sampling steps before the CSP process is terminated. The `Continue` tag determines if it's a continuous job or not. If `Continue = True`, the program will pick up the results from the previous CSP run and inherit its current structure configuration, current step number and temperature.

2.3.3. Molecule/rigid Body Files. The molecule/rigid body files are required to determine the atom types for CSP simulations and to set up rigid bodies. The format of these files is discussed in Section 2.3.2. It is important to note that the names of these files should correspond to the names specified in the `MolFiles` tag in the input.txt file.

2.3.4. First Principles Calculation Inputs. The input files required for first-principles calculations correspond to the first-principles package being used. For example, if VASP is the chosen package, the necessary input files include INCAR, POTCAR (POTCAR for each element), and KPOINTS.

2.3.5. Determine the Combination Number. The first step for the CSP simulaton procedure is to analysis the symmetry of rigid bodies and determine viable combinations in selected space groups as illustrated in Figure 2.1. Here is how this function is implemented in the code.

Input:

```
# Molecules in the structures are MOL_1.xyz and MOL_2.xyz
MolFiles = MOL_1.xyz MOL_2.xyz
# Ratio between these two molecules is 1:3
MolNumber = 1 3
# Maximum number of molecules can be 3:9
RatioMaxMultiplier = 3
```

```
# Space groups to search for possible combinations
SpaceGroup = 1 4 7 28-31
```

Code:

```
from readconfig import ReadConfig
from combination import WyckoffCombinations

# Reading the configurations from input.txt
config = ReadConfig()
sg = config.space_group()
mol_list = config.mol_list()
mol_number = config.mol_number()
z = config.max_multiplier()

# calling the function to search for combinations
comb = WyckoffCombinations(mol_list)
# Loop in selected space groups
for i in sg:
    comb_number = comb.mol_ratio_comb_list_sg(mol_number, i, z)
```

Output:

```
P1 (1), Z = 1, checking: 1/1, valid: 1, sites: [['1a'], ['1a', '1a', '1a
    ']]
P1 (1), Z = 2, checking: 1/1, valid: 1, sites: [['1a', '1a'], ['1a', '1a
    ', '1a', '1a', '1a', '1a']]
P1 (1), Z = 3, checking: 1/1, valid: 1, sites: [['1a', '1a', '1a'], ['1a
    ', '1a', '1a', '1a', '1a', '1a', '1a', '1a', '1a']]
```

```
P21 (4), Z = 2, checking: 1/1, valid: 1, sites: [['2a'], ['2a', '2a', '2a']]
Pc (7), Z = 2, checking: 1/1, valid: 1, sites: [['2a'], ['2a', '2a', '2a']]
Pma2 (28), Z = 2, checking: 39/39, valid: 39, sites: [['2c'], ['4d', '2c']]
Pnc2 (30), Z = 2, checking: 12/12, valid: 12, sites: [['2b'], ['4c', '2b']]
Pmn21 (31), Z = 2, checking: 2/2, valid: 2, sites: [['2a'], ['4b', '2a']]
```

The code will first search for combinations of WPs that satisfy the ratio between the rigid bodies (single atoms are treated as rigid bodies with one element, as discussed above). Next, the symmetry of the rigid bodies will be examined to see if it is compatible with the site symmetry of the assigned WP. This process is indicated by messages in the output: `checking:  39/39, valid:  39`.

2.3.6. CSP Simulation. The CSP simulation starts by executing `main.py`. An example output is as follows:

```
                  |\__/,|   (`
                _.|o o  |_   ) )
               -(((---(((--------
    Space Group Restricted Crystal Structure Prediction

-------------------(version 0.0.1 )--------------------

A Python package for symmerty restricted crystal structure prediction
The source code is available at https://github.com/ColdSnaap/sgrcsp.git
Developed by ColdSnaap
```

```
Routine: routine3
Spacegroup: 4
Molecule1: PS4
Molecule2: Li
Molecule3: Zn
Sites: [['2a'], ['2a'], ['2a']]
Number: [2, 2, 2]
Volume: 0.75
LoopNumber: 50
MaxStep: 10000

Generating structure ...
Increasing volume, volume factor: 0.8
Structure generated

Relaxing the initial structure ...
VASP running ...
VASP(relax=True) job has completed successfully.
Initial structure generated

Applying perturbation ...
Uniform perturbation
Trail: 1
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
Step: 1, Temperautre: 2.0
Applying perturbation ...
Uniform perturbation
Trail: 1
----- Calculating energy of current state -----
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
----- Calculating energy of new state -----
```

```
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
-------------------------------------------------
Current state: -62.53701 eV, -4.46693 eV/atom
New state: -62.15503 eV, -4.43965 eV/atom
Accept: True
Acceptance rate: 100.0%


Step: 2, Temperautre: 2.0
Applying perturbation ...
Uniform perturbation
Trail: 1
----- Calculating energy of current state -----
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
----- Calculating energy of new state -----
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
-------------------------------------------------
Current state: -62.15503 eV, -4.43965 eV/atom
New state: -60.77522 eV, -4.34109 eV/atom
Accept: True
Acceptance rate: 100.0%

...
```

Explanation:

```
                  |\__/,|   (‘
                _.|o o  |_   ) )
```

```
               -(((---(((--------
    Space Group Restricted Crystal Structure Prediction

-------------------(version 0.0.1 )--------------------

A Python package for symmerty restricted crystal structure prediction
The source code is available at https://github.com/ColdSnaap/sgrcsp.git
Developed by ColdSnaap
```

The package version and public Github reposotory address is listed.

```
Routine: routine3
Spacegroup: 4
Molecule1: PS4
Molecule2: Li
Molecule3: Zn
Sites: [[’2a’], [’2a’], [’2a’]]
Number: [2, 2, 2]
Volume: 0.75
LoopNumber: 50
MaxStep: 10000
```

The important information about the system is listed, including the space group in which the CSP is restricted, the molecules/rigid bodies involved along with their assigned WPs, and the maximum step number.

```
Generating structure ...
Increasing volume, volume factor: 0.8
Structure generated

Relaxing the initial structure ...
VASP running ...
```

```
VASP(relax=True) job has completed successfully.
Initial structure generated
```

The initial random structure is generated and then relaxed using VASP. The resultant structure serves as the initial structure for entering the random sampling phase.

```
Step: 1, Temperautre: 2.0
Applying perturbation ...
Uniform perturbation
Trail: 1
----- Calculating energy of current state -----
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
----- Calculating energy of new state -----
CHGNet v0.3.0 initialized with 412,525 parameters
CHGNet will run on cpu
-------------------------------------------------
Current state: -62.53701 eV, -4.46693 eV/atom
New state: -62.15503 eV, -4.43965 eV/atom
Accept: True
Acceptance rate: 100.0%
```

Upon entering the random sampling phase, the current state's energy is compared to that of the subsequent state after applying perturbation to determine whether to accept the new state. The decision is made by comparing a random number n drawn from zero to one to the output of the acceptance probability function $p(T, \Delta E)$ (Equation 2.2). If $p(T, \Delta E) > 0.5$, which means the energy of the new state is lower than that of the current state, the energetically favored structure is instantly accepted. On the other hand, if $n > p(T, \Delta E)$, the new state is accepted; otherwise, it is rejected.

3. CRYSTAL STRUCTURE PREDICTION OF TERNARY METAL CHALCOGENIDES

3.1. INTRODUCTION

Ternary metal chalcogenides are a class of compounds that have shown great potential in various fields. In the quest for improved energy storage solutions, the focus on anodes is crucial due to their significant impact on battery performance. Materials like $LiLnSe_2$ and $NaFeS_2$ have shown high storage capacities and efficient energy cycling in recent study[47, 127], with $NaFeS_2$ showcasing excellent cycling stability as a Li-ion battery anode, suggesting potential for enhanced battery lifespan and reliability. Optoelectronically, Ternary metal chalcogenides such as $NaSbS_2$ offer strong absorption in the visible spectrum, and their affordability, abundance, and non-toxicity make them attractive for solar energy applications[105]. Furthermore, in thermoelectric applications, materials like $CsAg_5Te_3$ exhibit promising mid-temperature performance with high ZT values[70], thanks to their intrinsically low thermal conductivity[75, 76, 88]. Ternary chalcogenides have also played a crucial role in enabling the synthesis of more complex multinary chalcogenides in a rational manner[8]. Many of these compounds have demonstrated significant technological potential. For instance, outstanding electric thermal properties have been observed in materials like $Na_{0.95}Pb_{19}SbTe_{22}$, achieving a ZT value greater than 1 across one of the broadest temperature ranges (475 to 650 K) reported for any single material[93]. Additionally, a study on $Na_{1-x}K_xAsQ_2$ (Q = S,Se) highlighted its potential applications in signal processing and data transmission, attributed to its nonlinear optical properties[45].

To test our method, we carried out the crystal structure predictions for ternary metal chalcogenide system: Li_3PS_4 and $Na_6Ge_2Se_6$, wherein PS_4 and Ge_2Se_6 form a tetrahedral and an ethane-like dimer rigid block respectively, as illustrated in Figure 3.1. Li_3PS_4 is a thoroughly studied system, with three distinct phases experimentally identified[42] and subsequently verified via computational studies[50]. Regrettably, two of these phases, the

α- and β-phases, exhibit fractional occupancy of the crystallographic sites, rendering them unidentifiable through our methodology, thus we expect to find only a ground state γ-phase. $Na_6Ge_2Se_6$ has gained lesser attention in the research community and only a single phase has been experimentally observed[23]. In summary, our method has not only successfully found the γ-phase Li_3PS_4 and experimentally observed $Na_6Ge_2Se_6$ phase, but also yielded several candidates for metastable structures. Notably, we identified several structures for $Na_6Ge_2Se_6$ that possesses lower energy than that of the observed structure.

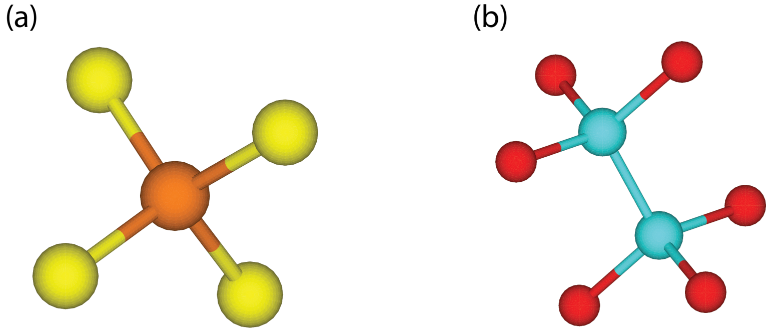


Figure 3.1. Illustrations of a tetrahedral rigid body PS_4 and Ethane-like dimer rigid body Ge_2Se_6.

3.2. CALCULATION DETAILS

The CSP process involves two types of structure energy evaluations. The first type is the calculation of energy for unrelaxed structures, which is required for the structure update procedure of SA. In addition to first principles methods, we offer the option of using the MLP CHGnet[22] (version 0.3.0) for faster energy calculations. The second type is the energy calculation with structure relaxation. For this part, we prioritize accuracy and therefore use the first principles method implemented with VASP[62, 63] (version 5.4.4).

The reported results are a combination of using VASP and CHGnet for the SA process. For calculations performed with VASP which employs Projector-Augmented-Wave method for the treatment of the core electrons[15, 61]. The Perdew-Burke-Ernzerhof

(PBE)[89] functional was used. Integration over the Brillouin zone was carried out using the tetrahedron method with Blöchl corrections (ISMEAR=-5)[16]. For structure relaxation, a $3 \times 3 \times 3$ K-point mesh and plane wave basis sets with an energy cutoff of 520 eV were employed. The convergence threshold for the electronic self-consistent loop was set at 2×10^{-5} eV. For single point calculations, the K-point mesh was reduced to $2 \times 2 \times 2$, with an energy cutoff of 420 eV, and the convergence criteria were loosened to 2×10^{-4} eV.

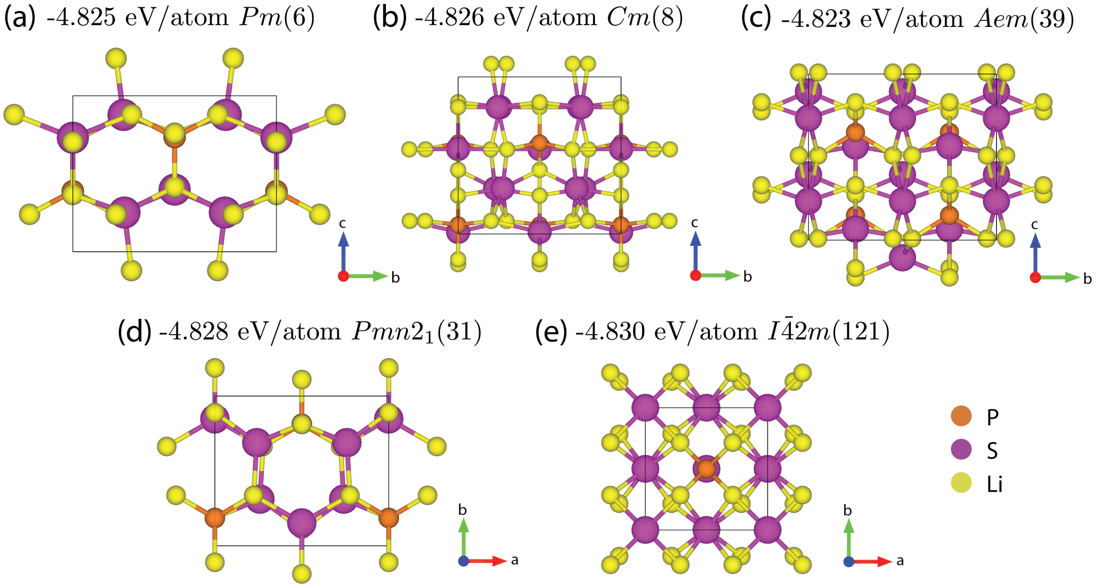
3.3. RESULTS

3.3.1. Li_3PS_4. Our analysis starts with the Li_3PS_4 system, featuring a rigid block that is the previously mentioned tetrahedron composed of phosphorus and sulfur atoms. As its symmetries have been thoroughly addressed in Section 2.2.1, we will not delve into them here. Instead, for a detailed account of potential combinations for arranging up to three stoichiometric units in various space groups, we refer the reader to Table 2.2. Space groups not included in the table are those where allocating the PS_4 tetrahedron and lithium atoms among the WPs is not viable. Some groups, like 10, 16, 25, and 47, present an exceptionally high number of combinations. To maintain practicality, we set a cap of 20 as the maximum number of combinations for our calculations (highlighted space groups), thus omitting space groups that surpass this limit and focusing instead on their subgroups.

Table 3.1 enumerates the relative energy compared to the lowest energy values found within each space group setting, while Figure 3.2 illustrates five low-energy structures along with their corresponding space groups. The lowest energy recorded across all space groups is -4.838 eV/atom, observed in space groups $Pmn2_1(31)$, $Pc(7)$, $P2_1(4)$, and $P1(1)$. These structures are identical to one another, representing the γ-phase Li_3PS_4 illustrated in Figure 3.2(d). From Figure 3.3, it can be observed that the simulation identified the γ-phase in the first 200 relaxed structures (10k steps of SA) in $Pmn2_1(31)$, $P2_1(4)$ and $P1(1)$, with the exception of $Pc(7)$. Consequently, further simulations are necessary for $Pc(7)$ to ensure it visits the possible lowest energy phase (γ-phase). In the second trial with a new

Table 3.1. Lowest energy/atom above the ground state found in different space groups, Li_3PS_4.

Space group	E/Atom(meV)	Space group	E/Atom(meV)	Space group	E/Atom(meV)
1	0	4	0	5	11
7	0	8	13	11	12
18	23	26	57	30	32
31	0	32	71	34	53
35	41	38	87	44	66
75	76	85	45	86	47
90	33	94	50	100	67
101	86	102	43	114	54
121	8	129	32	134	46
137	53	144	56	145	185
146	62	151	137	152	70
153	56	154	63	159	55
160	146	171	216	172	156
173	117	185	93	186	82
197	67	201	106	208	74
215	145	217	104	218	96
224	55				


(a) -4.825 eV/atom $Pm(6)$ (b) -4.826 eV/atom $Cm(8)$ (c) -4.823 eV/atom $Aem(39)$

(d) -4.828 eV/atom $Pmn2_1(31)$ (e) -4.830 eV/atom $I\bar{4}2m(121)$

Figure 3.2. Low-energy structures identified for Li_3PS_4.

initial random structure, the simulation successfully identified the γ-phase Li_3PS_4, with no structure of lower energy found after another 200 relaxed structures. According to our

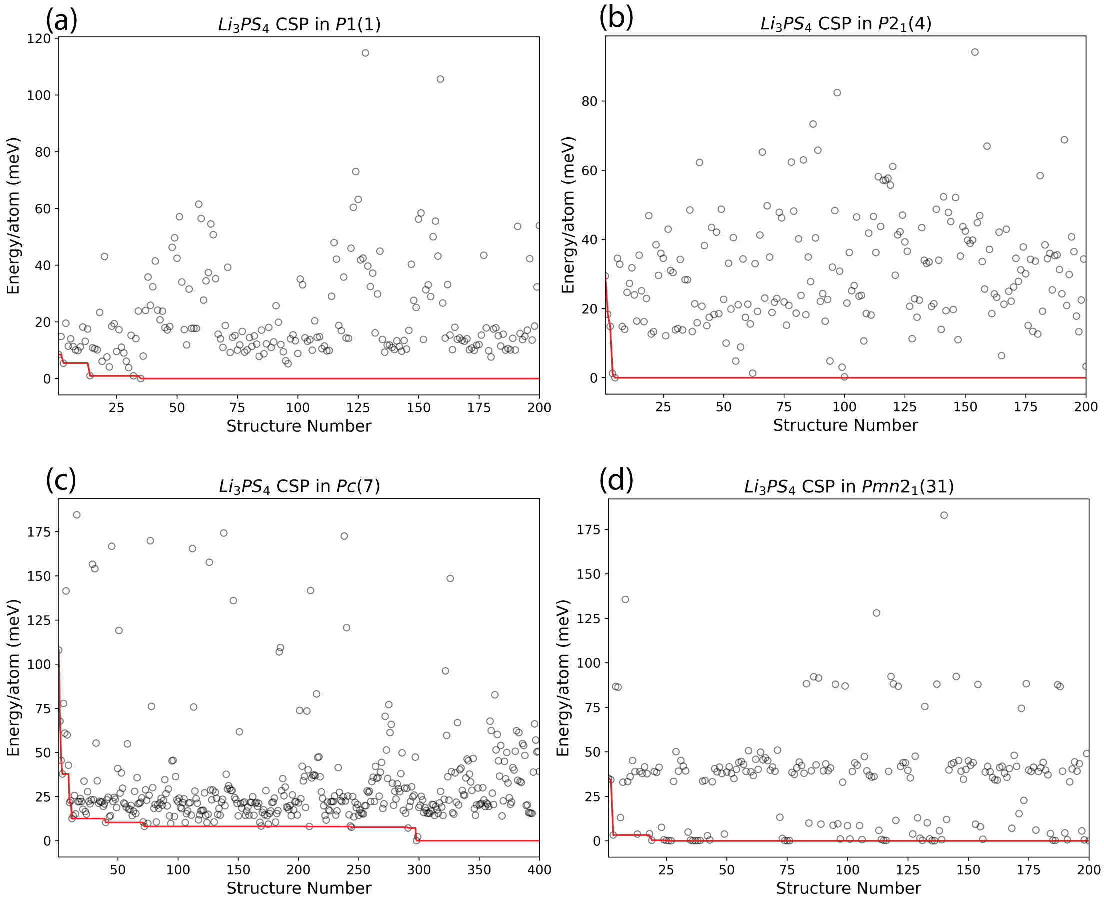


Figure 3.3. The relaxed structure numbers and their corresponding energies above the ground state for Li_3PS_4 are plotted for four different space groups:(a) $P1(1)$, (b) $P2_1(4)$, (c) $Pc(7)$ and (d) $Pmn2_1(31)$.

consistency check criteria, the fact that space group $Pmn2_1(31)$ found the same structure as its subgroups ($Pc(7)$, $P2_1(4)$ and $P1(1)$) adds another layer of confidence that the γ-phase is the lowest energy configuration in $Pmn2_1(31)$ and possibly the ground state of this chemical composition.

In addition to the γ-phase, we identified several metastable candidates as illustrated in Figure 3.2(a), (b), (c) and (e). The next lowest energy structure is found 8 meV/atom above the γ-phase and crystallizes in space group $I\bar{4}2m(121)$. This structure represents the well-known stannite structure, common in the M_xBX_4 compound family including systems

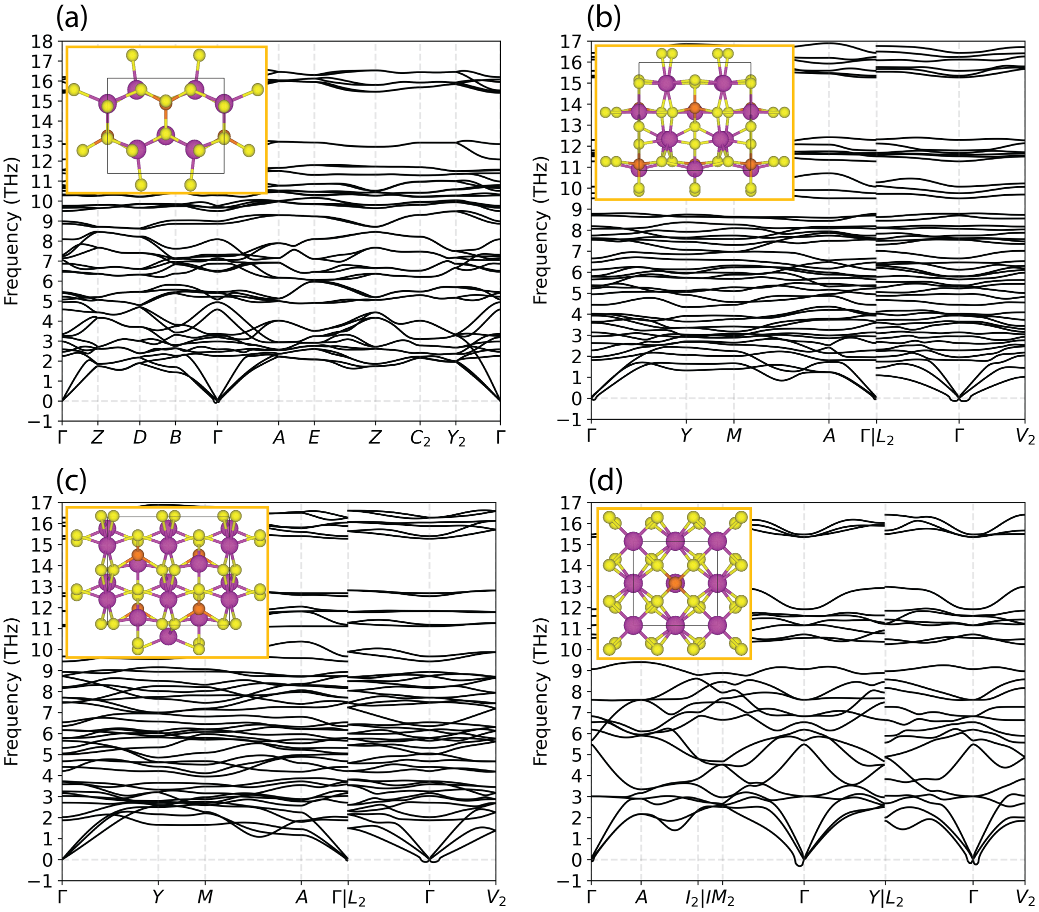


Figure 3.4. The phonon band structure of the predicted Li_3PS_4 structures.

like Cu_2ZnSnS_4[44, 55, 121], Zn_2CuGaS_4[13], $Cu_2HgSnSe_xTe_{4-x}$[80], and Cu_2FeSnS_4[37]. The subsequent lowest energy structure is found in $Cm(8)$, 12 meV/atom above the γ-phase, followed by structures in $Pm(6)$ and $Aem2(39)$, with 13 meV/atom and 15 meV/atom above the γ-phase respectively. The phonon band structures for these predicted structures are shown in Figure 3.4 to demonstrate their dynamical stability. With only a small amount of imaginary frequency around the Γ point due to accuracy numerical accuracy issues, these predicted Li_3PS_4 phases are dynamically stable at $T = 0$ K. However, it is possible for a low-energy structure to be dynamically unstable. For instance, the lowest energy structure found in space group $P2_1/m(11)$ is unstable, even though it is only 12 meV/atom above the γ-phase.

3.3.2. $Na_6Ge_2Se_6$. We now turn our attention to $Na_6Ge_2Se_6$. This compound features an ethane-like, dumbbell-shaped rigid block composed of germanium and selenium atoms as depicted in Figure 3.1. The symmetry elements of this rigid body include mirror planes, 2-fold and 3-fold rotation axes, 3-fold rotoinversion and inversion symmetry. The stoichiometric unit of $Na_6Ge_2Se_6$ comprises 14 atoms, exceeding the 8 atoms in Li_3PS_4, leading us to limit the unit cell size to two stoichiometric units for this compound. Details of the possible combinations are shown in Table 3.2, and Table 3.3 presents the lowest energies identified for each space group where calculations were conducted.

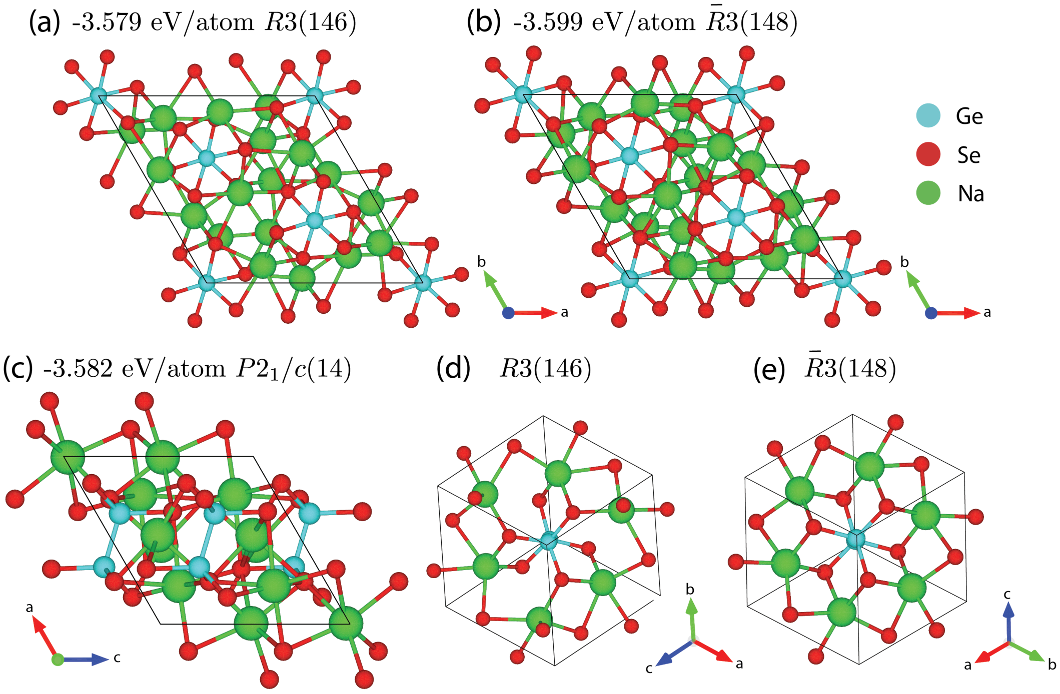


Figure 3.5. Low-energy structures identified for $Na_6Ge_2Se_6$.

The CSP results for space groups $P1(1)$, $P2_1(4)$, $Pc(7)$ and $P2_1/c(14)$ are presented in Figure 3.6. The lowest energy structures identified in groups $P2_1(4)$, $Pc(7)$, and $P2_1/c(14)$ are identical and it is illustrated in Figure 3.5(c). This $Na_6Ge_2Se_6$ phase

Table 3.2. Numbers of combinations for $Na_6Ge_2Se_6$ in different space groups with a maximum of two Ge_2Se_6 rigid bodies in the unit cell.

Space group	Combinations	Space group	Combinations	Space group	Combinations
1	2	2	>1000	3	>1000
4	1	5	32	6	228
7	1	8	4	10	>1000
11	140	13	632	14	16
16	>1000	17	520	18	32
25	>1000	26	32	27	520
28	150	30	32	31	4
32	32	34	32	49	>1000
51	>1000	53	228	55	228
58	136	75	84	77	150
81	828	83	874	84	616
143	>1000	147	151	149	>1000
150	244	156	>1000	157	62
158	117	159	24	162	247
163	120	164	247	165	22
168	27	173	24	174	>1000
176	30	177	272	182	124
182	124	183	37	185	8
186	26	187	>1000	188	417
189	213	190	31	193	18
194	32				

Note: We have applied the proposed CSP method to the highlighted space groups, which have 20 or fewer combinations. Combinations exceeding 1000 are marked as '> 1000'. Space groups not included in the table have no viable cases.

Table 3.3. Lowest energy/atom above the ground state found in different space groups, $Na_6Ge_2Se_6$

Space group	E/Atom(meV)	Space group	E/Atom(meV)	Space group	E/Atom(meV)
1	0	4	17	7	18
8	34	14	17	31	36
185	32	193	27		

was experimentally observed in 1985[23] and has an energy of -3.582 eV/atom. Notably, the CSP results for space group $P1(1)$ indicates the presence of several structures with significantly lower energy than the observed phase. Most of these structures crystallize in space group $P\bar{1}(2)$ following symmetry analysis (Figure 3.8). The lowest energy phase found from $P1(1)$ crystallizes in space group $R\bar{3}(148)$ as illustrated in Figure 3.5(b), with

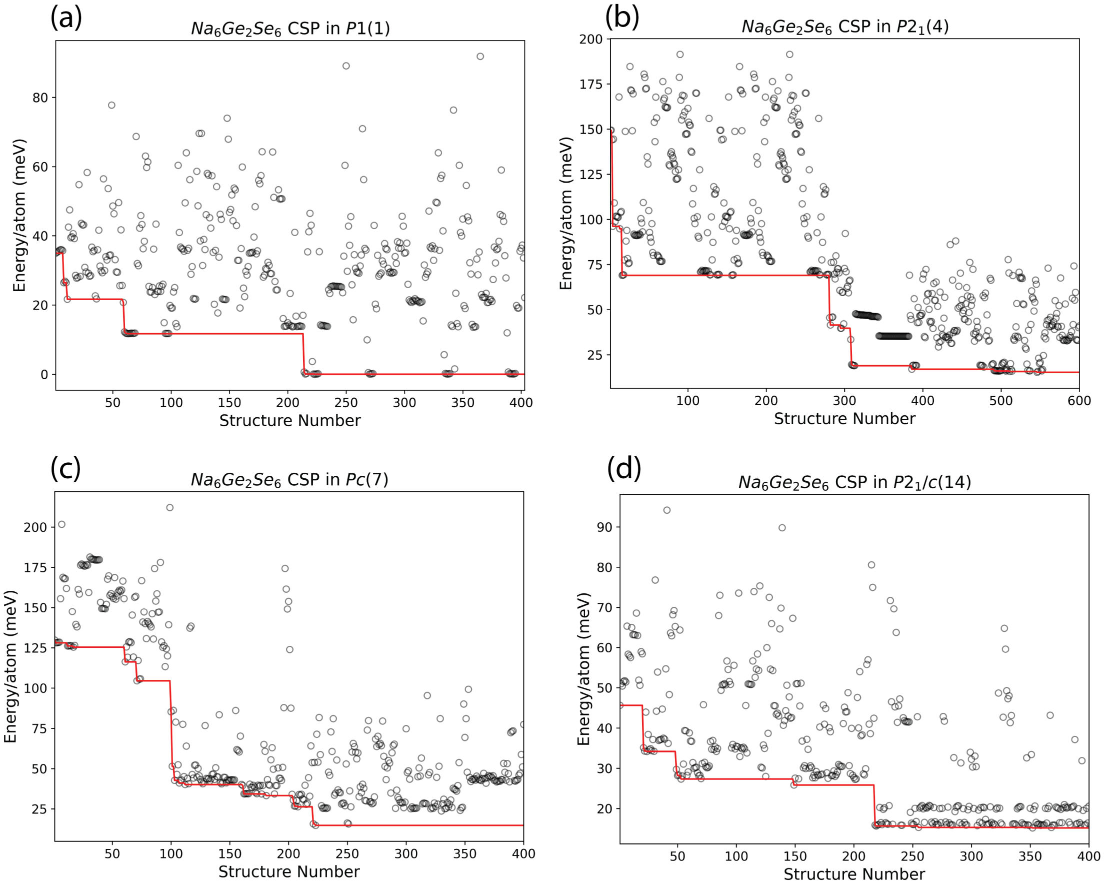


Figure 3.6. The relaxed structure numbers and their corresponding energies above the ground state of $Na_6Ge_2Se_6$ are plotted for four different space groups:(a) $P1(1)$, (b) $P2_1(4)$, (c) $Pc(7)$ and (d) $P2_1/c(14)$.

an energy 17 meV/atom lower than the observed phase. Additionally, another structure crystallizes in space group $R3(146)$ as shown in Figure 3.5(a), is a metastable phase candidate with an energy 3 meV/atom higher than the observed phase. The primitive cells of these two phases are illustrated in Figures 3.5(d) and 3.5(e). The calculation of their phonon band structures, presented in Figure 3.7, indicates their dynamical stability. The lowest energy phase in $R\bar{3}(148)$ differs markedly from the observed one, suggesting that the experimentally observed structure may be metastable. To ensure this finding is not a simulation artifact, we conducted additional relaxations with VASP using the PBE0 hybrid

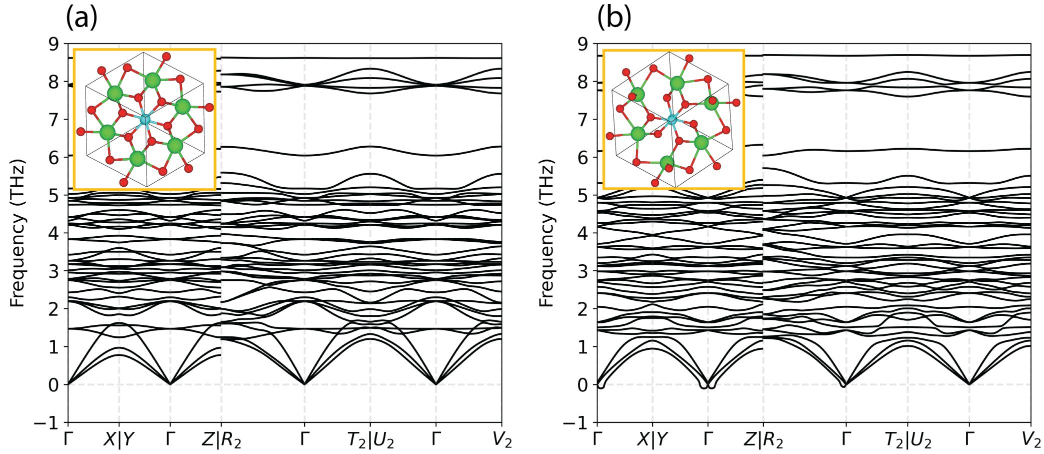
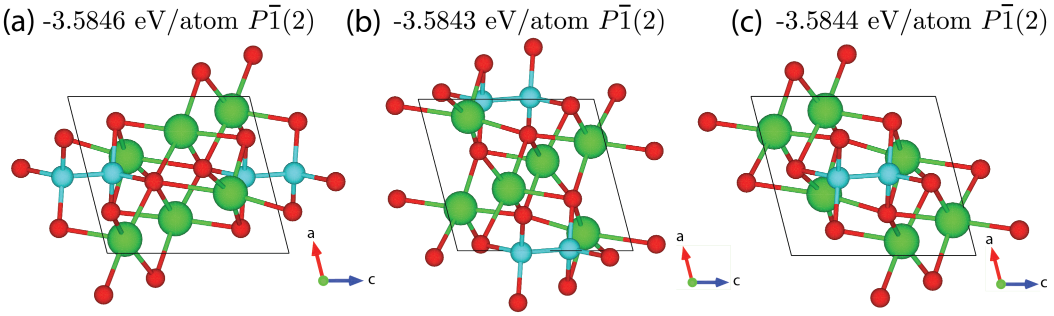


Figure 3.7. The phonon band structures of the predicted $Na_6Ge_2Se_6$ phases.


(a) -3.5846 eV/atom $P\bar{1}(2)$ (b) -3.5843 eV/atom $P\bar{1}(2)$ (c) -3.5844 eV/atom $P\bar{1}(2)$

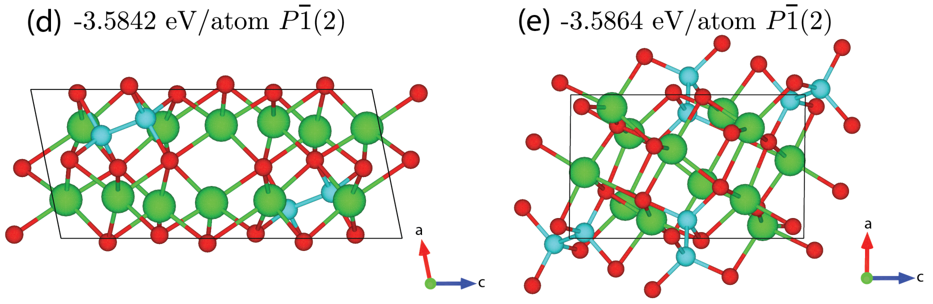

(d) -3.5842 eV/atom $P\bar{1}(2)$ (e) -3.5864 eV/atom $P\bar{1}(2)$

Figure 3.8. The low energy structure found crystallized in space group $P\bar{1}(2)$.

functional [3, 89], which provides a higher level of accuracy. The results indicate that its energy is 15 meV/atom lower than that of the observed phase, reaffirming the validity of our findings.

3.4. DISCUSSION

The maximum stoichiometric units for Li_3PS_4 and $Na_6Ge_2Se_6$ are three and two respectively in our simulations. However, post-simulation symmetry analysis can exceed these limits. For example, the Li_3PS_4 phase in Figure 3.2(c) contains four stoichiometric units, yet it emerged as a low-energy structure from the CSP results in $P1(1)$ with only two stoichiometric units. Using the Pymatgen symmetry analyzer, we found that the atoms of the relaxed structure with two Li_3PS_4 stoichiometric units occupy special WPs in a higher symmetry group with larger multiplicities. Similarly, the $Na_6Ge_2Se_6$ structures in Figures 3.5(a) and 3.5(b) have three stoichiometric units, despite the simulation cell containing only two. Although we limited the number of atoms in our simulations, it is still possible to identify low-energy structures with a larger number of atoms.

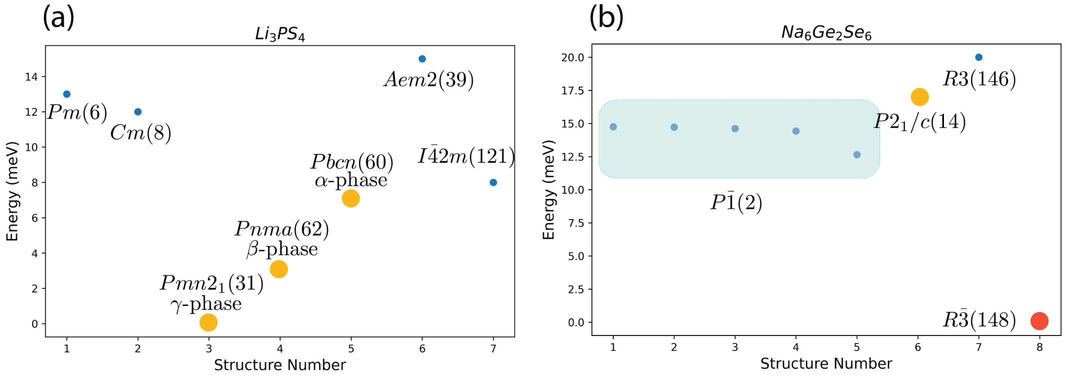


Figure 3.9. Energy comparison between predicted and observed phases for (a) Li_3PS_4 and (b) $Na_6Ge_2Se_6$.

Figure 3.9 illustrates the energy comparison between experimentally observed and predicted phases for Li_3PS_4 and $Na_6Ge_2Se_6$. For Li_3PS_4, we include the calculated energies of the experimentally observed α- and β-phases of Li_3PS_4 as reported[50] at 0 K. Notably, all candidates for metastable states possess higher energies than the the α-, β-, and γ-phase, aligning with the fact that these are the only phases observed to date. The next lowest energy structure in space group $I\bar{4}2m(121)$ is merely 1 meV/atom higher than the α-phase. For $Na_6Ge_2Se_6$, the predicted phase in $R\bar{3}(148)$ has a significantly lower energy than the observed phase. Its dynamical stability suggests that this phase might be experimentally observable. Further exploration of other properties of this structure is warranted, although it lies outside the scope of this paper.

3.4.1. Comparison to USPEX Results. The proposed method successfully identified the experimentally observed phases for Li_3PS_4 and $Na_6Ge_2Se_6$ along with several metastable candidates, demonstrating that focusing on the most promising regions of the phase space by fixing rigid blocks is an effective strategy for complex chalcogenides. There are other CSP tools that also support this feature, such as USPEX which utilizes evolutionary algorithms. One of the main differences between our methods is that we restrict atoms to move only within their designated WPs. Consequently, the CSP simulation is strictly confined to the pre-determined space group. As shown in Figure 3.3, this restriction significantly increases the chances of visiting the lowest energy phase in space group $Pmn2_1(31)$ compared to $P1(1)$, $P2_1(4)$, and $Pc(7)$, which have only one general WP without any restrictions on atomic positions. However, the WP restriction prohibits the situation where a lower energy phase requires a transition between WPs for atoms. For instance, Figure 3.6 illustrates that although the CSP simulation in space group $P2_1/c(14)$ frequently visited the observed $Na_6Ge_2Se_6$ phase, it was unable to identify even lower energy structures that were found in $P1(1)$.

While we impose a WP restriction in our method, USPEX employs mutation operations to deviate from the current state and explore more regions of the PES. This is one of the key aspects contributing to the performance of a CSP method. However, there may be limitations in identifying metastable phases compared to the WP restrictions we imposed in our method. We conducted the same simulations using USPEX for the Li_3PS_4 and $Na_6Ge_2Se_6$ systems with fixed PS_4 and $Na_6Ge_2Se_6$ rigid blocks. The results are presented in Figure 3.10. The USPEX simulations demonstrate a more systematic approach to evolving towards the lowest energy phase compared to the random sampling scheme we use. For Li_3PS_4, USPEX found the same two lowest energy structures as we did, depicted in Figures

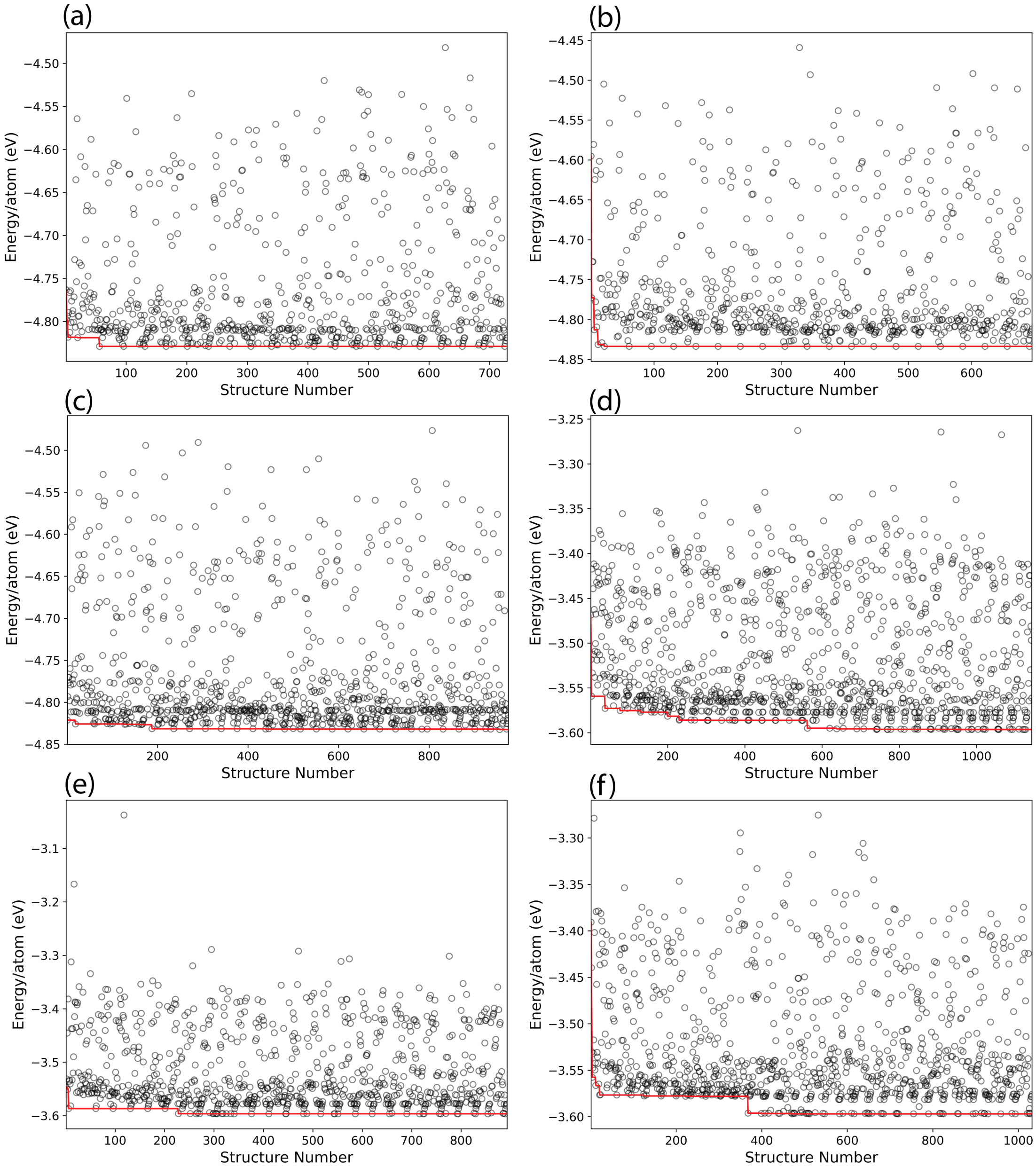


Figure 3.10. The USPEX CSP results for Li_3PS_4 and $Na_6Ge_2Se_6$.

3.2(d) and 3.2(e). However, it found the structure in $I\bar{4}2m$ (121) with a higher energy in only one out of three trials. In contrast, our method identified this structure multiple times in the CSP results within the space groups $P1(1)$, $P2_1(4)$ and $Pc(7)$, which are subgroups of $I\bar{4}2m(121)$. More importantly, for $Na_6Ge_2Se_6$, USPEX failed to find the experimentally observed phase in all three trials, despite all trials converging to the same lower energy structure illustrated in Figure 3.5(b). In contrast, our method identified the observed phase not only in the CSP results within space group $P2_1/c(14)$ but also in its subgroups $P2_1(4)$ and $Pc(7)$.

The possible reason for that is the $Na_6Ge_2Se_6$ system features many structures crystallizing in space group $P\bar{1}(2)$ with lower energy than the observed phase (Figure 3.8). The success of evolutionary algorithms depends on their ability to inherit the best structural patterns from their parent structures. With so many lower energy structures distinct from the observed one, the likelihood of visiting the configurations of the observed phase on the PES is significantly reduced when relying mainly on random structure generations and mutations. On the other hand, since we imposed WP restrictions in our method, it is prohibited to visit those phases with lower symmetry in $P\bar{1}(2)$ while conducting CSP simulations in space groups $P2_1(4)$, $Pc(7)$ and $P2_1/c(14)$. Consequently, this increases the likelihood of identifying the observed phase with higher energy.

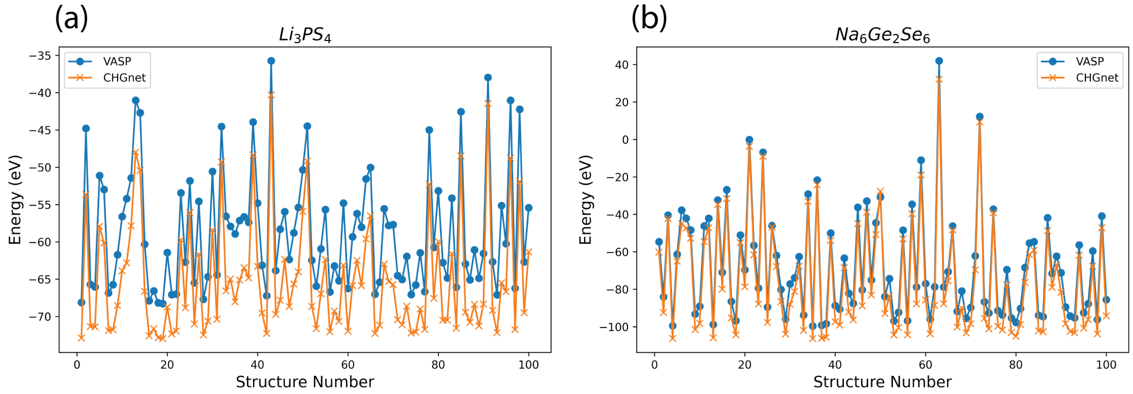


Figure 3.11. The energy of randomly generated structures comparison between VASP and CHGnet for (a) Li_3PS_4 and (b) $Na_6Ge_2Se_6$.

3.4.2. Performance of CHGnet. We also evaluated the accuracy of MLP CHGnet[22] in predicting energy of structures. For unrelaxed structures, We randomly generated a hundred structures for Li_3PS_4 and $Na_6Ge_2Se_6$, then evaluated their energy using VASP and CHGnet. The results, shown in Figure 3.11, demonstrate the remarkable performance of

CHGnet. Accurately determining the energy difference between two consecutive states is crucial for SA. CHGnet achieved 96% and 100% accuracy in predicting the lower energy state among two consecutive states for Li_3PS_4 and $Na_6Ge_2Se_6$, respectively. Regarding the absolute energy difference values, CHGnet managed to predict within a 16% difference compared to VASP. For the Li_3PS_4 system, the percentage difference increased to 37%. However, CHGnet's accuracy is not sufficient in differentiating the lower energy phase in relaxed structures. For $Na_6Ge_2Se_6$, it still maintains relatively high performance, successfully predicting the lowest energy phase. However, it predicted the energy of number three $Na_6Ge_2Se_6$ structure in Figure 3.9(b) to be 0.12 meV/atom lower than that of number four, whereas VASP results indicates it is 0.21 meV/atom higher. For Li_3PS_4, CHGnet predicted a lower energy for the $I\bar{4}2m(121)$ phase than for the γ-phase by 14 meV/atom, contrary to the VASP results. The outstanding performance of CHGnet in predicting the energies of unrelaxed structures has led us to incorporate it into the latest implementation of our method as an option for calculating the energy of unrelaxed structures during the SA process.

4. COMPARATIVE ANALYSIS OF TWO $Na_6Ge_2Se_6$ PHASES

4.1. INTRODUCTION

A new $Na_6Ge_2Se_6$ structure was computationally predicted to crystallize in space group $R\bar{3}(148)$, as discussed in Section 3.3.2. This compound represents a new phase distinct from the experimentally observed compound reported in 1985 with the same composition ($Na_6Ge_2Se_6$)[23]. The latter has been utilized recently in synthesizing a potential nonlinear optical material, $Na_8Mn_2(Ge_2Se_6)_2$[8]. Notably, the newly predicted $Na_6Ge_2Se_6$ phase shows a lower total energy compared to the observed phase from first-principles calculations, suggesting it as a potential meta-stable state of $Na_6Ge_2Se_6$ that has not been reported before. Furthermore, the observed $Na_6Ge_2Se_6$ phase has received little attention, with no comprehensive analysis conducted to the best of our knowledge. This gap in research has motivated us to perform a comparative analysis of the structural, elastic, electronic, phonon, thermal, and optical properties of these two compounds, which could serve as a valuable reference for future studies.

4.2. COMPUTATIONAL METHODOLOGY

In this study, we examined the ground state properties of both predicted and observed $Na_6Ge_2Se_6$ phases using density functional theory (DFT) as implemented in the Vienna Ab initio Simulation Package (VASP)[62, 63]. This involved solving the Kohn-Sham equations[60] to determine the ground state energies of a crystalline system. Given the close proximity of the ground state energies between the predicted and observed phases, calculations were performed using two distinct approaches for the exchange-correlation potential: the Generalized Gradient Approximation (GGA) within the Perdew-Burke-Ernzerhof framework (GGA-PBE)[89] and the hybrid functionals from Heyd-Scuseria-Ernzerhof (HSE06)[64]. In both $Na_6Ge_2Se_6$ structures, Integration over the Brillouin zone was executed via the tetrahedron method with Gaussian smearing, employing a $5 \times 5 \times 5$ Monkhorst–Pack k-point mesh[78]. The energy cutoff for the plane wave basis set was set to 600 eV. The Ge $3d^{10}4s^{3}4p^{2}$, Se $4s^{2}p^{4}$, Na $2p^{6}3s^{1}$ electrons are treated as valence electrons. convergence thresholds were 10^{-8} eV for the electronic self-consistent loop and 10^{-7} eV for structure optimization.

Further, we explored the elastic and dynamical properties of optimized structures using first-principle calculations with the GGA-PBE functional within VASP. This analysis allowed for the calculation of elastic constants C_{ij} and several elastic properties, including the bulk modulus B, Young's modulus E and shear modulus G, along with optical properties such as the dielectric function $\epsilon(\omega)$, absorption coefficient $\alpha(\omega)$ and conductivity $\sigma(\omega)$ using the VASPKIT[120] tool. The Phonopy package[107, 108] facilitated the calculation of phonon dispersion relations, phonon density of states, and related thermal properties. Considering the tendency of GGA-PBE to underestimate band gaps in semiconductors and insulators, the HSE06 functional[64] was applied to investigate electronic and optical properties. The electronic band structures and density of states for both phases were visualized using the open-source Python library Pymatgen[84].

4.3. RESULTS

4.3.1. Crystal Structure. We analyzed the $Na_6Ge_2Se_6$ structures using two different exchange-correlation energy functionals: GGA-PBE and the HSE06 hybrid functional. The calculated energies per atom are presented in Table 4.1. Both PBE and HSE06 functionals indicated lower energies for the predicted $Na_6Ge_2Se_6$ phase by 16 meV/atom. To confirm the convergence of our DFT calculations, we conducted two additional structure relaxations with the GGA-PBE functional: one with a 400 eV energy cutoff and a 3x3x3 k-point mesh, and another with a 500 eV energy cutoff and a 4x4x4 k-point mesh. For the predicted phase, both of these settings produced energy per atom within 0.006% of the most accurate result listed in the Table 4.1. For the observed phase, the difference is even smaller, within 0.002%. The energy difference between the predicted and observed phases listed above is an order of magnitude larger than 0.06%, thus numerical accuracy is sufficient to make conclusions about the energy order of two structures. This leads us to conclude that the predicted $Na_6Ge_2Se_6$ phase is energetically more favorable than the observed one at 0 K.

Table 4.1. Crystal data and DFT evaluated energy of predicted and observed $Na_6Ge_2Se_6$.

	Predicted	Observed
Empirical formula	$Na_6Ge_2Se_6$	$Na_6Ge_2Se_6$
Crystal system	Trigonal	Monoclinic
Space group	$R\bar{3}$	$P2_1/c$
Unit cell	$a = 11.349$ Å	$a = 8.445$ Å
	$b = 11.349$ Å	$b = 12.038$ Å
	$c = 10.900$ Å	$c = 8.316$ Å
	$\alpha = 90°$	$\alpha = 90°$
	$\beta = 90°$	$\beta = 118.877°$
	$\gamma = 120°$	$\gamma = 90°$
Volume (Å^3)	$V = 405.27$	740.20
Z	1	2
Density(ρ)	3.075 g/cm^{-3}	3.367 g/cm^{-3}
Energy (eV/atom)		
GGA-PBE	−3.598	−3.582
HSE06	−4.268	−4.252
Bond length		
Ge-Ge	2.49 Å	2.47 Å
Ge-Se	2.39 Å	2.38 Å

The predicted $Na_6Ge_2Se_6$ crystallizes in the space group $R\bar{3}$ (No.148), whereas the observed phase is found in the space group $P2_1/c$ (No.14). Both structures exhibit identical ethane-like selenide [Ge_2Se_6] units, featuring a Ge-Ge bond length of 2.49 Å for the predicted phase and 2.47 Å for the observed phase, and an average Ge-Se bond length of 2.39 Å as detailed in Table 4.1. These bond lengths are slightly greater than the originally observed $Na_6Ge_2Se_6$ data[23], which recorded a Ge-Ge bond length of 2.43 Å and a Ge-Se bond length of 2.33 Å. The [Ge_2Se_6] ethane-like dimer configuration bears similarity to the [P_2Se_6] in $Mg_2P_2Se_6$[48], and [Si_2Se_6] in $Na_4MgSi_2Se_6$[124]. For both $Na_6Ge_2Se_6$ phases, the formed [Ge_2Se_6] dimers are isolated from each other, similar to the cases in

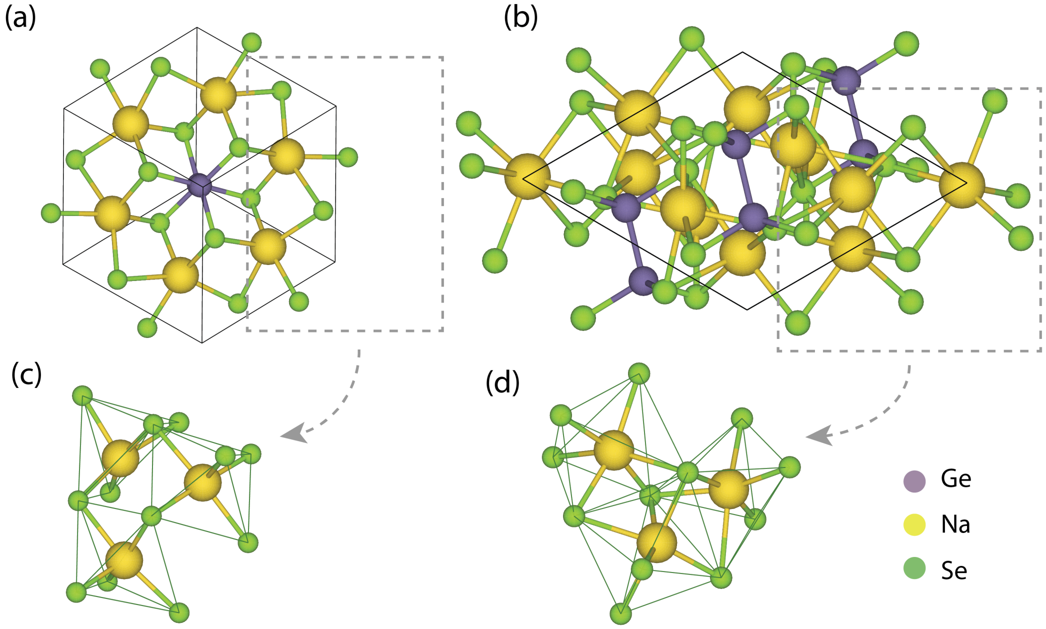


Figure 4.1. Crystal structures of both $Na_6Ge_2Se_6$ phases.

$K_6Ge_2Se_6$[24] and $Cs_6Ge_2Se_6$[99]. The predicted $Na_6Ge_2Se_6$ phase exhibits cell parameters of $a = 11.349$ Å, $b = 11.349$ Å, $c = 10.900$ Å, $\alpha = 90°$, $\beta = 90°$ and $\gamma = 120°$, while the observed $Na_6Ge_2Se_6$ has $a = 8.445$ Å, $b = 12.038$ Å, $c = 8.316$ Å, $\alpha = 90°$, $\beta = 118.877°$ and $\gamma = 90°$ with a smaller unit cell volume. The predicted $Na_6Ge_2Se_6$ features only one crystallographically unique Na atom, one Ge atom, and one Se atom, contrasting with the observed $Na_6Ge_2Se_6$, which has three unique Na atoms, one Ge atom, and three Se atoms. In the predicted $Na_6Ge_2Se_6$, each Na atom coordinates with five Se atoms to form a [$NaSe_5$] pyramid. This pyramid shares corners with its neighboring [$NaSe_5$] pyramids, whereas in the observed phase, each Na atom bonds with six Se atoms to form a [$NaSe_6$] slightly distorted octahedron. These octahedra are connected by face-sharing with their neighboring [$NaSe_6$] octahedron, as illustrated in Figure 4.1(c) and Figure 4.1(d).

4.3.2. Elastic Properties. Analyzing the elastic properties offers insights into the mechanical behavior of compounds, including their stability, ductility, and brittleness. Such information is especially valuable in industrial and device manufacturing sectors, enhancing our understanding of the forces acting in solids. To calculate elastic constants, we employed the energy-strain method[67], as facilitated by VASPKIT[120] based on DFT calculations with GGA-PBE functional.

For the predicted $Na_6Ge_2Se_6$ phase, which crystallizes in the trigonal crystal system, There are five independent elastic constants: C_{11}, C_{33}, C_{44}, C_{12}, C_{13} with the added relation

$$C_{66} = (C_{11} - C_{12})/2 \tag{4.1}$$

This contrasts with its observed counterpart, which crystallizes in monoclinic crystal system with lower symmetry that has thirteen independent elastic constants: C_{11}, C_{12}, C_{13}, C_{15}, C_{22}, C_{23}, C_{25}, C_{33}, C_{35} , C_{44}, C_{46} ,C_{55} and C_{66}. The configurations of these constants constitute the elastic matrices for each respective phase as follows:

$$C_{predicted} = \begin{pmatrix} 32.06 & 10.86 & 13.31 & \cdot & \cdot & \cdot \\ \cdot & 32.06 & 13.31 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 21.06 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 12.09 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 12.09 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 10.60 \end{pmatrix} \tag{4.2}$$

$$C_{observed} = \begin{pmatrix} 28.28 & 10.82 & 15.82 & \cdot & 2.45 & \cdot \\ \cdot & 32.08 & 15.21 & \cdot & 1.55 & \cdot \\ \cdot & \cdot & 40.55 & \cdot & 4.76 & \cdot \\ \cdot & \cdot & \cdot & 10.43 & \cdot & 2.81 \\ \cdot & \cdot & \cdot & \cdot & 10.67 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 10.75 \end{pmatrix} \tag{4.3}$$

These elastic constants are pivotal for assessing other elastic properties. A crystal must satisfy the Born-Huang criteria[79] to be deemed mechanically stable. For the predicted structure, the necessary and sufficient Born-Huang criteria include:

$$C_{11} > |C_{12}|, C_{44} > 0, C_{66} > 0 \tag{4.4}$$

$$2C_{13}^2 < C_{33}(C_{11} + C_{12}) \tag{4.5}$$

For monoclinic crystals, the stability criteria are more complex due to the lower symmetry of the crystal structure. They are listed as follows:

$$C_{11} > 0, C_{22} > 0, C_{33} > 0, C_{44} > 0, C_{55} > 0, C_{66} > 0 \tag{4.6}$$

$$C_{11} + C_{22} + C_{33} + 2(C_{12} + C_{13} + C_{23}) > 0 \tag{4.7}$$

$$C_{11} + C_{22} - 2C_{12} > 0 \tag{4.8}$$

$$C_{11} + C_{33} - 2C_{13} > 0 \tag{4.9}$$

$$C_{22} + C_{33} - 2C_{23} > 0 \tag{4.10}$$

$$C_{11}C_{22}C_{33} + 2C_{12}C_{13}C_{23} - C_{11}C_{23}^2 - C_{22}C_{13}^2 - C_{33}C_{12}^2 > 0 \tag{4.11}$$

The calculated elastic constants for both $Na_6Ge_2Se_6$ phases fulfill their respective stability criteria, indicating mechanical stability for both the predicted and observed phases.

Constants C_{11}, C_{22} and C_{33} quantify the response to uniaxial strain along the three principal axes. The differences between these constants measure the extent of anisotropy in linear compressibility. From matrix $C_{predicted}$ and $C_{observed}$, we can see that for predicted phase, $C_{11} = C_{22} > C_{33}$ which means linear compressibility is isotropic in the x- and y-directions and greater than in the z-direction. For the observed phase, $C_{33} > C_{22} >$

C_{11}, indicating that the observed $Na_6Ge_2Se_6$ phase is characterized by anisotropic linear compressibility. While both phases have similar resistance to deformation along the x- and y-directions, the observed phase is significantly stiffer in the z-direction.

On the other hand, C_{44}, C_{55} and C_{66} assess the material's resistance to shear deformation about these axes. Examination of both elastic constant matrices reveals that for the predicted $Na_6Ge_2Se_6$ phase, it exhibits similar C_{44}, C_{55} and C_{66}, indicating its shear modulus is relatively weakly anisotropic. For the predicted phase, the resistance to shear deformation is isotropic about the x- and y-directions and greater than in the z-direction. Comparing both phases, the observed phase shows slightly higher resistance to shear deformation about the principal axes.

The off-diagonal terms indicate coupling between different deformation modes. For instance, C_{12} correlates the stress applied in the x-direction to the resultant strain in the y-direction. Analysis of the matrices suggests that both phases have a similar response to deformation in the y-direction when stress is applied in the x-direction. However, the predicted phase shows greater shear strain in the z-direction induced by stress in both the x- and y-directions.

Utilizing the independent elastic constants, we can derive key mechanical properties such as the bulk modulus (B), shear modulus (G), Young's modulus (E), and Poisson's ratio (v) through the following relations, with the calculated values presented in Table 4.2:

$$B = \frac{B_V + B_R}{2} \tag{4.12}$$

$$G = \frac{G_V + G_R}{2} \tag{4.13}$$

$$E = \frac{9BG}{3B + G} \tag{4.14}$$

$$v = \frac{3B - 2G}{2(3B + G)} \tag{4.15}$$

Here, B_V, B_R and G_V, G_R represent the Voigt[117] and Reuss[97] values of the bulk modulus and shear modulus respectively. Furthermore, we evaluated Kleinman's parameter (ζ) using VASPKIT[120].

Table 4.2. Calculated mechanical properties of polycrystal for both predicted and observed $Na_6Ge_2Se_6$ (all in GPa).

	B	E	G	ν	B/G	ζ
Predicted	17.543	24.407	9.623	0.268	1.82	0.59
Observed	19.770	25.731	10.027	0.283	1.97	0.65

The bulk and shear moduli offer insights into a crystal's mechanical behavior, indicating its resistance to volumetric and shape deformations, respectively. The data in the Table 4.2 reveal that the observed $Na_6Ge_2Se_6$ phase possesses a greater resistance to deformation. This is expected as the unit-cell volume of the predicted phase is greater than that of the observed phase. Pugh's ratio (B/G) serves as a gauge for a material's ductility or brittleness, with a threshold value of 1.75 distinguishing between the two behaviors where the material is claimed to be brittle, if $B/G < 1.75$ and classified to be ductile if $B/G > 1.75$[94]. Our findings suggest that both phases exhibit ductility. The stiffness and thermal shock resistance of these materials are inferred from Young's modulus (E)[58], indicating superior stiffness in the observed phase. Conversely, a lower Young's modulus hints at enhanced thermal shock resistance[53] for the predicted phase. Poisson's ratio (ν) sheds light on the nature of bonding forces within a crystal and aids in evaluating its mechanical properties, including stability against shear[95]. Values for the predicted and observed phases are 0.268 and 0.283, respectively, falling within the central force solids' expected range of 0.25 to 0.50[7]. This metric also predicts the material's ductility or brittleness, classifying those with $\nu > 0.26$ as ductile and those below 0.26 as brittle[53, 115]. Thus, our analysis corroborates both phases' ductility, aligning with conclusions drawn from Pugh's ratio. Lastly, the Kleinman parameter (ζ), varying between 0 and 1, indicates the dominance of bond stretching or bending in resisting external stress[59] where the lower value of the Kleinman parameter ζ indicates a minimal role of bond bending in resisting external stress, whereas a higher value suggests a negligible contribution of bond stretching or contracting to resist externally applied stress. For both $Na_6Ge_2Se_6$ phases, a $\zeta > 0.5$ indicates a major role for bond bending in its mechanical strength.

Figure 4.2 and 4.3 shows the 3D and 2D representations of Young's modulus and linear compressibility for predicted and observed $Na_6Ge_2Se_6$ phases respectively. For the predicted phase, the Young's modulus and linear compressibility in the xy-plane form a circle, as shown in Figure 4.2(a) and (c), suggesting both properties are isotropic within the xy-plane. However, both properties show significant deviation from a circular form within the yz- and xz-planes, indicating that Young's modulus and linear compressibility are anisotropic overall for the predicted $Na_6Ge_2Se_6$. For the observed phase, Figure 4.3 shows that both properties are anisotropic, with the 2D representation of Young's modulus in the xy- and yz-planes and linear compressibility in the xy-plane more closely resembling a circle. This suggests relative isotropy for both properties within the corresponding planes.

4.3.3. Electronic Properties. We present the electronic band structure along with the partial and total density of states (DOS) in Figure 4.4. These calculations were performed using optimized lattice parameters along high-symmetry directions in the first Brillouin zone. The partial density of states (PDOS) on orbitals and elements are depicted in Figure 4.5. The band structures and DOS for both $Na_6Ge_2Se_6$ phases were determined using the HSE06 hybrid exchange-correlation functional[64] to ensure enhanced accuracy[40]. The energy reference point is set at the top of the valence band. The band structure analysis reveals that the predicted $Na_6Ge_2Se_6$ phase possesses an indirect band gap, with the conduction band minimum (CBM) located at the X-point and the valence band maximum (VBM) at the Z-point, resulting in a band gap of 2.972 eV. In contrast, the observed $Na_6Ge_2Se_6$ phase exhibits a direct band gap of 2.932 eV, with both the CBM and VBM situated at the Y_2-point.

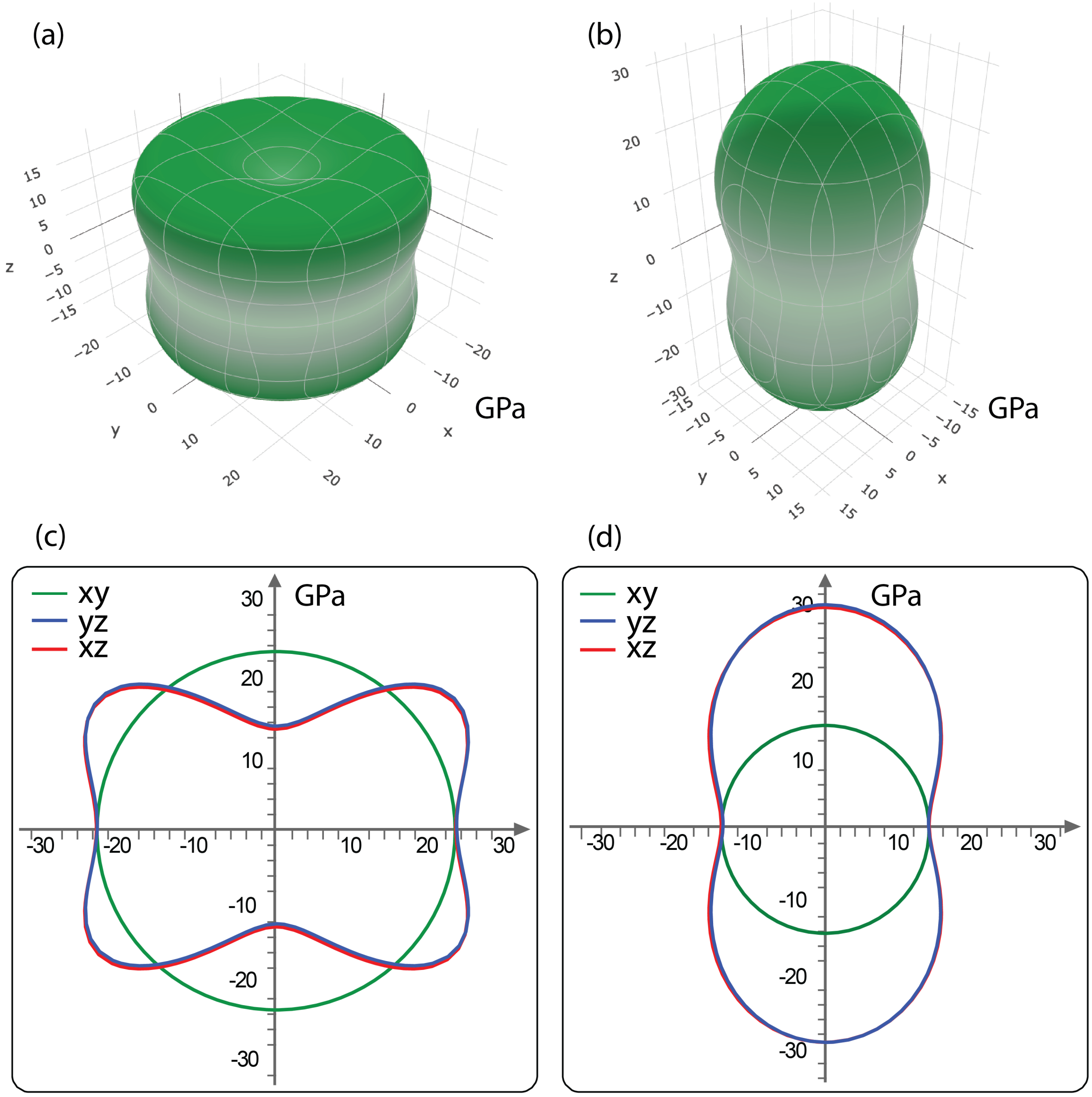


Figure 4.2. 3D representations of the spatial dependence of the (a) Young's modulus and (b) linear compressibility for predicted $Na_6Ge_2Se_6$ phase.

From Figure 4.5, it can be seen that both phases share similar orbital and element contributions to the total DOS. Regions close to the Fermi level are predominantly influenced by contributions from Se-p and Ge-s orbitals, while Na's involvement near the Fermi level is minimal. This observation suggests that the band gaps in both $Na_6Ge_2Se_6$ phases are largely attributable to the $[Ge_2Se_6]$ dimers. Within the valence band, ranging from -4 to 0

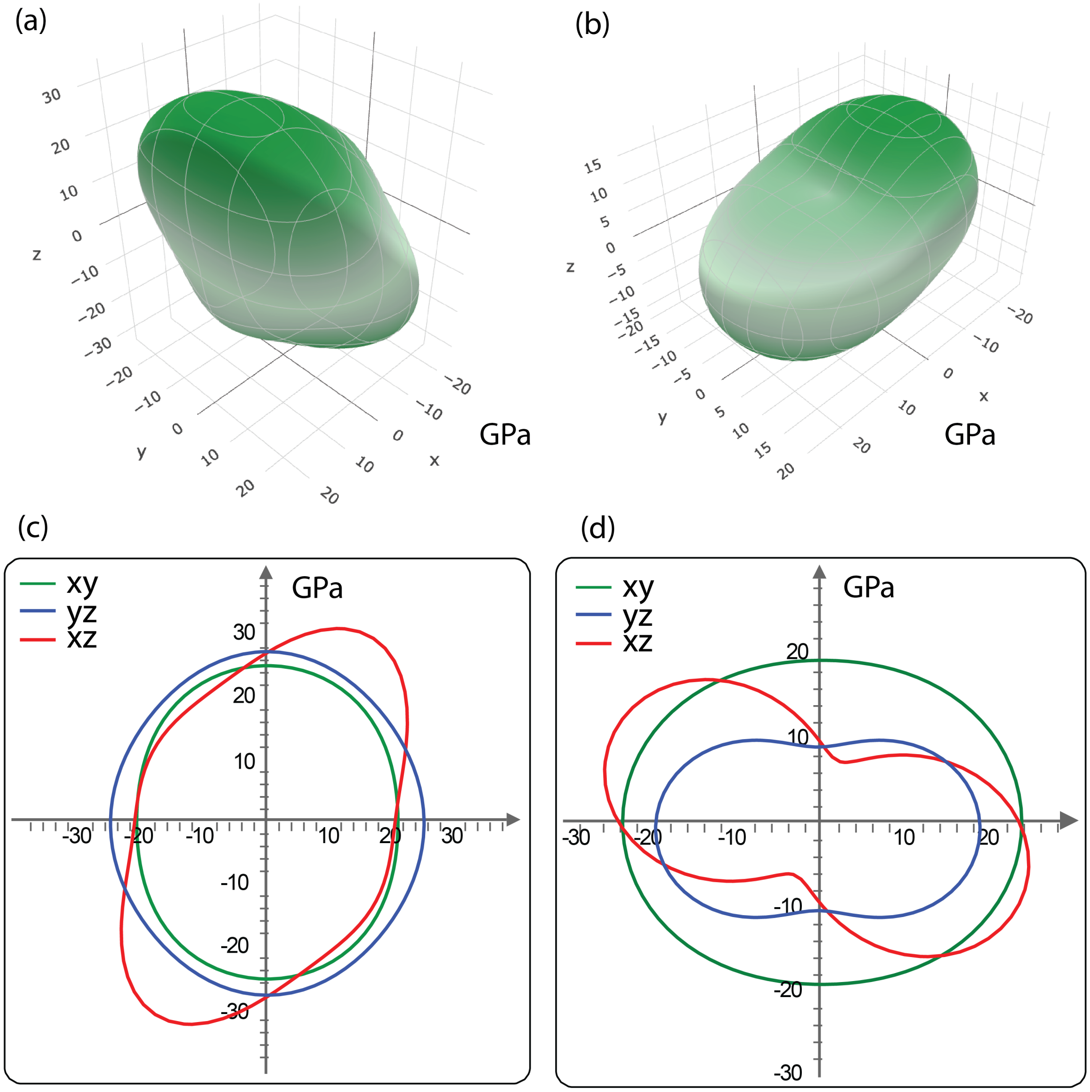


Figure 4.3. 3D representations of the spatial dependence of the (a) Young's modulus and (b) linear compressibility for observed $Na_6Ge_2Se_6$ phase.

eV, the primary contributions come from Se-p and Ge-p orbitals in both phases. This trend extends to the lower regions of the conduction band, from 4 to 8 eV, where Na-p displays a slight contribution to the total DOS in this region.

4.3.4. Phonon and Thermal Properties. Investigating phonon properties is fundamental for understanding crystalline materials, as it provides insights into structural stability, phase transitions, and how vibrations influence their thermal behavior. In our

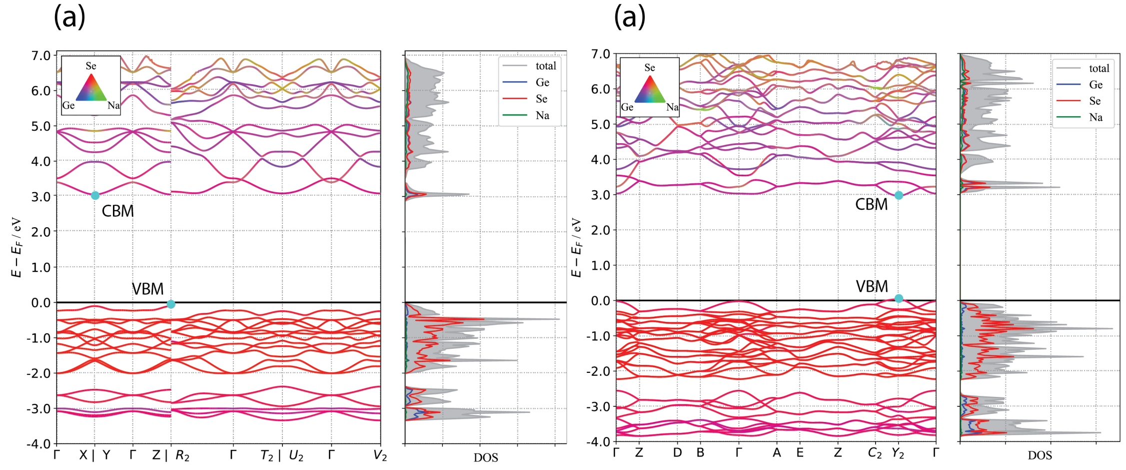


Figure 4.4. Electronic band structures (BS) and density of states (DOS) of both $Na_6Ge_2Se_6$ compounds.

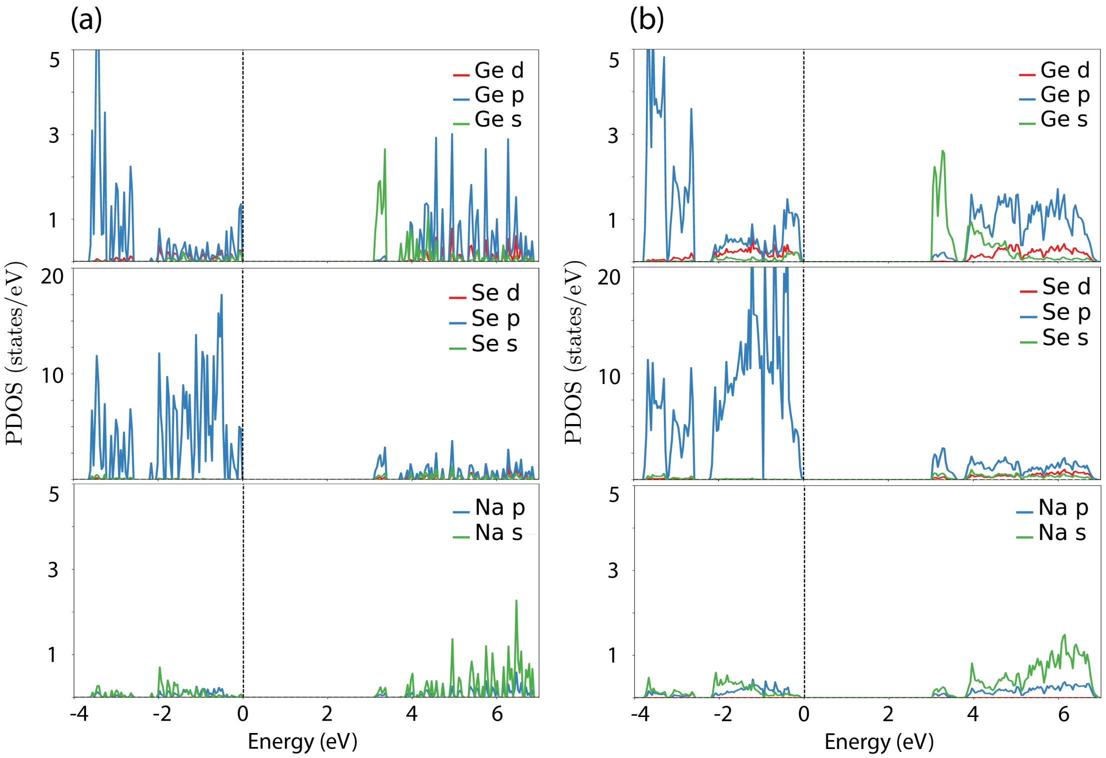


Figure 4.5. The partial density of states (PDOS) on orbitals and elements for (a) the predicted $Na_6Ge_2Se_6$ and (b) the observed $Na_6Ge_2Se_6$.

study of the phonon dispersion and phonon band structures for both $Na_6Ge_2Se_6$ phases, we employed the Phonopy package[107, 108] to create a series of $2 \times 2 \times 2$ supercell structures with various displacements. We then conducted force calculations using VASP, employing the GGA-PBE exchange-correlation functional. The phonon frequencies and eigenvectors were determined from the dynamical matrices, which were calculated based on the force constants derived using Phonopy.

The phonon dispersion relations along high-symmetry paths for both phases are depicted in Figure 4.6. The unit cell of the predicted phase comprises 14 atoms, resulting in 42 phonon branches, while the observed phase has 28 atoms, resulting in 82 branches, which include three acoustic and the rest optical branches. The frequency spectrum of these modes spans from 0 to 10 THz, without any discernible gap between the acoustic and optical modes for either phase. In the lower frequency domain, below 7 THz, Na and Se atoms predominantly contribute to the optical branches. In the higher frequency range, the contributions mainly stem from Ge and Se. This distribution underscores the strong bonding present within the $[Ge_2Se_6]$ dimers, reflecting their structural integrity across both phases of $Na_6Ge_2Se_6$. The lack of negative frequency branches within the dispersion plots affirms the dynamic stability of these phases at zero pressure.

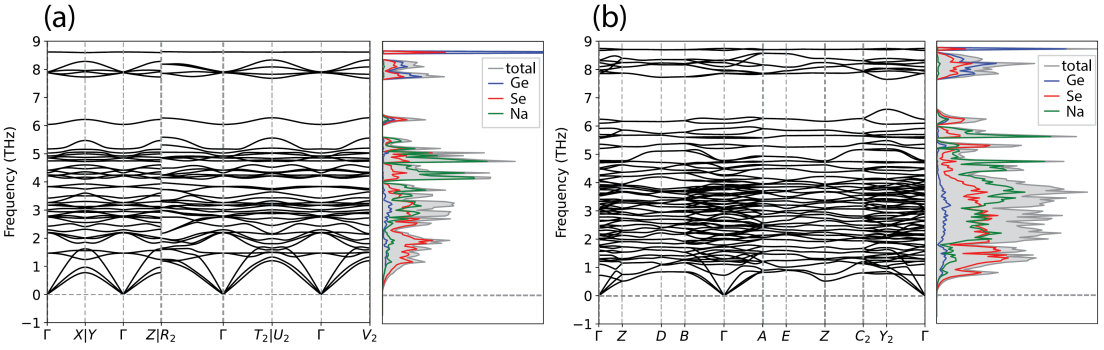


Figure 4.6. Phonon density of states and band structure of both $Na_6Ge_2Se_6$ compounds for (a) observed $Na_6Ge_2Se_6$ phase and (b) predicted $Na_6Ge_2Se_6$ phase.

The Debye temperature (θ_D) represents the temperature corresponding to the highest energy vibrational mode in a solid, derived from the equation[6]:

$$\theta_D = \frac{h}{k_B}\left[\frac{3n}{4\pi V_0}\right]^{\frac{1}{3}} v_m \tag{4.16}$$

Here, h is the Planck's constant, k_B is the Boltzmann's constant, n is the number of atoms in the unit cell, V_0 is the unit cell's equilibrium volume, and v_m is the material's average speed of sound. The average speed of sound v_m can be determined from the material's mass density (ρ), along with its bulk (B) and shear modulus (G), through the equations for longitudinal (v_l) and transverse (v_t) sound speeds:

$$v_m = \left[\frac{1}{3}\left(\frac{3}{v_t^3} + \frac{1}{v_l^3}\right)\right]^{\frac{1}{3}} \tag{4.17}$$

$$v_t = \sqrt{\frac{G}{\rho}} \tag{4.18}$$

$$v_l = \sqrt{\frac{3B + 4G}{3\rho}} \tag{4.19}$$

The calculated values for v_t, v_l, v_m and θ_D are provided in Table 4.3. These values indicate that the predicted $Na_6Ge_2Se_6$ phase exhibits a slower wave speed in the longitudinal direction but a higher wave speed in the transverse direction and overall average speed. The Debye temperature of the predicted phase is approximately 0.89% lower than that of its counterpart.

Table 4.3. Calculated v_t, v_l, v_m and θ_D for both predicted and observed $Na_6Ge_2Se_6$.

	v_t(m/s)	v_l(m/s)	v_m(m/s)	θ_D (K)
Predicted	1761.41	3129.30	1959.71	190.0
Observed	1725.51	3136.88	1923.28	191.7

Phonon calculations offer insights into additional thermal properties such as entropy (S), constant-volume heat capacity (C_v), and Helmholtz free energy (F), which are depicted in Figure 4.7. For heat capacity, it's notable from Figure 4.7(b) that the curves for both phases exhibit a similar trend, starting from zero and rising exponentially with increasing temperature up to around 300 K. Beyond this point, the rate of increase slows as they approach the Dulong-Petit limit[**?**], which is identified to be around 696 J/(mol K) for both phases.

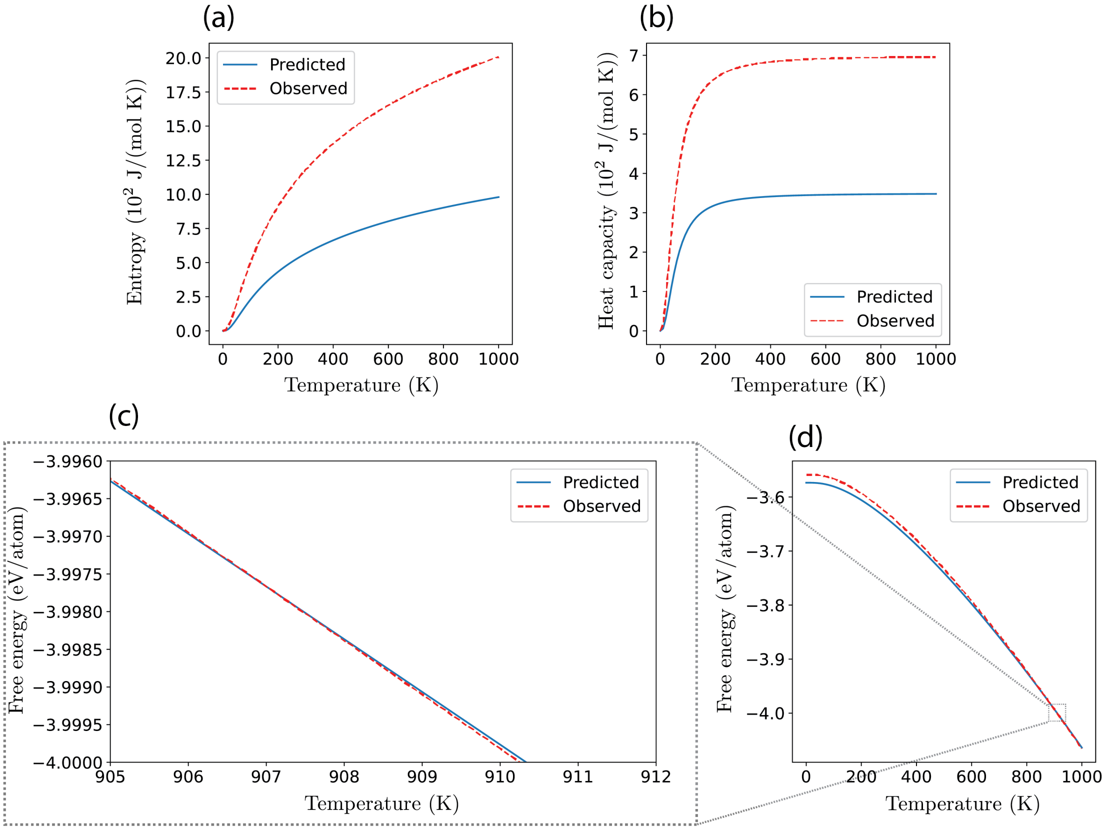


Figure 4.7. Thermal properties comparison between predicted and observed $Na_6Ge_2Se_6$ phases.

The calculation of the Helmholtz free energy (F) within the harmonic approximation[29] facilitates an evaluation of the thermal contributions to the relative stability of two phases, as described by the following equation:

$$F = E_{total} + \frac{1}{2}\sum_{qv}\hbar\omega(qv) + k_B T\sum_{qv}\ln\left[1 - exp(-\hbar\omega(qv)/k_B T\right] \tag{4.20}$$

Here, E_{total} denotes the total energy of the crystal, available in Table 4.1. The summed terms represent the Helmholtz free energy attributable to phonons[109], with the initial sum reflecting the zero-point energy (ZPE) that is independent of temperature. The predicted phase's ZPE is approximately 1 meV/atom higher than that of the observed phase. The subsequent sum accounts for the temperature-dependent term referring to the thermally induced occupation of phonon modes. The Helmholtz free energies, plotted in Figure 4.7(d) for temperatures ranging from 0 to 1000 K, with a closer look at the range of 905 to 912 K in Figure 4.7(c), illustrate that the predicted phase possesses lower free energy from 0 to approximately 907 K. Since its rate of decrease with temperature is more gradual than that of the observed phase, the free energy of the observed phase becomes lower when the temperature exceeds 907 K. This suggests that the predicted phase is more thermodynamically stable than the observed phase in the temperature range from 0 to 907 K, indicating that this predicted phase is very likely to be experimentally observable.

4.3.5. Optical Properties. A material's optical behavior is characterized by several energy/frequency-dependent parameters, such as the dielectric function $\epsilon(\omega)$, absorption coefficient $\alpha(\omega)$, conductivity $\sigma(\omega)$, energy-loss function $L(\omega)$, reflectivity $R(\omega)$ and refractive index $n(\omega)$. The outcomes of these calculated properties are depicted in Figure 4.5, based on formulas integrated within VASPKIT[120].

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega) \tag{4.21}$$

$$\alpha(\omega) = \frac{\sqrt{2}\omega}{c}\left[\sqrt{\epsilon_1^2 + \epsilon_2^2} - \epsilon_1\right]^{\frac{1}{2}} \tag{4.22}$$

$$L(\omega) = \frac{\epsilon_2}{\epsilon_1^2 + \epsilon_2^2} \tag{4.23}$$

$$n(\omega) = \left[\frac{\sqrt{\epsilon_1^2 + \epsilon_2^2} + \epsilon_1}{2}\right]^{\frac{1}{2}} \tag{4.24}$$

$$R(\omega) = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2} \tag{4.25}$$

To examine the anisotropic characteristics, we calculated the optical parameters along the three principal axes and presented the optical spectra for incident photon energies up to 25 eV. For the predicted phase, these parameters are isotropic in the x- and y-directions. Therefore, Figure 4.8 displays the optical parameters and the corresponding derived optical properties plotted only in the y- and z-directions for the predicted phase.

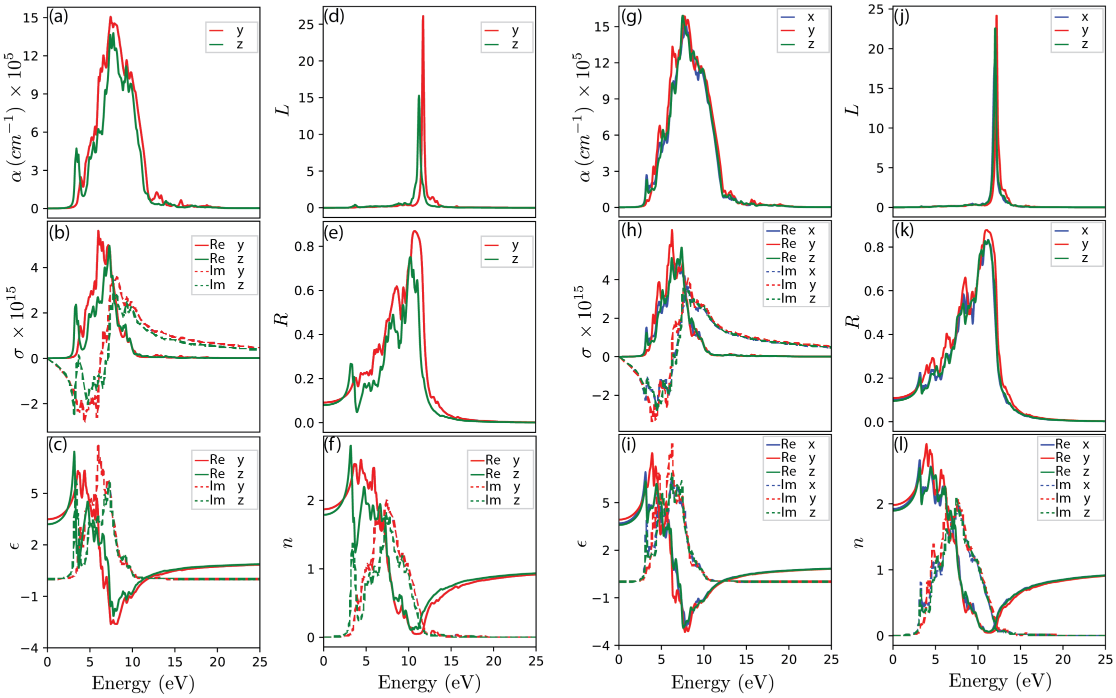


Figure 4.8. Optical properties for predicted $Na_6Ge_2Se_6$ and observed $Na_6Ge_2Se_6$.

The complex dielectric function is essential for describing a material's optical properties, as it serves as the foundation from which other energy-dependent optical constants are derived. We illustrate the dielectric functions for both predicted and observed $Na_6Ge_2Se_6$ phases in Figure 4.8(c) and (i), respectively. The real part of the dielectric function (ϵ_1) initially increases, reaches a peak, then sharply decreases, eventually dropping below zero. The imaginary part (ϵ_2), indicative of dielectric loss, closely aligns with the optical absorption coefficient. From the figure we can see ϵ_1 and ϵ_2 peak in different directions. For the predicted phase, ϵ_1 peaks in z-direction, and ϵ_2 peaks in y-direction. This contrasts with the observed phase, where both ϵ_1 and ϵ_2 peak in the same y-direction. The material's electronic characteristics can be inferred from the absorption coefficient ($\alpha(\omega)$), depicted in Figure 4.8(a) and (g). Both phases begin to absorb at photon energies of approximately 3.1 eV and 2.9 eV respectively, indicating their optical band gaps. The absorption coefficients in three different directions follow a similar trend, with peaks at 7.4 eV for the predicted phase and 9.5 eV for the observed phase, both within the ultraviolet spectrum. The material's reflection and absorption properties, related to its loss function, are illustrated in Figure 4.8(e)(k) and (d)(j). The edge trailing of the material's reflection and absorption spectra matches the peak in the loss function, showing a sharp increase around 12 eV for both phases. This spike indicates enhanced absorption and reflection of electromagnetic waves in the ultraviolet region, with the predicted phase exhibiting much notable anisotropy along y- and z-directions with the maximum appearing in y-direction for both phase. Optical conductivity which represents electrical conductivity over specific photon energy ranges is shown in Figure 4.8(b) and (h). Conductivity sharply increases from zero at photon energies of 3.1 eV and 2.9 eV for the predicted and observed $Na_6Ge_2Se_6$ phases respectively, mirroring the optical band gap observed in the absorption curves in Figure 4.8(a)(g). In summary, the predicted and observed phases display similar optical properties, with the predicted phase exhibiting a higher degree of optical anisotropy in the y- and z-directions, while being isotropic in the x- and y-directions.

4.4. CONCLUSION

A comprehensive comparison was conducted between two $Na_6Ge_2Se_6$ phases: a computationally predicted phase and an experimentally observed phase. This comparison spans their structural, elastic, electronic, phonon, thermal, and optical properties, employing first-principles methods.

The predicted $Na_6Ge_2Se_6$ phase is the energetically favored structure at zero pressure and temperature based on first-principles calculations. Both phases feature the identical $[Ge_2Se_6]$ dimer within their crystal structures, with significant differences observed in the bonding environment between Se and Na atoms (five Na atoms bond with one Se atom in the predicted phase, whereas six Na atoms bond with one Se atom in the observed phase).

Regarding elastic properties, both phases meet the criteria for mechanical stability and share similar characteristics. These include their resistance to volumetric and shape deformations, ductility, nature of bonding forces (central force), and bond bending as a major role in resisting external stress. The differences arise from their anisotropic characteristics. While the predicted phase is isotropic in the x- and y-directions, it exhibits greater anisotropy in the y- and z-directions compared to the observed phase.

Electronic structure analyses reveal a closely matched energy band gap between the two phases, primarily contributed by the $[Ge_2Se_6]$ dimers, with only a 1.3% difference. However, the electronic band structure indicates that the predicted phase exhibits an indirect band gap, while the observed phase has a direct band gap.

The optical behavior between the phases is similarly aligned, with both exhibiting a comparable optical band gap (3.1 eV for the predicted phase and 2.9 eV for the observed phase) and enhanced absorption and reflection of electromagnetic waves in the ultraviolet region. Although the predicted phase is isotropic in the x- and y-directions, it displays a greater degree of optical anisotropy in the y- and z-directions, aligning with the anisotropic nature identified in the elastic properties.

Phonon dispersion analyses affirm the dynamic stability of both phases. Further calculations of thermal properties indicate that the predicted phase is more thermodynamically stable in the temperature range from 0 to 907 K, suggesting a high likelihood of experimental observation.

5. CRYSTAL STRUCTURE PREDICTION OF QUATERNARY METAL CHALCOGENIDES

5.1. INTRODUCTION

We now take a step further to test the proposed method on quaternary chalcogenides, specifically $LiMnPS_4$ and $LiZnPS_4$, by replacing two Li^+ cations with one Mn^{2+} or Zn^{2+} cation. These compounds belong to a subclass known as metal thiophosphates, which contain phosphorus, sulfur, and metal cations. This class of compounds has been found suitable for various technical applications. For example, Na_3PSe_4[17], $Li_{10}GeP_2S_{12}$[51], and $Li_5PS_4Cl_2$[129] can be used as superionic solid electrolytes in all-solid-state rechargeable batteries. Cu_3PS_4 and Cu_3PSe_4 are favorable materials for application in solar cells[102]. $LiCd_3PS_6$[25] is a good candidate for nonlinear optical conversion.

For these two specific chemical compositions, $LiMnPS_4$ and $LiZnPS_4$, researchers have successfully synthesized these compounds. For $LiMnPS_4$, the only experimentally observed structure was identified in 2021, synthesized using a solid-state metathesis route with γ-phase Li_3PS_4 and $MnCl_2$[9]. $LiZnPS_4$ was first synthesized in 2002[49] and was later found to fulfill the criteria for use as a mid-infrared nonlinear optical material (mid-IR NLO), making it an excellent candidate for generating mid-infrared light[52].

5.2. CALCULATION DETAILS

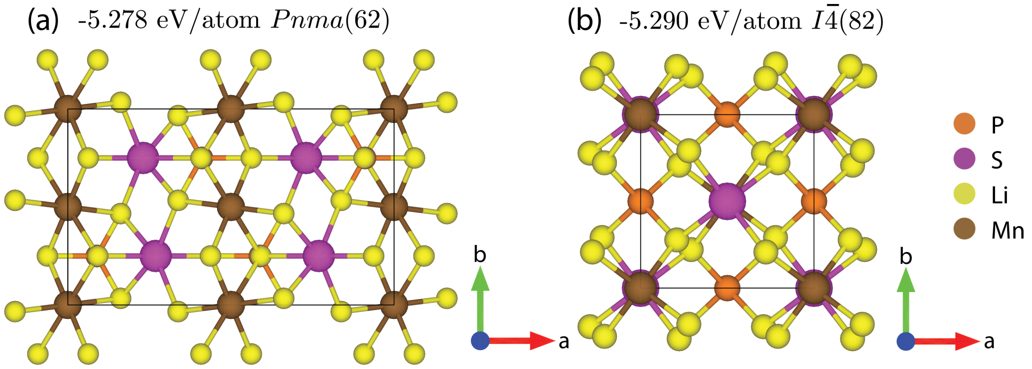
For the CSP conducted on the Li_3PS_4 and $Na_6Ge_2Se_6$ systems, we set the maximum stoichiometric unit to three and two, respectively. Based on the results presented in Sections 3.3.2 and 3.3.1, although CSP was conducted in many space groups with various combinations of WPs, all the structure candidates from Figures 3.2 and 3.5 were identified from the results of three lower symmetry space groups that have only one WP: $P1(1)$, $P2_1(4)$, and $Pc(7)$. Consequently, there is only one way to distribute atoms by placing them in the only general WP, which greatly simplifies the CSP procedure and requires fewer computational resources. Other than $P1(1)$, space groups $P2_1(4)$ and $Pc(7)$ impose symmetry restrictions on the atoms during simulations, which greatly reduces the system's degrees of freedom. Benefiting from the group and subgroup relationship discussed in Section 2.2.2, higher symmetry groups are also effectively sampled. With these considerations, for the CSP of quaternary systems, the simulations were conducted only in space groups $P1(1)$, $P2_1(4)$, and $Pc(7)$ with a greater maximum number of stoichiometric units, which was set at four. For the details of the computational settings, please refer to Section 3.2.

5.3. RESULTS

2×10^4 SA steps were conducted for every WP combination in space groups $P1(1)$, $P2_1(4)$, and $Pc(7)$ with a maximum of four stoichiometric units for both $LiMnPS_4$ and $LiZnPS_4$. The results indicate that most of the lower and lowest energy structures identified are from the simulations where the stoichiometric unit is two for both compounds. Therefore, the analysis in the following sections will primarily focus on the results with two stoichiometric units.

5.3.1. $LiMnPS_4$. The only experimentally observed $LiMnPS_4$[9], crystallized in space group $Pnma(62)$, is illustrated in Figure 5.1(a). Unfortunately, the CSP results from the three space groups where simulations were conducted did not find this phase. However, a metastable candidate was found with a significantly lower energy from first-principles calculations, 12 meV lower. This newly predicted structure crystallizes in space group $I\bar{4}(82)$ and is depicted in Figure 5.1(b). This structure is very similar to the synthesized $LiZnPS_4$ compound, which also crystallizes in $I\bar{4}(82)$ with Zn^{2+} cations replaced by Mn^{2+} cations. The $[PS_4]$ tetrahedra are arranged in cubic close packing, while Li and Mn atoms reside in the tetrahedral voids with 100% occupancy. Each $[LiS_4]$ tetrahedron connects to four [MnS_4] tetrahedra by corner-sharing to form a [$LiMn_4S_{16}$] layer. Similarly, each [MnS_4] tetrahedron shares corners with four [LiS_4] tetrahedra, forming a [Li_4MnS_{16}] layer adjacent to the [$LiMn_4S_{16}$] layer.

After relaxing the predicted $LiMnPS_4$ and the observed $LiZnPS_4$ phases with the same parameters from first-principles calculations, it shows that they share almost identical P-S bond lengths (2.063 Å in $LiMnPS_4$, 2.062 Å in $LiZnPS_4$) and Li-S bond lengths (2.461 Å in $LiMnPS_4$, 2.455 Å in $LiZnPS_4$), while the difference in bond lengths between Mn-S and Zn-S is relatively greater (2.408 Å for Mn-S, 2.362 Å for Zn-S).


(a) -5.278 eV/atom $Pnma(62)$

(b) -5.290 eV/atom $I\bar{4}(82)$

Figure 5.1. Crystal structures for (a) experimentally observed $LiMnPS_4$ phase and (b) predicted phase.

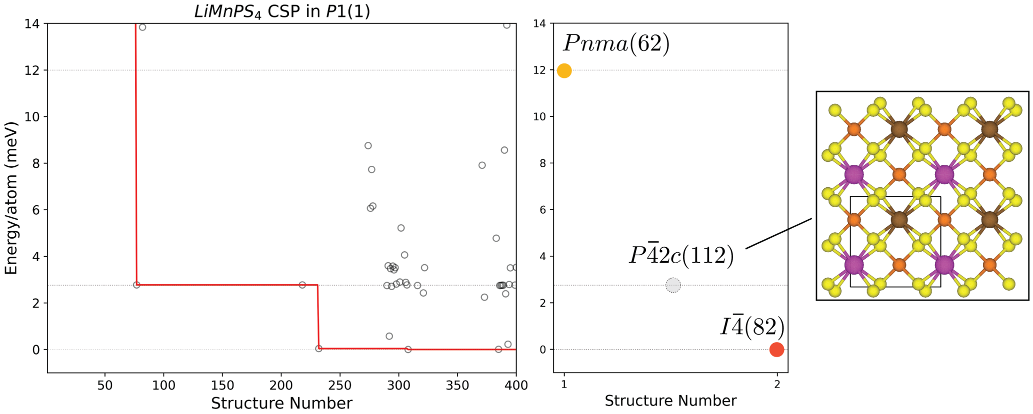


Figure 5.2. The energy of the sampled structure between 0-14 meV respect to the lowest energy phase.

The energy of the sampled structures is shown in Figure 5.2, ranging from 0 to 14 meV above the lowest energy phase identified in the CSP. From the graph, it is evident that between the $Pnma(62)$ and $I\bar{4}(82)$ phases, there are many intermediate states in terms of energy. Noticeably, a $P\bar{4}2c(112)$ phase (determined by the Pymatgen symmetry analyzer) occurs multiple times during the CSP. Its crystal structure, illustrated in Figure 5.2, shares a similar structure to the lowest energy $I\bar{4}(82)$ phase. More calculations are needed to determine if the $P\bar{4}2c(112)$ phase will converge to the $I\bar{4}(82)$ phase with tighter local optimization.

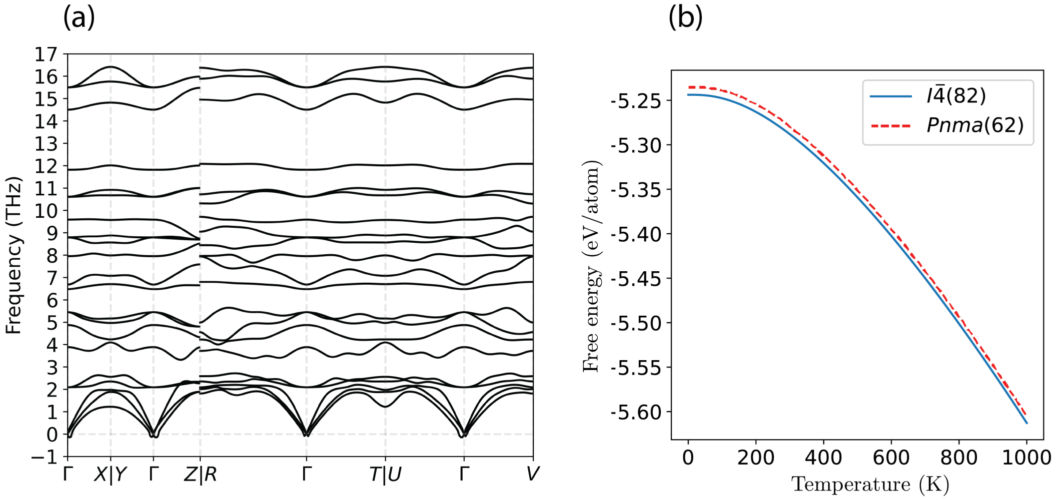


Figure 5.3. Phonon and thermal properties for $I\bar{4}(82)$ $LiMnPS_4$ phase.

The dynamic stability of the predicted $LiMnPS_4$ phase in space group $I\bar{4}(82)$ is evaluated by calculating the phonon band structure, as illustrated in Figure 5.3(a). There are only minor imaginary frequencies around the Γ points, which are caused by numerical errors, thus indicating the dynamical stability of the predicted phase. The Helmholtz free energy (F) from 0 to 1000 K is also evaluated and shown in Figure 5.3(b). Please refer to Equation 4.20 for the equation of the Helmholtz free energy. The result shows a lower free energy throughout the entire temperature range (0-1000 K) compared to the experimentally observed phase, indicating a possibility of experimental observability.

5.3.2. $\mathbf{LiZnPS_4}$. For the $LiZnPS_4$ system, the CSP successfully identified the observed phase, which crystallizes in space group $I\bar{4}(82)$, and it is also the lowest energy phase among all the candidate structures, as shown in Figure 5.4(a). Two other metastable candidate structures are shown in Figures 5.4(b) and 5.4(c). The predicted $Aem2(39)$ phase shows an energy close to the $I\bar{4}$ (82) phase, being only 8 meV greater, whereas the $P2_1/c(14)$ phase displays a larger energy gap, being 27 meV greater than the $I\bar{4}(82)$ phase.

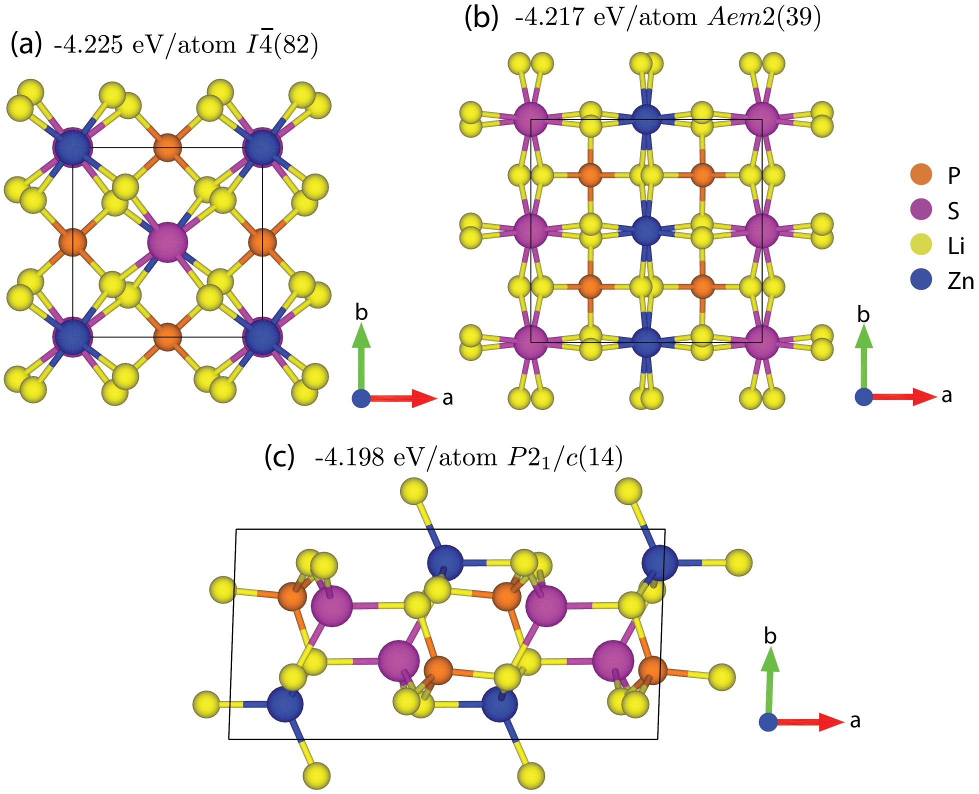


Figure 5.4. CSP results for $LiZnPS_4$.

Figure 5.5 shows the energy of the sampled structures from the CSP simulation in the energy range from 0 to 28 meV above the $I\bar{4}(82)$ phase. It shows a similar situation where a candidate structure in another space group, this time in $P\bar{4}2_1m(113)$, shares a similar crystal structure to that of the lowest energy phase in $I\bar{4}(82)$. However, the $P\bar{4}2_1m(113)$ phase possesses a much higher energy than the $I\bar{4}(82)$ phase (23 meV higher compared to 2.6 meV for the $LiMnPS_4$ system). The $P\bar{4}2_1m(113)$ and $I\bar{4}(82)$ phases are separated by

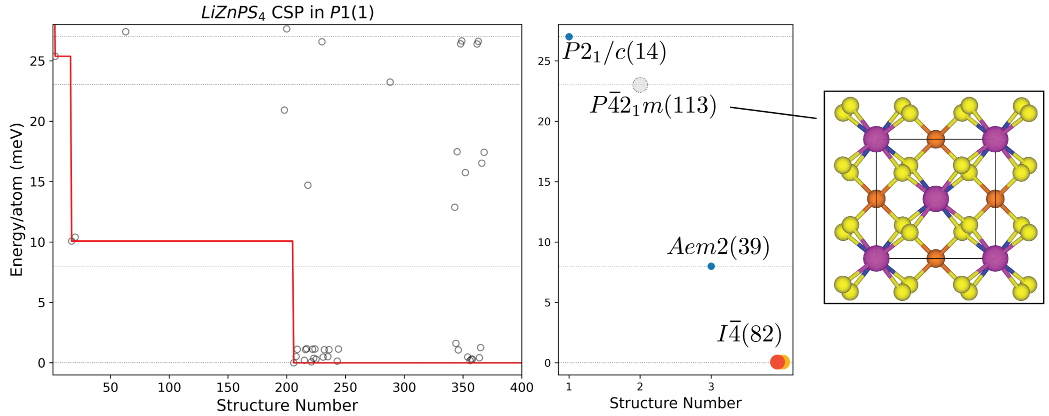


Figure 5.5. The energy of the sampled structure between 0-28 meV respect to the lowest energy phase are shown.

the $Aem2(39)$ phase. From the low-energy sampled structures shown in Figure 5.5, a large number cluster around the energy of the $I\bar{4}(82)$ phase; however, not a single phase was found between the $Aem2(39)$ and $I\bar{4}(82)$ phases. All these observations indicate that the $Aem2(39)$ and $I\bar{4}(82)$ phases are different structures with close energies. It is questionable whether the $P\bar{4}2_1m(113)$ phase would converge to the $I\bar{4}(82)$ phase, despite their similar crystal structures. More in-depth analysis is needed to draw a more definitive conclusion.

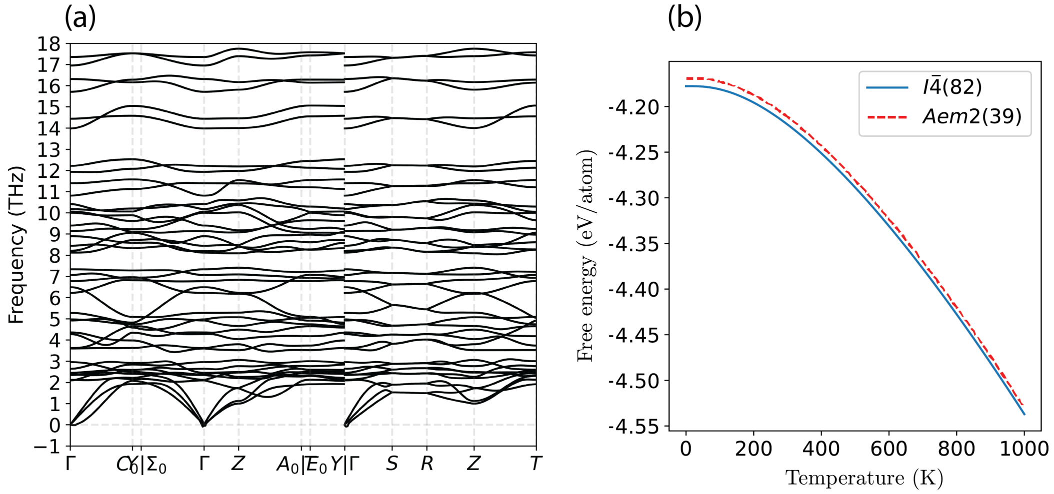


Figure 5.6. Phonon and thermal properties for $Aem2(39)$ $LiZnPS_4$ phase.

The dynamical stability of the $P2_1/c(14)$ and $Aem2(39)$ phases is evaluated in the same way as for the $LiMnPS_4$ system. The phonon band structure plot in Figure 5.6(a) shows no imaginary frequencies for the $Aem2(39)$ phase, indicating its dynamical stability. On the other hand, the $P2_1/c(14)$ phase is proven to be dynamically unstable with an excessive amount of imaginary frequencies; therefore, its phonon band structure is not shown, and its free energy at elevated temperatures is not evaluated. In terms of Helmholtz free energy, unlike the $LiMnPS_4$ case, the predicted $Aem2(39)$ phase shows a higher Helmholtz free energy than the observed $I\bar{4}(82)$ phase in the temperature range from 0 to 1000 K.

5.4. DISCUSSION

The simulation results for both $LiMnPS_4$ and $LiZnPS_4$ systems yield promising outcomes with a relatively small amount of first-principles calculations required. For the $LiMnPS_4$ system, although the experimentally observed phase has not been identified, a lower energy phase crystallizing in space group $I\bar{4}(82)$ has been predicted. The dynamical stability of this newly predicted $LiMnPS_4$ phase has been confirmed through phonon calculations. Additionally, thermal properties analysis indicates that this phase exhibits higher thermodynamic stability than its observed counterpart in the temperature range from 0 to 1000 K.

On the other hand, the results for the $LiZnPS_4$ system identified the same lowest energy phase, crystallizing in space group $I\bar{4}(82)$, corresponding to the experimental results. Interestingly, for both systems, the lowest energy phases are in space group $I\bar{4}(82)$, with the $LiMnPS_4$ phase being newly predicted and the $LiZnPS_4$ phase already synthesized. These two phases share similar structural characteristics, with Mn^+ ions occupying the sites of Zn^+ ions. This suggests that the currently observed $LiMnPS_4$ phase, crystallizing in space group $P2_1/c(14)$, is likely not the global minimum for this chemical composition. Additionally, a new dynamically stable metastable candidate was identified for the $LiZnPS_4$ system in space group $Aem2$(39), possessing only 8 meV/atom higher energy than the lowest energy phase but showing lower thermodynamic stability.

Table 5.1. Numbers of combinations for $LiXPS_4$ (X = Mn, Zn) in different space groups with a maximum of four PS_4 rigid bodies in the unit cell.

Space group	Combinations	Space group	Combinations	Space group	Combinations
1	4	2	>1000	3	>1000
4	2	5	72	6	>1000
7	2	8	9	10	>1000
11	193	12	267	13	>1000
14	19	15	21	16	>1000
17	>1000	18	72	19	1
20	8	21	>1000	22	24
23	996	24	27	25	>1000
26	72	27	>1000	28	370
29	1	30	72	31	9
32	72	33	1	34	72
35	224	36	1	37	27
38	224	39	27	40	8
41	1	42	1	44	133
45	8	46	8	47	>1000
48	>1000	49	>1000	50	>1000
51	>1000	52	28	53	280
54	69	55	280	56	28
58	153	59	243	60	7
62	7	63	7	65	984
66	40	67	105	71	684
72	12	74	21	75	216
76	1	77	370	78	1
79	4	80	1	81	>1000

Note: Highlighted space groups have 20 or fewer combinations.Combinations exceeding 1000 are marked as '> 1000'. Space groups not included in the table have no viable cases.

This study on these two quaternary systems aims to test the validity of the simplified simulation procedure. However, more comprehensive simulations, similar to those conducted for the ternary systems illustrated in Section 3, are still possible. The possible Wyckoff position combinations are shown in the Table 5.1 and 5.2 with four maximum PS_4 rigid bodies.

Table 5.2. Numbers of combinations for $LiXPS_4$ (X = Mn, Zn) in different space groups with a maximum of four PS_4 rigid bodies in the unit cell.

Space group	Combinations	Space group	Combinations	Space group	Combinations
82	140	83	>1000	84	>1000
85	104	85	104	86	58
87	7	89	>1000	90	196
91	27	92	1	93	>1000
94	112	95	27	96	1
97	14	99	592	100	36
101	133	102	28	103	16
104	4	105	539	106	8
107	4	108	4	108	4
109	1	111	>1000	112	>1000
113	139	114	20	115	>1000
116	689	117	448	118	448
119	140	120	24	121	34
123	>1000	124	50	125	102
126	14	127	210	128	7
129	104	130	6	131	>1000
132	688	134	94	135	12
136	122	137	20	138	47
139	7	140	6	143	>1000
144	1	145	1	146	1
147	212	149	>1000	150	381
151	8 152	8	153	8	
154	8	156	>1000	157	100
158	243	159	35	160	1
162	181	163	68	164	212
165	16	168	45	171	8
172	8	173	35	174	>1000
175	84	176	68	177	136
180	24	181	24	182	68
183	45	184	4	185	9
186	35	187	>1000	188	567
189	425	190	68	191	84
193	13	194	68	195	25
196	24	198	1	215	25
216	24				

Note: Highlighted space groups have 20 or fewer combinations.Combinations exceeding 1000 are marked as '> 1000'. Space groups not included in the table have no viable cases.

6. CONCLUSIONS

A new crystal structure prediction method tailored for predicting structures featuring recurring motifs or rigid bodies is proposed in this dissertation. The entire procedure has been automated with the open-source Python package, Space Group Restricted Crystal Structure Prediction (SGRCSP). Its validity and effectiveness have been demonstrated through its application in predicting ternary and quaternary metal chalcogenides.

Despite the success stories presented, there are significant limitations that need to be addressed in future work. With the current implementation of the method, it only supports predicting structures with pre-determined chemical formulas. However, it is equally crucial to determine which chemical composition is the most promising for identifying structural phases with our desired properties. Many synthesized compounds feature partial occupation or defects, and currently, no crystal structure prediction method is specifically designed for predicting these types of structures. It remains an open question whether it is possible to formulate such a CSP method. Exploring this direction is worthwhile.

REFERENCES

[1] Nadine L Abraham and Matthew IJ Probert. A periodic genetic algorithm with real-space representation for crystal structure and polymorph prediction. *Physical Review B*, 73(22):224104, 2006.

[2] Sadao Adachi. *Earth-abundant materials for solar cells: Cu2-II-IV-VI4 semiconductors*. John Wiley & Sons, 2015.

[3] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of chemical physics*, 110(13):6158–6170, 1999.

[4] Rachmat Adhi Wibowo, Eun Soo Lee, Badrul Munir, and Kyoo Ho Kim. Pulsed laser deposition of quaternary cu2znsnse4 thin films. *physica status solidi (a)*, 204(10):3373–3379, 2007.

[5] Maximilian Amsler and Stefan Goedecker. Crystal structure prediction using the minima hopping method. *The Journal of chemical physics*, 133(22), 2010.

[6] Orson L Anderson. A simplified method for calculating the debye temperature from elastic constants. *Journal of Physics and Chemistry of Solids*, 24(7):909–917, 1963.

[7] Orson L Anderson and Harold H Demarest Jr. Elastic constants of the central force model for cubic structures: Polycrystalline aggregates and instabilities. *Journal of Geophysical Research*, 76(5):1349–1369, 1971.

[8] Srikanth Balijapelly, Andrew J Craig, Jeong Bin Cho, Joon I Jang, Kartik Ghosh, Jennifer A Aitken, Aleksandr V Chernatynskiy, and Amitava Choudhury. Building-block approach to the discovery of na8mn2 (ge2se6) 2: A polar chalcogenide exhibiting promising harmonic generation signals with a high laser-induced damage threshold. *Journal of Alloys and Compounds*, 900:163392, 2022.

[9] Srikanth Balijapelly, Kartik Ghosh, Aleksandr V Chernatynskiy, and Amitava Choudhury. Discovery of an olivine-type lithium manganese thiophosphate, limnps 4, via a building block approach. *Chemical Communications*, 57(97):13182–13185, 2021.

[10] Philip Ball. In retrospect: On the six-cornered snowflake, 2011.

[11] Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M Elena, Dávid P Kovács, Janosh Riebesell, Xavier R Advincula, Mark Asta, William J Baldwin, Noam Bernstein, et al. A foundation model for atomistic materials chemistry. *arXiv preprint arXiv:2401.00096*, 2023.

[12] Ante Bilić, Julian D Gale, Mark A Gibson, Nick Wilson, and Kathie McGregor. Prediction of novel alloy phases of al with sc or ta. *Scientific reports*, 5(1):9909, 2015.

[13] Luca Bindi and John A Jaszczak. Richardsite, zn2cugas4, a new gallium-essential member of the stannite group from the gem mines near merelani, tanzania. *Minerals*, 10(5):467, 2020.

[14] Peter Blaha, Karlheinz Schwarz, Georg KH Madsen, Dieter Kvasnicka, Joachim Luitz, et al. wien2k. *An augmented plane wave+ local orbitals program for calculating crystal properties*, 60(1), 2001.

[15] Peter E Blöchl. Projector augmented-wave method. *Physical review B*, 50(24):17953, 1994.

[16] Peter E Blöchl, Ove Jepsen, and Ole Krogh Andersen. Improved tetrahedron method for brillouin-zone integrations. *Physical Review B*, 49(23):16223, 1994.

[17] Shou-Hang Bo, Yan Wang, Jae Chul Kim, William Davidson Richards, and Gerbrand Ceder. Computational and experimental investigations of na-ion conduction in cubic na3pse4. *Chemistry of Materials*, 28(1):252–258, 2016.

[18] Jacilynn A Brant, Daniel J Clark, Yong Soo Kim, Joon I Jang, Ashley Weiland, and Jennifer A Aitken. Outstanding laser damage threshold in li2mnges4 and tunable optical nonlinearity in diamond-like semiconductors. *Inorganic Chemistry*, 54(6):2809–2819, 2015.

[19] Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

[20] In Chung and Mercouri G Kanatzidis. Metal chalcogenides: a rich source of nonlinear optical materials. *Chemistry of Materials*, 26(1):849–869, 2014.

[21] David M Deaven and Kai-Ming Ho. Molecular geometry optimization with a genetic algorithm. *Physical review letters*, 75(2):288, 1995.

[22] Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.

[23] B Eisenmann, J Hansa, et al. Oligoselenidogermanate (iii): Zur kenntnis von na6ge2se6 und na8ge4se10. *Materials research bulletin*, 20(11):1339–1346, 1985.

[24] Brigitte Eisenmann, Emma Kieselbach, Herbert Schäfer, and Heike Schrod. On thio-, selenido-, and telluridogermanates (iii): K 6 ge 2 s 6, k 6 ge 2 se 6, and na 6 ge 2 te 6. *Zeitschrift fuer Anorganische und Allgemeine Chemie (1950)*, 516, 1984.

[25] Jianghe Feng, Chun-Li Hu, Bingxuan Li, and Jiang-Gao Mao. Liga2ps6 and licd3ps6: molecular designs of two new mid-infrared nonlinear optical materials. *Chemistry of Materials*, 30(11):3901–3908, 2018.

[26] Carlos Fiolhais, Fernando Nogueira, and Miguel AL Marques. *A primer in density functional theory*, volume 620. Springer Science & Business Media, 2003.

[27] Bruno Focassio, Luis Paulo Mezzina Freitas, and Gabriel R Schleder. Performance assessment of universal machine learning interatomic potentials: Challenges and directions for materials' surfaces. *arXiv preprint arXiv:2403.04217*, 2024.

[28] Scott Fredericks, Kevin Parrish, Dean Sayre, and Qiang Zhu. Pyxtal: A python library for crystal structure generation and symmetry analysis. *Computer Physics Communications*, 261:107810, 2021.

[29] Brent Fultz. Vibrational thermodynamics of materials. *Progress in Materials Science*, 55(4):247–352, 2010.

[30] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741, 1984.

[31] Paolo Giannozzi, Oliviero Andreussi, Thomas Brumme, Oana Bunau, M Buongiorno Nardelli, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Matteo Cococcioni, et al. Advanced capabilities for materials modelling with quantum espresso. *Journal of physics: Condensed matter*, 29(46):465901, 2017.

[32] S Girard, P Pullumbi, C Mellot-Draznieks, and G Férey. 07-p-21-application of the aasbu method to the prediction of inorganic structures built exclusively of sodalite cages. *Studies in Surface Science and Catalysis*, 135:254, 2001.

[33] Colin W Glass, Artem R Oganov, and Nikolaus Hansen. Uspex—evolutionary crystal structure prediction. *Computer physics communications*, 175(11-12):713–720, 2006.

[34] William A Goddard. Classical force fields and methods of molecular dynamics. *Computational Materials, Chemistry, and Biochemistry: From Bold Initiatives to the Last Mile: In Honor of William A. Goddard's Contributions to Science and Engineering*, pages 1063–1072, 2021.

[35] Xavier Gonze, Bernard Amadon, P-M Anglade, J-M Beuken, François Bottin, Paul Boulanger, Fabien Bruneval, Damien Caliste, Razvan Caracas, Michel Côté, et al. Abinit: First-principles approach to material and nanosystem properties. *Computer Physics Communications*, 180(12):2582–2615, 2009.

[36] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4:1–17, 2012.

[37] J Niklas Hausmann, Eva M Heppke, Rodrigo Beltrán-Suito, Johannes Schmidt, Martin Mühlbauer, Martin Lerch, Prashanth W Menezes, and Matthias Driess. Stannites–a new promising class of durable electrocatalysts for efficient water oxidation. *ChemCatChem*, 12(4):1161–1168, 2020.

[38] Andreas Hermann, NW Ashcroft, and Roald Hoffmann. High pressure ices. *Proceedings of the National Academy of Sciences*, 109(3):745–750, 2012.

[39] Andreas Hermann, Roald Hoffmann, and NW Ashcroft. Condensed astatine: monatomic and metallic. *Physical Review Letters*, 111(11):116404, 2013.

[40] Jochen Heyd, Gustavo E Scuseria, and Matthias Ernzerhof. Hybrid functionals based on a screened coulomb potential. *The Journal of chemical physics*, 118(18):8207–8215, 2003.

[41] P Hohenberg and WJKW Kohn. Phys rev 136: B864. *Kohn W, Sham LJ (1965) Phys Rev*, 140:A1133, 1964.

[42] Kenji Homma, Masao Yonemura, Takeshi Kobayashi, Miki Nagao, Masaaki Hirayama, and Ryoji Kanno. Crystal structure and phase transitions of the lithium ionic conductor li3ps4. *Solid State Ionics*, 182(1):53–58, 2011.

[43] Jürg Hutter, Marcella Iannuzzi, Florian Schiffmann, and Joost VandeVondele. cp2k: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(1):15–25, 2014.

[44] Kentaro Ito, editor. John Wiley & Sons, Ltd, 2014.

[45] Abishek K Iyer, Jeong Bin Cho, Hye Ryung Byun, Michael J Waters, Shiqiang Hao, Benjamin M Oxley, Venkat Gopalan, Christopher Wolverton, James M Rondinelli, Joon I Jang, et al. Structure tuning, strong second harmonic generation response, and high optical stability of the polar semiconductors na1−x k x as q 2. *Journal of the American Chemical Society*, 143(43):18204–18215, 2021.

[46] Huanhuan Jia, Linfeng Peng, Chuang Yu, Li Dong, Shijie Cheng, and Jia Xie. Chalcogenide-based inorganic sodium solid electrolytes. *Journal of Materials Chemistry A*, 9(9):5134–5148, 2021.

[47] Ning Jia, Mengqi Zhang, Bo Li, Chuanchuan Li, Yangyang Liu, Yupeng Zhang, Tongtong Yu, Yang Liu, Deliang Cui, and Xutang Tao. Ternary chalcogenide liinse2: A promising high-performance anode material for lithium ion batteries. *Electrochimica Acta*, 320:134562, 2019.

[48] Stefan Joergens and Albrecht Mewis. Die kristallstrukturen von hexachalcogenohypodiphosphaten des magnesiums und zinks. *Zeitschrift für anorganische und allgemeine Chemie*, 630(1):51–57, 2004.

[49] Stefan Jörgens, Dirk Johrendt, and Albrecht Mewis. Motive dichtester kugelpackungen: Die verbindungen zn3 (ps4) 2 und liznps4. *Zeitschrift für anorganische und allgemeine Chemie*, 628(8):1765–1769, 2002.

[50] Ronald L Kam, KyuJung Jun, Luis Barroso-Luque, Julia H Yang, Fengyu Xie, and Gerbrand Ceder. Crystal structures and phase stability of the li2s–p2s5 system from first principles. *Chemistry of Materials*, 35(21):9111–9126, 2023.

[51] Noriaki Kamaya, Kenji Homma, Yuichiro Yamakawa, Masaaki Hirayama, Ryoji Kanno, Masao Yonemura, Takashi Kamiyama, Yuki Kato, Shigenori Hama, Koji Kawamoto, et al. A lithium superionic conductor. *Nature materials*, 10(9):682–686, 2011.

[52] Lei Kang, Molin Zhou, Jiyong Yao, Zheshuai Lin, Yicheng Wu, and Chuangtian Chen. Metal thiophosphates with good mid-infrared nonlinear optical performances: a first-principles prediction and analysis. *Journal of the American Chemical Society*, 137(40):13049–13059, 2015.

[53] AMM Tanveer Karim, MA Hadi, MA Alam, F Parvin, SH Naqib, and AKMA Islam. Newly synthesized mgal2ge2: A first-principles comparison with its silicide and carbide counterparts. *Journal of Physics and Chemistry of Solids*, 117:139–147, 2018.

[54] Hironori Katagiri, Kazuo Jimbo, Satoru Yamada, Tsuyoshi Kamimura, Win Shwe Maw, Tatsuo Fukano, Tadashi Ito, and Tomoyoshi Motohiro. Enhanced conversion efficiencies of cu2znsns4-based thin film solar cells by using preferential etching technique. *Applied physics express*, 1(4):041201, 2008.

[55] Hironori Katagiri, Kotoe Saitoh, Tsukasa Washio, Hiroyuki Shinohara, Tomomi Kurumadani, and Shinsuke Miyajima. Development of thin film solar cell based on cu2znsns4 thin films. *Solar Energy Materials and Solar Cells*, 65(1-4):141–148, 2001.

[56] Johannes Kepler. *The six-cornered snowflake*. Paul Dry Books, 2010.

[57] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[58] Charles Kittel and Paul McEuen. *Introduction to solid state physics*. John Wiley & Sons, 2018.

[59] Leonard Kleinman. Deformation potentials in silicon. i. uniaxial strain. *Physical Review*, 128(6):2614, 1962.

[60] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, 140(4A):A1133, 1965.

[61] G Kresse. G. kresse and d. joubert, phys. rev. b 59, 1758 (1999). *Phys. Rev. B*, 59:1758, 1999.

[62] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational materials science*, 6(1):15–50, 1996.

[63] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996.

[64] Aliaksandr V Krukau, Oleg A Vydrov, Artur F Izmaylov, and Gustavo E Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of chemical physics*, 125(22), 2006.

[65] Alessandro Laio, Antonio Rodriguez-Fortea, Francesco Luigi Gervasio, Matteo Ceccarelli, and Michele Parrinello. Assessing the accuracy of metadynamics. *The journal of physical chemistry B*, 109(14):6714–6721, 2005.

[66] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Du lak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, et al. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017.

[67] Yvon Le Page and Paul Saxe. Symmetry-general least-squares extraction of elastic coefficients from ab initio total energy calculations. *Physical Review B*, 63(17):174103, 2001.

[68] Jonathan W Lekse, Beth M Leverett, Charles H Lake, and Jennifer A Aitken. Synthesis, physicochemical characterization and crystallographic twinning of li2znsns4. *Journal of Solid State Chemistry*, 181(12):3217–3222, 2008.

[69] Jonathan W Lekse, Meghann A Moreau, Katie L McNerny, Jeongho Yeon, P Shiv Halasyamani, and Jennifer A Aitken. Second-harmonic generation and crystal structure of the diamond-like semiconductors li2cdges4 and li2cdsns4. *Inorganic chemistry*, 48(16):7516–7518, 2009.

[70] Hua Lin, Gangjian Tan, Jin-Ni Shen, Shiqiang Hao, Li-Ming Wu, Nicholas Calta, Christos Malliakas, Si Wang, Ctirad Uher, Christopher Wolverton, et al. Concerted rattling in csag5te3 leading to ultralow thermal conductivity and high thermoelectric performance. *Angewandte Chemie International Edition*, 55(38):11431–11436, 2016.

[71] Min-Ling Liu, I-Wei Chen, Fu-Qiang Huang, and Li-Dong Chen. Improved thermoelectric properties of cu-doped quaternary chalcogenides of cu2cdsnse4. *Advanced Materials*, 21(37):3808–3812, 2009.

[72] David C Lonie and Eva Zurek. Xtalopt version r7: An open-source evolutionary algorithm for crystal structure prediction. *Computer Physics Communications*, 182(10):2305–2306, 2011.

[73] Jian Lv, Yanchao Wang, Li Zhu, and Yanming Ma. Predicted novel high-pressure phases of lithium. *Physical Review Letters*, 106(1):015503, 2011.

[74] Andriy O Lyakhov, Artem R Oganov, Harold T Stokes, and Qiang Zhu. New developments in evolutionary structure prediction algorithm uspex. *Computer Physics Communications*, 184(4):1172–1182, 2013.

[75] Ni Ma, Fei Jia, Lin Xiong, Ling Chen, Yan-Yan Li, and Li-Ming Wu. Cscu5s3: promising thermoelectric material with enhanced phase transition temperature. *Inorganic Chemistry*, 58(2):1371–1376, 2019.

[76] Ni Ma, Yan-Yan Li, Ling Chen, and Li-Ming Wu. α-cscu5se3: discovery of a low-cost bulk selenide with high thermoelectric performance. *Journal of the American Chemical Society*, 142(11):5293–5303, 2020.

[77] John Maddox. Crystals from first principles. *Nature*, 335(6187):201–201, 1988.

[78] Hendrik J Monkhorst and James D Pack. Special points for brillouin-zone integrations. *Physical review B*, 13(12):5188, 1976.

[79] Félix Mouhat and François-Xavier Coudert. Necessary and sufficient elastic stability conditions in various crystal systems. *Physical review B*, 90(22):224104, 2014.

[80] J Navrátil, V Kucek, T Plecháček, E Černošková, F Laufek, Č Drašar, and P Knotek. Thermoelectric properties of cu 2 hgsnse 4-cu 2 hgsnte 4 solid solution. *Journal of electronic materials*, 43:3719–3725, 2014.

[81] John Arthur Niesse and Howard R Mayne. Global geometry optimization of atomic clusters using a modified genetic algorithm in space-fixed coordinates. *The Journal of chemical physics*, 105(11):4700–4706, 1996.

[82] Artem R Oganov and Colin W Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *The Journal of chemical physics*, 124(24), 2006.

[83] Travis E Oliphant et al. *Guide to numpy*, volume 1. Trelgol Publishing USA, 2006.

[84] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L Chevrier, Kristin A Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

[85] Mohnish Pandey and Karsten W Jacobsen. Promising quaternary chalcogenides as high-band-gap semiconductors for tandem photoelectrochemical water splitting devices: A computational screening approach. *Physical Review Materials*, 2(10):105402, 2018.

[86] Jean Pannetier, J Bassas-Alsina, Juan Rodriguez-Carvajal, and Vincent Caignaert. Prediction of crystal structures from crystal chemistry rules by simulated annealing. *Nature*, 346(6282):343–345, 1990.

[87] Robert G Parr. Density functional theory. In *Electron Distributions and the Chemical Bond*, pages 95–100. Springer, 1982.

[88] Yanling Pei, Cheng Chang, Zhe Wang, Meijie Yin, Minghui Wu, Gangjian Tan, Haijun Wu, Yuexing Chen, Lei Zheng, Shengkai Gong, et al. Multiple converged conduction bands in k2bi8se13: a promising thermoelectric material with extremely low thermal conductivity. *Journal of the American Chemical Society*, 138(50):16364–16371, 2016.

[89] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[90] John P Perdew, Karla Schmidt, V Van Doren, C Van Alsenoy, and P Geerlings. Density functional theory and its application to materials. *Van Doren, V*, pages 1–20, 2001.

[91] Chris J Pickard and RJ Needs. High-pressure phases of silane. *Physical review letters*, 97(4):045504, 2006.

[92] Chris J Pickard and RJ Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

[93] Pierre FP Poudeu, Jonathan D'Angelo, Adam D Downey, Jarrod L Short, Timothy P Hogan, and Mercouri G Kanatzidis. High thermoelectric figure of merit and nanostructuring in bulk p-type na1- xpbmsbytem+ 2. *Angewandte Chemie International Edition*, 45(23):3835–3839, 2006.

[94] SF Pugh. Xcii. relations between the elastic moduli and the plastic properties of polycrystalline pure metals. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 45(367):823–843, 1954.

[95] P Ravindran, Lars Fast, P A_ Korzhavyi, B Johansson, J Wills, and Os Eriksson. Density functional theory for calculation of elastic properties of orthorhombic crystals: Application to tisi 2. *Journal of Applied Physics*, 84(9):4891–4904, 1998.

[96] Michelle D Regulacio, Dan-Thien Nguyen, Raymond Horia, and Zhi Wei Seh. Designing nanostructured metal chalcogenides as cathode materials for rechargeable magnesium batteries. *Small*, 17(25):2007683, 2021.

[97] András Reuß. Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 9(1):49–58, 1929.

[98] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE transactions on neural networks*, 20(1):61–80, 2008.

[99] J. Schlirf and H. J. Deiseroth. Crystal structure of hexacesium hexaselenidodigermanate(iii), cs6ge2se6. *Zeitschrift für Kristallographie - New Crystal Structures*, 216(1-4):27–28, 2001.

[100] Jonathan Schmidt, Mário RG Marques, Silvana Botti, and Miguel AL Marques. Recent advances and applications of machine learning in solid-state materials science. *npj computational materials*, 5(1):83, 2019.

[101] C Sevik and T Çağın. Ab initio study of thermoelectric transport properties of pure and doped quaternary compounds. *Physical Review B*, 82(4):045202, 2010.

[102] Tingting Shi, Wan-Jian Yin, Mowafak Al-Jassim, and Yanfa Yan. Structural, electronic, and optical properties of cu3-v-vi4 compound semiconductors. *Applied Physics Letters*, 103(15), 2013.

[103] Yixuan Shi, Cheryl Sturm, and Holger Kleinke. Chalcogenides as thermoelectric materials. *Journal of Solid State Chemistry*, 270:273–279, 2019.

[104] Brad A Steele and Ivan I Oleynik. Sodium pentazolate: A nitrogen rich high energy density material. *Chemical Physics Letters*, 643:21–26, 2016.

[105] Wei-Chi Sun, Siti Utari Rahayu, and Ming-Way Lee. Eco-friendly nasbs 2 quantum dot-sensitized solar cells. *IEEE Journal of Photovoltaics*, 8(4):1011–1016, 2018.

[106] Harold Szu and Ralph Hartley. Fast simulated annealing. *Physics letters A*, 122(3-4):157–162, 1987.

[107] Atsushi Togo. First-principles phonon calculations with phonopy and phono3py. *J. Phys. Soc. Jpn.*, 92(1):012001, 2023.

[108] Atsushi Togo, Laurent Chaput, Terumasa Tadano, and Isao Tanaka. Implementation strategies in phonopy and phono3py. *J. Phys. Condens. Matter*, 35(35):353001, 2023.

[109] Atsushi Togo and Isao Tanaka. First principles phonon calculations in materials science. *Scripta Materialia*, 108:1–5, 2015.

[110] Atsushi Togo and Isao Tanaka. Spglib: a software library for crystal symmetry search. *arXiv preprint arXiv:1808.01590*, 2018.

[111] Giancarlo Trimarchi and Alex Zunger. Global space-group optimization problem: Finding the stablest crystal structure without constraints. *Physical Review B*, 75(10):104113, 2007.

[112] Issei Tsuji, Yoshiki Shimodaira, Hideki Kato, Hisayoshi Kobayashi, and Akihiko Kudo. Novel stannite-type complex sulfide photocatalysts ai2-zn-aiv-s4 (ai= cu and ag; aiv= sn and ge) for hydrogen evolution under visible-light irradiation. *Chemistry of Materials*, 22(4):1402–1409, 2010.

[113] VS Urusov. Crystal chemical conditions for occupying regular point systems. *Vestnik Moskovskogo universiteta. Seriya 4. Geologiya [Moscow University Geology Bulletin]*, 4:3–19, 1991.

[114] VS Urusov and TN Nadezhina. Frequency distribution and selection of space groups in inorganic crystal chemistry. *Journal of Structural Chemistry*, 50:22–37, 2009.

[115] G Vaitheeswaran, V Kanchana, Axel Svane, and Anna Delin. Elastic properties of mgcni3—a superconducting perovskite. *Journal of Physics: Condensed Matter*, 19(32):326214, 2007.

[116] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[117] Woldemar Voigt. *Lehrbuch der kristallphysik:(mit ausschluss der kristalloptik)*, volume 34. BG Teubner, 1910.

[118] David J Wales and Jonathan PK Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997.

[119] Oswald J Walker. August kekulé and the benzene problem. *Annals of Science*, 4(1):34–46, 1939.

[120] Vei Wang, Nan Xu, Jin-Cheng Liu, Gang Tang, and Wen-Tong Geng. Vaspkit: A user-friendly interface facilitating high-throughput computing and analysis using vasp code. *Computer Physics Communications*, 267:108033, 2021.

[121] Wei Wang, Mark T Winkler, Oki Gunawan, Tayfun Gokmen, Teodor K Todorov, Yu Zhu, and David B Mitzi. Device characteristics of cztsse thin-film solar cells with 12.6% efficiency. *Advanced energy materials*, 4(7):1301465, 2014.

[122] Yanchao Wang, Jian Lv, Li Zhu, and Yanming Ma. Crystal structure prediction via particle-swarm optimization. *Physical Review B*, 82(9):094116, 2010.

[123] ScottáM Woodley, PeteráD Battle, JulianáD Gale, and CáRichard A Catlow. The prediction of inorganic crystal structures using a genetic algorithm and energy minimisation. *Physical Chemistry Chemical Physics*, 1(10):2535–2542, 1999.

[124] Kui Wu, Zhihua Yang, and Shilie Pan. Na4mgm2se6 (m= si, ge): the first noncentrosymmetric compounds with special ethane-like [m2se6] 6–units exhibiting large laser-damage thresholds. *Inorganic Chemistry*, 54(21):10108–10110, 2015.

[125] Yansun Yao, S Tse John, and Kaori Tanaka. Metastable high-pressure single-bonded phases of nitrogen predicted via genetic algorithm. *Physical Review B*, 77(5):052103, 2008.

[126] Haochen Yu, Matteo Giantomassi, Giuliana Materzanini, and Gian-Marco Rignanese. Systematic assessment of various universal machine-learning interatomic potentials. *arXiv preprint arXiv:2403.05729*, 2024.

[127] Jian Zhang, Tianlin Li, Biao Li, Shuoxiao Zhang, Yibo Dou, Qingyan Yuan, Yunjia Wu, and Jingbin Han. Erdite nafes2 as a new anode material for lithium-ion batteries. *ACS Sustainable Chemistry & Engineering*, 10(32):10666–10674, 2022.

[128] Li Zhu, Hanyu Liu, Chris J Pickard, Guangtian Zou, and Yanming Ma. Reactions of xenon with iron and nickel are predicted in the earth's inner core. *Nature chemistry*, 6(7):644–648, 2014.

[129] Zhuoying Zhu, Iek-Heng Chu, and Shyue Ping Ong. Li3y (ps4) 2 and li5ps4cl2: new lithium superionic conductors predicted from silver thiophosphates using efficiently tiered ab initio molecular dynamics simulations. *Chemistry of Materials*, 29(6):2474–2484, 2017.

VITA

Qi Zhang was born in Jinhua, China. He attended China Jiliang University in 2011 and received his Bachelor's degree in Applied Physics in 2015. In 2016, he was admitted to the Master's program in Physics at Missouri University of Science and Technology. In 2017, he transferred to the PhD program and began working with Dr. Alexksandr Chernatynskiy in computational physics until he earned his PhD in Physics in July 2024.