UNIVERSITY OF BELGRADE

FACULTY OF PHYSICS

Milan Pejić

Energy Landscapes, Crystal Structure Prediction, and Modeling of Rare Earth Ternary Compounds

Doctoral Dissertation

Belgrade, 2026

УНИВЕРЗИТЕТ У БЕОГРАДУ

ФИЗИЧКИ ФАКУЛТЕТ

Милан Пејић

Енергетски пејзажи, предвиђање кристалних структура и моделовање тернарних једињења елемената ретких земаља

Докторска дисертација

Београд, 2026

Thesis Defense Committee

Thesis advisor:

Dr. Dejan Zagorac
Research Professor
Vinča Institute of Nuclear Sciences
University of Belgrade

Committee Members:

Dr. Đorđe Spasojević
Full Professor
Faculty of Physics
University of Belgrade

Dr. Goran Poparić
Full Professor
Faculty of Physics
University of Belgrade

Dr. Božidar Nikolić
Associate Professor
Faculty of Physics
University of Belgrade

Dr. Jelena Zagorac
Research Professor
Vinča Institute of Nuclear Sciences
University of Belgrade

Dr. Vladimir Jovanović
Associate Research Professor
Institute for Multidisciplinary Research
University of Belgrade

Abstract

Title: Energy Landscapes, Crystal Structure Prediction, and Modeling of Rare Earth Ternary Compounds

Abstract:

This doctoral dissertation investigates energy landscapes, crystal structures, thermodynamic stability, and selected electronic, magnetic, and vibrational properties of five rare-earth ternary compounds: lanthanum oxyiodide (LaOI), cerium oxynitride (Ce_3O_3N), scandium oxychloride (ScOCl), holmium fluoroselenide (HoFSe), and lanthanum fluorosulfide (LaFS). The work combines crystal structure prediction, crystallographic analysis, and first-principles calculations to identify stable and metastable polymorphs and connect their structural features with functional properties.

The computational workflow combines global optimization and prototype-based data mining to generate large sets of structure candidates, followed by symmetry analysis, structure-type classification, energy ranking, DFT-based structural optimization, and calculations of electronic, magnetic, and vibrational properties. One methodological contribution of the thesis is the development of the STyX software package for processing large structure datasets, extracting symmetry and structure-type information, idealizing structures, and selecting candidates for subsequent DFT local optimization. A second methodological contribution is AI-ELX, a graph neural network model trained on system-specific DFT data and applied to the LaFS system in order to guide DFT calculations in the extreme-pressure region.

The results show that broad energy-landscape exploration can recover known ground-state structures and predict plausible polymorphs that are stable or metastable under different thermodynamic conditions. For LaOI system, the known α phase is confirmed as the ground state modification and six additional low-energy modifications are predicted. For Ce_3O_3N, a data-mined structure is identified as the main global-minimum candidate, and spin-polarized DFT+U calculations show that it is an antiferromagnetic semiconductor. For ScOCl, the known ground-state α phase is recovered and a pressure-induced transition to β-ScOCl is predicted; both phases are indirect-band-gap semiconductors. For HoFSe system, three experimentally observed modifications are recovered and seventeen additional polymorphs are identified, including two high-pressure phases, while DFT+U calculations identify an antiferromagnetic semiconductor as the ground state. For LaFS system, the known *PbFCl*-type phase is recovered, a pressure-induced transition to a *TiNiSi*-type phase is predicted, and it is shown that all investigated polymorphs are semiconducting, while the AI-ELX model identifies additional extreme-pressure candidates.

The thesis demonstrates that broad energy-landscape exploration, automated structure analysis, and first-principles property calculations provide a coherent framework for studying rare-earth mixed-anion systems. The STyX software package and the AI-ELX neural network model further extend this approach toward more systematic exploration of complex crystalline materials.

Keywords: energy landscapes, crystal structure prediction, density functional theory, rare-earth elements, condensed matter, polymorphism, electronic structure, magnetism, machine learning, neural networks

Scientific field: Physics

Scientific subfield: Condensed Matter Physics and Statistical Physics

Сажетак

Наслов докторске дисертације: Енергетски пејзажи, предвиђање кристалних структура и моделовање тернарних једињења елемената ретких земаља

Сажетак:

Ова докторска дисертација истражује енергетске пејзаже, кристалне структуре, термодинамичку стабилност и изабрана електронска, магнетна и вибрациона својства пет тернарних једињења елемената ретких земаља: лантан-оксијодида (LaOI), церијум-оксонитрида (Ce_3O_3N), скандијум-оксохлорида (ScOCl), холмијум-флуороселенида (HoFSe) и лантан-флуоросулфида (LaFS). У раду су комбиновани предвиђање кристалних структура, кристалографска анализа и *ab initio* прорачуни ради идентификације стабилних и метастабилних модификација и повезивања њихових структурних карактеристика са функционалних својстава.

Примењена методологија обједињује глобалну оптимизацију и претраживање структурних прототипова у базама података за генерисање великих скупова структурних кандидата, након чега следе анализа симетрије, класификација структурних типова, рангирање по енергији, структурна оптимизација заснована на *DFT* прорачунима и прорачуни електронских, магнетних и вибрационих својстава. Један методолошки допринос дисертације је развој софтверског пакета *STyX* за обраду великих скупова структура, издвајање информација о њиховој симетрији и структурном типу, идеализацију структура и избор кандидата за локалну оптимизацију на *DFT* нивоу. Други методолошки допринос је *AI-ELX* модел неуронских мрежа трениран на DFT подацима везаним за систем који се истражује, који је примењен на LaFS кристални систем ради навођења *DFT* прорачуна на области екстремних притисака.

Резултати показују да опсежно истраживање енергетског пејзажа може да репродукује познате структуре основног стања и предвиди вероватне модификације које су стабилне или метастабилне под различитим термодинамичким условима. За LaOI је потврђено да је позната α-фаза модификација основног стања и предвиђено је шест додатних нискоенергетских модификација. За Ce_3O_3N је идентификована структура која је добијена претраживањем база података као главни кандидат за глобални минимум, а спин-поларизовани *DFT+U* прорачуни су показали да је та структура антиферомагнетни полупроводник. За ScOCl потврђена је позната α-фаза основног стања и предвиђен је фазни прелаз под притиском у β-ScOCl; обе фазе су индиректни полупроводници. За HoFSe су репродуковане три експериментално добијене модификације и идентификовано је седамнаест додатних модификација, укључујући две фазе под високим притиском, док су *DFT+U* прорачуни идентификовали антиферомагнетно полупроводно стање као основно. За LaFS је репродукована позната фаза *PbFCl* структурног типа, предвиђен је прелаз под притиском у фазу структурног типа *TiNiSi* и показано је да су све предвиђене модификације полупроводници, док је *AI-ELX* модел идентификовао додатне структурне кандидате у области екстремних притисака.

Дисертација показује да широко истраживање енергетских пејзажа, аутоматизована анализа структура и прорачуни својстава из првих принципа чине конзистентан оквир за проучавање тернарних система ретких земаља. Софтверски пакет *STyX* и *AI-ELX* модел неуронских мрежа проширују дати приступ ка систематичнијем истраживању сложених кристалних материјала.

Кључне речи: енергетски пејзажи, предвиђање кристалних структура, теорија функционала густине, елементи ретких земаља, кондензована материја, полиморфизам, електронска структура, магнетизам, машинско учење, неуронске мреже

Научна област: Физика

Ужа научна област: Физика кондензоване материје и статистичка физика

Acknowledgments

The work presented in this thesis was carried out under the supervision of Dr. Dejan Zagorac. I am deeply grateful for his guidance and support throughout this research, as well as for his advice on many aspects of my academic and professional development.

I would particularly like to thank Prof. Dr. Johann Christian Schön for his constant willingness to discuss ideas, for his enthusiasm for my research topic, and for being an invaluable collaborator on most of my publications.

I would also like to thank the members of my thesis committee: Prof. Dr. Đorđe Spasojević, Prof. Dr. Goran Poparić, and Prof. Dr. Božidar Nikolić, Dr. Jelena Zagorac, and Dr. Vladimir Jovanović. I am grateful to Professors Spasojević, Poparić, and Nikolić for being excellent teachers during my studies, and to Dr. Jelena Zagorac and Dr. Vladimir Jovanović for being supportive colleagues.

I extend my thanks to all my colleagues at the Vinča Institute of Nuclear Sciences with whom I collaborated over the years, especially Dr. Branko Matović, Dr. Zoran Ivić, Dr. Tamara Škundrić, Dr. Vesna Maksimović, and Dr. Ivana Cvijović-Alagić.

Some of the people I want to thank supported me even before my scientific career truly began: my late elementary-school physics teacher Branko Rajčić, my high-school physics teacher Ivan Zornić, and my aunt Radosija Pejić, who was my mathematics teacher throughout elementary school.

Finally, my deepest thanks go to my family and friends for their patience, support, and understanding, and especially to my life partner, Biljana Glumac, who has stood by me throughout this journey – it has not always been easy.

Contents

1. Introduction

Rare-earth elements (REEs) comprise the lanthanides together with scandium and yttrium [1-3]. Their compounds underpin a wide range of functional materials used in catalysis, optics, phosphors, high-performance magnets, electronics, energy conversion and storage, and advanced ceramics [1, 2]. Contrary to their historical name, REEs are not intrinsically scarce in the Earth's crust. Their practical "rarity" is instead rooted in their geological occurrence in the same minerals and ores and in the difficulty of separating chemically similar trivalent ions into individual elements or high-purity compounds [1, 4]. For these reasons, and because many technologically important optical, electronic, and magnetic properties emerge in compound form, rare-earth research is concerned at least as much with oxides, halides, chalcogenides, and other compounds as with the elemental metals themselves [2, 5].

Beyond the structural, electronic, magnetic, and vibrational characterization of the five selected compounds, the thesis also focuses on applying global exploration methods and on improving the computational workflow itself, including automated post-processing and structural refinement of very large structure sets and the development and first application of AI models for targeted energy-landscape exploration.

A central motivation of this thesis is therefore the physics of ternary mixed-anion solids, i.e., solids that incorporate more than one anionic species, their chemical and crystallographic analysis, and the use and improvement of computational methods for theoretical exploration of their energy landscapes, structural types, and calculation of their properties. By combining anions that differ in charge, size, electronegativity, polarizability, and hard/soft character, it becomes possible to stabilize coordination environments and structure types that are not simple derivatives of single-anion compounds, while simultaneously tuning band structure, transport, and optical response [6, 7]. In rare-earth systems this strategy is especially powerful because ionic bonding remains dominant, yet subtle differences in anion species can still produce major changes in dimensionality, local coordination, and electronic structure [6, 7].

The purpose of this thesis is to combine energy-landscape exploration, crystal-structure prediction, first-principles calculations, and machine learning models to study selected rare-earth ternary compounds. The thesis focuses on lanthanum oxyiodide (LaOI) [8], scandium oxychloride (ScOCl) [9], cerium oxynitride (Ce_3O_3N) [10], holmium fluoroselenide (HoFSe) [11], and lanthanum fluorosulfide (LaFS) [12].

1.1. Energy Landscapes

An energy landscape is a representation of all possible configurations of a system, along with the energy assigned to each configuration. For crystal systems, this means that every point in configuration space corresponds to a single arrangement of atoms, and the value on the landscape at that point is the potential energy or, at fixed pressure, the enthalpy. The energy landscape concept provides a compact way of connecting crystal-structure prediction, metastability, and the dynamical behavior of condensed-matter systems [13, 14]. An energy landscape can be specified by three components: the configuration space, a scalar function defined on that space, and a local-connectivity relation between neighboring configurations. In crystals, neighboring configurations are accessed by small continuous changes in atomic positions and, if the unit cell is allowed to change, by small changes in the lattice vectors as well. Crystal-structure prediction can then be viewed as the search for the local minima on this high-dimensional landscape and for the relations between the basins in which these minima are located.

A fully quantum-mechanical description would require the electronic wave function and, in magnetic systems, spin degrees of freedom. In practical crystal-structure exploration, this problem is usually simplified by the Born-Oppenheimer separation of electronic and nuclear degrees of freedom. The nuclei are treated as classical particles moving on an effective potential-energy surface, while the electronic degrees of freedom are incorporated into the potential energy. In an empirical search, this potential is approximated by an analytic model of the interatomic interactions. In an *ab initio* calculation, the electronic structure problem is solved for each fixed nuclear geometry, and the resulting total electronic energy plus ion-ion interactions defines the ground state energy of that configuration. In this approximation, an arrangement of N atoms is described by the 3N-dimensional coordinate vector

$$X = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N), \qquad \mathbf{r}_i = (x_i, y_i, z_i) \in \mathbb{R}^3 \tag{1.1}$$

where the index i labels the atoms. The corresponding velocities and momenta are represented by the vectors $V = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N)$ and $P = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$. The full classical state is therefore a point in phase space, while the structural part relevant for crystal-structure prediction is a point in configuration space. For isolated molecules, this configuration space is typically defined solely by atomic coordinates, but for periodic solids, it must also include the unit cell because changes in the lattice vectors correspond to physically distinct crystal structures.

For a nonrelativistic classical system, the Hamiltonian separates into potential and kinetic contributions,

$$E(X, V) = E_{\text{pot}}(X) + E_{\text{kin}}(V), \qquad E_{\text{kin}}(V) = \frac{1}{2}\sum_{i=1}^{N} m_i\, |\mathbf{v}_i|^2 \tag{1.2}$$

where $E_{\text{pot}}(X)$ defines the static potential-energy landscape and m_i is the mass of atom i. Equivalently, in terms of momenta one can write $E_{\text{kin}}(P) = \sum_i |\mathbf{p}_i|^2 / (2m_i)$. The forces acting on the atoms determine their trajectories and are given by the negative gradient of the potential energy,

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i} = \mathbf{v}_i, \qquad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i = -\nabla_{\mathbf{r}_i} E_{\text{pot}}(X). \tag{1.3}$$

In principle, for a smooth potential and specified initial positions and momenta, Newton's equations determine a unique trajectory in phase space, and this deterministic trajectory is the basis of molecular dynamics. Crystal-structure prediction, however, usually does not rely on one long physical trajectory to discover structures because the trajectory depends on the initial conditions, and low-energy minima may be separated by large energy barriers, and a single trajectory may remain trapped in one basin for a time much longer than the available simulation time. Global searches, therefore, combine stochastic exploration and local optimization to sample many distinct regions of configuration space.

The most important geometrical and dynamical features of an energy landscape are local minima, basins of attraction, transition regions, and energy barriers (Figure 1.1). A local minimum corresponds to a configuration that is stable against sufficiently small displacements at zero temperature. The basin of attraction of a minimum is the set of configurations that relax into that minimum under a specified local optimization or quench procedure. Transition regions are the parts of configuration space that a system passes through as it moves between neighboring basins. In the simplest picture, these regions contain saddle points or transition states; for large systems, they may involve many alternative pathways. The energy barrier is the minimum increase in energy that must be overcome along a relevant path to escape from one basin and reach another. Thus, energy landscape exploration identifies the lowest-energy structure and also describes how many competing structural modifications exist, how close they are in energy, and whether they are separated by barriers large enough to produce kinetic trapping on the relevant time scale [13, 14].

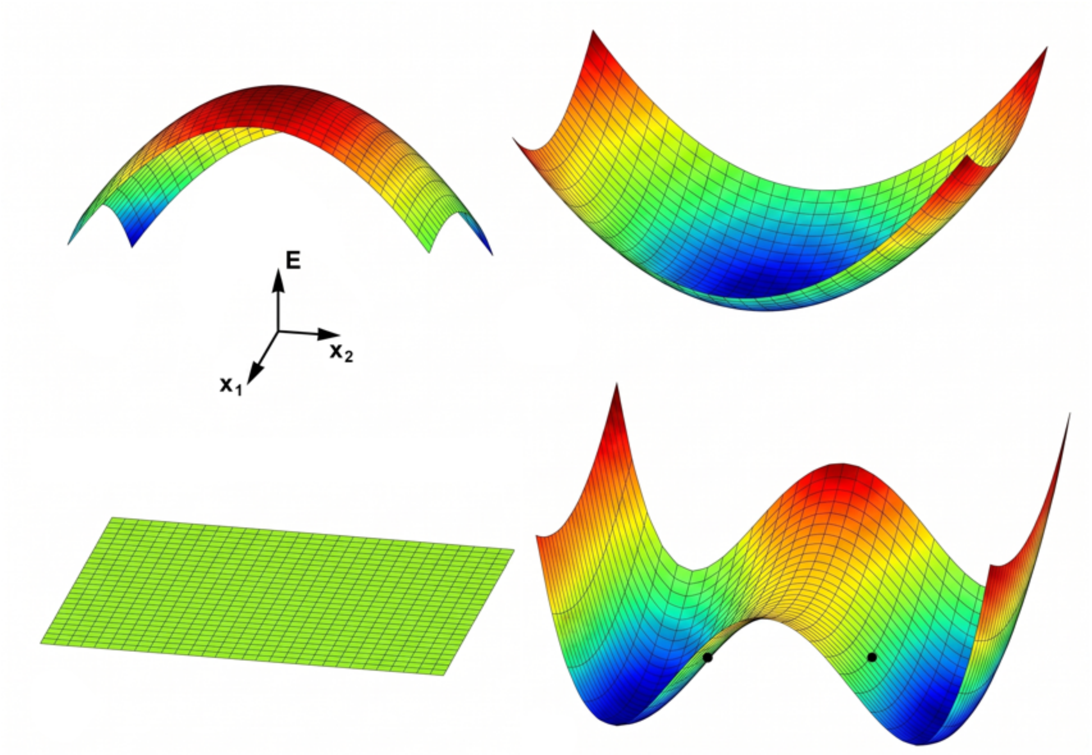

Figure 1.1. Characteristic elements of an energy landscape, projected onto two coordinates x_1 and x_2. Local minima (energy valley or basin) correspond to possible stable or metastable structure candidates (top right), while the barriers (energy peaks) between them influence kinetic accessibility (top left). A flat energy path (bottom left), which shows no energy change, and two neighboring minima with a barrier between them (bottom right) are also shown.

The thermodynamic interpretation of a local region of the landscape is commonly expressed in terms of local ergodicity. A region R of configuration space is locally ergodic, for a chosen observable and a specified temperature, if the system samples configurations inside R many times before it escapes from R. Let $t_{\mathrm{rel}}(R,T)$ be the intra-basin relaxation time, t_{obs} the observation or measurement time, and $t_{\mathrm{esc}}(R,T)$ the escape time from the region. The condition for local ergodicity on the observation time scale can then be written as

$$t_{\mathrm{rel}}(R,T) \ll t_{\mathrm{obs}} \ll t_{\mathrm{esc}}(R,T). \tag{1.4}$$

Here, the observation time is the time scale over which an experiment or simulation probes the system. If the Eq. (2.4) is fulfilled, measured properties are reproducible for the metastable region R, even though the system is not in global equilibrium of the entire landscape. This is the physical basis of kinetic stability: a metastable modification is not the absolute minimum of the whole landscape, but it behaves as a well-defined structure because it remains trapped in a basin that has reached local equilibrium for the relevant observation time. Diamond at ambient conditions is the classic example: although graphite is the thermodynamic ground state of carbon, the barriers to reconstructing the diamond lattice are so large that diamond persists for 10^{80} years under ordinary conditions.

For an observable O, the time average along a trajectory in phase space can be written schematically as

$$\langle O \rangle_t = \frac{1}{t_{\mathrm{obs}}} \int_{t_1}^{t_2} O\left(X(t), V(t)\right) dt, \qquad t_{\mathrm{obs}} = t_2 - t_1. \tag{1.5}$$

If local relaxation inside the region R is fast enough, this time average can be replaced, within the desired accuracy, by a canonical ensemble average restricted to the same region R of phase space,

$$\langle O \rangle_R(T) = \frac{\int_R O\,(X, V)\ \mathrm{e}^{-E(X,V)/k_B T} dX\, dV}{\int_R \mathrm{e}^{-E(X,V)/k_B T} dX\, dV}. \tag{1.6}$$

This relationship is useful only when local relaxation inside region R is fast compared with the observation or simulation time, and escape from R is slow on that same time scale. It justifies treating individual basins as separate modifications, each with its own local free energy and physical properties, such as vibrational spectra, electronic structure, magnetic behavior, optical properties, etc. It also clarifies why not every mathematical local minimum is physically relevant: the minimum must be sufficiently low in energy and sufficiently kinetically isolated from other basins on the time scale of interest.

The barrier height around a locally ergodic region determines the escape time. At low and moderate temperatures, thermally activated escape processes are often described qualitatively by an Arrhenius-type relation,

$$t_{\mathrm{esc}} \approx \tau_0 \exp\left(\frac{E_{\mathrm{barrier}}}{k_{\mathrm{B}} T}\right), \tag{1.7}$$

where E_{barrier} is an effective barrier height and τ_0 is a pre-exponential factor. The exponential factor explains why kinetic stability is extremely sensitive to the height of the barrier: a structure may remain stable for macroscopic times despite not being the global minimum if the escape barrier is high. Conversely, a shallow local minimum may be physically unobservable if barriers toward neighboring lower-energy basins are too small.

External thermodynamic variables are control variables imposed by the environment, and temperature and pressure are the most important examples for the present context. For a periodic solid at fixed external pressure and low temperature, the appropriate static objective function is the enthalpy

$$H_p(X) = E_{\mathrm{pot}}(X) + pV_{\mathrm{cell}}(X), \tag{1.8}$$

where $V_{\mathrm{cell}}(X)$ is the unit-cell volume associated with configuration X. During the structural relaxation at fixed external pressure, the cell volume adjusts together with the atomic coordinates. After local optimization of a particular structural modification j at pressure p, the enthalpy used to compare phases is

$$H_j(p) = E_j\left(V_j(p)\right) + pV_j(p). \tag{1.9}$$

Therefore, pressure-dependent crystal-structure prediction can be viewed as the exploration of a family of pressure-modified landscapes. The same idea underlies the construction of $E(V)$ curves and $H(p)$ relations used later in this thesis: the relevant stable phase at a given pressure is the structure with the lowest optimized enthalpy, while metastable phases correspond to other local minima of the same pressure-dependent landscape.

1.2. Rare Earth Elements and their Compounds

Rare-earth elements (REEs) are conventionally defined as the fifteen lanthanides together with scandium and yttrium [1-3]. The historical term “rare earths” reflects the difficulty of their discovery, separation, and purification rather than true geochemical scarcity: several REEs (such as Ce, Y, Nd, and La) are relatively abundant in the crust, but they usually occur together in minerals and exhibit strongly similar chemistry, which makes their separation laborious and expensive [1, 2, 4].

From the standpoint of atomic structure and condensed matter theory, the lanthanides are distinguished by the progressive filling of the inner 4*f* shell, consistent with the Pauli principle and the electronic structure sequence across the series [15]. In most atoms and condensed-matter environments, the 4*f* orbitals are spatially compact and are shielded by the filled 5*s* and 5*p* subshells. As a result, direct overlap between 4*f* and ligand orbitals is limited, most lanthanide bonding remains predominantly ionic, and coordination chemistry changes more smoothly across the series than in many *d*-block families [5, 15]. At the same time, the localized 4*f* electrons are responsible for several characteristic properties of rare-earth materials, notably sharp atom-like optical transitions and local-moment magnetism with strong anisotropy and complex ordering behavior in solids [5, 16].

Lanthanide magnetic behavior is often remarkably complex. In elemental lanthanides, magnetic order emerges from a competition among indirect 4*f*-4*f* exchange mediated by conduction electrons, crystal-field anisotropy, and spin-orbit coupling, which together generate ferromagnetic, antiferromagnetic, helical, conical, fan-like, and field-induced spin-reorientation phases across the series [16, 17]. In rare-earth intermetallic compounds, the same physical mechanisms, reshaped by composition, crystal structure, and band filling, produce an even broader spectrum of magnetic behavior and magnetic anisotropy, which is one reason why rare-earth systems have long served as model materials for studying exchange interactions and collective many-body magnetic behavior [16, 18].

One of the most important systematic trends is the lanthanide contraction, i.e., the gradual decrease in ionic radius across the series as effective nuclear charge increases [2, 5]. This contraction influences lattice parameters, preferred coordination numbers, structural stability, and separation chemistry because even small differences in ionic radius affect complex formation equilibria and ion-exchange selectivity [4, 5]. Much of the structural behavior and many physical properties of rare-earth compounds can therefore be understood as a balance between overall ionic similarity and these subtle radius-driven variations.

Although the trivalent state dominates the chemistry of rare-earth compounds, several important exceptions exist. Cerium is well known for an accessible tetravalent state, whereas europium and ytterbium most commonly stabilize in divalent states; under appropriate conditions, related non-trivalent behavior can also occur for other lanthanides [5]. Such deviations from the dominant $^{3+}$ state can alter stoichiometry, defect chemistry, transport, phase stability, conductivity, optical gaps, and magnetic response, and they are central to the unusual reduction-oxidation chemistry, optical, and magnetic behavior of many rare-earth materials [5].

Because lanthanide cations are hard Lewis acids (electron-pair acceptors with a strong preference for hard donor anions), oxides and fluorides are among the most classical and extensively studied families of rare-earth compounds [5]. The rare-earth oxides are especially important as reference systems: the Ln_2O_3 oxides occur in several well-known polymorphic families, while mixed-valent systems such as cerium oxides display additional non-stoichiometry and defect-related structural complexity [19]. For this reason, oxide chemistry provides a natural starting point for exploring polymorphism, thermodynamic stability, and the relationship between local coordination and crystal structure in rare-earth solids [19].

Rare-earth halides, i.e., compounds containing halogen anions such as F^-, Cl^-, Br^-, or I^-, form another foundational family. Besides their intrinsic structural interest, lanthanide trihalides are among the most important synthetic precursors in solid-state rare-earth chemistry [20]. Their structures vary systematically with both halogen size and cation radius, and the corresponding fluoride chemistry is particularly relevant for mixed-anion materials because fluoride anions often stabilize robust ionic frameworks within oxide-fluoride, fluoride-chalcogenide, and related systems [20]. The broader rare-earth halide literature also illustrates how even nominally ionic systems can generate substantial structural diversity once more than one cation or anion type is introduced [20]. Comprehensive thermodynamic treatments of lanthanide fluorides and chlorides also show that these comparatively simple halide systems provide one of the best-quantified reference frames for assessing phase stability, phase equilibria, and thermodynamic trends in rare-earth mixed-anion systems [21].

Compounds containing softer, more polarizable anions, such as chalcogenides and pnictides, connect rare-earth chemistry to semiconducting behavior, transport anomalies, many-body electronic behavior, and structure-property trends that are less evident in classical oxide systems [22, 23]. Besides oxides, rare-earth sulfides, selenides, and tellurides constitute a large family of binary and higher compounds with multiple stoichiometries and structural types [23]. Rare-earth pnictides, including nitrides, are similarly important as model systems for understanding how changes in anion species influence electronic structure, magnetism, bonding character, and synthesis routes [22]. A classic example is the europium chalcogenide family EuO, EuS, EuSe, and EuTe, which established rare-earth chalcogenides as model magnetic semiconductors in which exchange interactions, band structure, and $4f$–$5d$ interactions are tightly coupled [24].

Once one moves from binaries to ternaries and higher systems, the structural diversity increases dramatically. Systematic treatments of rare-earth ternary chalcogenide systems show that the accessible structure types are governed by a combination of cation size, preferred coordination geometry, anion polarizability, and synthesis conditions such as temperature, pressure, and flux environment [25]. From the viewpoint of materials discovery, this is particularly important because many multicomponent rare-earth systems can contain several competing low-energy arrangements rather than a single overwhelmingly dominant structure type [25]. A survey of ternary and quaternary rare-earth chalcogenides with Si, Ge, Sn, Pb, and In reinforces this picture by documenting a wide range of stoichiometries, coordination environments, and crystal architectures in higher-order chalcogenide systems [26].

In modern solid-state physics and materials science, mixed-anion compounds are widely recognized as an important route to functional materials [6], i.e., materials whose electronic, magnetic, optical, or transport properties are central to their use. Unlike simple single-anion analogues, many mixed-anion phases contain crystallographically distinct anion sites, mixed-anion coordination polyhedra, and site-ordering phenomena that strongly affect structure and properties [6, 7]. Introducing an anion that is less electronegative and more polarizable than oxygen into an oxide-derived framework can, for example, shift valence-band and conduction-band edges, modify local symmetry, and stabilize structure types that are inaccessible in single-anion systems [6]. Rare-earth oxychalcogenides provide a particularly instructive subset of mixed-anion chemistry because their structures vividly reflect the different bonding preferences of oxide and chalcogenide anions [7].

Within this broad landscape, rare-earth oxyhalides and related oxide-halide compounds occupy a special place. Their layered structural motifs commonly place chemically distinct anions in different local environments, so that bonding within the layers and interactions between the layers can differ strongly, and this structural anisotropy is often accompanied by rich polymorphic behavior under changing thermodynamic conditions [6]. In this thesis, lanthanum oxyiodide (LaOI) [8] and scandium oxychloride (ScOCl) [9] serve as representative oxide-halide systems through which these general mixed-anion principles can be studied in concrete chemical settings.

Rare-earth oxynitrides represent the oxide-nitride branch of mixed-anion rare-earth chemistry. Because the N^{3-} anion differs from the O^{2-} anion in charge, size, and bonding character, incorporation of nitrogen can significantly alter both electronic structure and structural stability relative to oxide analogues [6]. The not-yet-synthesized Ce_3O_3N compound, studied here computationally as a candidate stoichiometry in the Ce–O–N system, combines an oxide-nitride anion framework with the flexibility of cerium oxidation state [5, 10].

Another rare-earth compound studied in this thesis belongs to rare-earth fluoride chalcogenides, in which a hard fluoride anion with low polarizability coexists with a softer, more polarizable chalcogenide anion. Classical representatives are the rare-earth fluorosulfides LnFS, which have attracted interest as thermally robust color pigments and as model systems for linking mixed-anion coordination to optical response, because local coordination directly affects crystal fields, band structure, and dielectric behavior [27, 28]. At fixed 1:1:1 stoichiometry, the LnFS family spans multiple structure types across the rare-earth series, illustrating how the competition between cation size and anion coordination can generate genuine polymorphic diversity even in compositionally simple ternary systems [29].

Lanthanum fluoroselenide chemistry provides a closely related and experimentally well-anchored reference frame for fluoride-chalcogenide systems. The polymorphism of LaFSe, including phases with PbFCl-type and CeHSe-type modifications, has been investigated by combining high-temperature synthesis, diffraction, spectroscopy, and crystal-structure prediction [30]. Subsequent work on sulfur substitution in $LaFSe_{1-x}S_x$ showed that even partial anion replacement can induce systematic, and in some cases non-intuitive changes in band-gap magnitude, gap character, and overall electronic structure [31]. These results directly motivate the present focus on fluoride-chalcogenide energy landscapes, where locating new structures in different pressure and temperature regions can provide a broader basis for future property studies of more complex substituted and higher-order systems. Holmium fluoride selenide (HoFSe) extends the fluoride-chalcogenide family into the heavier lanthanides, where partially filled 4*f* shells make local-moment magnetism of particular interest. A combined theoretical and experimental study has shown that HoFSe exhibits several polymorphs accessible via high-temperature synthesis and that these can be consistently understood within a broad energy-landscape exploration framework based on structure prediction and first-principles optimization [11].

Finally, lanthanum fluoride sulfide (LaFS) is a mixed-anion semiconductor and an important representative of the rare-earth fluoride sulfides. Earlier crystallographic work already demonstrated that rare-earth fluoride sulfides can adopt distinct structure types at fixed stoichiometry [29]. In the thesis, LaFS is used as a key system to extend energy-landscape exploration beyond structure identification toward broader property exploration, including electronic structure, pressure response, vibrational stability, and AI-assisted ultra-high-pressure screening [12]. This makes LaFS particularly suitable for linking crystal-structure prediction to experimentally relevant questions of phase stability: enthalpy calculations identify which polymorphs are thermodynamically favored at different pressure and temperature conditions, while phonon dispersions test whether selected structures are dynamically stable.

1.3. Research Objectives

The general objectives of this thesis are to: (i) improve the use of existing energy-landscape exploration methods and develop new tools that make large-scale structural searches more efficient and reproducible; (ii) apply these methods to ternary rare-earth mixed-anion compounds in order to discover and classify feasible crystal structures within their energy landscapes; and (iii) investigate the thermodynamic, electronic, magnetic, and vibrational properties of the most relevant structure types obtained in these energy landscape explorations.

One objective of this thesis is to advance structure-property understanding in rare-earth and rare-earth-related mixed-anion compounds by combining systematic exploration of crystal-structure energy landscapes with first-principles property calculations and, where available, direct comparison with experiment. At the most general level, the thesis examines which crystalline modifications are feasible for a given composition, how closely they compete energetically, how pressure affects their relative stability, and which key electronic, vibrational, and magnetic properties follow from the resulting structural differences.

A second objective is to refine and extend the underlying methodology. The thesis develops and applies a reproducible multistep workflow in which large-scale structure generation via global optimization is complemented by prototype-based data mining, followed by post-processing, structural refinement, and, finally, *ab initio* relaxation, energetic ranking, and property evaluation. In particular, the workflow is extended by STyX post-processing software package, which makes symmetry analysis, duplicate removal, structure-type clustering, database tracking, and representative selection reproducible for very large structure sets, and by AI-ELX, which introduces a graph-neural-network-assisted route for faster screening of targeted regions of the energy landscape. These additions address the central scale problem when global exploration generates hundreds of thousands or millions of candidate structures, but only a compact, crystallographically analyzed, and physically meaningful subset can be passed to first-principles calculations.

For lanthanum oxyiodide (LaOI), the main objectives are threefold: first, to recover the experimentally known ground-state structure, thereby validating the search strategy; second, to identify additional low-energy polymorphs with symmetry-distinct local coordinations; and third, to characterize their electronic structure and pressure response in order to determine how structural differences affect stability and physical properties [8].

For scandium oxychloride (ScOCl), the objectives mirror those for LaOI and involve another nonmagnetic oxide-halide system that is compositionally related. The aims are to reproduce the known ScOCl modification, to map additional feasible minima that may correspond to high-pressure or high-temperature polymorphs, and to evaluate their relative stability through $E(V)$ and $H(p)$ relationships. In this way, ScOCl provides an internal comparison system for interpreting how much of mixed-anion polymorphism is controlled by ionic size and anion species alone [9]. At the same time, exploring a scandium compound broadens the study toward the broader rare-earth family, rather than strictly lanthanide systems.

For cerium oxynitride, Ce_3O_3N, a possible stoichiometry in the ternary Ce–O–N system, the objectives are explicitly predictive. The first goal is to enumerate low-energy structure candidates by combining global search and data mining, and then to refine and rank them after *ab initio* optimization. The second goal is to establish how pressure affects the competition among fixed-composition Ce_3O_3N candidates and to identify which modification is thermodynamically preferred within that composition. The third goal is to connect the predicted structures to plausible electronic and magnetic behavior in a Ce-containing mixed-anion system whose properties are strongly influenced by cerium's accessible nontrivial oxidation states [5, 10]. In the thesis, this research is extended beyond the published paper by including magnetic order: antiferromagnetic solutions are expected to be semiconducting, whereas nonmagnetic calculations indicate metallic behavior.

For holmium fluoride selenide (HoFSe), which was explored through a combined theoretical and experimental study, the thesis aims to connect theory and experiment as closely as possible. The objectives are to rationalize experimentally observed polymorphs in terms of their relative energetic ordering, to analyze the electronic structure of these polymorphs, and to treat magnetism as an essential component of heavy-lanthanide fluoride-chalcogenide behavior, including the role of magnetic ordering and anisotropy when spin-orbit coupling becomes important [11, 16]. The structural and electronic parts are anchored in the published HoFSe paper [11], whereas the magnetic property calculations are additionally included in the thesis.

For lanthanum fluorosulfide (LaFS), the objectives are designed to extend the fluoride-chalcogenide family to a mixed-anion semiconductor system and to explicitly include vibrational stability of its high-pressure phases. The goals are to map the LaFS energy landscape, identify a set of low-energy polymorphs, compute their electronic structures and pressure-dependent stability relations, and finally test their dynamical stability through phonon calculations [12, 29]. The LaFSe/LaFS chemical neighborhood provides additional context for these goals. The LaFSe polymorphism offers experimentally anchored structure types against which the predictive workflow can be tested and calibrated, while the study of anion substitution in $LaFSe_{1-x}S_x$ demonstrates directly how changes within the chalcogen sublattice can alter electronic properties in a systematic but not always intuitive way [30, 31].

Across all these explored systems, a unifying practical objective is to translate structural predictions into experimentally useful guidance and to identify properties that can be tuned before costly experimental work is undertaken. This is achieved by combining equations of state, relative enthalpies, and pressure-dependent energetic ordering to identify plausible transition pressures and stability windows for competing polymorphs. Electronic structure, band-gap trends, magnetic ordering, and vibrational spectra are then used to compare candidate phases on a physically meaningful basis and to guide experimental synthesis toward materials with technologically favorable properties. Because mixed-anion systems often contain several low-lying competing structures, the thesis emphasizes comparative ranking and pressure response rather than isolated single-structure predictions [8, 9, 12]. This emphasis is also consistent with broader high-pressure studies of rare-earth compounds, in which pressure serves as a central control parameter governing structural, electronic, and magnetic competition [32].

Finally, the thesis is intended to serve as a transferable framework for exploring energy landscapes in rare-earth mixed-anion systems and other crystal systems. By treating oxide-halide, oxide-nitride, and fluoride-chalcogenide compounds within a single coherent framework, the thesis demonstrates how robust structure-generation methods, first-principles evaluation, new post-processing tools, and AI-assisted energy screening can be combined in a computational program that can later be extended to multiple-cation, doped-ternary, quasi-ternary, and quaternary systems. In this context, the AI model developed for targeted LaFS exploration is part of the broader convergence of computational physics, data-driven modeling, and high-throughput first-principles workflows. The importance of artificial neural networks and energy landscape for physics was recognized at the highest level by the 2024 Nobel Prize in Physics, awarded to John J. Hopfield and Geoffrey Hinton for foundational discoveries in computational design and structural prediction, and inventions that enable machine learning with artificial neural networks.

1.4. Outline

The thesis is divided into nine chapters. In Chapter 1: "Introduction", the motivation of the thesis is introduced through the concepts of energy landscapes, structure prediction, overview of rare-earth elements and their compounds, and the overview of research objectives.

Chapter 2: "Theoretical Background" summarizes the theoretical background needed for the work: crystallographic symmetry, space groups, structure types, phonons, density functional theory, magnetism in solids, equations of state, pressure-induced phase stability, and the basic ideas of machine learning and graph neural networks used later in the thesis.

Chapter 3: "Computational Methodology in Energy Landscape Exploration" describes the computational methodology used in the thesis. The first part introduces the established energy-landscape tools used in this work, including global optimization, data mining, and local optimization. The second part presents the methodological developments that form an important original contribution of the thesis: the STyX workflow for symmetry analysis, structure-type comparison, clustering, and ranking of large structure sets, and the AI-ELX model for graph-neural-network-based energy prediction, structural relaxation, and equation-of-state-based high-pressure region screening.

Chapters 4-8 contain the results for five rare-earth systems. Chapter 4 investigates LaOI, where global energy-landscape exploration reproduces the known α-LaOI ground state and predicts additional metastable polymorphs, with thermodynamic stability and electronic properties are also being investigated. Chapter 5 explores Ce_3O_3N system, combining structure prediction with PBE+U magnetic and electronic-structure calculations for the ground-state candidate. Chapter 6 investigates ScOCl structural landscape, including its pressure-dependent stability and electronic properties of two predicted phases. Chapter 7 presents the HoFSe study, connecting structure prediction with pressure response, magnetic ordering, non-collinear spin-orbit-coupled calculations, and electronic structure. Chapter 8 focuses on LaFS, combining large-scale structure candidate prediction, investigation of electronic and vibrational properties, pressure-induced phase stability, and an AI-guided search for high-pressure phases.

Chapter 9 summarizes the main scientific results and methodological improvements, and the appendices A1-A8 collect data that support the main discussion, including optimized structural parameters, electronic band structures, and settings for DFT calculations and AI model training.

2 Theoretical Background

2.1 Symmetry, Space Groups, Structure Types, and Lattice Vibrations

Symmetry represents a central principle in condensed-matter physics: it identifies transformations that leave a structure, Hamiltonian, or excitation problems invariant, and therefore determines degeneracies, selection rules, allowed couplings, and the labels used for electronic, vibrational, and magnetic states [33-35]. In this thesis, symmetry provides the crystallographic framework for comparing predicted and observed structures by their space groups, Bravais lattices, Wyckoff positions, and structure types. The concept of symmetry is also needed for phonons, because lattice vibrations are classified by wave vector and by the symmetry of their displacement patterns. The same principle applies to electronic structure, where Bloch states, band degeneracies, and band crossings are labeled by crystal momentum and by the irreducible representations of the relevant little groups, and to magnetic arrangements, where ordering wave vectors, spin directions, and magnetic symmetry determine which magnetic configurations are symmetry-distinct. The following sections introduce these ideas only to the depth needed in the thesis for the structure-prediction and property-analysis chapters that follow.

2.1.1 Symmetry Operations, Groups, and Representations

A symmetry operation is an operation that leaves the relevant physical object invariant, i.e. unchanged in the sense relevant to the problem. In quantum mechanics, the object may be the Hamiltonian, in crystallography, it is the infinite periodic arrangement of nuclei, and in spectroscopy, it is the dynamical problem consisting of the states, operators, and allowed transitions. The set of all such operations forms a group, where all group axioms are satisfied: closure, associativity, the existence of an identity, and the existence of an inverse for every element [33, 36]. In quantum mechanics, symmetry may be formulated more precisely as a transformation of the state space that preserves the dynamical law. In operator form, this implies that the symmetry operators commute with the Hamiltonian of the system [35].

In the Seitz notation, a symmetry element is written as a combined point operation and translation,

$$g \equiv \{R \mid \tau\}, \qquad \mathbf{r}' = R\mathbf{r} + \tau. \tag{2.1}$$

Here, R is a rotation, reflection, inversion, or rotoinversion acting on spatial coordinates, while τ is a translation that may be either primitive, i.e. an integer combination of the primitive lattice vectors, or fractional, i.e. only a fraction of such a lattice translation. In general, rotational and translational parts are intertwined because operations such as screw axes and glide planes combine a point operation with a nonprimitive translation, which is why the full symmetry group of a crystal is a space group rather than an ordinary point group [36].

Physical states, displacement patterns, spin configurations, or basis functions transform under representations of the symmetry group, and the irreducible representations provide the natural labels of symmetry-distinct states [33]. In practice, this is the origin of symmetry-based degeneracy counting, compatibility relations between symmetry labels at different points of the Brillouin zone, and selection rules for matrix elements. If an eigenspace of the Hamiltonian transforms according to an irreducible representation of dimension d, then a d-fold degeneracy is symmetry-allowed unless further perturbations split it [33].

2.1.2 Bravais Lattices, Basis, Primitive and Wigner-Seitz Cell, and the Crystallographic Space Groups

A crystal is specified by a Bravais lattice together with a basis. The Bravais lattice describes the periodic translational framework of the crystal, while the basis specifies which atoms occupy each repeat unit. The discrete lattice translations may be written as:

$$T(n_1, n_2, n_3) = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \qquad n_i \in \mathbb{Z}. \tag{2.2}$$

The primitive vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ generate the full translation group. The primitive cell is not unique, but choices such as the Wigner-Seitz cell, which perfectly reflects the symmetry of the underlying Bravais lattice (see Figure 2.1), are often especially informative because they reflect the metric symmetry of the lattice, i.e. the symmetry encoded in lattice lengths and angles, as directly as possible [36, 37].

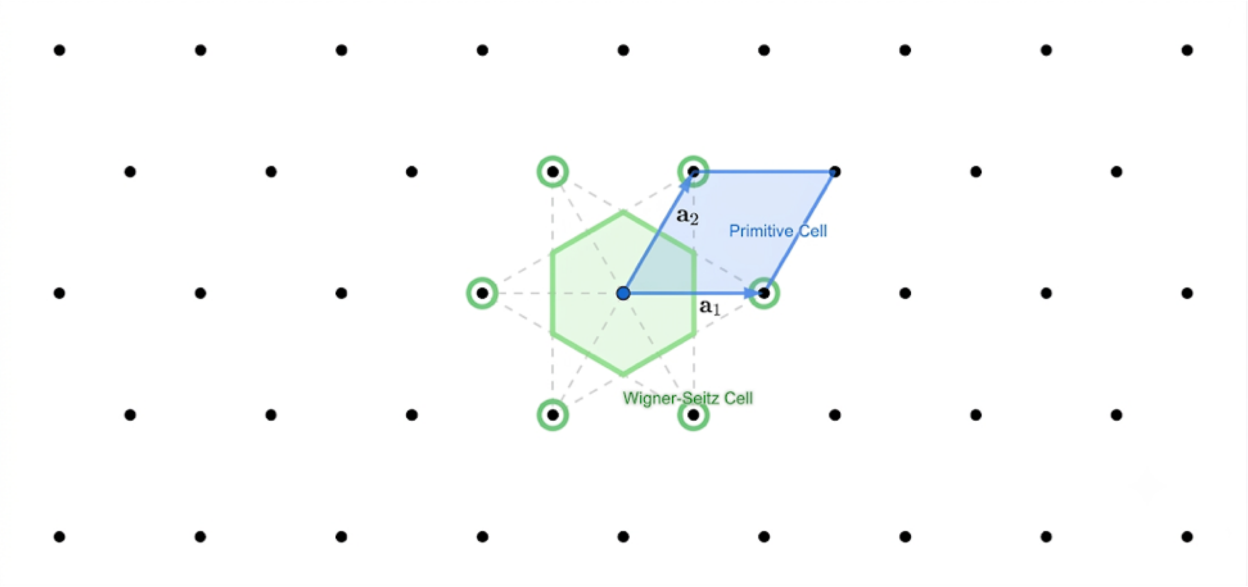

Figure 2.1. Conventional primitive cell (blue rhombus) and Wigner-Seitz cell (green hexagon) for a 2D hexagonal lattice.

Not every point-group rotation is compatible with a periodic lattice. The crystallographic restriction theorem states that a rotation compatible with a periodic lattice must map the lattice onto itself, which in three dimensions leaves only 1-, 2-, 3-, 4-, and 6-fold rotational axes; for example, 5-fold rotations can occur in finite objects and in quasiperiodic order, but not in a three-dimensional periodic lattice. This restriction leads to the seven crystal systems, fourteen Bravais lattices, thirty-two crystallographic point groups, and two hundred and thirty crystallographic space groups [33, 36]. The 230 space groups are grouped and conventionally numbered as follows: triclinic (No. 1-2), monoclinic (No. 3-15), orthorhombic (No. 16-74), tetragonal (No. 75-142), trigonal (No. 143-167), hexagonal (No. 168-194), and cubic (No. 195-230).

2.1.3 Space Groups, Wyckoff Positions, and Symmorphic and Nonsymmorphic Groups

The translation group T is an Abelian normal (invariant) subgroup of every crystallographic space group S. The point part of the crystal symmetry is recovered by factoring out the translations, so that

$$\frac{S}{T} \cong P, \tag{2.3}$$

where P is the crystallographic point group associated with the lattice. In nonsymmorphic space groups, some coset representatives contain fractional translations associated with screw axes or glide planes. Ordinary point-group operations also impose selection rules and degeneracies, but these fractional translations add characteristic nonsymmorphic effects such as band connectivity and zone-boundary degeneracies [36, 37].

If all fractional translations can be removed by an appropriate choice of the origin, the group is symmorphic and may be written as a semidirect product of the translation subgroup with the point group,

$$S = T \rtimes P \qquad \text{(symmorphic case).} \tag{2.4}$$

Among the 230 crystallographic space groups, 73 are symmorphic, while the rest are nonsymmorphic [36] (the full classification is summarized in Appendix A2 Table A2.1).

A Wyckoff position is the set of symmetry-equivalent points generated from one point under the action of the space group, characterized by a multiplicity and a site-symmetry group. Two structures may share the same space group but differ in occupied Wyckoff positions, local site symmetries, or internal coordinate parameters; for that reason, the space-group label alone is rarely sufficient to fully characterize a crystal structure.

The same logic is applicable to the reciprocal space. Translational symmetry labels states by crystal momentum **k**, while point-group symmetry groups the symmetry-equivalent wave vectors into a star of **k** (**k**-star is the complete set of distinct, non-equivalent wavevectors, obtained from **k** by the allowed point-group operations).

At a generic point in the Brillouin zone, the little group (the subgroup that leaves **k** invariant up to a reciprocal lattice vector) is small, but at high-symmetry points or lines it becomes larger, which is why special **k** points can support extra degeneracies and more restrictive selection rules [34, 35].

The Bloch relation

$$\psi_{n\mathbf{k}}(\mathbf{r} + \mathbf{T}) = e^{i\mathbf{k}\cdot\mathbf{T}}\psi_{n\mathbf{k}}(\mathbf{r}) \tag{2.5}$$

is the most compact manifestation of translational symmetry for electronic states, and its analogues reappear for phonons and other crystal excitations [37]. It originates from the commutation of the crystal Hamiltonian with all lattice-translation operators, so eigenstates can be chosen simultaneously as eigenstates of translation and therefore acquire only a phase under lattice translation.

2.1.4 Structure Types, Prototypes, and Polymorphism

A structure type identifies a recurring symmetry and connectivity pattern that can be realized by many different compositions. In practice, a structure type or prototype denotes a family of compounds sharing the same symmetry setting, Wyckoff-position pattern, and topological arrangement of coordination polyhedra, even if the lattice parameters and bond lengths differ because of chemistry, pressure, or temperature. Prototype names used within this thesis, such as NaCl, PbFCl, TiNiSi, and ZrBeSi, remain useful precisely because they condense a large amount of structural information into a short label. Since some of the same motifs recur across rare-earth mixed-anion compounds and reappear in LaOI-, LaFSe-, HoFSe-, and LaFS-related studies, structures found in energy-landscape exploration are best interpreted as members of known structure type families [8, 11, 29-31].

Conversely, two structures may share the same space group but still belong to different structure types if the occupied Wyckoff positions, coordination topology, or polyhedron connectivity differ. This is why a good structural description should normally report, at a minimum, the space group, the prototype label when one is known, the occupied Wyckoff positions, and the characteristic local coordination environment. In energy-landscape work, these descriptors help separate genuinely distinct minima from small distortions, cell transformations, or coordinate relabelings of the same underlying structure type.

From the viewpoint of polymorphism and relative phase stability, structure types are also a compact way to discuss structural diversity. A single composition may have several locally stable arrangements that belong to different prototype families and are stabilized at different pressure or temperature conditions. A familiar example is water ice, which has more than a dozen known crystalline modifications across different pressure-temperature ranges.

2.1.5 Symmetry, Normal Modes, and Phonons

In lattice dynamics, for a finite system with N atoms, the full displacement space has dimension $3N$ and transforms as a dynamical representation, where the symmetry operations act on the $3N$-dimensional displacement vector space of the nuclei. To isolate true vibrations, one subtracts the pure translations and rigid rotations, so that

$$\Gamma_{\mathrm{vib}} = \Gamma_{3N} - \Gamma_{\mathrm{trans}} - \Gamma_{\mathrm{rot}}. \tag{2.6}$$

Here, Γ_{vib} denotes the vibrational representation that remains after the translational and rotational degrees of freedom are removed, while Γ_{3N} , Γ_{trans} , and Γ_{rot} denote, respectively, the full displacement representation, the representation of rigid translations, and the representation of rigid rotations.

This decomposition is the basis of normal-mode classification in molecular and cluster problems. The irreducible decomposition of Γ_{vib} determines the symmetry labels of the normal modes, their degeneracies, and, once the transformation behavior of the dipole vector or polarizability tensor is known, whether they are infrared- or Raman-active [33, 35] Infrared-active modes couple to the electric dipole moment, whereas Raman-active modes modulate the polarizability tensor.

For crystals, translational periodicity changes the classification because the normal modes are labeled not only by their transformation under a point group but also by a wave vector $\mathbf{q}$ in the Brillouin zone, which specifies the phase relation of the periodic displacement pattern from cell to cell. The Bloch form used below makes this explicit: neighboring cells differ by the phase factor e^{iqR_l}. At Γ ($q = 0$), one recovers the familiar decomposition into irreducible representations of the crystal point group or, more precisely, the little group of Γ. Away from Γ, the symmetry is controlled by the little group of $\mathbf{q}$ (the subgroup that leaves that particular wave vector invariant up to a reciprocal lattice vector). Acoustic branches are the long-wavelength continuation of the three translational degrees of freedom of the crystal, so in a three-dimensional crystal with N atoms in the primitive-cell atomic there are three acoustic branches whose frequencies vanish as $\boldsymbol{q} \to 0$, whereas remaining $3N - 3$ branches are optical branches, corresponding to internal relative motion of atoms within that primitive-cell basis [34, 37].

A convenient methodological starting point is the Born–Oppenheimer approximation, in which the electrons are assumed to adjust essentially instantaneously to the slower ionic motion. The nuclei therefore move on the electronic ground-state potential-energy surface. Near an equilibrium structure, this surface can be expanded in atomic displacements up to second order, yielding the harmonic approximation. In that limit, lattice dynamics reduces to a set of coupled harmonic oscillators.

Let $u_{l\kappa\alpha}$ be the displacement of atom κ in cell l along Cartesian direction α, measured from equilibrium. Expanding the total energy around the equilibrium gives

$$E \approx E_0 + \frac{1}{2}\sum_{l\kappa\alpha}\sum_{l'\kappa'\beta} \Phi_{l\kappa\alpha,l'\kappa'\beta}\, u_{l\kappa\alpha}u_{l'\kappa'\beta} \tag{2.7}$$

where $\Phi_{l\kappa\alpha,l'\kappa'\beta}$ are the harmonic force constants, i.e. the second derivatives of the energy with respect to atomic displacements. The linear term vanishes because the expansion is taken around an equilibrium position [38].

Newton's equations are then

$$M_\kappa \ddot{u}_{l\kappa\alpha} = -\sum_{l'\kappa'\beta} \Phi_{l\kappa\alpha,l'\kappa'\beta}\, u_{l'\kappa'\beta}. \tag{2.8}$$

For a periodic crystal, one seeks wave-like solutions of Bloch form,

$$u_{l\kappa\alpha}(t) = \frac{1}{\sqrt{M_\kappa}}\, e_{\kappa\alpha}(\mathbf{q}\nu)\, e^{i(\mathbf{q}\cdot\mathbf{R}_l - \omega_{\mathbf{q}\nu}t)}, \tag{2.9}$$

where $\mathbf{q}$ is the phonon wave vector, ν is the branch index, $e_{\kappa\alpha}$ is the polarization vector, and $\omega_{\mathbf{q}\nu}$ is the angular frequency. Substituting this ansatz into the equations of motion yields the eigenvalue problem of lattice dynamics,

$$\sum_{\kappa'\beta} D_{\kappa\alpha,\kappa'\beta}\,(\mathbf{q})\, e_{\kappa'\beta}(\mathbf{q}\nu) = \omega_{\mathbf{q}\nu}^2\, e_{\kappa\alpha}(\mathbf{q}\nu), \tag{2.10}$$

with the dynamical matrix

$$D_{\kappa\alpha,\kappa'\beta}(\mathbf{q}) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}}\sum_{l'} \Phi_{0\kappa\alpha,l'\kappa'\beta}\, e^{i\mathbf{q}\cdot(\mathbf{R}_{l'}+\boldsymbol{\tau}_{\kappa'}-\boldsymbol{\tau}_\kappa)}. \tag{2.11}$$

Thus, phonon frequencies are obtained as the square roots of the dynamical-matrix eigenvalues, and the eigenvectors describe the atomic displacement patterns of the modes. In practical first-principles work, the force constants entering $D(\mathbf{q})$ are computed either by finite-displacement ("frozen-phonon") supercell calculations or by linear-response / density-functional perturbation theory.

The set of frequencies $\omega_{\mathbf{q}\nu}$ plotted as a function of the wave vector $\mathbf{q}$ is the phonon dispersion relation. It shows how vibrational energy varies with wavelength and propagation direction in the crystal. Because $\mathbf{q}$ is defined in reciprocal space, dispersions are usually shown along selected high-symmetry paths in the first Brillouin zone. An example of the phonon dispersion relation of diamond-structure Si is shown in Figure 2.2 while the representative displacement patterns for TA, TO, LOA, and LTO modes are shown in Figure 2.3. In this notation, L means longitudinal, T transverse, A acoustic, and O optical; LTO denotes a longitudinal/transverse optical mode at the zone center, while LOA denotes a longitudinal branch with mixed optical/acoustic character along the plotted path [39].

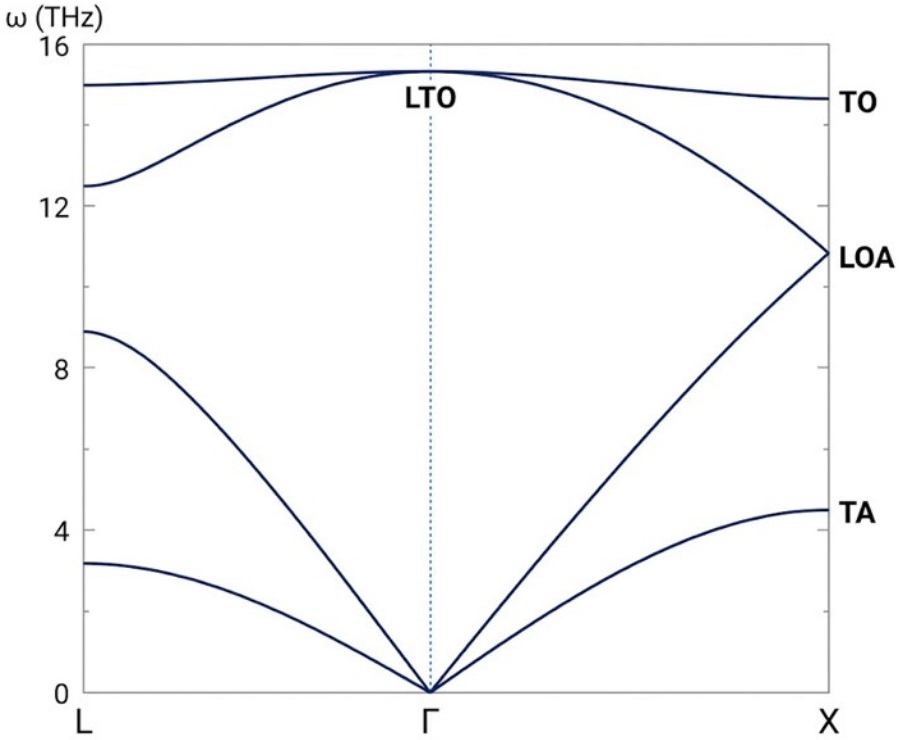

Figure 2.2. The phonon spectrum of Si along the chosen high-symmetry path (adapted from E. Kaxiras [39]).

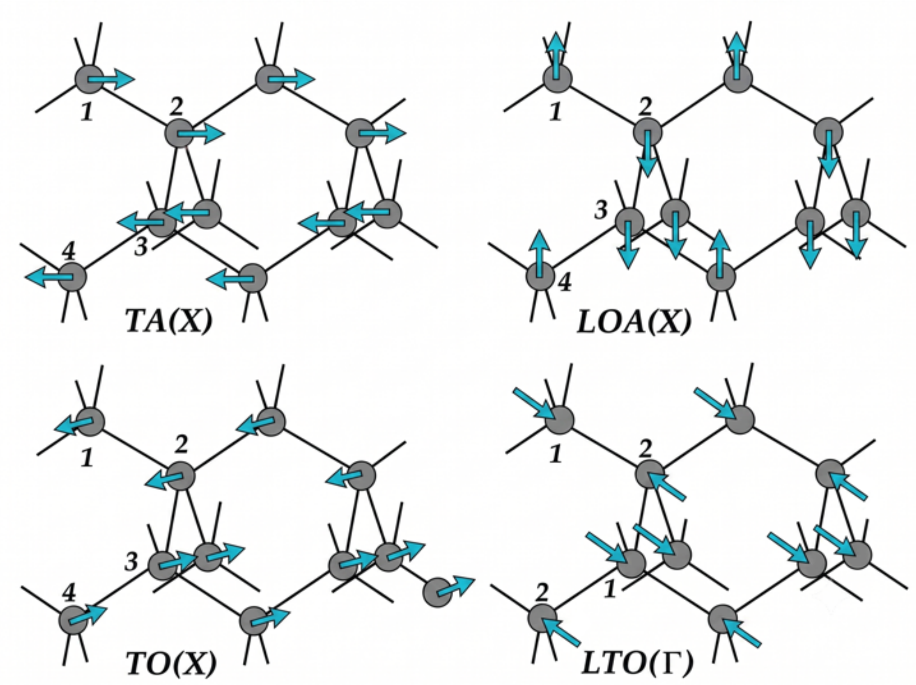

Figure 2.3. Representative atomic displacement patterns for selected phonon modes of diamond-structure Si at the Γ and X points in the Brillouin zone (adapted from E. Kaxiras [39]).

2.2 Density Functional Theory

2.2.1 Density as the Basic Variable

Within the Born–Oppenheimer approximation, the nuclei are regarded as fixed parameters while the electronic problem is solved in the external potential generated by those nuclei and by any external fields [37]. The electronic Hamiltonian in atomic units ($\hbar = m_e = e = 4\pi\varepsilon_0 = 1$) can be written as

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{ee}, \tag{2.12}$$

$$\hat{T} = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2, \tag{2.13}$$

$$\hat{V}_{\text{ext}} = \sum_{i=1}^{N} v_{\text{ext}}(\mathbf{r}_i), \tag{2.14}$$

$$\hat{V}_{ee} = \frac{1}{2}\sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.15}$$

where $\hat{T}$ is the many-electron kinetic-energy operator, $\hat{V}_{\text{ext}}$ is the sum of one-electron external potentials $v_{\text{ext}}(\mathbf{r}_i)$ which contain the electron–ion interaction (and any external fields) of the i-th electron, and $\hat{V}_{ee}$ is the electron-electron Coulomb repulsion. The ion–ion interaction $\hat{V}_{\text{II}}$ can be omitted from the electronic Hamiltonian because, for fixed nuclear positions (under Born–Oppenheimer approximation), it only results in a constant shift E_{II} in the total energy. However, after the electronic problem is solved, the full total energy can be obtained as $E = \langle \Psi | \hat{H} | \Psi \rangle + E_{II}$ [37].

The exact ground state minimizes the variational energy functional

$$E_0^{\text{el}} = \min_{\Psi} \langle \Psi | \hat{H} | \Psi \rangle, \tag{2.16}$$

over antisymmetric, normalized N-electron wavefunctions $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. The corresponding one-particle density is

$$n(\mathbf{r}) = N \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \, |\Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)|^2, \tag{2.17}$$

which is a scalar field in ordinary three-dimensional space [37]. Density functional theory (DFT) is founded on the statement that the exact ground-state density determines the exact ground-state properties, so that the central variable is $n(\mathbf{r})$ rather than the full many-body wavefunction [37, 40]. The ground-state formalism was extended to finite-temperature equilibrium [41], thereby providing the finite-temperature generalization of DFT [37].

2.2.2 Hohenberg–Kohn Theorems

The first Hohenberg–Kohn (HK) theorem states that for a nondegenerate ground state, the ground-state electron density $n_0(\mathbf{r})$ uniquely determines the external potential $V_{\text{ext}}(\mathbf{r})$ up to an additive constant, and therefore determines the Hamiltonian and all ground-state observables [37, 40].

A standard proof proceeds by contradiction [37, 40]. Let us assume that two different external potentials $V_{\text{ext}}^{(1)}$ and $V_{\text{ext}}^{(2)}$ define two different Hamiltonians $\widehat{H}^{(1)}$ and $\widehat{H}^{(2)}$, with ground states Ψ_1 and Ψ_2, but with the same ground-state density $n_0(\mathbf{r})$. Since Ψ_2 is not the ground state of $\widehat{H}^{(1)}$, the variational principle gives

$$E^{(1)} = \langle \Psi^{(1)} | \widehat{H}^{(1)} | \Psi^{(1)} \rangle < \langle \Psi^{(2)} | \widehat{H}^{(1)} | \Psi^{(2)} \rangle. \tag{2.18}$$

Using the relation $\widehat{H}^{(1)} = \widehat{H}^{(2)} + \sum_i \left[v_{\text{ext}}^{(1)}(\mathbf{r}_i) - v_{\text{ext}}^{(2)}(\mathbf{r}_i) \right]$ and the fact both states yield the same density n_0 we obtain

$$E^{(1)} < E^{(2)} + \int d^3r \left[v_{\text{ext}}^{(1)}(\mathbf{r}) - v_{\text{ext}}^{(2)}(\mathbf{r}) \right] n_0(\mathbf{r}), \tag{2.19}$$

and, vice-versa (by symmetry),

$$E^{(2)} < E^{(1)} + \int d^3r \left[v_{\text{ext}}^{(2)}(\mathbf{r}) - v_{\text{ext}}^{(1)}(\mathbf{r}) \right] n_0(\mathbf{r}). \tag{2.20}$$

Adding those two inequalities leads to the contradiction $E^{(1)} + E^{(2)} < E^{(1)} + E^{(2)}$. Therefore, the assumption is false and the mapping $n_0 \mapsto v_{\text{ext}}$ from the ground-state density to the external potential is one-to-one (up to an additive constant) [37, 40].

The second Hohenberg-Kohn theorem constructs an energy functional and establishes a variational principle over densities. It can be written that the total-energy functional

$$E_{\text{HK}}[n] = F_{\text{HK}}[n] + \int d^3r \, v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + E_{II}, \tag{2.21}$$

$$F_{\text{HK}}[n] = T[n] + E_{\text{int}}[n], \tag{2.22}$$

where functional $F_{\text{HK}}[n]$ is universal, i.e., independent of particular external potential v_{ext} of the system. Then the exact ground-state energy E_0 is the global minimum of $E_{\text{HK}}[n]$ over all densities with fixed particle number $\int n = N$, and it is minimized at the exact ground-state electron density $n_0(\boldsymbol{r})$ [37, 40].

In the original Hohenberg-Kohn formulation, the functional is defined on densities that are ground-state densities of some external potential (V-representable densities). The Levy–Lieb (LL) constrained-search formulation removes this restriction by defining the universal functional for any N-representable density (any density $n(\mathbf{r})$ that can be generated by at least one N-electron wavefunction, written as $\Psi \to n$), [37, 42, 43]

$$F_{\text{LL}}[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle, \tag{2.23}$$

and then minimizing to obtain the ground state energy

$$E_0 = \min_n \left\{ F_{\text{LL}}[n] + \int d^3r \, v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + E_{II} \right\}. \tag{2.24}$$

This formulation is also valid in the presence of degeneracies and provides the modern mathematical foundation of ground-state DFT [37, 42, 43].

At finite temperature, Mermin generalized the formalism by replacing the ground-state energy functional with a thermodynamic free-energy (or grand-potential) functional whose minimum yields the equilibrium density [37, 41]. This extension provides the foundation of the finite-temperature DFT, which becomes important when thermal electronic excitations are non-negligible.

2.2.3 Kohn–Sham Equations

While Hohenberg-Kohn and Levy-Lieb theorems establish the existence of an exact density functional $E[n]$ (or $F[n]$), they do not provide a direct practical way to construct it. The practical solution is obtained by using the Kohn–Sham (KS) auxiliary system, which replaces the interacting many-electron problem with an auxiliary noninteracting one that reproduces the same ground-state electron density [37, 44].

The auxiliary single-particle Hamiltonian for spin σ can be written as

$$\hat{H}^{\sigma}_{\text{aux}} = -\frac{1}{2}\nabla^2 + \hat{V}^{\sigma}(\mathbf{r}), \tag{2.25}$$

and its ground state is a Slater determinant built from orthonormal single-particle orbitals ψ_i^{σ} [37, 44]. The electron density is

$$n(\mathbf{r}) = \sum_{\sigma} n_{\sigma}(\mathbf{r}) = \sum_{\sigma}\sum_{i=1}^{N_\sigma} |\psi_i^{\sigma}(\mathbf{r})|^2, \tag{2.26}$$

and the auxiliary (noninteracting, single-particle) kinetic energy term can be written as

$$T_s[n] = -\frac{1}{2}\sum_{\sigma}\sum_{i=1}^{N_\sigma}\langle \psi_i^{\sigma}|\nabla^2|\psi_i^{\sigma}\rangle = \frac{1}{2}\sum_{\sigma}\sum_{i=1}^{N_\sigma}\int d^3r\, |\nabla\psi_i^{\sigma}(\mathbf{r})|^2 \,. \tag{2.27}$$

where i labels the occupied KS orbital $\psi_i^{\sigma}(\mathbf{r})$, σ is the spin index, and N_σ is the number of electrons with spin σ [37, 44].

The Hartree (Coulomb) energy of the electron density $n(\mathbf{r})$ is

$$E_H[n] = \frac{1}{2}\int d^3r\, d^3r'\, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.28}$$

The Kohn–Sham total-energy functional can then be written as

$$E_{\text{KS}}[n] = T_s[n] + \int d^3r\, v_{\text{ext}}(\mathbf{r})n(\mathbf{r}) + E_H[n] + E_{II} + E_{xc}[n]. \tag{2.29}$$

The exchange–correlation (xc) term can be written in terms of the Hohenberg-Kohn functional as

$$E_{xc}[n] = F_{\text{HK}}[n] - (T_s[n] + E_H[n]) \tag{2.30}$$

$$= \langle\hat{T}\rangle - T_s[n] + \langle\hat{V}_{ee}\rangle - E_H[n], \tag{2.31}$$

where $F_{\text{HK}}[n]$ is the exact universal functional and the expectation values refer to the interacting ground state that yields the density $n(\mathbf{r})$ [37, 44]. Therefore, the exchange-correlation functional $E_{xc}[n]$ contains both the many-body correction to the kinetic energy and the nonclassical part of the interaction energy.

Minimization of $E_{\mathrm{KS}}[n]$ with respect to the orbitals, subject to the orthonormality constraints $\langle\psi_i^\sigma|\psi_j^{\sigma'}\rangle = \delta_{ij}\delta_{\sigma\sigma'}$, leads to the Kohn–Sham equations [37, 44]:

$$\left(\hat{H}_{\mathrm{KS}}^\sigma - \varepsilon_i^\sigma\right)\psi_i^\sigma(\mathbf{r}) = 0, \tag{2.32}$$

$$\hat{H}_{\mathrm{KS}}^\sigma = -\frac{1}{2}\nabla^2 + V_{\mathrm{KS}}^\sigma(\mathbf{r}), \tag{2.33}$$

where the effective Kohn–Sham potential is

$$V_{KS}^\sigma(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}^\sigma(\mathbf{r}), \tag{2.34}$$

$$V_H(\mathbf{r}) = \int d^3r' \, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{2.35}$$

$$V_{xc}^\sigma(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n_\sigma(\mathbf{r})}. \tag{2.36}$$

The Hartree term $V_H(\mathbf{r})$ is the classical electrostatic potential of the full electron density, whereas $V_{xc}^\sigma(\mathbf{r})$ is the exchange-correlation potential obtained as the functional derivative of E_{xc} with respect to the spin density $n(\mathbf{r},\sigma)$. The equations are coupled because V_H and V_{xc}^σ depend on the density, while the density itself is built from the Kohn–Sham orbitals [37, 44]. In practice the problem is solved self-consistently: an input density is used to construct V_{KS}, the one-particle equations are solved, a new density is formed, and the procedure is iterated until the convergence is achieved [37, 45].

For magnetic systems, one allows different electron spin densities, $n_\uparrow(\mathbf{r}) \neq n_\downarrow(\mathbf{r})$, so that the Kohn–Sham problem becomes explicitly spin dependent. Spin–orbit coupling can be incorporated by adding the appropriate one-electron relativistic term to the Kohn–Sham Hamiltonian without changing the overall self-consistent structure of the method [37]. These extensions are essential for practical calculations of magnetic compounds such as Ce_3O_3N and HoFSe, which are discussed later in the thesis.

Although the Kohn–Sham eigenvalues ε_i arise as Lagrange multipliers, only selected ones have a strict physical interpretation [37, 46]. For the exact functional in a finite system, the highest occupied eigenvalue equals minus the first ionization energy, but the full Kohn–Sham spectrum is not, in general, an exact excitation spectrum [37, 46]. In solids, the fundamental band gap differs from the Kohn–Sham eigenvalue gap by the derivative discontinuity of the exact exchange–correlation potential, and approximate local or semilocal functionals often underestimate the gap even further [37, 46], which motivates the use of generalized KS schemes (hybrid and meta-GGA functionals) and many-body corrections beyond DFT.

2.2.4 Exchange–Correlation Functionals

All practical accuracy in ground-state DFT depends on the approximation chosen for the exchange-correlation functional $E_{xc}[n]$. It is convenient to express it as an integral over an exchange–correlation energy per particle ε_{xc},

$$E_{xc}[n] = \int d^3r \, n(\mathbf{r}) \, \varepsilon_{xc}([n],\mathbf{r}). \tag{2.37}$$

The interaction component of ε_{xc} can be written in terms of the exchange–correlation hole $n_{xc}(\mathbf{r},\mathbf{r}')$ (which measures how the conditional probability of finding a second electron at $\mathbf{r}'$ is reduced around an electron fixed at $\mathbf{r}$ relative to the uncorrelated Hartree picture [37]),

$$\varepsilon_{xc}^{\text{int}}([n],\mathbf{r}) = \frac{1}{2}\int d^3r' \, \frac{n_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.38}$$

Both the Pauli exchange and the Coulomb interactions reduce the probability of finding another electron near a reference electron. Since the total electron density integrates to the total number of electrons, whereas the conditional density integrates to the total number of electrons minus one, because the reference electron at r is already fixed, the difference satisfies the exact sum rule

$$\int d^3r' \, n_{xc}(\mathbf{r},\mathbf{r}') = -1. \tag{2.39}$$

A particularly important exact relation is the adiabatic-connection formula [37]:

$$E_{xc}[n] = \int_0^1 d\lambda \left\langle \Psi^\lambda \left| \frac{d\hat{V}_{ee}^\lambda}{d\lambda} \right| \Psi^\lambda \right\rangle - E_H[n] = \frac{1}{2}\int d^3r \, n(\mathbf{r}) \int d^3r' \frac{\bar{n}_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{2.40}$$

with

$$\bar{n}_{xc}(\mathbf{r},\mathbf{r}') = \int_0^1 d\lambda \, n_{xc}^\lambda(\mathbf{r},\mathbf{r}'). \tag{2.41}$$

The adiabatic connection refers to the continuous increase in the interaction strength from the noninteracting Kohn–Sham system ($\lambda = 0$) to the fully interacting physical system ($\lambda = 1$), while the density is constrained to remain unchanged throughout the connection [37].

The local spin-density approximation (LSDA) assumes that the exchange–correlation energy density at each point is the same as in a homogeneous electron gas with the same local spin densities [37, 44]:

$$E_{xc}^{\text{LSDA}}[n_\uparrow, n_\downarrow] = \int d^3r \, n(\mathbf{r}) \, \varepsilon_{xc}^{\text{hom}}\big(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})\big). \tag{2.42}$$

Here $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$ are the local spin-up and spin-down densities, $n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r})$ is the total density, and $\varepsilon_{xc}^{\text{hom}}$ is the exchange–correlation energy per particle of the homogeneous electron gas evaluated at the local spin density values [37].

Accurate correlation energies for the homogeneous electron gas were obtained from Quantum Monte Carlo (QMC) calculations [47], and practical LDA parametrizations were subsequently constructed using these many-body benchmark data [37, 48].

Generalized-gradient approximations (GGAs) improve on LSDA by allowing the functional to depend on density gradients [37, 49]:

$$E_{xc}^{\text{GGA}}[n_\uparrow, n_\downarrow] = \int d^3r \, n(\mathbf{r}) \, \epsilon_{xc}(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow, \dots). \tag{2.43}$$

For exchange, many GGAs are conveniently written in enhancement-factor form [37]:

$$E_x^{\text{GGA}}[n] = \int d^3r \, n(\mathbf{r}) \, \varepsilon_x^{\text{hom}}(n) \, F_x(s), \qquad s = \frac{|\nabla n|}{2k_F n} = \frac{|\nabla n|}{2(3\pi^2)^{1/3} n^{4/3}}, \tag{2.44}$$

where $F_x(s)$ is the exchange enhancement factor and s is the dimensionless reduced gradient [37]. This form is used, for example, in PBE [49] and also underlies the interpretation of earlier gradient-corrected exchange functionals such as B88 [50].

The quantity entering the Kohn–Sham equations is always the functional derivative $V_{xc}^{\sigma}(\mathbf{r}) = \delta E_{xc}/\delta n_{\sigma}(\mathbf{r})$ [37]. For the LSDA, one obtains

$$V_{xc}^{\sigma}(\mathbf{r}) = \left[\varepsilon_{xc}^{\mathrm{hom}} + n\,\frac{\partial \varepsilon_{xc}^{\mathrm{hom}}}{\partial n_{\sigma}}\right]_{\mathbf{r},\sigma}, \tag{2.45}$$

whereas for a GGA, the explicit gradient dependence leads, after integration by parts, to

$$V_{xc}^{\sigma}(\mathbf{r}) = \left[\varepsilon_{xc} + n\,\frac{\partial \varepsilon_{xc}}{\partial n_{\sigma}} - \nabla\cdot\left(n\,\frac{\partial \varepsilon_{xc}}{\partial \nabla n_{\sigma}}\right)\right]_{\mathbf{r},\sigma}. \tag{2.46}$$

Beyond LSDA and GGA, there exist other approaches such as: meta-GGA functionals, which add dependence on kinetic-energy density; hybrid functionals, which mix a fraction of Hartree–Fock exchange in a generalized Kohn-Sham framework; and orbital-dependent corrections such as DFT+U for localized d- and f-electrons [37]. These extensions are directly relevant for this thesis because semilocal, hybrid, and DFT+U calculations are used to compare crystal structures, energies, and electronic band structures.

2.2.5 Basis Sets and Pseudopotential Method

To solve the Kohn–Sham equations numerically, the orbitals must be represented in a finite basis. In a periodic supercell of volume Ω, a complete orthonormal plane-wave basis is given by $\Omega^{-1/2}e^{i\mathbf{q}\cdot\mathbf{r}}$, where $\mathbf{q}$ is a reciprocal-lattice vector [37, 45]. An orbital is expanded as

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{q}} c_{i\mathbf{q}}\,\frac{1}{\sqrt{\Omega}}e^{i\mathbf{q}\cdot\mathbf{r}}, \qquad \langle\mathbf{q}'|\mathbf{q}\rangle = \delta_{\mathbf{q}'\mathbf{q}}. \tag{2.47}$$

Projection onto this basis gives an eigenvalue problem in reciprocal space. The kinetic operator is diagonal,

$$\langle\mathbf{q}'\left|-\frac{1}{2}\nabla^2\right|\mathbf{q}\rangle = \frac{q^2}{2}\,\delta_{\mathbf{q}'\mathbf{q}}, \tag{2.48}$$

whereas the potential couples different plane waves through its Fourier components [37, 45]. In practice, the basis is truncated by an energy cutoff $E_{\mathrm{cut}} = 1/2\,q_{\mathrm{max}}^2$ (only the wave vectors with $|\mathbf{q}| \leq q_{\mathrm{max}}$ are chosen for the basis), which makes the method systematically improvable by simply increasing E_{cut} [37, 45].

Localized basis functions, including Gaussian-type orbitals and numerical atomic orbitals, are often advantageous for molecules, low-symmetry systems, and all-electron treatments near nuclei. Augmented-plane-wave methods combine extended functions in the interstitial region with atom-centered radial functions multiplied by spherical harmonics within the augmentation spheres [37]. In the DFT calculations presented in this thesis, CRYSTAL [51, 52] uses localized Gaussian basis functions, whereas VASP [53-55] uses a plane-wave basis together with PAW datasets.

Plane waves are inefficient for calculating all-electron Coulomb potentials because the strong nuclear attraction generates rapid oscillations and near-nuclear cusps in the valence wavefunctions [37, 45]. This motivates the use of pseudopotentials, which replace the ionic potential and frozen-core effects with an effective ionic potential designed to reproduce the valence-scattering properties [37, 56].

In the standard semilocal representation, the pseudopotential operator is

$$\hat{V}_{SL} = \sum_{lm} |Y_{lm}\rangle\, V_l(r)\, \langle Y_{lm}|\,, \tag{2.49}$$

where Y_{lm} are spherical harmonics, l and m are the angular-momentum (orbital and magnetic) quantum numbers, and $V_l(r)$ is the radial pseudopotential for the orbital quantum number l. This form is called semilocal because it is local in the radial coordinate r for a fixed angular momentum number l, but it acts differently on different angular-momentum components of the wavefunction and is therefore nonlocal in the full three-dimensional sense [37].

For norm-conserving pseudopotentials, one chooses a reference atomic configuration and imposes conditions that reproduce the corresponding valence scattering properties [37, 56, 57]. The most important conditions are the equality of reference eigenvalues,

$$\varepsilon_l^{PS} = \varepsilon_l^{AE}, \tag{2.50}$$

the matching of radial functions outside the core radius r_c,

$$\phi_l^{PS}(r) = \phi_l^{AE}(r), \qquad r \geq r_c, \tag{2.51}$$

and norm conservation inside the core,

$$\int_0^{r_c} d\, r\, r^2 \left|\phi_l^{PS}(r)\right|^2 = \int_0^{r_c} d\, r\, r^2 \left|\phi_l^{AE}(r)\right|^2. \tag{2.52}$$

Here, ϕ_l^{PS} and ϕ_l^{AE} are the pseudo and all-electron radial partial waves for angular momentum l [37, 56, 57]. These conditions ensure that the pseudo-orbital has the correct valence behavior outside the core region and that the integrated charge inside the core is preserved [37, 56]. Norm conservation also implies matching of the first energy derivative of the logarithmic derivative at r_c, which is the key reason for good first-order transferability around the reference energy [37, 56].

The practical purpose of these conditions is transferability: a pseudopotential generated in an isolated atom should still reproduce valence scattering reasonably well in ions, molecules, and solids [37, 56]. In many plane-wave calculations, however, ultrasoft pseudopotentials and PAW datasets are often preferred because they achieve a given accuracy with lower cutoffs, especially for first-row and transition-metal elements [37, 53, 58, 59]. Norm-conserving pseudopotentials nevertheless remain useful when a simpler operator form, cleaner response-function formalism, or tighter control of transferability is desired [37, 57].

For computational efficiency, the semilocal pseudopotential is often transformed into a separable Kleinman–Bylander form [37, 60]:

$$\hat{V}_{NL} = V_{\text{local}}(r) + \sum_{ij} |\beta_i\rangle D_{ij} \langle\beta_j|. \tag{2.53}$$

The projector functions β_i are localized inside the core region, and the coefficients D_{ij} are determined from the underlying semilocal potential [37, 60]. In this form, the action of the pseudopotential is reduced to overlaps with a small set of projectors, which is much more efficient than applying the full semilocal operator directly in plane-wave implementations [37, 45].

Ultrasoft pseudopotentials go one step further by relaxing norm conservation and introducing a generalized overlap operator [37, 58]. The resulting one-particle problem has the form

$$\hat{H}\, |\tilde{\psi}_n\rangle = \varepsilon_n \hat{S}\, |\tilde{\psi}_n\rangle, \qquad \hat{S} = 1 + \sum_{ij} |\beta_i\rangle q_{ij} \langle\beta_j|. \tag{2.54}$$

Here $|\tilde{\psi}_n\rangle$ is a smooth pseudo-wavefunction, ε_n is the Kohn–Sham eigenvalue, and $\hat{S}$ differs from the identity because the pseudo partial waves are no longer norm conserving. This relaxation of norm conservation permits smoother pseudo-wavefunctions with fewer Fourier components and allows ultrasoft methods to be used with substantially lower plane-wave cutoffs than conventional norm-conserving pseudopotentials [37, 58].

In the projector augmented-wave (PAW) method, the smooth auxiliary state is related to the reconstructed all-electron state by a linear transformation [37, 53, 59]:

$$\psi_n(\mathbf{r}) = \hat{T}\,\tilde{\psi}_n(\mathbf{r}). \tag{2.55}$$

Here $\tilde{\psi}_n(\mathbf{r})$ is the smooth auxiliary wavefunction, $\psi_n(\mathbf{r})$ is the reconstructed all-electron wavefunction, and $\hat{T}$ denotes the augmentation transformation built from smooth and all-electron partial waves together with localized projector functions [37, 53, 59]. The PAW method is therefore computationally efficient because it uses smooth auxiliary functions, while retaining access to reconstructed all-electron valence wavefunctions and on-site charge densities [37, 53, 59]. The ultrasoft pseudopotentials can be viewed formally as a linearized approximation to a PAW-like functional [53], which makes the two methods closely related.

Finally, practical DFT calculations involve two distinct convergence problems: self-consistency of the Kohn–Sham potential, and discretization errors associated with basis size, real-space grids, and Brillouin-zone sampling [37, 45]. Plane-wave methods calculations are converged by increasing the cutoff energy and refining the k-point mesh [37, 45], whereas localized-basis methods must be converged by improving the basis itself because there is no single universal cutoff parameter.

2.3 Magnetism in Solids: Concepts for First-Principles Calculations

Magnetism in crystalline solids encompasses both linear field-induced responses and collective effects characterized by spontaneous long-range order. The stability and orientation of an ordered magnetic state are governed by exchange interactions, spin–orbit coupling (SOC), crystal-field splitting, and the symmetry of the local coordination environment. Rare-earth materials are particularly useful for examining these phenomena because their partially filled $4f$ shells often give rise to large localized magnetic moments and strong single-ion anisotropy [16, 24, 61]. The scope of this section is limited to the introduction of the magnetic concepts needed to interpret spin-polarized, DFT+U, and spin-orbit-coupled first-principles calculations of the ground-state magnetic ordering in HoFSe and Ce_3O_3N crystal systems.

2.3.1 Introduction to Magnetism and Ordered States

The magnetic moment $\boldsymbol{\mu}$ is the microscopic magnetic dipole moment associated with a specified electronic state, atomic shell, atom or ion. The corresponding macroscopic quantity is the magnetization $\mathbf{M}$, defined as the magnetic moment per unit volume,

$$\mathbf{M} = \frac{1}{V}\sum_i \boldsymbol{\mu}_i. \tag{2.56}$$

Here V is the volume and $\boldsymbol{\mu}_i$ are the microscopic magnetic moments associated with the atoms, ions, or electronic states included in that volume. The magnetic susceptibility is, in general, a tensor that measures the linear response of the magnetization to an applied magnetic field $\mathbf{H}$,

$$\chi_{\alpha\beta} = \left(\frac{\partial M_\alpha}{\partial H_\beta}\right)_{\mathbf{H}\to 0}, \tag{2.57}$$

where α and β denote Cartesian components. Long-range magnetic ordering develops below a certain critical temperature, called the Curie temperature T_C for ferromagnets and the Néel temperature T_N for antiferromagnets [62]. Diamagnetism and paramagnetism are magnetic responses without long-range magnetic order. Ferromagnetism corresponds to a state with parallel moments and a nonzero spontaneous magnetization, antiferromagnetic materials have ordered sublattices with opposite moments and zero net magnetization, while ferrimagnetic materials also contain oppositely moment-oriented sublattices, but their moments do not cancel completely [62].

A local moment is associated with a given atomic site or localized electronic shell, whereas magnetic order specifies how such moments are arranged throughout the unit cell and the whole crystal lattice. In order to determine the magnetic ordering of a crystal structure in *ab initio* calculations, different initial magnetic moments are assigned, but the final local moments and total energy are determined self-consistently. The same crystal structure can therefore be tested in ferromagnetic (FM), antiferromagnetic (AFM), ferrimagnetic (FiM), or various non-collinear arrangements, and the lowest total energy determines the preferred magnetic arrangement.

2.3.2 Local Moments, Exchange, and Itinerant Magnetism

The electronic origin of magnetism in solids is the angular momentum of electrons, where the spin angular momentum always contributes to the electronic magnetic moment, while the orbital angular momentum can also contribute when the crystal environment does not reduce it strongly. In many $3d$ transition-metal compounds, the crystal field lifts the degeneracy of the d orbitals and forms combinations for which the expectation value of the orbital angular momentum is small; the orbital contribution is then often described as strongly reduced, and the magnetism is approximated mainly in terms of spin. In rare-earth compounds, this spin-only approximation is less appropriate because the $4f$ states are spatially localized and shielded by the outer $5s$ and $5p$ shells, so the orbital contribution remains an important part of their magnetic description [61]. The electron quantum numbers used to label shells, subshells, orbital projections, and spin projections are given in Table 2.1.

The interaction that favors parallel or antiparallel alignment of moments is called exchange. In a localized-moment model, it is commonly represented by the Heisenberg exchange Hamiltonian,

$$\mathcal{H}_{\text{ex}} = -\sum_{i<j} J_{ij}\, \mathbf{S}_i \cdot \mathbf{S}_j. \tag{2.58}$$

Here $\mathcal{H}_{\text{ex}}$ is the exchange part of the Hamiltonian, $\mathbf{S}_i$ and $\mathbf{S}_j$ are spin operators or classical spin vectors on magnetic sites i and j, and J_{ij} is the exchange coupling parameter between these sites. The notation $i < j$ indicates that each pair is counted only once. With the sign convention used here, positive J_{ij} favors parallel alignment and negative J_{ij} favors antiparallel alignment. Other sign conventions are also used in the literature, so the Hamiltonian definition must always be stated explicitly.

Table 2.1. Electronic quantum numbers.

Quantum number	Symbol	Values	Description
Principal	n	1, 2, 3, ...	Labels the shell.
Orbital / azimuthal / angular momentum	l	0, 1, ..., n - 1	Labels the subshell; s, p, d, f correspond to l = 0, 1, 2, 3, respectively
Magnetic	m_l	$-l$, ..., 0, ..., $+l$	Labels orbital orientation within a subshell.
Spin	m_s	+1/2, -1/2	Labels the spin projection of the electron.

A useful distinction in solid-state magnetism is between localized-moment and itinerant magnetism [62]: in localized-moment systems, the electrons carrying most of the moment remain associated mainly with particular atoms or ions, while in itinerant systems, magnetism originates from spin-polarized Bloch bands, so the same electronic states can participate in bonding, transport, and magnetic polarization. Rare-earth magnetism is usually closer to the localized limit because the $4f$ electrons are spatially localized, whereas the more extended $5d$ and $6s$ states participate more strongly in bonding and can mediate interactions between $4f$ moments [61]. This contrasts with many $3d$ magnetic materials, where $3d$ states are more spatially extended and often hybridize substantially with ligand $2p$ states; in such systems, the boundary between local-moment and itinerant descriptions is often less sharp [37, 62].

2.3.3 Spin, Orbital Moment, Spin–Orbit Coupling, and Magnetic Anisotropy

For rare-earth ions, the angular momenta relevant for magnetism are the orbital angular momentum **L**, the spin angular momentum **S**, and the total angular momentum **J**. Hund's rules determine, within Russell–Saunders coupling (i.e. LS coupling) and for the free-ion configuration, the ground multiplet: the total spin S is maximized first, then the orbital angular momentum L is maximized for that spin state, and finally $J = |L - S|$ is selected for shells that are less than half filled, whereas $J = L + S$ is selected for shells that are more than half filled. These rules are especially useful for lanthanides because progressive filling of the $4f$ shell produces strongly element-dependent values of L, S, and J [61].

In rare-earth compounds, the moment is often discussed as an ionic moment because the local configuration is frequently approximated by a trivalent rare-earth ion, Ln^{3+}, with atomic-like $4f$ states. If **J** is treated as an angular-momentum operator (in atomic units), the magnetic moment operator in the ground-state multiplet is

$$\hat{\boldsymbol{\mu}} = -g_J \mu_B \hat{\mathbf{J}} \ , \tag{2.59}$$

where μ_B is the Bohr magneton and g_J is the Landé factor,

$$g_J = g_L \frac{J(J+1) - S(S+1) + L(L+1)}{2J(J+1)} + g_S \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} , \tag{2.60}$$

which for $g_L = 1$ and the approximation $g_S = 2$ reduces to

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \tag{2.61}$$

In the high-temperature Curie response, the effective paramagnetic moment is

$$\mu_{\mathrm{eff}} = g_J \mu_B \sqrt{J(J+1)}. \tag{2.62}$$

The orbital contribution is large in many rare-earth ions because their $4f$ states have orbital angular momentum number $l = 3$, and Hund's rules often lead to ground states with nonzero L. Then the moment is not purely spin-like, and spin-orbit coupling must be considered explicitly when anisotropy or the direction of the magnetic moment matters [61].

Spin-orbit coupling is the relativistic interaction that couples the electron spin to its orbital motion in an electric potential (equivalently, in the electron rest frame, the electric field associated with the atomic or crystal potential appears as an effective magnetic field that acts on the spin). In an atomic central-field description, the spin-orbit part of the Hamiltonian is commonly written as

$$\mathcal{H}_{\mathrm{SOC}} = \xi(r)\, \mathbf{L} \cdot \mathbf{S}, \tag{2.63}$$

where $\xi(r)$ is the radial spin-orbit coefficient. In a common convention for a central potential-energy function $V(r)$,

$$\xi(r) = \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{dV(r)}{dr}, \tag{2.64}$$

so the magnitude of SOC is controlled by the radial derivative of the effective potential and becomes larger near heavy nuclei [37]. In the Pauli form used for electrons in a general potential,

$$\mathcal{H}_{\mathrm{SO}} = -\frac{\hbar}{4m_e^2 c^2} \boldsymbol{\sigma} \cdot [\mathbf{p} \times \nabla V_0(\mathbf{r})], \tag{2.65}$$

where m_e is the electron mass, $\mathbf{p}$ is the momentum operator, $\boldsymbol{\sigma}$ is the vector of Pauli matrices, and $V_0(\mathbf{r})$ is the potential entering the one-electron Hamiltonian [63].

The crystal field is the electrostatic field produced by the surrounding ions and electrons. It lowers the spherical symmetry of an isolated ion and splits the degeneracy of the free-ion multiplet according to the local point symmetry. For rare earth ions, the crystal field Hamiltonian is often expressed using Stevens operator equivalents,

$$\mathcal{H}_{\mathrm{CEF}} = \sum_{k,q} B_k^q \; O_k^q(\hat{\mathbf{J}}). \tag{2.66}$$

Here B_k^q are crystal field parameters determined by the local coordination environment, electrostatics, and covalency corrections, and $O_k^q(\hat{\mathbf{J}})$ are Stevens operators acting within a given total angular momentum multiplet [61, 64].

Because the $4f$ shell is localized and shielded by the outer 5s and 5p shells, its orbital contribution is not suppressed as strongly as in many transition-metal compounds. The crystal field determines the preferred spatial orientation of the $4f$ charge distribution, while SOC couples this orbital orientation to the spin part of the moment. If the minimum energy occurs when the moment is aligned along a particular crystallographic direction, the material has an easy axis. Similarly, if the minimum energy corresponds to moment orientations within a crystallographic plane, the material has an easy plane. The resulting orientation dependence is quantified by the magnetocrystalline anisotropy energy (MAE),

$$E_{\mathrm{MAE}} = E(\mathbf{M} \parallel \hat{\mathbf{n}}_1) - E(\mathbf{M} \parallel \hat{\mathbf{n}}_2), \tag{2.67}$$

where $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are two chosen crystallographic directions [16], i.e., MAE measures the total-energy difference between magnetic orientations of the same crystal structure.

2.3.4 Magnetic Configurations in First-Principles Calculations

The magnetic arrangements used in first-principles calculations are model states designed to sample physically plausible orderings. Collinear calculations assume that all moments are parallel or antiparallel to a common spin quantization axis. Ferromagnetic and many antiferromagnetic configurations can be treated in this way. Non-collinear calculations allow the moment direction to vary from site to site, which is necessary to model canted structures, spin spirals, helical and cycloidal states, or any magnetic arrangement in which the moments cannot be represented by a single common axis. In practice, energy differences between magnetic configurations can be interpreted through models such as the Heisenberg model.

In spin-density functional theory, magnetism is introduced through the spin dependence of the electronic density. In collinear calculations, this is commonly represented by spin-up and spin-down densities, whereas non-collinear calculations use a spin-density matrix or local magnetization vector. Local magnetic moments are then obtained by integrating the spin density in atomic regions or projection spheres, while total magnetic moments are obtained from the cell-integrated spin magnetization [37, 65].

For localized d and f states, LDA or GGA functionals may underestimate the energetic separation between correlated localized orbitals and more delocalized valence or conduction bands, and they may also over-delocalize the correlated states. The DFT+U correction adds an orbital-dependent on-site interaction for selected localized states, improving the description of their relative energies and occupations [37]. In the commonly used Dudarev formulation, which is a simplified rotationally invariant DFT+U approach depending only on the effective parameter U_{eff} ,

$$E_U = \frac{U_{\text{eff}}}{2} \sum_{I,\sigma} \text{Tr}\,[\mathbf{n}^{I\sigma}(1 - \mathbf{n}^{I\sigma})], \tag{2.68}$$

where $U_{\text{eff}} = U - J$ (U and J are parameters in the full Hubbard formulation), and $\mathbf{n}^{I\sigma}$ is the occupation matrix for the correlated shell on site I and spin index σ, and the trace is over the localized orbitals [66]. This correction is often essential for a physically meaningful description of localized $3d$ and $4f$ states.

The practical workflow used in the thesis follows the hierarchy of magnetic energy scales. Spin-polarized calculations with collinear spins (without SOC) are used to compare candidate collinear FM and AFM arrangements and to identify the dominant exchange-driven preference. The SOC calculations are then used when the ground-state energy depends on the direction of the moment relative to the crystal lattice [67], i.e., when magnetocrystalline anisotropy or non-collinear behavior is relevant.

2.4. Thermodynamics, Equations of State and Pressure-Induced Phase Transitions

For a crystalline solid, the equation of state (EOS) expresses how the thermodynamic potential changes with volume or pressure. In electronic-structure calculations, EOS is usually chosen to be the total ground-state energy $E(V)$, because the volume can be imposed directly and the internal atomic coordinates can then be relaxed at that volume. Volume and pressure are among the most fundamental thermodynamic variables in the condensed matter: changing V modifies bonding distances, band widths, orbital overlap, and ultimately the relative stability of competing crystal structures [37].

At finite temperature T, the stable phase at given pressure p is determined by the Gibbs free energy G,

$$G(p,T) = H(p,T) - TS(p,T), \tag{2.69}$$

where H is the enthalpy and S is the entropy, whereas at fixed volume V, it is natural to use the Helmholtz free energy F,

$$F(V,T) = E(V,T) - TS(V,T). \tag{2.70}$$

In the zero-temperature static-lattice approximation used in most structure-prediction calculations, the entropic term is neglected and the stability criterion reduces to the enthalpy,

$$H = E + pV\,, \tag{2.71}$$

which is very useful for comparing the thermodynamic stability of polymorphs.

From the microscopic point of view, the EOS is the connection between the electronic ground state and the macroscopic mechanical response of the crystal. For each imposed volume, the electrons are solved in the Born–Oppenheimer picture and the nuclei are relaxed to the lowest-energy configuration compatible with that volume. Thermal and zero-point vibrational effects can be added later through phonon or quasiharmonic calculations, but in the static-lattice treatment used for most crystal-structure prediction studies, the resulting EOS and transition pressures should be interpreted as zero-temperature approximations [37].

2.4.1 Equation of State

In practice, the equation of state is obtained by computing total energies at cell volumes around the optimized equilibrium volume. For each selected volume, the atomic positions and the cell shape are relaxed while the volume is held fixed. The result is a discrete set of $E(V)$ points and fitting these points to an analytic EOS provides a smooth function from which the equilibrium volume V_0, equilibrium energy E_0, bulk modulus B_0, and first derivative of bulk modulus with respect to pressure $B_0{}'$ can be obtained.

The pressure and bulk modulus follow directly from derivatives of the energy curve,

$$p = -\frac{dE}{dV}, \tag{2.72}$$

$$B = -V\frac{dp}{dV} = V\frac{d^2E}{dV^2}. \tag{2.73}$$

Thus, the minimum of $E(V)$ gives the zero-pressure equilibrium volume, while the curvature at the minimum measures the resistance to compression: a flat minimum corresponds to a more compressible phase, while a steep minimum corresponds to a stiffer one. The fitted EOS also suppresses small numerical noise in the calculated total energies and permits different phases to be compared on the same pressure scale.

A commonly used form is the third-order Birch–Murnaghan equation of state written for the total energy,

$$E(V) = E_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^3 B_0{}' + \left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^2 \left[6 - 4\left(\frac{V_0}{V}\right)^{2/3}\right]\right\}. \tag{2.74}$$

Other EOS forms can be used, but the physical procedure is the same: the fitted curve transforms a finite set of total-energy calculations into a differentiable thermodynamic description of the phase.

2.4.2 Enthalpy vs Pressure Curves

Once $E(V)$ (EOS) is known and consequently also the $p(V)$ relation, the $E(V)$ can be converted into an enthalpy curve $H(p)$ by changing from a volume-based description to a pressure-based one: for each chosen pressure p, the EOS is used to find the volume $V(p)$ that satisfies $p(V) = p$. The enthalpy is then calculated as

$$H(p) = E\big(V(p)\big) + p\,V(p)\,. \tag{2.75}$$

In practice, this can be done either by solving the EOS directly for $V(p)$, or by evaluating the EOS on a fine volume grid and interpolating the resulting $H(p)$ values.

The $E(V)$ curves are useful for visually inspecting the phase density and pressure response, while the $H(p)$ curves are more directly connected to phase stability, because at fixed pressure, the stable phase is the one with the lowest enthalpy. A metastable phase at ambient pressure may become stable under compression if its smaller volume causes the pV contribution to decrease its enthalpy relative to a larger-volume structure. Conversely, an open framework may become favored at effective negative pressure, created under conditions that expand the lattice.

2.4.3 Pressure-Induced Phase Transitions

For two competing phases α and β, the transition pressure p_t is obtained from the condition that their enthalpies are equal,

$$H_\alpha(p_t) = H_\beta(p_t), \tag{2.76}$$

In terms of the $E(V)$ representation, this is equivalent to the construction of the common tangent. The two tangent points (V_α, E_α) and (V_β, E_β) on curves $E_\alpha(V)$ and $E_\beta(V)$ (Figure 2.4), respectively, satisfy

$$p_t = -\left.\frac{dE_\alpha(V)}{dV}\right|_{V_\alpha} = -\left.\frac{dE_\beta(V)}{dV}\right|_{V_\beta} \tag{2.77}$$

and

$$E_\alpha(V_\alpha) + p_tV_\alpha = E_\beta\big(V_\beta\big) + p_tV_\beta. \tag{2.78}$$

The same condition appears graphically as a crossing of the $H(p)$ curves. Below the crossing, the lower-enthalpy phase is the ground state; above it, the other phase becomes thermodynamically stable. If no crossing occurs in the investigated pressure interval, the candidate remains metastable with respect to the lowest-enthalpy phase over that range.

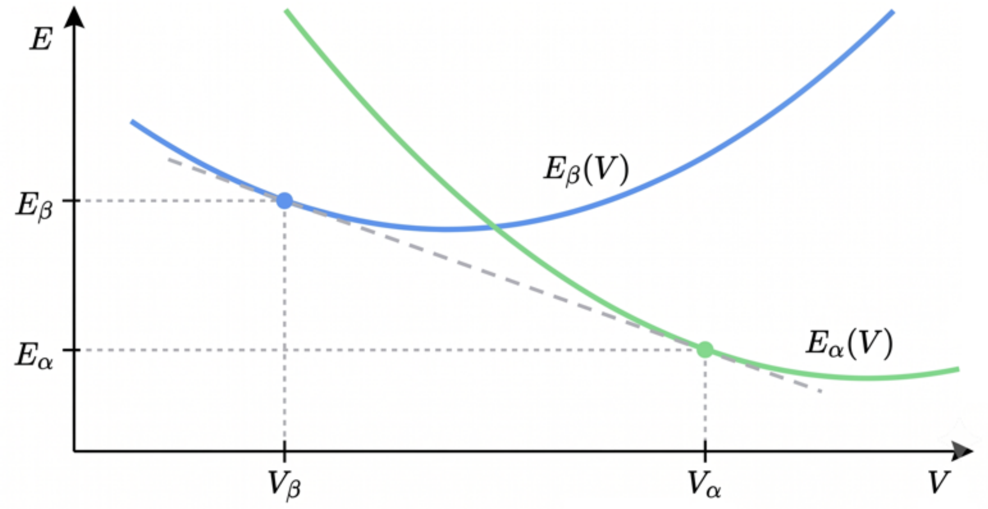

Figure 2.4. Energy vs volume diagram illustrating a first-order phase transition between phases α and β using the common tangent.

Because tangent points generally occur at different volumes, pressure-induced structural transitions between distinct polymorphs are typically first-order. In the static pressure-driven case, the relevant first-derivative discontinuity is the volume jump at p_t, while the enthalpies of the two phases are equal at the transition pressure. More generally, crystalline solid-solid phase transitions illustrate the same thermodynamic principle: pressure or temperature changes the relative stability of different crystalline polymorphs, and first-order transitions are characterized by discontinuities in first derivatives of volume and entropy (at finite temperature, this is also associated with latent heat) [68]. The graphite-diamond boundary in the carbon pressure–temperature diagram is a classic example of such a transition (Figure 2.5).

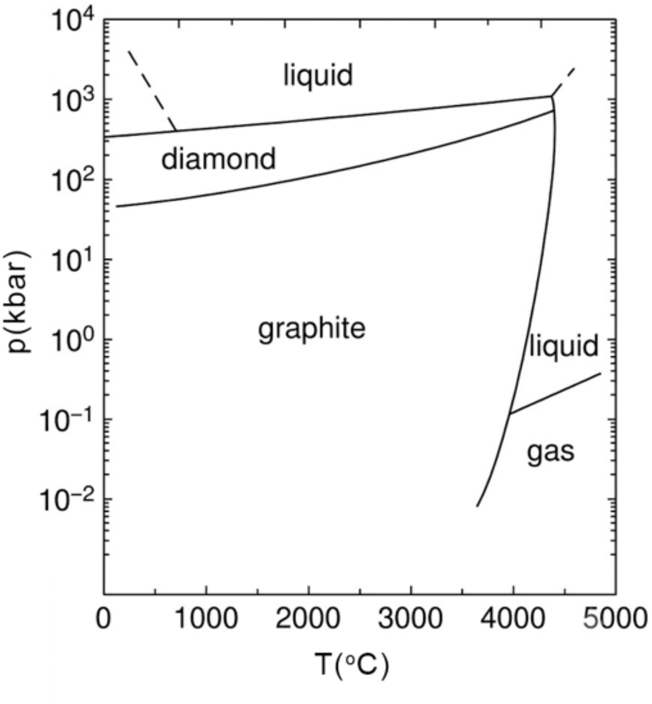

Figure 2.5. Carbon pressure–temperature phase diagram showing the graphite and diamond stability regions (adapted from S. Cheng [68]).

The zero-temperature ground-state phase diagram is constructed by applying the enthalpy comparison to all relevant candidate structures. At $p = 0$, the structure with the total-energy minimum defines the ambient-pressure ground state. At finite pressure, one evaluates the enthalpy of each candidate phase at that pressure and follows the branch with the lowest $H(p)$. When another phase has a lower enthalpy, the current pressure defines a pressure-induced phase transition; if no crossing occurs within the investigated interval, the candidate remains metastable in that pressure range. In this way, $E(V)$ curves and the corresponding $H(p)$ curves connect the local minima of the energy landscape with their pressure stability, metastability ranges, and possible pressure-induced polymorphism.

2.5 Machine Learning and Artificial Neural Networks

In the broadest sense, artificial intelligence refers to the field of constructing systems that perform tasks associated with intelligent behavior. Machine learning is the subset of artificial intelligence in which predictive rules are inferred from data rather than prescribed explicitly, and deep learning is the subset of machine learning based on multilayer differentiable models, most prominently neural networks [69]. In this chapter, the term artificial neural network (ANN) is used in the broad sense to refer to trainable layered function approximators, while graph neural network (GNN) denotes the ANN subclass designed for relational data represented as graphs [70, 71]. The scope of the present section is to establish the minimum theoretical background needed for the later discussion of the AI-ELX model.

2.5.1. Historical Development and Conceptual Foundations

The conceptual starting point of neural computation is the idealized formal neuron [72], which showed that networks of threshold units can implement nontrivial logical operations. Then the perceptron supplied the first trainable linear-threshold model and established the idea that network weights can be adapted from examples [73]. The important improvement came with multilayer networks trained by backpropagation, in which gradients of the loss function with respect to trainable parameters are obtained efficiently by repeated application of the chain rule for derivatives of composite functions [74]. Together with the backpropagation work [74], early gradient-based convolutional networks [75] helped establish deep learning as a practical trainable modeling framework.

Universal-approximation theorems showed that even comparatively simple feed-forward networks can approximate broad classes of continuous functions on compact subsets of $\mathbb{R}^n$ [76-78]. This guarantees that, in principle, the approximation error can be made arbitrarily small. In practical scientific modeling, issues of efficient training, data efficiency, numerical stability, and physical interpretability are governed by the representation of the input, the architecture, the optimization procedure, and how well the model conforms to the symmetries of the underlying problem [69].

The distinction between formal approximation power and problem-adapted structure is especially important in computational physics and chemistry, where many important datasets are not naturally arranged as fixed-length vectors or images on a regular grid. Molecules and crystals are described by relations among atoms, changing local neighborhoods, and symmetries. The transition from ordinary neural networks to graph neural networks is a necessary upgrade in model representation to better represent structured physical systems [70, 71].

2.5.2 Feed-Forward Neural Networks

2.5.2.1 Perceptron, Multilayer Perceptron, and Layered Composition

In the simplest case, a perceptron maps an input vector x to an output through an affine transformation followed by a nonlinear activation. The affine part is defined by a weight vector w and a bias term b, where the bias shifts the decision boundary independently of the input coordinates. In binary classification, this produces a linear separator in feature space; in regression or hidden-layer computation, the same algebra is used as a trainable transformation without restricting the output to a strict threshold classifier [69, 73].

$$\hat{y} = \sigma(w^{\mathrm{T}}x + b) \tag{2.79}$$

Once several perceptrons are stacked in layers, a multilayer perceptron (MLP), or feed-forward neural network, is constructed. Each hidden layer forms a learned intermediate representation, and the full model becomes a nested nonlinear map. This layered composition is what gives neural networks their practical flexibility. Figure 2.6 illustrates the basic idea of a fully connected network with two hidden layers, in which each unit in one layer is connected to all units in the next layer.

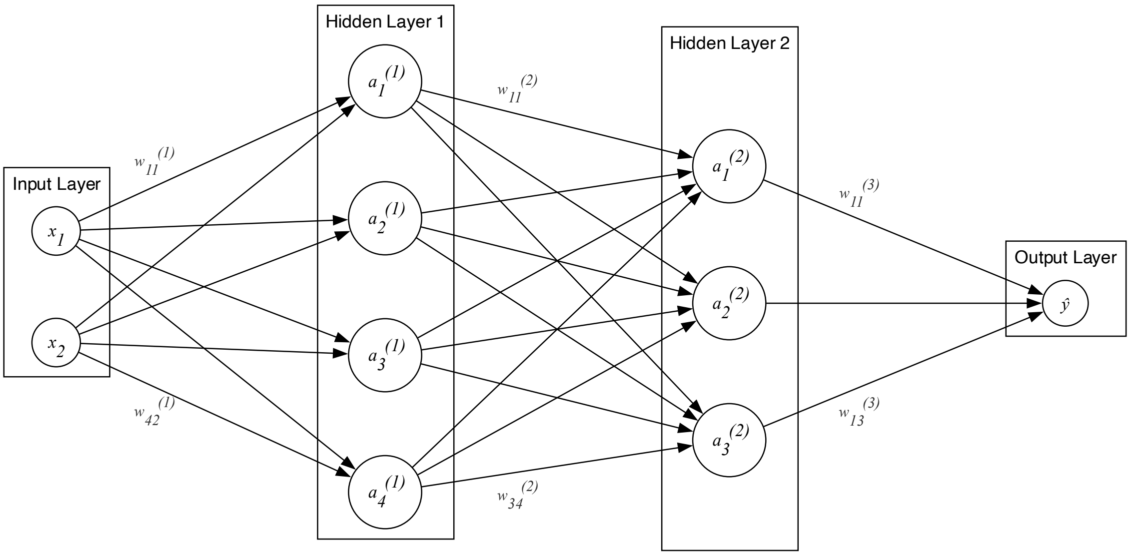

Figure 2.6. A fully connected multilayer perceptron with two hidden layers.

2.5.2.2 Matrix Formulation of the Forward Pass

For a network with layers indexed by l, it is convenient to write the forward pass in matrix form. Let $h^{(l-1)}$ denote the vector of activations entering layer l, $W^{(l)}$ the corresponding weight matrix, and $b^{(l)}$ the bias vector of that layer. The pre-activation $z^{(l)}$ is obtained by an affine transformation, and the activation $h^{(l)}$ is obtained after the application of a nonlinear activation function $\varphi^{(l)}$.

$$z^{(l)} = W^{(l)}\, h^{(l-1)} + b^{(l)} \tag{2.80}$$

$$h^{(l)} = \varphi^{(l)}(z^{(l)}) \tag{2.81}$$

In a network with two (or more) hidden layers, the repeated structure of the forward pass becomes immediately visible: each layer applies the same two-step (or multi-step) pattern of affine and nonlinear transformations. The resulting algebra is simple enough to be written explicitly, yet powerful enough to illustrate how a network builds complex nonlinear dependencies through iterated composition rather than a single large linear map [69].

$$a^{(1)} = \varphi\big(W^{(1)}x + b^{(1)}\big), \tag{2.82}$$

$$a^{(2)} = \varphi\big(W^{(2)}a^{(1)} + b^{(2)}\big), \tag{2.83}$$

$$\hat{y} = W^{(3)}a^{(2)} + b^{(3)}. \tag{2.84}$$

2.5.2.3 Activation Functions and Smoothness Considerations

If the activation functions were linear, any combination of layers would collapse to a single affine map and the benefit of depth would disappear. Nonlinear activations are therefore essential because they allow successive layers to build progressively more complex representations. Common smooth-function choices include the logistic sigmoid and the hyperbolic tangent, while modern practice often favors piecewise-linear or smooth rectifier-type functions because they allow more robust training in deep networks [69].

$$sigmoid(x) = \frac{1}{1+e^{-x}} \tag{2.85}$$

$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.86}$$

$$ReLU(x) = \max(0, x) \tag{2.87}$$

The choice of activation function affects both approximation properties and gradient flow during training. Saturating functions such as $sigmoid$ and $tanh$ can lead to very small derivatives far from the origin, which in deep networks may slow the learning. Rectifier-type activations reduce this problem but introduce their own trade-offs, such as non-differentiability at isolated points or inactive ("dead") units under inadequate initialization. In practical model design, activation functions are therefore chosen not only for representational power but also for the stability of gradient-based optimization [69].

2.5.2.4 Loss Function, Backpropagation and Optimization

In supervised learning, the network parameters θ are adjusted so that the prediction $\hat{y}$ approaches a target value y over a training set. This discrepancy is quantified by a loss function. For regression, a mean-squared error is the standard choice, whereas for probabilistic classification, a cross-entropy loss is more appropriate. The main point is that the loss provides a scalar objective whose gradient with respect to all trainable parameters can be computed and used for optimization [69, 74].

$$L_{MSE}(\theta) = \frac{1}{N}\sum_{n=1}^{N} \|\mathrm{y}_n - \hat{\mathrm{y}}_n\|^2 \tag{2.88}$$

$$L_{CE}(\theta) = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c} y_{nc}\log(p_{nc}) \tag{2.89}$$

Here, L_{MSE} and L_{CE} denote the mean-squared-error and cross-entropy loss functions, respectively, both written as functions of the trainable parameters θ, where y_n is the target value for sample n, $\hat{\mathrm{y}}_n$ is the predicted value, y_{nc} is the target label for sample n and class c, and p_{nc} is the probability assigned by the network to class c, usually after a softmax transformation of the output logits.

Backpropagation propagates derivatives from the output layer back through the network by repeatedly applying the chain rule to composite functions. If $\delta^{(l)}$ denotes the error signal at layer l, then the weight gradient is obtained from the outer product of $\delta^{(l)}$ and the activations of the previous layer, while the gradient with respect to the bias vector $b^{(l)}$ is the error signal itself, or the corresponding batch sum or average in minibatch training. In this sense the bias term is another trainable parameter that shifts the affine map at each layer [74].

$$\delta^{(l)} = ((W^{(l+1)})^T\, \delta^{(l+1)}) \odot \varphi'(z^{(l)}) \tag{2.90}$$

$$\partial L/\partial W^{(l)} = \delta^{(l)}(a^{(l-1)})^T \tag{2.91}$$

$$\partial L/\partial b^{(l)} = \delta^{(l)} \tag{2.92}$$

The full differentiable training loop (forward pass, loss evaluation, backpropagation, and parameter update) is summarized schematically in Figure 2.7. In practice, an optimizer such as stochastic gradient descent or Adam [79] repeatedly applies this loop. Training is usually monitored using both the training objective and a separate validation metric, i.e., an error measure evaluated on held-out data not used for parameter updates. Because gradients are propagated through the complete network, all trainable parameters can be optimized together [69].

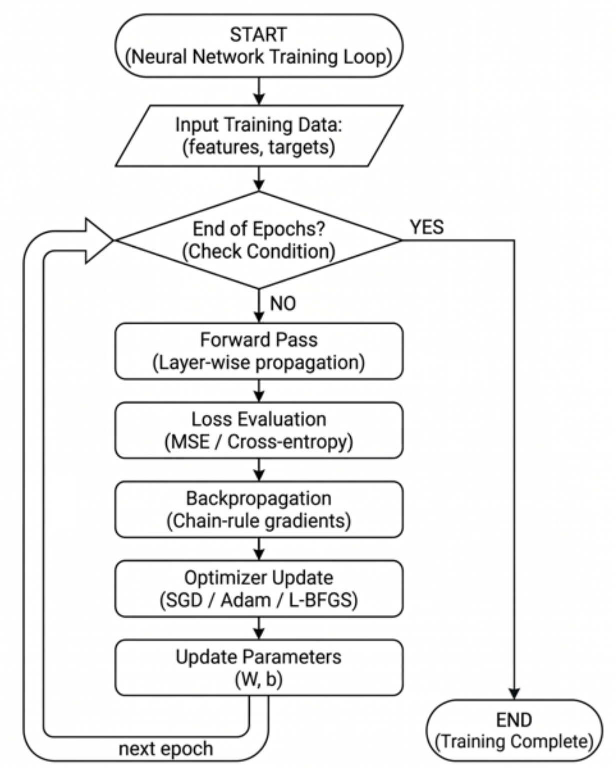

Figure 2.7. Schematic training loop for a differentiable feed-forward neural network.

2.5.3 Graph Neural Networks

Ordinary feed-forward networks assume that the input can be embedded naturally in a fixed Euclidean coordinate system or in a descriptor vector of fixed size. This assumption is adequate for tabular data, regular grids, or images, but it is too restrictive for molecules and crystals. Atomistic structures may contain different numbers of atoms, different local neighborhoods, and no unique preferred ordering of atoms. Moreover, physically equivalent structures should not change their predicted property because the atom indexing has been permuted.

One solution to this problem is to engineer symmetry-preserving local descriptors and then use an ordinary neural network on top of those descriptors, as in Behler–Parrinello-type neural-network potentials, or in related learned interatomic models based on local environments [80, 81]. A second solution is to make the representation itself relational: instead of forcing a crystal structure into a fixed vector, one allows the model to operate directly on atoms and their neighborhoods. This step leads naturally to graph neural networks.

A graph $G = (V, E)$ consists of a set of vertices V and a set of edges E describing relations among them. For graph learning, the central symmetry requirement is permutation invariance, i.e. relabeling the vertices must not change the physical content of the input. Consequently, graph-level predictions should be invariant under node relabeling (i.e., changing the arbitrary indices assigned to atoms or graph vertices without altering connectivity or geometry), whereas node-level outputs should transform equivariantly. The original graph neural network model already recognized that graph learning requires recursive state updates over local neighborhoods while respecting this relational structure [70, 82].

2.5.3.1 Graph Notation and Message Passing

Let A denote the adjacency matrix of a graph, D the diagonal degree matrix, and X the matrix of node features. At message-passing step k, node v carries an embedding $h_v^{(k)}$, i.e. a learned feature vector summarizing the information currently stored at that node. A generic message-passing layer first constructs messages from neighboring nodes and edges, then aggregates them, and finally updates the node state. This separates the architecture into local interaction rules, neighborhood aggregation, and node update [71, 83].

$$m_v^{(k)} = \mathrm{AGG}_{u \in \mathcal{N}(v)} M_k\left(h_v^{(k-1)}, h_u^{(k-1)}, e_{uv}\right), \tag{2.93}$$

$$h_v^{(k)} = U_k\left(h_v^{(k-1)}, m_v^{(k)}\right), \tag{2.94}$$

$$\hat{y} = R\left(\left\{ h_v^{(K)} : v \in V \right\}\right). \tag{2.95}$$

Here, $\mathcal{N}(v)$ is the set of neighbors of node v, e_{uv} denotes the edge attribute connecting nodes u and v, M_k is the learnable message function at step k, AGG is a permutation-invariant aggregation operation such as a sum, mean, or maximum, U_k is the node-update function, R is the graph-level readout, and K is the total number of message-passing steps.

For molecules and crystals, the nodes correspond to atoms, the edges encode neighborhood relations, and the edge attributes may include bond lengths, bond directions, or other geometric features. After K rounds of message passing, the embedding $h_v^{(K)}$ contains information from atoms within the K-hop graph neighborhood of node v, and the permutation-invariant readout R maps the collection of node embeddings to a graph-level prediction such as energy, band gap, or any other property. In atomistic energy models, the physical meaning of a K-hop neighborhood depends on how the graph is built, for example, by a distance cutoff or by a k-nearest-neighbor rule.

2.5.3.2 Representative GNN Architectures

Representative GNN architectures differ in how they aggregate neighborhood information, how they update node states, and how much inductive bias they impose. Table 2.2 summarizes four influential families that are conceptually useful for the later discussion of the AI-ELX model: GCN, GraphSAGE, GAT, and GIN [84-87].

Table 2.2. Comparison of representative graph-neural-network architectures

Architecture	Main update idea	Main advantages	Typical limitation
GCN [84]	Degree-normalized neighborhood aggregation followed by a linear map	Simple and efficient baseline; strong for homophilic graph learning	Deep stacks may oversmooth node representations
GraphSAGE [85]	Node update from the current node state plus an aggregated sample of neighbor features	Naturally inductive and scalable to large graphs	Expressivity depends strongly on the chosen aggregator
GAT [86]	Neighborhood aggregation with learned attention coefficients on the incident edges	Adaptive weighting of different neighbors	Attention can become costly or noisy on dense or high-degree graphs
GIN [87]	Sum aggregation followed by an MLP with a trainable self-weight	High discriminative power close to the Weisfeiler–Lehman test	Needs additional geometric features for strongly geometry-dependent tasks

2.5.3.3 Relevance to Molecules, Crystals, and Learned Interatomic Models

Molecules and crystals can be viewed as relational systems in which local coordination environments, bond lengths, bond directions, and chemical species determine target quantities such as energy, forces, charges, or descriptors related to the electronic band structure. A successful atomistic model must therefore respect permutation symmetry and, depending on the representation, may need to treat translations, rotations, and periodicity in a controlled way [83, 88, 89].

In molecular and materials modeling, message-passing architectures learn directly from local environments rather than from fixed, manually constructed global descriptors or explicitly prescribed analytical forms. Physics-informed neural networks are a complementary approach in which known equations or physical constraints are incorporated into the model or loss function, whereas atomistic graph models primarily encode the physical inductive bias through the graph representation, periodicity, and geometric features [90, 91]. Message-passing neural networks formalize this viewpoint directly [83], crystal-graph models such as CGCNN show how it can be applied to periodic materials [89], while continuous-filter atomistic models demonstrate how geometric information can be injected through distance-dependent learned filters [88]. These developments make GNNs a natural framework for learned energy models in computational physics, chemistry, or materials science.

The AI-ELX model presented in Chapter 3 belongs to the graph-based family in whose implementation crystal structures are converted into periodic k-nearest-neighbor graphs. Geometric information is expressed in reduced, volume-scaled form so that local environments are compared on a dimensionless scale, and graph-level conditioning variables derived from the volume per atom are supplied to the network together with the graph itself. The directly learned quantity is energy per atom, and once the learned energy is treated as a differentiable function of fractional coordinates and cell parameters, the same model can be inserted into a gradient-based optimization loop. This enables, in principle, fast predictions of DFT-level-precision energies when the model is trained on a sufficiently large and representative dataset.

3. Computational Methodology in Energy Landscape Exploration

3.1. Energy Landscape Exploration – Existing Methods and Techniques

The presence of an extremely large number of local minima on the energy landscape of multicomponent crystal systems presents a central practical difficulty in their energy landscape exploration. Direct *ab initio* evaluation of every trial configuration is impossible in practice, particularly when both atomic coordinates and cell parameters are varied. Global exploration, therefore, often begins with a choice of a fast approximate energy function, such as simple empirical potentials, which are not expected to reproduce all physical properties with high quantitative accuracy, but their role is to generate chemically reasonable structure candidates and to sample the topography of the low-energy landscape efficiently [13].

Using more elaborate energy models, such as shell models, breathing-ion models, semiempirical tight-binding approaches, or fitted machine-learning potentials, may improve the description of a selected class of structures, but they can also become computationally more expensive and more strongly dependent on fitted system-specific parameters. For a true structure prediction problem, the stable structure is not known in advance, so an overly specialized potential may bias the search toward the structure type on which it was fitted and may suppress important competing minima. For this reason, broad global searches commonly favor robust, computationally inexpensive potentials during the exploratory stage, followed by accurate local optimization and *ab initio* energy ranking. In this two-step strategy, minima found on the approximate empirical landscape are treated as structure candidates rather than final predictions. Empirical exploration identifies candidate basins, *ab initio* calculations test whether these basins persist when the computational methodology is improved, and finally, equations of state, phonon dispersions, and electronic-structure calculations determine the stability and properties of the candidates. This hierarchy is the main part of the structure-prediction workflows used in the thesis.

3.1.1. Global Optimization

Global optimization (GO) is the general mathematical task of finding the minimum value of a cost function over a large search space. In crystal structure prediction, the cost function is usually the energy or, under external pressure, the enthalpy of the crystal system, while each point in the search space represents one possible periodic arrangement of atoms. For an unconstrained crystal-structure search, a configuration includes the lattice vectors, the atomic coordinates, and the assignment of the specific elements to the atomic positions. The aim of a global search is to sample the energy landscape broadly enough to identify the lowest-energy basin and other low-energy basins that may correspond to metastable polymorphs [13].

In the energy landscape exploration workflows used in the thesis, the global search is performed on an empirical energy or enthalpy landscape. The exploration of the empirical landscape, which serves as a fast exploratory model, is necessary because direct first-principles sampling of many thousands or even millions of trial configurations would be computationally infeasible. The role of GO is to generate a diverse collection of plausible structural candidates and to identify the main low-energy regions of the landscape, but the final quantitative stability ranking is obtained from subsequent *ab initio* calculations [92].

In the general case, for a periodic system under external pressure, the empirical objective function can be written as

$$H_{\mathrm{emp}} = E_{\mathrm{emp}} + pV, \tag{3.1}$$

where E_{emp} is the empirical potential energy, p is the external pressure, and V is the unit-cell volume. At zero pressure, the objective function reduces to the empirical potential energy. The empirical potential energy used in the global searches is a fast two-body potential composed of a damped Coulomb term and a Lennard-Jones term, so that

$$E_{\mathrm{emp}} = \sum_{i<j} V_{ij}\left(r_{ij}\right), \tag{3.2}$$

and

$$V_{ij}\left(r_{ij}\right) = V_{ij}^{\mathrm{dC}}\left(r_{ij}\right) + V_{ij}^{\mathrm{LJ}}\left(r_{ij}\right). \tag{3.3}$$

Here r_{ij} is the distance between atoms i and j. The damped Coulomb contribution represents the long-range electrostatic part in a computationally stable approximate form

$$V_{ij}^{\mathrm{dC}}\left(r_{ij}\right) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \mathrm{e}^{-d_{ij} r_{ij}} \tag{3.4}$$

where q_i and q_j are the formal ionic charges, d_{ij} is the damping factor, r_{ij} is the distance between atoms i and j, and ϵ_0 is the permittivity of vacuum, while the Lennard-Jones contribution represents the short-range repulsion and attraction terms. It can be written as

$$V_{ij}^{\mathrm{LJ}}\left(r_{ij}\right) = \epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]. \tag{3.5}$$

The repulsive r^{-12} term prevents unphysically short interatomic distances, whereas the attractive r^{-6} term gives a simple representation of short-range attraction. The distance parameter can be expressed in terms of ionic radii of the interacting species as

$$\sigma_{ij} = r_s\left(r_i + r_j\right), \tag{3.6}$$

where r_i and r_j are the effective ionic radii and r_s is a scaling factor.

In addition to the radii, the empirical model contains formal charges q_i and q_j, the Lennard-Jones energy parameter ϵ_{ij}, and damping/scaling parameters such as the Coulomb damping factor d_{ij} and the radius-scaling factor r_s. These parameters make the potential useful as a fast model for large-scale sampling, but they also limit the transferability of the empirical ranking, i.e., the ability of one empirical parameter set to rank different structures, coordination environments, compositions, or pressure regimes with the same quantitative reliability. One set of empirical parameters should therefore not be expected to describe all structure types with equal accuracy.

One of the global-search methods used in practice is simulated annealing [93], implemented as a stochastic Monte Carlo [94] random walk on the empirical energy landscape. Starting from a configuration x_i, which describes the complete current state of the periodic model, including cell parameters, atomic positions, and atomic species, a trial configuration x_{i+1} is generated by applying a move from a predefined move class. In periodic crystal searches, the move class includes changes of atomic coordinates and cell degrees of freedom: it can include atomic displacements, cell-shape or cell-volume changes, combined atom-and-cell moves, and, when it is chemically meaningful, exchange of atomic positions. The energy change associated with the trial move is

$$\Delta E = E(x_{i+1}) - E(x_i). \tag{3.7}$$

If $\Delta E \leq 0$, the move is accepted. If $\Delta E > 0$, the move may still be accepted according to the Metropolis [94] acceptance probability,

$$P_{\mathrm{acc}} = \min\left[1, \mathrm{e}^{-\frac{\Delta E}{k_{\mathrm{B}}T}}\right], \tag{3.8}$$

where T is the current effective temperature and k_{B} is the Boltzmann constant. The acceptance of some energetically unfavorable moves is the mechanism that allows the walker to cross barriers between local minima: at high temperature, uphill moves are accepted often enough to explore multiple regions of the landscape, while at lower temperature, such moves become less probable and the trajectory becomes concentrated near low-energy minima.

The effective temperature is reduced according to a cooling schedule. A common choice in global optimization workflows is a geometric schedule,

$$T_n = T_0\gamma^n, \tag{3.9}$$

where T_0 is the initial temperature, n is the annealing level, and $\gamma < 1$ is the cooling factor. Values of $\gamma = 0.95$ to 0.995 are commonly used in practice. The cooling factor controls the balance between exploration and convergence, i.e., between broad sampling of different regions of the landscape (including barrier crossings and visits to structurally unrelated minima) and concentrating the search on low-energy regions to extract useful candidate minima. If the cooling is too rapid, the search can become trapped in a high-energy local minimum. If it is too slow, the calculation becomes inefficient because too many steps are spent sampling configurations that are ultimately not useful.

Simulated annealing in crystal-structure prediction is not normally used to obtain a single final structure; instead, many runs and periodic quenches are used to generate a collection of candidate structures at different local minima. A quench is a local descent from the current finite-temperature configuration to a nearby local minimum on the energy landscape. Repeating the process for many walkers, pressures, cell sizes, and random seeds gives a large set of candidate structures. In the systems studied in this thesis, the number of candidates ranges from thousands to millions, depending on the composition, the number of formula units in the simulation cell, and the pressure sampling values. Depending on the global optimization goal, a pressure scan can be important because the stable or metastable structure types at 0 GPa may differ from those favored in high-pressure or expanded-volume regions.

The output of an unconstrained global search is necessarily redundant: many structures are duplicates, distorted versions of the same underlying type, or low-symmetry representatives of a higher-symmetry structural family. However, the repeated occurrence of a structural motif across independent trajectories is still useful information, as it suggests that the corresponding minimum has a robust basin of attraction on the empirical landscape. Because global optimization can generate a very large number of structural candidates, the final goal cannot be to pass all generated minima to the *ab initio* stage, but rather to select a representative subset that covers the low-energy part of the landscape without being dominated by duplicates. The final candidate set is evaluated using empirical energy or enthalpy, frequency of occurrence, symmetry, and structural diversity [8, 10, 95].

The structural refinement phase, which includes symmetry and structure-type analysis between the empirical search and the *ab initio* stage, is therefore essential. The KPLOT program provides several algorithms for this purpose: SFND (Ger. *Symmetrie FiNDer*) identifies the symmetry operations of a periodic structure from the atomic coordinates; RGS (Ger. *Raum Gruppen Sucher*) then determines the corresponding space group from the detected symmetry operations; and CMPZ (Ger. *CoMPare Zellen*) compares pairs of periodic structures and determines whether they are equivalent within selected tolerances [96-98]. These operations refine the structural information by recovering symmetry, eliminating duplicates, and grouping candidate minima into structure families.

3.1.2. Data Mining

Another search route is prototype-based data mining, a crystallographic prototype search in which known structures from databases are selected, transformed, and used as starting structural models for the target systems. The general logic follows the knowledge discovery in databases (KDD) workflow: data selection, preprocessing, transformation, mining, and interpretation/evaluation [99]. The data source is usually the Inorganic Crystal Structure Database (ICSD), which contains experimental inorganic crystal structures with unit cell parameters, atomic coordinates, space groups, bibliographic information, and, in many cases, structure type assignments and formula classifications (e.g. ABX, AXY, or ANX-type classes) [100].

The goal of data mining is to reduce a very large set of structures in the database to a small set of prototype candidates suitable for the system under study. Thus, the search is first restricted to compositions and formula classes relevant to the target system, and the retained structures are then filtered by stoichiometry, chemical analogy, structure type, and uniqueness, so that only structure-type-distinct prototypes are retained in the final candidate set. In the next step, the atoms in a selected prototype are remapped to the elements of the target system. The candidate structures transformed in this way are not assumed to be stable, and they are only a starting point for the local optimizations, after which the structure may remain close to the starting prototype, transform into a different structure type, or become unstable. Even though data mining does not guarantee that the stable modification and the global minimum will be found, it is valuable because it injects chemically plausible, structurally diverse, and often high-symmetry starting models that may be difficult to reach by random global exploration alone [10, 11, 101].

Data mining and global optimization are complementary methods: global optimization is not biased toward known structure types and can discover unexpected motifs, but it can also miss narrow or symmetry-rich basins that are rarely reached by the stochastic trajectories or may require search over large unit cells, where the number of obtained structure candidates can be limited; data mining is biased toward known structure types, but that bias is useful when chemically related compounds contain prototypes likely to be transferable to the target composition. In the results of Chapters 4-8, this complementary approach appears repeatedly: many data-mined candidates may remain high in energy, but in some systems they relax into the ground state or into important low-energy competitors.

After data-mined structures are generated, they are treated in the same post-processing workflow as global-search candidates: their symmetries are determined, structures are grouped into structure types, duplicate structures are removed, leaving only unique types, and in the end, only a representative subset is selected for the *ab initio* calculations. For selecting the final subset for *ab initio* calculations, the same criteria used for global optimization are applied: low energy, structural diversity, high symmetry, and recurrence of the structure type across the energy landscape.

3.1.3. Local optimization

After the structure candidates are selected through global optimization, data mining, or both, they must be locally optimized using more accurate, *ab initio* methods, such as density-functional theory. Local optimization is a deterministic numerical minimization for a fixed starting structure and fixed computational settings. Unlike global optimization, it does not attempt to explore the entire landscape; instead, it follows the local gradient or an equivalent minimization procedure until it reaches a nearby local minimum. In crystal-structure prediction, this stage tests whether a global search or data-mined candidate remains a stable crystal structure when treated at the *ab initio* level.

In many previous energy landscape explorations, the CRYSTAL software package [52, 102-104] has been used extensively for local optimization. CRYSTAL uses localized Gaussian-type basis functions and can optimize both atomic positions and lattice parameters using analytical gradients. A range of exchange-correlation functionals can be used, such as LDA, GGA-PBE, or hybrid functionals.

To check whether a structure candidate obtained on an empirical landscape remains a local minimum under a first-principles description, the optimization is usually performed without enforcing the space group suggested by the empirical search, because relaxing the structure without symmetry constraints allows the structure to distort, change its symmetry, and transform into a nearby minimum.

The VASP [53-55] software package provides the complementary plane-wave PAW route used in later stages of the thesis. In addition to structural optimization, it is used to calculate electronic, magnetic, thermodynamic, and vibrational properties of the optimized structures. That makes local optimization the methodological bridge between empirical candidate generation and selection and obtaining physically interpretable results.

3.2 Energy Landscape Exploration – New Methods and Techniques

The established hierarchical energy-landscape exploration workflow includes: global optimization, which samples the empirical landscape broadly and generates many structure candidates; data mining, which supplies additional prototype-based starting structures from crystallographic databases; structural refinement that uses KPLOT-based functions for post-processing; and local optimization at the *ab initio* level that refines the selected candidates and provides the basis for thermodynamic, electronic, magnetic, and vibrational property calculations.

At the same time, this workflow has one bottleneck: the management of scale in the structural refinement step. Unconstrained global exploration can generate very large numbers of candidate minima, going up to the millions. For small datasets, manual inspection and existing post-processing algorithms may be sufficient, but for datasets containing hundreds of thousands or millions of structures, a more automated and reproducible analysis workflow becomes necessary. This need motivates the development of new or improved methods presented in Section 3.2.

One software package developed in this thesis, STyX (Structure Type eXplorer), addresses the problem of structural refinement of large output sets by streamlining symmetry determination, structural comparison, clustering, database tracking, and representative selection from very large structure sets. Another software package, AI-ELX (Artificial Intelligence Energy Landscape eXplorer), addresses a different but related limitation: when the target structure is in a specified thermodynamic region, such as an extreme-pressure phase, a graph-neural-network-assisted workflow can guide the exploration more efficiently than broad direct DFT screening. These two new methods extend the established GO / DM / DFT workflow by improving candidate selection, managing very large search outputs, and enabling more targeted exploration of regions of the energy landscape that would otherwise be difficult to sample systematically.

3.2.1 Structure Type Exploration Software Package (STyX)

The software package STyX[1] (Structure Type eXplorer) was developed as a post-processing and structure refinement layer between large-scale crystal-structure generation and first-principles calculations. Its purpose is to import the very large, highly redundant output of empirical global search and prototype-based data-mining runs and transform it into a compact, crystallographically analyzed, and physically interpretable set of candidate structure types, making this limited number of distinct structure type representatives input structures for the next, local optimization phase. A simplified description of the STyX algorithm and its use in the energy landscape exploration workflow is shown in Figure 3.1.

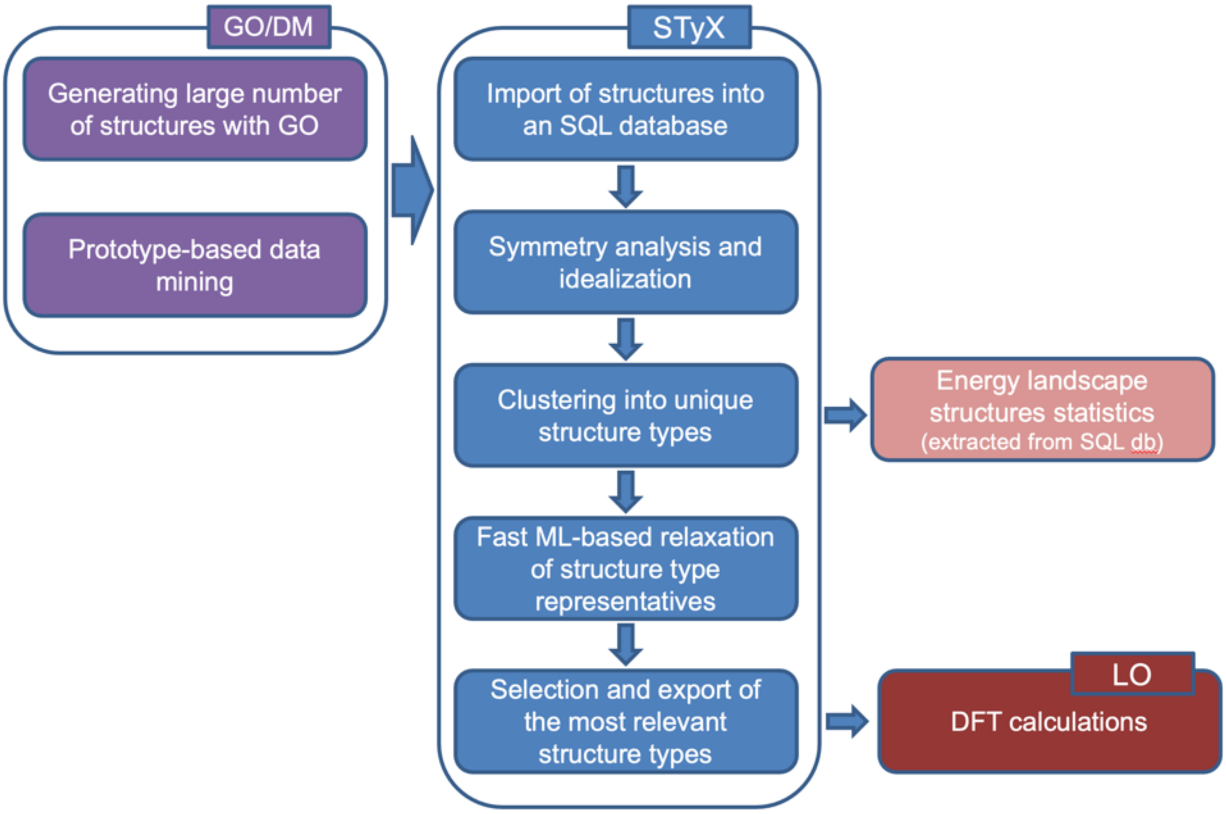

Figure 3.1. The STyX algorithm within a broader energy landscape exploration workflow.

The primary input of STyX is a collection of structures arranged in user-defined structure sources (Figure 3.1). In practice, these sources are suitably labeled directories containing Crystallographic Information File (CIF) files produced by empirical global optimization, by prototype mapping, or related structure-generation workflows. Each source can correspond to a particular search condition, such as a pressure value, a radius-scaling setting, a simulation cell size, or a prototype family. The

---

[1] https://github.com/milanpejic/styx

workflow is controlled through an input JSON (JavaScript Object Notation) settings file that specifies the database name, source patterns, export directory, scratch directory, structural-matching tolerances, and the selection of stages to be executed. In the current implementation, these stages are: source import; structure import and metadata parsing; relative empirical energy calculation; symmetry analysis and idealization; structure-type search; machine learning (ML) relaxation of non-*P*1 representatives; additional post-ML-relaxation duplicate removal; relaxation and duplicate checking for a selected number of *P*1 structures; and export of idealized and relaxed structure type representatives. During the import phase, crystal structure information together with metadata, such as energy, pressure, temperature, source label, and filename are also imported.

Internally, STyX stores the imported information in an SQLite database, which contains separate source and structure tables, where each structure record includes the original CIF file, empirical energy, relative energy (calculated within each source, i.e. directory), pressure, temperature, number of atoms, space group, idealized (symmetrized) and ML-relaxed CIF structures, ML-calculated energies, error flags, and the parent-child relations used for clustering the structures into structure types. This design is relatively simple from a database engineering perspective, yet powerful enough for high-throughput exploratory work. It ensures reproducibility because every selected representative can be traced back to a source and an original structure file; it ensures flexibility because the final DFT candidate set can be constructed by explicit SQL queries rather than by a single hard-coded rule; and it improves interpretability because the energy, symmetry, recurrence, and structural type are all stored in one common, easily accessible representation.

The first processing step after import is energy normalization and symmetry analysis. Energies may be renormalized within each source so that the lowest structure in that source defines the local zero of energy. This is useful because different empirical search branches may have different offsets or may have been generated under different conditions. Symmetry analysis and idealization are then performed using crystallographic tools from pymatgen and spglib libraries [105, 106]. Besides determining the space group of each input structure, the aim is also to remove small numerical distortions from structures that should be crystallographically equivalent. The symmetry tolerances are chosen to be large enough to absorb numerical noise and recover the underlying symmetry, but still small enough to preserve relevant structural details and to avoid assigning an unrealistically high symmetry to structures because of overly large tolerances. In the current implementation, the symmetry search uses a default distance tolerance of 0.3 Å and an angular tolerance of 3°. Structures found to possess symmetry higher than *P*1, i.e. higher than the triclinic group with no symmetry operations beyond lattice translations, are here denoted non-*P*1 structures. For these structures, STyX generates an idealized CIF representation and assigns the corresponding space-group number and symbol. Structures for which no higher symmetry is found remain classified as *P*1 and are handled separately.

The central algorithmic step in STyX is structure-type clustering, i.e. identification of the structural family to which each input structure belongs and the linking of equivalent inputs to the same underlying structure type representative. After idealization, candidate structures are first compared within, and then across, the sources using tolerance-controlled structural matching. The default comparison tolerances are length tolerance LTOL = 0.2 Å, site tolerance STOL = 0.3 Å, and angle tolerance ANGLE_TOL = 3.0°; these values can be changed in the settings file. In the first comparison pass, the first structure encountered within a given source and space-group subset is, by default, promoted to a parent structure representative. Every subsequent structure in that subset is then compared with the parent representatives identified thus far, and if it is equivalent to one of them, it is linked to that representative as a child structure; if it is not equivalent to any of them, it is then flagged as a new parent structure representative. In the second comparison pass, the parent structures within the same space group are compared again across all sources, and the representatives remaining after this pass become the final structure type representatives. In this way, STyX efficiently converts a large list of structures into a graph of unique structure types and their equivalents.

Recurrence information is important for the final selection of structure types because it complements pure energy ranking with a statistical measure of the robustness of structure candidates. A low empirical energy is an important indicator that a candidate warrants further attention, but it is not the only one. A structure that repeatedly appears in independent search runs, at different cell sizes, or under related pressure conditions is more likely to correspond to a broad and physically relevant region of the empirical energy landscape. For this reason, STyX supports a multi-criterion selection of favorable structure types based on symmetry, relative empirical energy, recurrence of the structure type across the whole landscape or within the selected conditions, and ML-calculated energy. The final choice of physically meaningful structures for subsequent DFT calculations is intentionally left flexible, because exploratory structure prediction benefits from expert judgment, and the database is designed so that this judgment is exercised through explicit, reproducible, and usually simple SQL queries, selecting candidates interactively rather than by ad hoc hard-coded filtering.

After clustering the idealized structures, STyX can perform fast ML-based relaxation of selected representatives of each structure type and calculate their equilibrium energies. In the current implementation, this is done using the MatGL M3GNet-MP-2021.2.8-PES model [107, 108]. The structural relaxation is applied to non-*P*1 structure-type representatives rather than to every structure occurrence of the same type, which keeps the workflow computationally manageable. While the idealization of a structure is a symmetry-restoring step (it removes small numerical distortions and replaces the original structure with the closest symmetry-consistent geometry within the chosen tolerances), the ML relaxation, in contrast, is an approximate energy-minimization step that can change the cell volume and, to a lesser degree, the cell shape and internal coordinates. In the present implementation, the standard (default) relaxation of non-*P*1 representatives preserves the starting space group, but not necessarily the full initial structure type; if needed, the stricter mode that keeps the original structure type can also be activated. After the ML-relaxation step, the relaxed structure representatives are compared again to eliminate duplicates that only become evident after the ML structural relaxation. In this way, the ML stage serves as an additional duplicate filter and a more refined ranking step, because the energies are calculated using the same method and can therefore be compared on a common scale across structure candidates from different sources.

The *P*1 structures require a different approach because their number can be extremely large, often well into the hundreds of thousands in large searches, and an exhaustive structure-by-structure comparison among all of them may not be practical. In typical applications, these *P*1 structures are not clustered into structure-type groups; instead, only a limited number of *P*1 structures per source (with the lowest empirical energy) are passed to the ML-calculation stage. This is a necessary practical compromise between completeness and computational feasibility because it prevents the often very large set of low-symmetry structures from dominating the workflow while still allowing physically relevant low-symmetry local minima to enter the final DFT candidate set.

The final output of STyX, besides the exported structure-type representatives, is the populated database itself, including the structure files and their associated metadata. The database stores the original imported structures, their idealized versions, ML-relaxed structures, empirical and ML-calculated energies, space-group information, relations between the structures, and the error flags needed to track problematic cases. The SQL queries can then be used to construct the final set of custom-selected DFT structure candidates. In the LaFS workflow, for example, the final selection was assembled from several SQL-defined subsets that emphasized the most frequently occurring structure types, low-energy representatives (both within specific sources and across the whole landscape), high-symmetry representatives, and a selected number of low-energy *P*1 structures. The export step then writes representative CIF files from which the subsequent first-principles input files can be generated.

The software is written in Python for ease of data handling, crystallographic and database operations, and reliable communication with specialized libraries. The application uses pymatgen and ASE (Atomic Simulation Environment) objects as its principal structure representation, spglib-based symmetry analysis for crystallographic classification, and MatGL for the ML relaxations [105-109]. The implementation is correspondingly modular: separate components handle structure input/output, symmetry analysis and structure comparison, database operations, and ML relaxation, which provides flexibility so that individual parts of the workflow can be easily improved without rewriting large parts of the code.

From a methodological point of view, STyX is best viewed as a high-throughput structure-reduction and selection workflow that converts a highly redundant representation of the explored energy landscape. It formalizes a workflow that is often performed informally in structure-prediction studies: sorting by energy, analyzing symmetry, removing duplicates, recognizing recurring motifs, and deciding which structures satisfy the final selection criteria. The advantage of STyX workflow is that these operations become reproducible, easily accessible, and scalable to datasets far larger than could be handled by manual inspection.

3.2.2 AI Model for Ground-State Energy Prediction, Structural Optimization, and DFT-Based Energy Landscape Exploration (AI-ELX)

This section describes a GNN-based AI-ELX[2] (Artificial Intelligence Energy Landscape eXplorer) model used in the machine-learning-assisted workflow for energy-landscape exploration. The model is used between broad structure generation and final density-functional-theory calculations: it predicts ground-state energies of structures at a near-DFT-level of precision, supports differentiable structural relaxation, and enables rapid $E(V)$-based screening before selected candidates are used in DFT calculations. The section is methodological, and the results of the AI-ELX application to LaFS are discussed later in the thesis, where AI-ELX is used as a targeted high-pressure search tool [12].

3.2.2.1 Role of AI-ELX Within Energy Landscape Exploration

The AI-ELX model is intended to make DFT calculations more efficient during crystal energy-landscape exploration. Conventional global optimization, data mining, and crystallographic post-processing still provide the initial pool of candidate structures, while DFT remains the standard method for final phase-stability and property calculations [12]. The AI model is inserted between these stages as a fast energy evaluator: it can relax and screen many more candidates than would be practical with direct DFT, both in global low-energy searches and in targeted searches such as the search for high-pressure modifications.

This role is summarized in Figure 3.2. Starting from an initial DFT data set, the model is trained to reproduce DFT energies per atom. It then predicts and relaxes candidate structures, selects promising structures for additional DFT calculations, and optionally retrains if the search enters an underrepresented region of the structural or volume space. In this way, the DFT calculations are concentrated where they are expected to be most informative.

---

[2] https://github.com/milanpejic/ai-elx

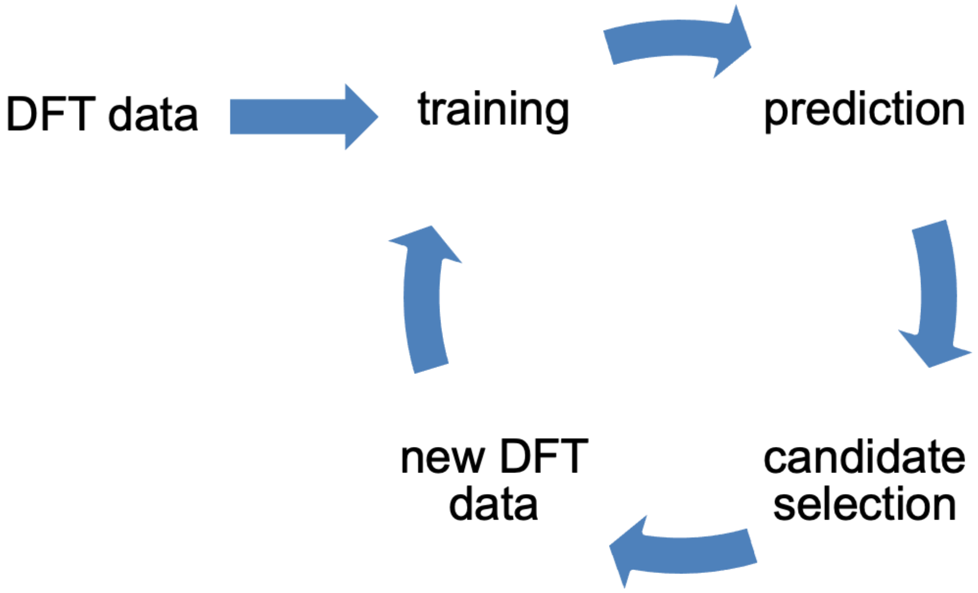

Figure 3.2. AI-ELX workflow for iterative energy-landscape exploration: DFT data generation, model training, energy prediction, candidate selection, new DFT calculations, and optional retraining.

The AI-ELX model should be distinguished from universal machine-learning potentials used as broad pre-screening tools (Table 3.1). A universal potential can be applied immediately to many structures and is therefore valuable for approximate ranking during large-scale filtering. Even though AI-ELX model is system-specific and must be trained on DFT data from the configurational space being explored, its advantage is the near-DFT-level precision that can be systematically improved by adding more DFT calculations to the training set, and also that the data set, volume range, and candidate-selection loop can be deliberately focused on the targeted region of the landscape, giving more control over the particular phase competition being investigated.

Table 3.1. Practical roles of DFT calculations, universal ML potentials, and the custom AI-ELX model in structural exploration.

Method	Primary use	Strengths	Limitations
DFT calculations	Accurate total energies, relaxed structures, and ground-state properties	Physically grounded and quantitatively reliable	Computationally expensive for exhaustive relaxation and volume scans across large candidate sets
Universal ML potential	Fast pre-screening and approximate ranking in broad workflows	Immediately usable for many crystal systems	May overlook system-specific phase competition or subtle enthalpy differences; final DFT validation is still required
AI-ELX model	Structural relaxation, volume scanning, and $E(V)$-based candidate selection	Can be retrained and focused on the relevant region of the landscape	Requires system-specific DFT training data and final DFT validation of selected candidates

3.2.2.2 Periodic Graph Representation and Volume-Aware Architecture

The AI-ELX model represents a crystal structure as a periodic graph $G = (V, E)$, where atoms are represented as vertices V and periodic neighbor relations as edges E. This representation avoids dependence on arbitrary atom ordering and is suitable for unit cells with different sizes and local environments. The approach follows the general idea of graph neural networks for relational data [70] and the message-passing formulation used in molecular and atomistic learning [83].

The implementation constructs a directed periodic k-nearest-neighbor graph. Each atom is described by its chemical species and by fractional coordinates inside the unit cell. The cell, written as a lattice matrix A, contains the three lattice vectors and converts fractional-coordinate differences into real-space displacements. To find nearest neighbors, translated periodic images of the unit cell are searched within a user-defined image range. For example, an image range of 2 2 2 means that translated copies from -2 to $+2$ along each lattice direction are available for neighborhood construction. The nearest neighbors are retained after distances are computed.

The volume-aware part of the model comes from reduced geometry and explicit global conditioning. If V_{cell} is the cell volume and N is the number of atoms in a unit cell, the characteristic length scale is

$$l_0 = (\frac{V_{cell}}{N})^{1/3}. \tag{3.10}$$

All geometric quantities used by the graph model are expressed relative to this scale, for example

$$r'_{ij} = \frac{r_{ij}}{l_0}, \quad A' = \frac{A}{l_0}, \quad c = [log(\frac{V_{cell}}{N}), l_0, \frac{1}{l_0}]. \tag{3.11}$$

Here r'_{ij} is the reduced displacement between atoms i and j, and A' is the reduced lattice matrix. The lattice matrix is stored in the row-vector convention, so the rows of A are the lattice vectors and a real-space vector is obtained from fractional coordinates by multiplication with A. The conditioning vector c is a small set of global cell descriptors supplied to the network. It holds the model information of volume per atom, through $log(V_{cell}/N)$, and the corresponding length scale, through l_0 and $1/l_0$, while the local graph uses dimensionless distances. This allows local environments at different volumes to be compared consistently while still retaining the absolute density information needed for variable-cell structure optimization and $E(V)$ modeling.

The model therefore predicts the energy of the structure as it is provided to the network. An equilibrium-like structure is obtained only after this energy predictor is connected to an optimizer. In practice, this means that the optimizer repeatedly changes the atomic positions and, if allowed, the cell parameters, asks the trained GNN for the predicted energy, and moves the structure in the direction that lowers it.

3.2.2.3 Training Data and Supervised Learning

The training set consists of periodic structures and matching DFT-calculated energies per atom. For energy-landscape exploration, in addition to the optimized minima, the training data set should also include distorted geometries, competing structure types, and volume-scaled configurations, which are needed for the model to learn the regions of the landscape it will be used in after training.

The DFT data set is split into training, validation, and test subsets. The training subset is the only part used to update the neural-network weights. The validation subset is used to monitor whether training is improving the model or it is beginning to overfit, and the test subset is reserved as an independent final check. Before training, the algorithm computes the mean and standard deviation of the conditioning vector c and of the energy target using only the training subset. The same scaling constants are then applied to validation and test structures. This prevents information leakage as validation and test data remain independent because their target values do not influence either the weights or the normalization parameters used during training. The loss function that is minimized during supervised learning is the mean-squared error between the standardized predicted and reference energies per atom:

$$L = \frac{1}{N_{train}} \Sigma_n (\hat{\mathrm{E}}_n{}^{std} - E_n{}^{std})^2. \tag{3.12}$$

The training algorithm uses AdamW optimization, validation monitoring, reduction of the learning rate on a validation plateau, gradient clipping, and saving of the best checkpoint according to the validation score. Model quality is judged by prediction error on structures outside the training set. In the next training loop (Figure 3.2) additional DFT-calculated data can be added if the search enters an underrepresented part of the configuration space.

3.2.2.4 Differentiable Relaxation and Variable-Cell Exploration

After training, the AI-ELX model is used for energy prediction and structural relaxation, when the workflow shifts from the training script to the optimizer script. The trained GNN receives a candidate structure and returns a predicted energy. Because this energy is differentiable in the model, the optimizer can calculate how the energy responds to changes of fractional atomic coordinates and selected cell parameters. The optimizer then updates those variables so that the predicted energy decreases, and the procedure is repeated until the convergence criteria are reached.

The optimizer separates atomic coordinates from cell degrees of freedom. Atomic positions are optimized in fractional coordinates, while cell changes can be restricted or enabled in different modes. Isotropic volume change is represented by a scalar logarithmic length parameter s. Shape change is represented by a symmetric strain-like matrix S whose trace is removed before being used in the transformations, so that the shape part changes the cell shape without changing volume. In full variable-cell mode, the current cell is reconstructed as

$$C(s, S) = e^s \, e^{dev(S)} \, C_0. \tag{3.13}$$

Here C_0 is the starting cell, e^s is an isotropic expansion or compression factor, and $e^{dev(S)}$ is a volume-preserving shape deformation where $dev(S)$ is a trace-free shape-deformation matrix. The optimizer also applies bounds on the step size, total displacement, cell volume change, and strain. These restrictions are important because the model should not be allowed to relax structures far outside the part of configuration space represented in the training data, where the prediction error rises rapidly.

3.2.2.5 From Structural Relaxation to E(V)-Based Screening

The AI-ELX model can also be used when candidate structures are not screened only by their predicted ground-state energies. A candidate structure can be evaluated over a range of cell volumes to obtain a predicted $E(V)$ curve: the predicted (E, V) points are fitted to a third-order Birch-Murnaghan equation of state [110, 111], after which the enthalpy vs pressure curves can be calculated

$$p(V) = -\frac{dE}{dV}, \quad H(p) = E(V) + pV. \tag{3.14}$$

The neural network model predicts energies per atom for the explored structure candidates, and thermodynamic quantities are obtained afterward by calculating $E(V)$ and $H(p)$ curves from the AI-predicted data. The AI-ELX model, therefore, proposes candidates through the same $E(V)$ and $H(p)$ reasoning used in the standard energy landscape exploration workflow [13, 112]. The search can then be focused on a selected part of the landscape, so if the goal is finding a high-pressure phase, the volume grid and candidate-selection criteria can be shifted toward high-density geometries, which are not necessarily in the low-energy part of the energy landscape.

In Chapter 8, this methodology is used to extend an already extensive LaFS energy-landscape exploration toward the extreme-pressure region, where the AI model helps identify additional candidate modifications [12].

4. Energy Landscape, Structure Prediction, Thermodynamic Stability, and Electronic Structure of Lanthanum Oxyiodide (LaOI)

4.1 Introduction

Lanthanum Oxyiodide belongs to a group of rare-earth oxyhalides, mixed-anion materials for which structural flexibility is one of the reasons why they are relevant in solid state physics and chemistry and materials science: related rare-earth oxyhalides have been studied in the context of luminescent host lattices, nanomaterials, ionic transport, photocatalytic activity, and medical applications [113-116]. The LaOI is especially interesting within this family because it contains the large and polarizable iodide anion, which can enhance structural anisotropy and make polymorphism, anion mobility, and layered structural types particularly important.

Experimentally, ground-state structure of LaOI is a tetragonal *P*4/*nmm* phase, first reported in early crystallographic work [117, 118], while more recent studies broadened the motivation for revisiting this system: ultrathin LaOI nanosheets have been investigated for biomedical theranostic applications, LaOI-based solids have shown unusually high iodide-ion mobility, and LaOI nanocrystals have been synthesized through non-hydrolytic sol-gel chemistry followed by post-synthetic halide-exchange reactions [114-116], highlighting LaOI as a functional rare-earth oxyhalide material whose structural variability could be relevant for experimental exploration of its properties.

One of the questions addressed in this chapter was if additional modifications exist on the LaOI energy landscape and whether they could represent (meta)stable polymorphs [8]. Since experimental synthesis alone may not easily reveal all possible polymorphs, especially the metastable ones, the system was investigated by global energy landscape exploration using simulated annealing, followed by first-principles local optimization. This computational workflow allows the structural search to move beyond known structure prototypes and to identify low-energy candidates within many different space groups and structure types.

This chapter focuses on three connected results: the theoretical confirmation of the experimentally known phase as the global minimum, the prediction of six additional low-energy modifications, and the analysis of their thermodynamic stability and electronic structures. The discussion follows the complete workflow from structure-type search and first-principles optimization to symmetry analysis, stability evaluation, and band-structure calculations. In this thesis, LaOI is used as the first example of how broad structure prediction can connect polymorphism with functional electronic properties in a rare-earth mixed-anion compound.

4.2 Computational Setup and Energy Landscape Exploration Overview

Crystal structure candidates of LaOI were generated in the global optimization phase on an empirical potential-energy landscape, followed by local optimization of the most promising candidates on the *ab initio* level [8]. In the first stage, the energy landscape was explored by simulated annealing Monte Carlo walks with periodic stochastic quenches, as implemented in the G42+ code [13, 93]. A fast, robust two-body empirical potential, consisting of a damped Coulomb term plus Lennard-Jones contributions, was employed so that the search could be performed for a very large number of trial structures [8]. No symmetry restrictions were imposed during the global exploration: both the periodic simulation cell and the atomic positions were allowed to vary.

In the global optimization stage, the LaOI system was explored for 4, 6, and 12 formula units per simulation cell, corresponding to 12, 18, and 36 atoms, respectively [8]. Total of 1,066,000 structure candidates, i.e. local minima, were generated. Although this is a very large number of structures, many candidates differ only by small changes in cell parameters or atomic coordinates and, within the chosen numerical tolerances, so they represent duplicates or closely related distorted versions of the much smaller number of structure types. The structure candidates were then analyzed crystallographically. The symmetries were determined using the SFND and RGS algorithms, duplicates were assigned to the same structure-type group using CMPZ, and all of these procedures were carried out within the KPLOT package [96-98, 119]. After eliminating duplicates, more than 7000 structure-type candidates with a space group higher than *P*1 remained, together with several hundred thousand *P*1 candidates. The most relevant structures, selected according to empirical energy, space-group symmetry, and frequency of occurrence during the global exploration, were then chosen for the *ab initio* stage calculations [8].

The DFT calculations were performed using CRYSTAL17 [51, 52], with full structural relaxation without symmetry restrictions in order to obtain the equilibrium structure and test the stability of each candidate during the structural relaxation run. Calculations were carried out using the GGA-PBE functional [49, 120], and for the most relevant low-energy structures were also recalculated with the screened hybrid HSE06 functional [121] for comparison and testing the robustness of calculations [8].

Large number of local minima are not important as physical candidates, but their statistical distribution shows which space groups recur often and which structure types dominate the landscape at different cell sizes (see Table 4.1). For the simulation cell containing four formula units (Z = 4), most structures were found in triclinic *P*1 (No. 1), followed by the tetragonal *P*4/*nmm* (No. 129) phase, which is also the space group of the experimentally known LaOI modification [8]. The global search with Z = 6 shows a similar trend, although with a significantly larger proportion of low-symmetry *P*1 structures, i.e. structures with no symmetry beyond the identity operation. After further increasing the cell size to Z = 12, the landscape becomes significantly more complex and most local minima appear as highly distorted or defect-containing *P*1 structures. Among the higher-symmetry candidates, the LaOI landscape contains recurring tetragonal *P*4/*nmm* (No. 129), monoclinic *C*2/*m* (No. 12), orthorhombic *Pnma* (No. 62), rhombohedral *R*-3*m* (No. 166), and cubic $P2_13$ (No. 198) arrangements, with *Pnma* (No. 62) becoming particularly prominent for Z = 12 [8].

Table 4.1. Summary of LaOI global exploration results with 4, 6, and 12 formula units, together with the combined distribution across all cell sizes.

4 LaOI		6 LaOI		12 LaOI		All cell sizes	
Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence
*P*1 (1)	221,441	*P*1 (1)	74,000	*P*1 (1)	438,675	*P*1 (1)	734,116
*P*4/*nmm* (129)	83,931	*P*4/*nmm* (129)	5,909	*Pnma* (62)	4,649	*P*4/*nmm* (129)	90,169
P-1 (2)	77,144	*P*-1 (2)	5,244	$P2_1$ (4)	2,528	*P*-1 (2)	83,589
Pnma (62)	34,609	$P2_1$ (4)	1,103	*Pm* (6)	2,329	*Pnma* (62)	39,258
$P2_12_12_1$ (19)	32,478	$P2_1/m$ (11)	1,087	*P*-1 (2)	1,201	$P2_12_12_1$ (19)	33,014
$P2_1/m$ (11)	19,582	*Pm* (6)	915	$P2_1/c$ (14)	1,163	$P2_1/m$ (11)	21,565
$P2_1$ (4)	11,847	*C*2/*m* (12)	567	$P2_1/m$ (11)	896	$P2_1$ (4)	15,478
$P2_1/c$ (14)	4,246	*C*2 (5)	184	*P*-62*m* (189)	653	*Pm* (6)	6,540
I-4 (82)	4,007	*Pmmn* (59)	92	$Pmn2_1$ (31)	567	$P2_1/c$ (14)	5,409
Pm (6)	3,296	*Cm* (8)	87	*F*-43*m* (216)	547	*I*-4 (82)	4,007
$Pmn2_1$ (31)	3,248	$Cmc2_1$ (36)	58	$\mathrm{P2_12_12_1}$ (19)	536	$Pmn2_1$ (31)	3,829
$Pmc2_1$ (26)	2,592	$P42_12$ (90)	50	$Pmc2_1$ (26)	383	$Pmc2_1$ (26)	3,019
*C*2 (5)	1,690	$Pmc2_1$ (26)	44	*C*2/*m* (12)	332	*C*2 (5)	1,874
OTHER	19,889	OTHER	660	OTHER	1,541	OTHER	15,290
TOTAL	520,000	TOTAL	90,000	TOTAL	456,000	TOTAL	1,066,000

4.3 Optimized Geometry - Symmetry and Structure Type Analysis

The most relevant candidates identified by the global exploration were locally optimized on the *ab initio* level using CRYSTAL17 software package with GGA-PBE and HSE06 functionals. After the calculations that resulted in the structural refinement and ground state energies of about a hundred structural candidates, seven low-energy LaOI modifications were chosen as the most relevant and were labeled α-, β-, γ-, δ-, ε-, ζ-, and η-LaOI types [8]. This final set spans tetragonal (α-LaOI and ε-LaOI), rhombohedral (β-LaOI), monoclinic (γ-LaOI and δ-LaOI), orthorhombic (ζ-LaOI), and cubic (η-LaOI) crystal systems.

Results including structural details and calculated total energies obtained with the GGA-PBE functional are summarized in Table 4.2, while the corresponding HSE06 energies and the complete optimized structural parameters for both used functionals, along with the tables with full structural information, are given in Appendix A4. The α-LaOI phase is the energetically most favorable structure regardless of the functional employed. The β-, γ-, δ-, ε-, ζ-, and η-modifications are low-energy minima on the LaOI landscape and represent the most plausible metastable polymorphs identified in the published study [8].

Table 4.2. Summary of the seven most relevant LaOI polymorphs' information after local optimization.

Phase	Space group	Lattice parameters (Å)	PBE Energy (eV/f.u.)	HSE06 Energy (eV/f.u.)	Characteristic La coordination / structural description
α-LaOI	*P4/nmm* (129)	a=4.233 c=9.312	-3224.2757	-3223.48043	4O + 4I; layered *matlockite/PrOI*-type polyhedron
β-LaOI	*R-3m* (166)	a=4.1965 c=33.8044	-3223.9247	-3223.06410	4O + 3I; layered rhombohedral polytype candidate
γ-LaOI	*C2/m* (12)	a=17.836 b=4.349 c=9.156 β=91.40	-3223.8647	-3223.02950	Two eightfold La local environments with 4O + 4I and 3O + 5I nearest neighbors
δ-LaOI	*C2/m* (12)	a=28.917 b=4.401 c=8.718 β=103.71	-3223.7956	-3222.95543	One 3O + 5I and two 4O + 3I local La environments
ε-LaOI	*I*-4 (82)	a=12.525 c=4.643	-3223.7419	-3222.91063	4O + 3I; tetragonal block-connected framework
ζ-LaOI	*Pnma* (62)	a=7.914 b=4.239 c=10.460	-3223.6419	-3222.78527	3O + 5I; orthorhombic framework
η-LaOI	$P2_13$ (198)	a=6.785	-3223.5309	-3222.65676	3O + 4I; cubic framework

The global minimum is the α-LaOI modification (Figure 4.1) in the tetragonal space group *P*4/*nmm* (No. 129), which is in agreement with experimental reports and previous theoretical studies [117, 118, 122-124]. The *P*4/*nmm* is a tetragonal nonsymmorphic space group with point symmetry D_{4h}, featuring a fourfold axis along z, mirror planes, inversion symmetry, and n-glide operations. This space group appears very frequently in the global optimization runs (see Table 4.1). The *α*-LaOI phase can be described as similar to the *PbFCl* or *matlockite* structure, but with one significantly longer La-I distance, so that the coordination is effectively reduced and the arrangement acquires a more pronounced layered character that is also associated with the *PrOI* structure type [8, 113]. Lanthanum has four oxygen and four iodine nearest neighbors, forming a coordination polyhedron reminiscent of a square antiprism. The La–O and La–I distances are 2.44 and 3.55 Å, respectively. In the crystal structure, the LaO_4I_4 polyhedra share oxygen edges, and the oxygen and iodine layers are stacked along the crystal's c axis [8].

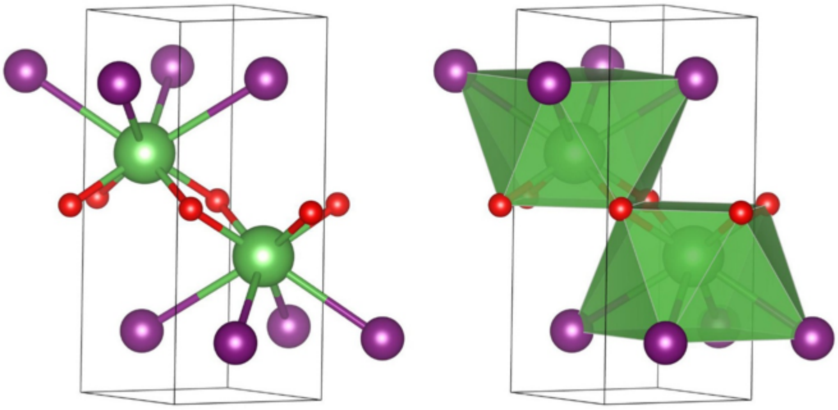

Figure 4.1. The *α*-LaOI structure (SG *P*4/*nmm*, No. 129) optimized using GGA-PBE functional.

The *β*-LaOI modification (Figure 4.2) crystallizes in the rhombohedral space group *R*-3*m* (No. 166) and also has layered character [8]. The *R*-3*m* space group has point symmetry D_{3d} and contains a threefold rotoinversion axis along the trigonal axis together with mirror symmetry. Lanthanum has sevenfold coordination with four oxygen and three iodine atoms as its nearest neighbors. Three oxygen atoms lie at a distance of 2.50 Å, the fourth oxygen is at 2.45 Å, and the three iodine atoms are at a distance of 3.34 Å from the central La atom. As in *α*-LaOI, the *β*-LaOI phase shows a layered arrangement along the c axis, consisting of alternating oxygen- and iodine-containing layers. This may indicate polytypism in the LaOI system, in analogy with polytypism discovered in previous global searches of oxides and oxysulfides [8, 92, 125]. Because the *β*-LaOI phase does not occur frequently in the global search, it could have been missed without the broad landscape exploration that is applied here. Related structure types are known in SmSI, YOF, and various lanthanide oxyfluorides [113, 126-128], which supports the experimental plausibility of this modification and suggests possible synthesis routes.

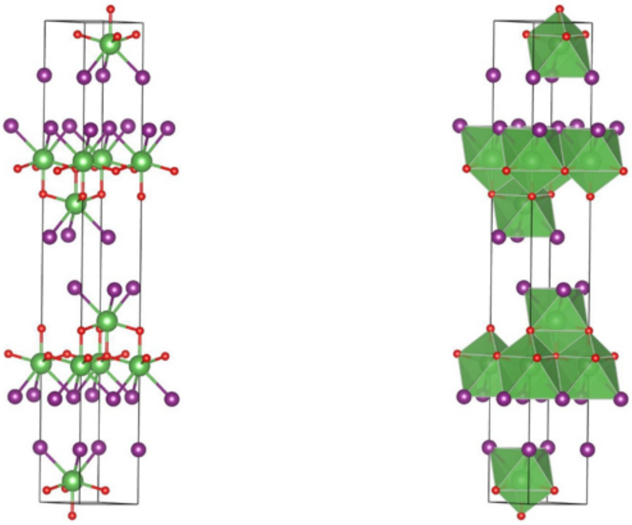

Figure 4.2. The *β*-LaOI structure (SG *R*-3*m*, No. 166) optimized using GGA-PBE functional.

Two relevant monoclinic modifications *γ*-LaOI and *δ*-LaOI (Figure 4.3) appear in space group *C*2/*m* (No. 12) [8]. The *C*2/*m* space group has point symmetry C_{2h}, which includes a twofold rotation axis along the b axis, a mirror plane perpendicular to b, and inversion symmetry; the full space group additionally includes C-centering translations. In the *γ*-LaOI phase, the lanthanum cations are eightfold coordinated, but two different local environments are present. In one coordination polyhedron, lanthanum is surrounded by four oxygen and four iodine atoms, with a mean cation–anion distance of approximately 3.02 Å. In the other, lanthanum is surrounded by three oxygen and five iodine atoms, and the mean cation–anion distance is approximately 3.21 Å [8]. In the *δ*-LaOI modification, the lanthanum local environment is even more varied because the coordination number also changes. One type of lanthanum site is eightfold coordinated (3O + 5I) with an average bond distance of approximately 3.09 Å, while the other two La environments have coordination number seven (4O + 3I) with average cation-anion distances of approximately 2.88 and 2.93 Å [8].

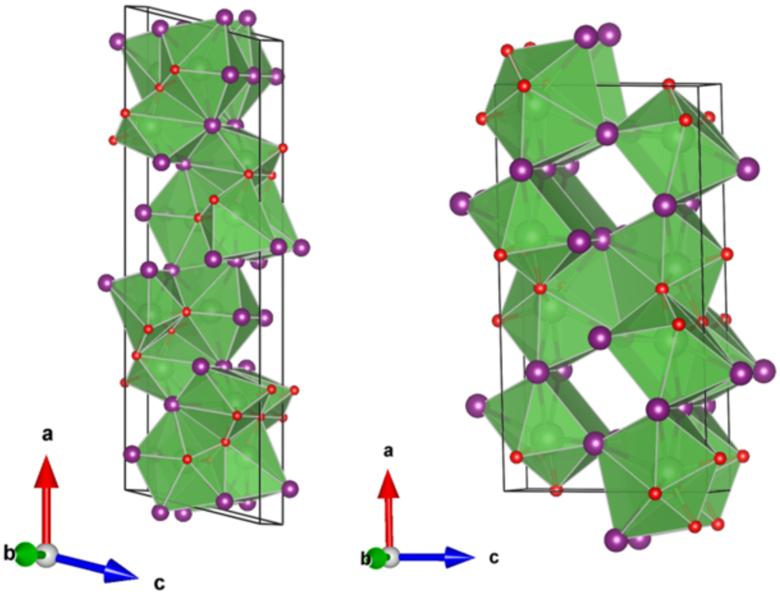

Figure 4.3. Monoclinic LaOI polymorphs: *γ*-LaOI (left) and *δ*-LaOI (right), both in *C*2/*m* (No. 12) space group, optimized using GGA-PBE functional.

The ε-LaOI modification crystallizes in the tetragonal space group *I*-4 (No. 82) [8]. This space group has point symmetry S_4, characterized by a fourfold rotoinversion axis parallel to the c axis, while the full space group additionally includes a body-centering translation. Lanthanum has coordination number 7, with four oxygen and three iodine atoms as its nearest neighbors. Four of these coordination polyhedra form a block by sharing oxygen edges, and these blocks are then linked through the remaining polyhedra by corner sharing (Figure 4.4). An important observation from the energy landscape exploration is that many low-energy *P*1 candidates relaxed into the ε-LaOI structure type during the *ab initio* relaxation run, which identifies this structure as a broad basin (minimum) on the LaOI energy landscape.

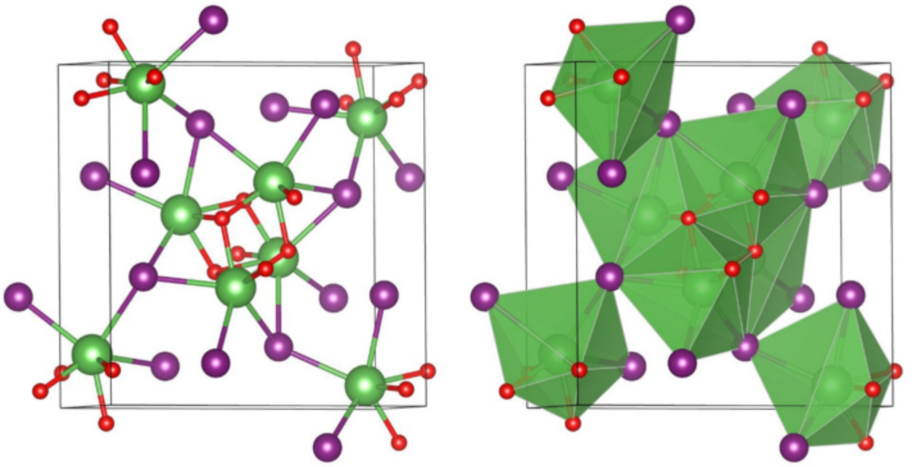

Figure 4.4. The ε-LaOI structure (SG *I*-4, No. 82) optimized using GGA-PBE functional.

Another low-energy modification is ζ-LaOI in the orthorhombic space group *Pnma* (No. 62) [8]. Its point group is D_{2h}, with three mutually perpendicular twofold rotational axes, three mirror planes, and inversion symmetry, while the full space group includes additional nonsymmorphic *n*- and *a*-glide operations. In the ζ-LaOI phase, lanthanum is eightfold coordinated (3O + 5I), with an average cation–anion distance of approximately 3.14 Å. The polyhedra are connected through a combination of edge and corner sharing. Among the higher-symmetry candidates, *Pnma* appears very frequently in the empirical global search, and among all the LaOI *Pnma* structures, the ζ-LaOI structure type is one of the most frequent [8]. Orthorhombic modifications are also known in related rare-earth oxyfluoride systems [113, 129-131], which suggests that ζ-LaOI is a plausible metastable phase. The structure is shown in Figure 4.5.

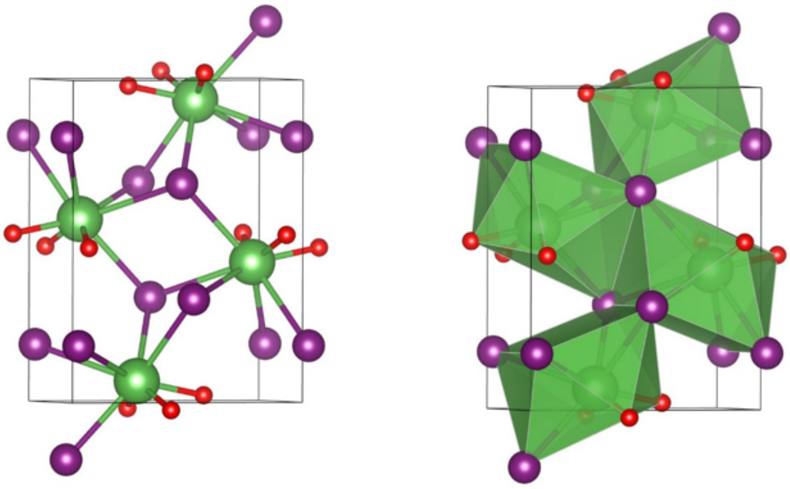

Figure 4.5. The *ζ*-LaOI structure (SG *Pnma*, No. 62) optimized using GGA-PBE functional.

The last low-energy phase, *η*-LaOI, crystallizes in the cubic space group $P2_13$ (No. 198) [8]. The associated point group is *T*, a chiral tetrahedral point group containing twofold and threefold rotations and no mirror planes, inversion center, or rotoinversion operations, while the full space group additionally includes 2_1 screw axes. Each lanthanum atom has three oxygen atoms as nearest neighbors at 2.43 Å, three iodine atoms at 3.47 Å, and one additional iodine atom at the shorter distance of 3.25 Å, giving an average cation–anion distance of approximately 2.99 Å. These polyhedra are connected by both edges and corners [8]. Although the *η*-LaOI phase is found only rarely during the global search, it is the highest-symmetry member of the selected low-energy set. Metastable cubic structures have been reported in LnOF systems [132, 133], and rhombohedral LnOF phases can transform to fluorite-type cubic structures at elevated temperature and pressure [134-136], which may suggest a possible route toward stabilization of *η*-LaOI. The structure is shown in Figure 4.6.

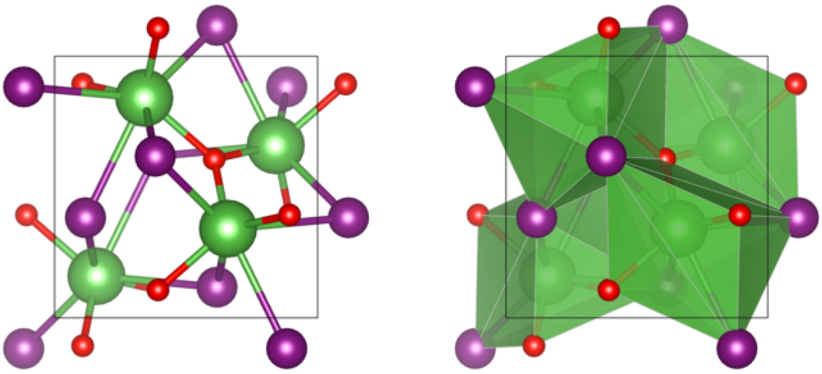

Figure 4.6. The *η*-LaOI structure ($P2_13$, No. 198) optimized using GGA-PBE functional.

The optimized structures show that the LaOI energy landscape contains a set of low-energy polymorphs with markedly different space groups, coordinations, and polyhedral connectivities. The *α*- and *β*-phases are layered or quasi-layered tetragonal and rhombohedral modifications, respectively; *γ*- and *δ*-LaOI represent more complex monoclinic arrangements; *ε*-LaOI is a tetragonal block-connected structure; *ζ*-LaOI is a very frequently-occurring orthorhombic modification, and *η*-LaOI is the cubic structure at the high-symmetry limit of the low-energy structure types [8].

4.4 Thermodynamic Stability and Pressure Response

The DFT energy rankings show that *α*-LaOI is the ground state, while *β*- through *η*-LaOI modifications form a group of low-energy local minima. The *α*-LaOI phase is the stable ambient-pressure structure, while the remaining phases are metastable candidates [8]. Both GGA-PBE and HSE06 preserve the same energy ordering. The thermodynamic discussion is based primarily on the ground-state energies of the chosen low-energy modifications and on their calculated *E(V)* curves, which provides the relevant information on equilibrium volumes, relative stability, and qualitative pressure trends.

4.4.1 The *E(V)* Equations of State

Figures 4.7 and 4.8 present the summary of the LaOI thermodynamic study and they establish the stability ranking of the seven low-energy structures at both GGA-PBE and HSE06 levels, respectively. The DFT-calculated *E(V)* curves show that *α*-LaOI has the deepest minimum regardless of whether the GGA-PBE or HSE06 functional is used and is therefore the global minimum on the LaOI energy landscape [8]. The good agreement between the two functionals is important because it shows that the identification of the ground state is robust and not an artifact of the chosen exchange-correlation treatment. Across the investigated volume interval, the *α*-LaOI *E(V)* curve remains significantly below the curves of the other six low-energy modifications. Therefore, no other low-energy modification becomes competitive enough to indicate a pressure-driven transition away from the α phase.

Among the six metastable candidates, *β*-LaOI is the closest low-energy competitor to the ground state. At the GGA-PBE level, their energy difference is approximately 120 meV/atom, corresponding to the thermal energy k_BT at approximately 1400 K [8], supporting *β*-LaOI as the most relevant elevated-temperature metastable phase. The remaining *γ*-, *δ*-, *ε*-, *ζ*-, and *η*-modifications lie higher in energy and, together with *β*-LaOI, may be stabilized under suitable high-temperature or nonequilibrium conditions. Additionally, the *η*-LaOI modification is the strongest candidate for possible stabilization under high pressure. This argument is based on its higher equilibrium density than the *α*-LaOI density and the analogy with related ternary rare-earth compounds, such as pressure-induced rhombohedral-to-cubic transitions reported for LnOF systems [134-136]. The actual experimental accessibility of these structures will depend not only on static energies, but also on synthesis route, kinetics, and the ability to trap metastable minima.

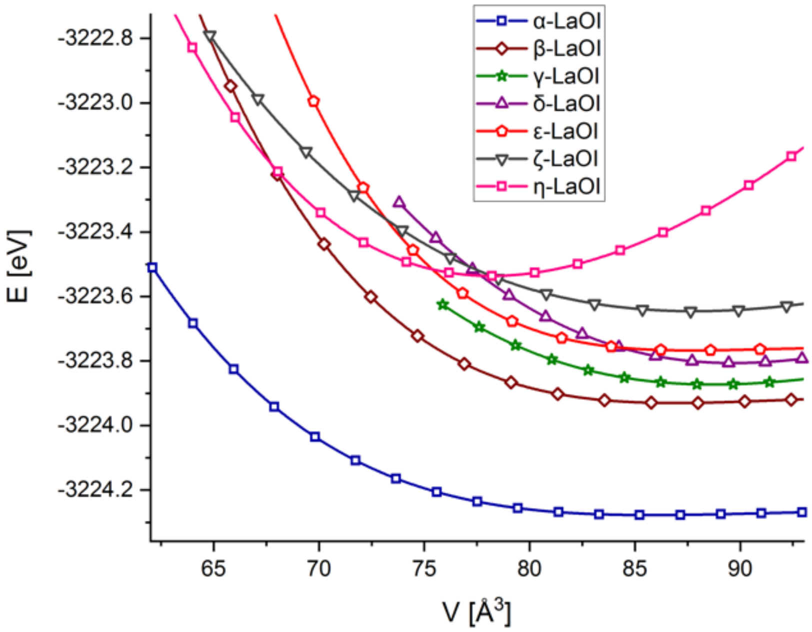

Figure 4.7. The *E(V)* curves for the low-energy LaOI polymorphs, calculated using GGA-PBE functional.

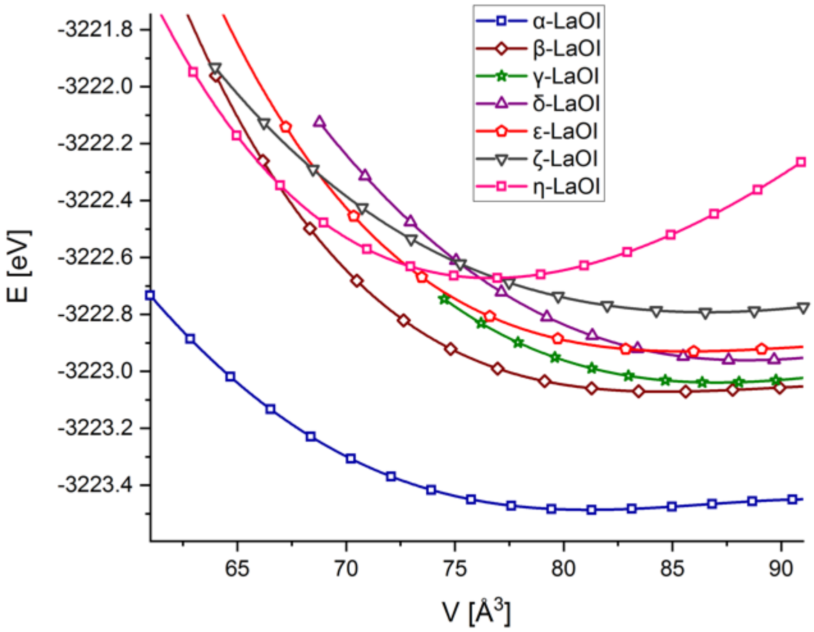

Figure 4.8. The *E(V)* curves for the low-energy LaOI polymorphs, calculated using HSE06 functional.

Explicit enthalpy vs pressure, *H(p)*, curves were not included in the LaOI thermodynamic stability study because the *E(V)* curves already show that *α*-LaOI remains far below the other selected low-energy polymorphs over the investigated volume range. The slopes of the *α*-LaOI equation of state at both lower and higher volumes, together with the large separation between the *α*-LaOI curve and the other *E(V)* curves at both GGA-PBE and HSE06 levels, do not indicate a competing high- or negative-pressure phase among the seven selected structures. Therefore, the LaOI study describes *α*-LaOI as the thermodynamically stable ground-state modification throughout the considered pressure range, with the other six structures as metastable structure candidates [8]. The optimized and literature structural parameters for α-LaOI are summarized in Table 4.3.

Table 4.3. Structural parameters for *α*-LaOI: calculated and literature data.

Method	a (Å)	c (Å)	V (Å^3)	Density (g cm^{-3})
GGA-PBE (present study)	4.233	9.312	166.83	5.61
HSE06 (present study)	4.218	9.082	161.57	5.79
Experiment [118]	4.144	9.126	156.71	5.97
Theory (LDA) [124]	4.14	8.85	151.45	6.18
Theory (GGA-PBE) [124]	4.21	10.00	178.23	5.32

4.5 Electronic Structure

The band gap values for all seven optimized polymorphs at both the GGA-PBE and HSE06 levels are summarized in Table 4.4. The results show a strong dependence of the gap on the structure type. At the HSE06 level, the gaps range from 3.14 eV in *β*-LaOI to 5.26 eV in the *ζ*-LaOI modification, while the ground-state *α*-LaOI has a band gap of 4.77 eV, which is in very good agreement with the experimental value of 4.82 eV reported for LaOI-based luminescent materials [123]. The GGA-PBE values are systematically smaller, as expected for semi-local functionals, whereas the screened hybrid HSE06 functional increases the calculated gaps through the inclusion of a fraction of Hartree-Fock exchange; this behavior is consistent with the general performance of screened hybrid functionals [121].

Table 4.4. Electronic band gaps of the low-energy LaOI polymorphs.

Modification and space group	Calculated PBE band gap (eV)	Calculated HSE06 band gap (eV)	Experimental band gap (eV) [123]
α-LaOI, *P*4/*nmm*	3.80	4.77	4.82
β-LaOI, *R*-3*m*	2.33	3.14	-
γ-LaOI, *C*2/*m*	3.70	4.70	-
δ-LaOI, *C*2/*m*	2.87	3.71	-
ε-LaOI, *I*-4	2.54	3.49	-
ζ-LaOI, *Pnma*	4.17	5.26	-
η-LaOI, *P*2$_1$3	3.73	4.90	-

The band structure and density of states (DOS) plots shown in Figures 4.9–4.12 for α- and *β*-LaOI modifications were calculated using VASP software package [53-55]. The GGA-PBE [49, 120] DFT+U used the Dudarev approach [66] with U_{eff} = 5 eV, while the HSE06 calculations follow the screened-hybrid-functional framework [121, 137].

Figures 4.9–4.12 show that both *α*-LaOI and *β*-LaOI are semiconductors, but with clearly different band gap sizes and gap characters. For *α*-LaOI, the plotted valence-band maximum and conduction-band minimum both appear at Γ, making it a direct band-gap semiconductor. The projected DOS at the PBE DFT+U level indicates that the upper valence region is dominated by I(*p*) and O(*p*) states, whereas the bottom of the conduction band consists mainly of La(*d*) states, with a smaller La(*f*) contribution. For *β*-LaOI, the similar broad separation between anion-p valence states and La-derived conduction states remains, but the gap is smaller than in the *α*-LaOI structure for both GGA-PBE and HSE06 calculations. The *β*-phase exhibits an indirect gap, with the valence-band maximum at Γ and the conduction-band minimum along the Γ–T path. In both *α*- and *β*-LaOI, the HSE06 functional preserves the overall band-dispersion pattern while simultaneously shifting the conduction band to higher energies and increasing the band gap relative to the PBE DFT+U results.

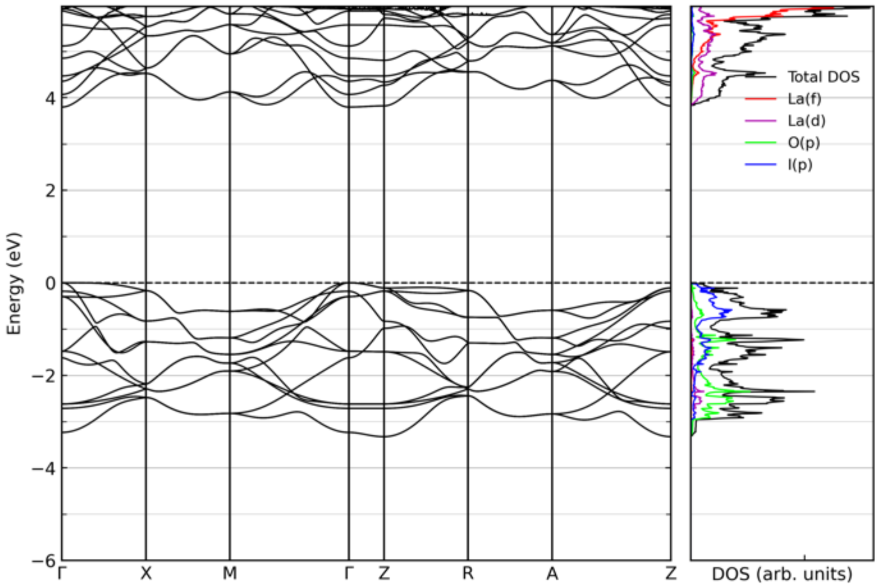

Figure 4.9. Band structure and DOS for *α*-LaOI calculated using PBE functional and DFT+U with U_{eff} = 5 eV for La 4*f* states.

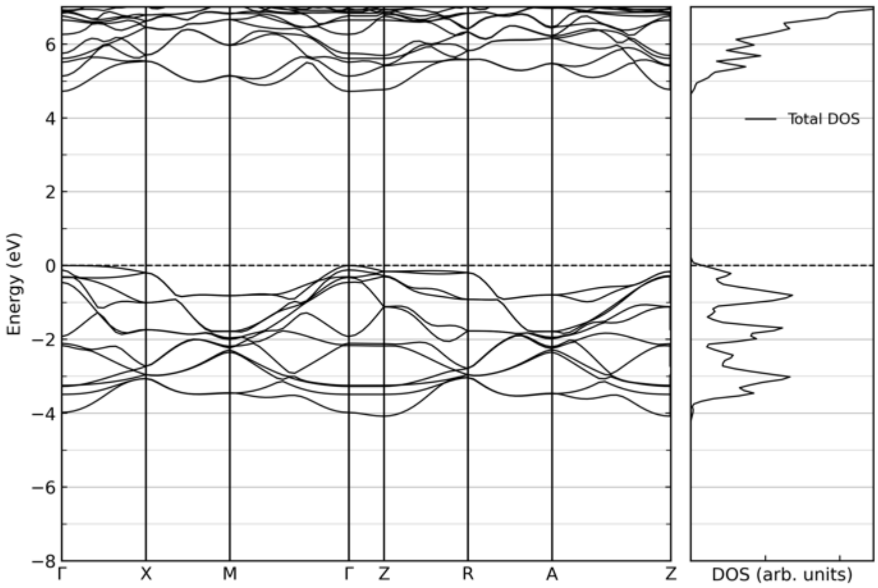

Figure 4.10. Band structure and DOS for *α*-LaOI calculated using HSE06 functional.

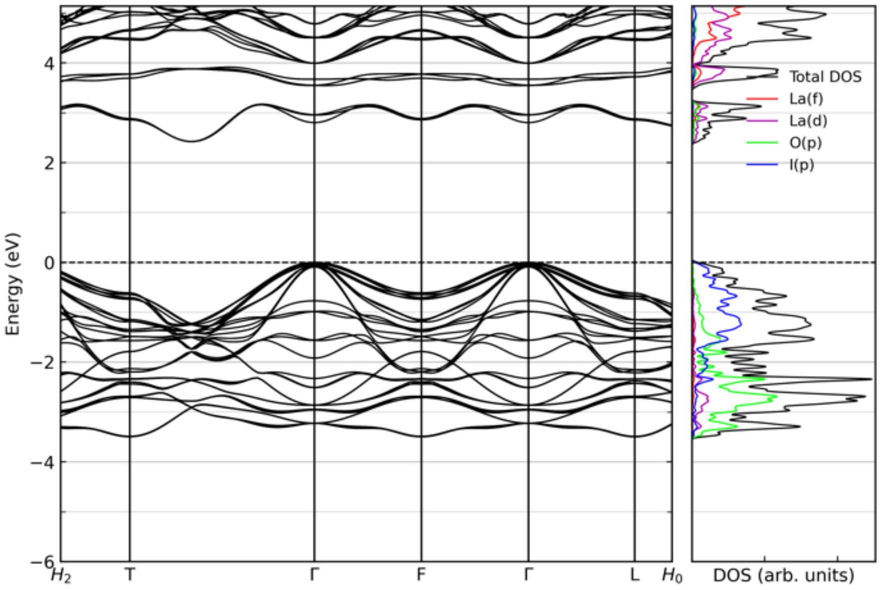

Figure 4.11. Band structure and DOS for *β*-LaOI calculated using PBE functional and DFT+U with U_{eff} = 5 eV for La 4*f* states.

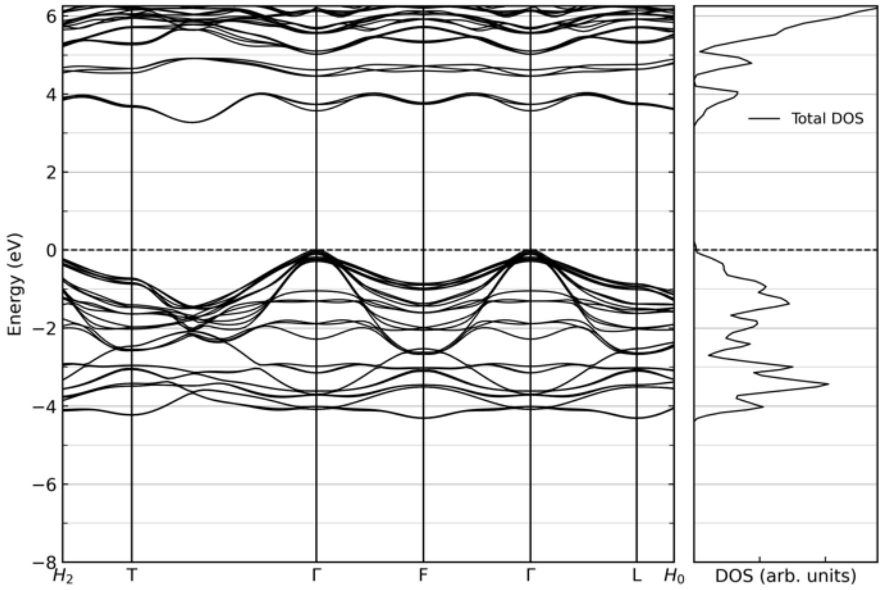

Figure 4.12. Band structure and DOS for β-LaOI calculated using HSE06 functional.

Additional electronic band structures for α-LaOI and β-LaOI obtained with the LDA, PBE, PBEsol, and hybrid PBE0 functionals are shown in Appendix A4 (Figures A4.1-A4.8). Comparison of the plain PBE and PBEsol band structures in the appendix with the PBE DFT+U results in Figures 4.9 and 4.11 shows no qualitative change in the band structure description: α-LaOI remains a direct-gap semiconductor at Γ, while β-LaOI remains an indirect-gap semiconductor. The main effect of the DFT+U scheme is to modify the relative position of the La-derived 4*f* conduction states and increase the gap, without changing the main orbital character near the band gap. Since La in LaOI is formally the La^{3+} cation with an empty 4*f* shell, the DFT+U correction is not essential for describing the valence band.

4.6 Conclusion

The LaOI study shows that the experimentally known tetragonal α-LaOI phase is the deepest minimum within a broader energy landscape containing several metastable local minima. The global search generated more than one million local minima, which were grouped and ranked by energy, symmetry, and frequency of occurrence on the landscape before the most relevant candidates were optimized at the *ab initio* level [8]. This large search space was essential because several predicted structures, in particular β-LaOI and η-LaOI, occur only rarely during the exploration and could easily have been missed in a narrower search.

The predicted β-, γ-, δ-, ε-, ζ-, and η-LaOI modifications illustrate the structural flexibility of the LaOI composition. Their space groups range from rhombohedral and monoclinic to tetragonal, orthorhombic, and cubic, and their La coordination environments include different La–anion coordinations. The structural relationships with known rare-earth oxyfluoride and oxyhalide phases provide possible guidance for future synthesis attempts, especially under non-equilibrium, high-temperature, or high-pressure conditions.

The thermodynamic results support the stability of the *α*-LaOI ground-state, with both GGA-PBE and HSE06 *E(V)* curves identifying α-LaOI as the lowest-energy phase stable over the whole inspected volume range, while the six other modifications remain metastable. All seven optimized LaOI phases are semiconductors: the PBE+U band gaps are in the range from 2.33 to 4.17 eV, while the HSE06 gaps are from 3.14 to 5.26 eV. In the projected DOS, the top of the valence band is dominated mainly by I(*p*) and O(*p*) states, whereas the bottom of the conduction band is dominated primarily by La(*d*) states, with a smaller La(*f*) contribution. The calculated HSE06 electronic band gap of *α*-LaOI agrees very well with the experimental optical gap, and the predicted metastable polymorphs show substantial changes in gap magnitude and its direct/indirect character. The results also define targets for future experimental work, especially for attempts to stabilize *β*-LaOI and other metastable modifications under high-temperature, non-equilibrium, or pressure-assisted synthesis conditions.

5. Structure Prediction, Thermodynamic Stability, Magnetic Ground State, and Electronic Structure of Cerium Oxynitride Ce_3O_3N

5.1 Introduction

The Ce_3O_3N system occupies a unique place within the ternary mixed-anion rare-earth compounds explored in this thesis because it was treated as a purely theoretical compound with no previously experimentally synthesized phase. Mixed-anion oxynitrides are attractive because the coexistence of oxide and nitride anions introduces additional degrees of freedom in anion ordering, Ce-centered coordination, and local bonding, which can modify both the topology of the Ce coordination polyhedra and the electronic structure. In the broader oxynitride literature, this mixed-anion flexibility has been associated with tunable optical and transport behavior, photocatalytic activity, and in some cases coupled magnetic responses [6, 138, 139]. For cerium compounds, the motivation is even stronger because the Ce–O system already supports multiple oxidation states and strongly correlated 4*f* states, so the proposed Ce_3O_3N composition offers a particularly complex test case for purely theoretical energy-landscape exploration [10] and physical properties calculation.

In this chapter the structure-type search of the Ce_3O_3N system using global optimization and complementary data mining is discussed first, followed by the thermodynamic analysis of the most feasible low-energy modifications. The last part examines different magnetic configurations of the predicted ground-state structure and their electronic structures.

5.2 Computational Setup and Energy Landscape Exploration Overview

The cerium oxynitride Ce_3O_3N study used a structural search strategy that combined global optimization on an empirical energy landscape with a complementary data-mining search among known inorganic prototypes [10]. In the global-search phase, simulated annealing on the empirical potential-energy landscape was used together with periodic stochastic quenches, as implemented in the G42+ framework [13, 93].

The global search was carried out for one, two, and three formula units in the simulation cell and for six pressure values between 0 and 160 GPa, with the enthalpy as the objective function. The move class of the stochastic random walk was deliberately broad so as to sample a wide region of the energy landscape. It included atom-only moves, local boxed displacements around a selected atom, joint atom-and-cell moves, and cell-shape and volume changes with or without simultaneous atomic motion. As in the related LaOI, ScOCl, HoFSe, and LaFS chapters (Chapters 4, 6, 7, and 8), the underlying idea is that global exploration must deliberately oversample configuration space in order to expose structurally distinct basins that can later be filtered by symmetry, recurrence on the energy landscape, and total-energy analyses. For Ce_3O_3N, this workflow generated more than 10,000 candidate structures across the full set of pressures and cell sizes [10]. To keep such exploration computationally feasible, a fast two-body empirical model consisting of Lennard–Jones and Coulomb terms was employed in the G42+ global-exploration package [13].

The data-mining phase served a complementary purpose. The ICSD-based [140] knowledge-discovery-in-databases procedure [99, 100, 141] started from more than 200,000 inorganic structures and was filtered down to 19 distinct ternary A_3B_3C prototypes, which served as starting structure geometries for Ce_3O_3N *ab initio* calculations [10].

The first-principles calculations were carried out mainly using the CRYSTAL17 code [102-104] at the DFT level within the local density approximation (LDA). For $\mathrm{Ce^{3+}}$ cations, a pseudopotential-based basis set was used, while oxygen and nitrogen were represented by Gaussian basis sets appropriate for oxide and nitride environments. The Ce basis choice followed earlier work on correlated cerium oxides [142]. Structural symmetries were determined with KPLOT using the SFND and RGS algorithms, and duplicate periodic structures were identified through CMPZ-based comparison [96-98, 119]. Auxiliary VASP calculations with the LDA functional were used to compare the two lowest-energy $\mathrm{Ce_3O_3N}$-*DM1* and $\mathrm{Ce_3O_3N}$-*GS1* structures with the $P2_1$ structure candidate proposed in the literature [143], These additional plane-wave calculations were introduced because the $P2_1$ structure candidate could not be successfully optimized with CRYSTAL17, showing that plane-wave DFT calculations can be useful when a local-basis optimization fails for a particular structural model.

In addition to structural relaxation and total-energy ranking, the thesis extends the published structure-prediction study by adding VASP DFT+U calculations for the magnetic and electronic properties of the DM1 ground-state modification. These calculations include three scalar-collinear PBE DFT+U magnetic configurations (AFM, FiM, and FM arrangements) from which the magnetic energy hierarchy, local magnetic moments, and band gaps were extracted, as well as a non-spin-polarized PBE reference calculation.

5.3 Optimized Geometry – Symmetry and Structure Type Analysis

The structural analysis begins on the empirical global search level, because the statistical distribution of symmetries already contains useful information about the energy landscape. For Z = 1 formula unit in a unit cell, most local minima are strongly distorted *P*1 structures, but several nontrivial symmetries occur with noticeable frequency. In particular, the monoclinic groups *Pm* (No. 6) and *Cm* (No. 8), the orthorhombic groups *Pmm*2 (No. 25) and *Amm*2 (No. 38), and the cubic group *Pm*-3*m* (No. 221) all appear repeatedly across the global search landscape. As pressure increases, the number of highly distorted structures decreases and some higher-symmetry candidates become relatively more prominent. On the other hand, the searches with Z = 2 and Z = 3 yielded far fewer additional nontrivial candidates after structure-type filtering, and none of those extra structures emerged in the subsequent *ab initio* refinement as low-energy structures. Therefore, the physically relevant low-energy set comes from the global optimization with Z = 1 formula units and data mining (Table 5.1) [10].

Table 5.1. Frequency of occurrence of the space groups obtained for the results of the global search performed at different pressures for Z = 1 of $\mathrm{Ce_3O_3N}$.

Pressure (GPa) / SG No.	1	3	5	6	8	10	16	25	38	44	47	99	115	146	155	160	207	221
0	379	–	–	32	11	–	–	7	–	2	–	1	–	–	–	–	–	–
0.016	376	2	1	27	7	–	–	10	2	–	–	1	1	–	1	–	1	3
0.16	368	–	–	38	3	1	1	5	–	1	1	1	–	–	–	–	3	10
1.6	373	1	2	26	9	–	–	7	3	–	–	1	–	–	1	–	2	7
16	370	2	3	25	10	–	–	–	7	2	–	–	–	–	–	–	6	7
160	332	–	3	36	20	1	1	5	16	1	11	–	–	4	–	2	–	–
Σ	2198	5	9	184	60	2	2	34	28	6	12	4	1	4	2	2	12	27

The data-mining search was included to introduce chemically and crystallographically diverse prototypes that may not be reached efficiently by the global optimization alone. The 19 starting prototypes used as data mining starting structures, together with their original literature references, are the following ternary A_3B_3C prototypes: Al_3ScC_3 [144], Ba_3FeN_3 [145], Ca_3PI_3 [146], Ca_3InP_3 [147], Cu_3SbS_3 [148], Fe_3TlTe_3 [149], Fe_3W_3C [150], Gd_3MnI_3 [151], K_3AlSe_3 [152], K_3BS_3 [153], $K_6Sn_2Te_6$ [154], KB_3H_3 [155], Na_3AsS_3 [156], Ni_3SmGe_3 [157], $NiSc_3Si_3$ [158], proustite [159], $Rh_4C_{12}O_{12}$ [160], Tl_3AsSe_3 [161], and Ag_3AsS_3 (xanthoconite) [162]. After local optimization, three data-mined structures remained in the final low-energy structures (Table 5.2).

The final step in the structural discovery phase is the comparison of *ab initio* ground state energies calculated during the local optimization. It identified Ce_3O_3N-*DM1* structure as the global minimum, followed by Ce_3O_3N-*GS1*, *GS2*, *GS3*, *GS4*, and *GS5* within approximately 0.03 Eh/f.u. (0.117 eV/atom) of the ground state. These six modifications were used for detailed thermodynamic stability analysis. The chosen threshold of 0.03 Eh/f.u., corresponding to approximately 1300 K on the thermal scale, was taken as a practical boundary for structurally realistic metastable candidates [10]. Table 5.2 gives the full LDA ranking of the optimized global search (GS) and data mining (DM) candidates, while highlighting the subset of structures that would be most relevant under realistic experimental conditions.

Table 5.2. The total-energy ranking of the optimized Ce_3O_3N modifications obtained from global search and data mining calculated within LDA approximation.

Modification	Origin	Space group	ΔE (Eh/f.u.)	ΔE (kcal/mol)
Ce_3O_3N-*DM1*	Data mining	*R*3*c* (161)	0	0
Ce_3O_3N-*GS1*	Global search	*P*2/*m* (10)	0.0062	3.89
Ce_3O_3N-*GS2*	Global search	*Amm*2 (38)	0.0190	11.92
Ce_3O_3N-*GS3*	Global search	*Imm*2 (44)	0.0248	15.56
Ce_3O_3N-*GS4*	Global search	*Pmmm* (47)	0.0287	18.01
Ce_3O_3N-*GS5*	Global search	*Amm*2 (38)	0.0299	18.76
Ce_3O_3N-*GS6*	Global search	*Pmmm* (47)	0.0366	22.97
Ce_3O_3N-*GS7*	Global search	*Pm*-3*m* (221)	0.0481	30.18
Ce_3O_3N-*DM2*	Data mining	*P*6$_3$/*m* (176)	0.0868	54.47
Ce_3O_3N-*DM3*	Data mining	*I*-43*m* (217)	0.2038	127.89

A very important result, which stresses the importance of DFT calculations and a broad structural landscape exploration, is that the Ce_3O_3N-*DM1* ground state is not a simple rescaled form of its data mining starting prototype. Although it originates from an Ag_3AsS_3 (proustite) prototype, local optimization transforms the initial structure into a structurally different *R*3*c* equilibrium structure. In the optimized structure, cerium has coordination number eight, with six oxygen and two nitrogen atoms as nearest neighbors, and a mean Ce–anion distance of approximately 2.48 Å. This coordination is notable because it differs from the six- and sevenfold motifs dominating the lowest-energy global-search structures. The fact that this denser eightfold environment becomes the ground state, while the empirical global search landscape more often favors lower coordination, may indicate that the ionic-radius balance in the empirical potential underestimates the relative stability of denser Ce-centered environments [10]. Figure 5.1 shows both the optimized DM1 structure and the *proustite*-type starting structure from which it emerged.

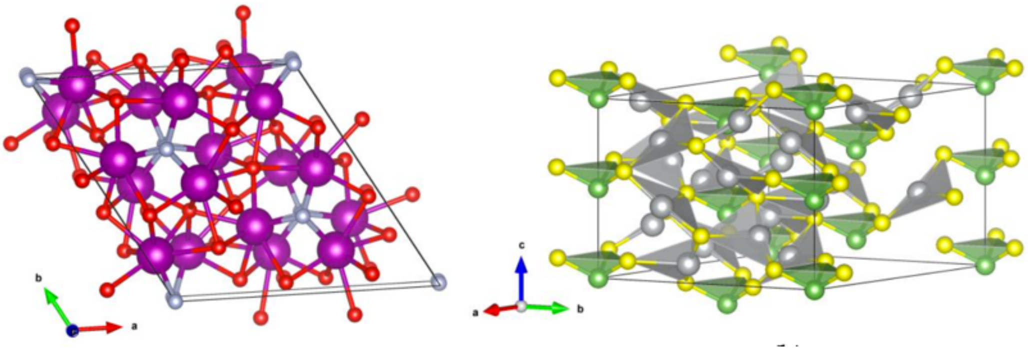

Figure 5.1. Optimized Ce_3O_3N-*DM1* ground-state candidate (left) together with the initial *proustite*-type starting model from which it emerged after local relaxation (right).

The lowest-energy structure found by global optimization is the Ce_3O_3N-*GS1* modification in the monoclinic space group *P*2/*m* (No. 10). This phase is the second-lowest-energy minimum overall, but it is also structurally important because it occupies the low-density side of the landscape and is the phase expected to become accessible under effective negative pressure. Such an expanded-volume metastable structure might be experimentally approached through routes that favor low-density or kinetically trapped products, for example crystallization from an amorphous phase deposited from the gas phase [10, 163]. Its cerium local environments are mixed: one Ce site has coordination number seven, whereas the second can be viewed as a distorted octahedral environment with coordination number six or, because of one longer Ce–O contact, as a 6 + 1 coordination.

The Ce_3O_3N-*GS2* modification, which crystallizes in the orthorhombic *Amm*2 space group (No. 38), is the third-lowest-energy structure candidate and the most stable orthorhombic modification. Here the cerium atoms are in distorted octahedral coordination, but the local nitrogen (and oxygen) content differs between the two octahedra. One octahedron contains one nitrogen atom and has a mean Ce–anion distance of approximately 2.36 Å, whereas the other contains two nitrogen atoms and has a larger mean Ce–anion distance of approximately 2.42 Å. This illustrates one of the recurring themes of the Ce_3O_3N landscape: even when the formal coordination number stays fixed, the oxygen/nitrogen distribution within the coordination shell can still sometimes change substantially, and this in turn modifies both equilibrium volume and total energy [10]. Figure 5.2 shows GS1 and GS2 structures, calculated using DFT with LDA functional.

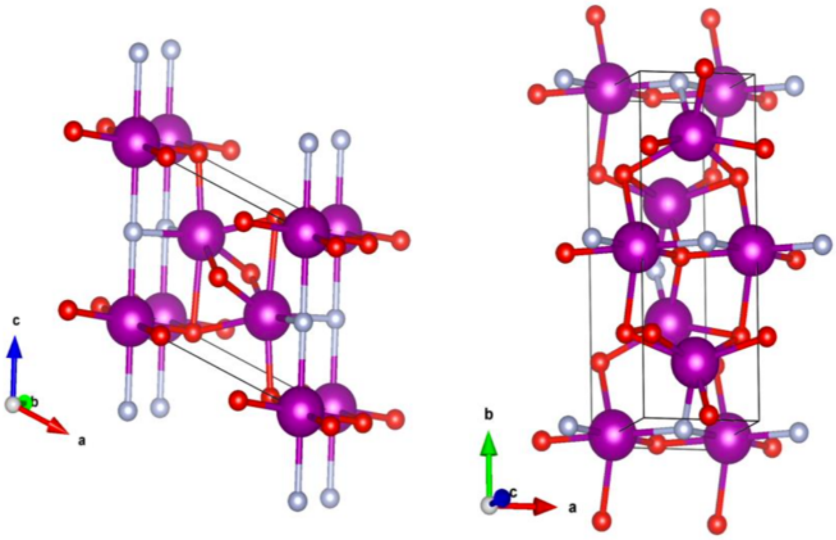

Figure 5.2. Structures of the low-energy global-search candidates Ce_3O_3N-*GS1* (left) and Ce_3O_3N-*GS2* (right).

The remaining low-energy orthorhombic candidates Ce_3O_3N-*GS3*, *GS4*, and *GS5*, show structural motifs that differ from one another and from the *GS2* phase (Figure 5.3). The Ce_3O_3N-*GS3* phase in space group *Imm*2 (No. 44) and Ce_3O_3N-*GS5* in *Amm*2 (No. 38) both contain mixtures of six- and sevenfold coordinated Ce local environments and appear as closely related polytype-like structures [10]. Their principal difference lies in the way the corresponding polyhedra are connected. In both phases, the separation of layers dominated by oxygen from layers containing mostly nitrogen is pronounced, but *GS5* differs from *GS3* in the connectivity between the 6- and 7-fold Ce-centered polyhedra from one layer to the other. The Ce_3O_3N-*GS4* modification (space group *Pmmm*, No. 47), by contrast, contains only octahedral Ce coordination with six nearest neighbors and represents a structurally simpler metastable orthorhombic structure type [10].

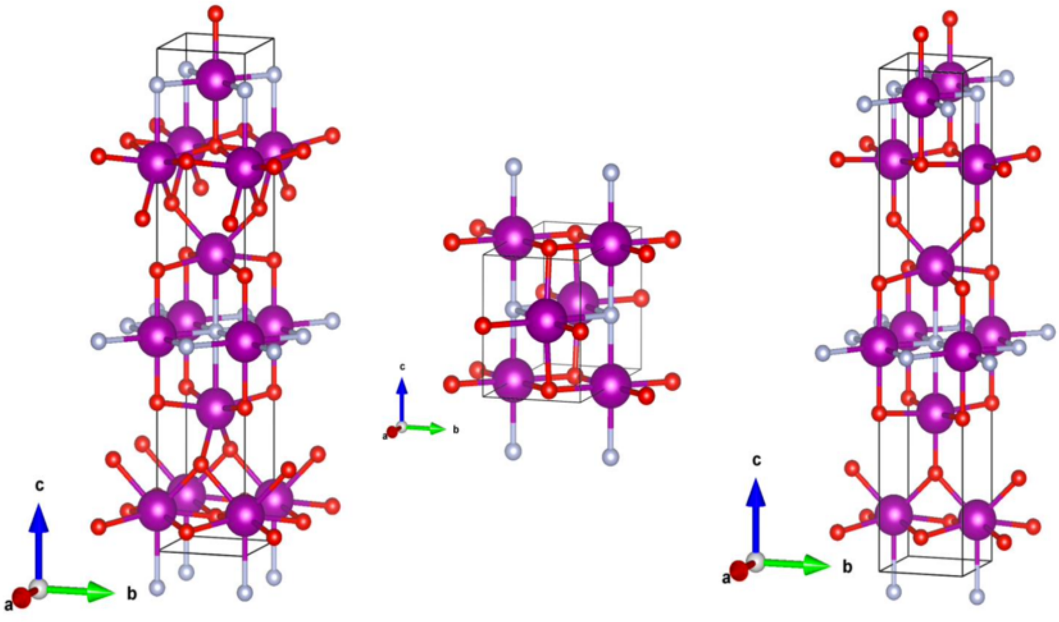

Figure 5.3. Orthorhombic low-energy global-search candidates Ce_3O_3N-*GS3* (left), Ce_3O_3N-*GS4* (middle), and Ce_3O_3N-*GS5* (right).

The higher-energy structures Ce_3O_3N-*GS6*, *GS7*, *DM2*, and *DM3* are not part of the main low-energy set considered in the thermodynamic analysis, but they remain structurally informative. The cubic Ce_3O_3N-*GS7* phase in space group *Pm*-3*m* (No. 221) is notable because local optimization of several distinct global-search candidates, initially exhibiting space group *R*3 (No. 146), *R*3*m* (No. 160), *P*4*mm* (No. 99), *P*432 (No. 207), and *R*32 (No. 155) symmetries, all converged to the same cubic GS7 modification. In this sense, the GS7 structural neighborhood represents a broad basin of attraction on the empirical energy landscape rather than an isolated high-symmetry local minimum. Since its relevant precursor structures appear mostly in the highest-pressure part of the search, it indicates that *Pm*-3*m* may be potentially relevant under extreme high-pressure conditions [10]. The data-mining candidates *DM2* and *DM3* remain too high in energy to compete with the main group of low-energy Ce_3O_3N modifications under the pressure or temperature conditions considered here. The crystallographic data of all the low-energy structures are shown in Appendix A5 Tables A5.1 and A5.2.

The literature-proposed $P2_1$ structure candidate [143] was compared with the two lowest-energy structures Ce_3O_3N-*DM1* and Ce_3O_3N-*GS1*, using VASP calculations within the LDA approximation. The VASP calculations resulted in the following energetic ordering: Ce_3O_3N-*DM1* is the lowest-energy structure, Ce_3O_3N-*GS1* lies higher, and the proposed literature $P2_1$ model [143] is highest in energy (Table 5.3). Thus, the plane-wave calculations additionally support the conclusion that Ce_3O_3N-*DM1* is the predicted ground-state structure of Ce_3O_3N [10].

Table 5.3. The total-energy comparison of the two lowest-energy Ce_3O_3N modifications with the literature proposed $P2_1$ structure, calculated within LDA approximation.

Modification	Space group (No.)	Total energy (eV)
Ce_3O_3N-*DM1*	*R*3*c* (161)	-65.8541
Ce_3O_3N-*GS1*	*P*2/*m* (10)	-65.7261
Literature [143]	$P2_1$ (4)	-65.5338

5.4 Thermodynamic Stability and Pressure Response

The thermodynamic stability of Ce_3O_3N was analyzed using the LDA energy-vs-volume, *E(V)*, curves of the most relevant low-energy modifications. These curves were sufficient to identify the principal pressure trends: the dense Ce_3O_3N-*DM1* structure remains the favored modification at standard conditions and under compression, whereas the expanded Ce_3O_3N-*GS1* structure becomes the only realistic competitor on the effective negative-pressure side of the landscape.

5.4.1 The *E(V)* Equations of State

The *E(V)* curves in Figure 5.4 show the six lowest-energy Ce_3O_3N structure candidates separating into three groups: the high-density Ce_3O_3N-*DM1* phase, which is the global minimum of the landscape; the Ce_3O_3N-*GS1* phase, which is the most favorable effective-negative-pressure modification and represents the principal low-density structure type; and the orthorhombic family *GS2*–*GS5* at much higher energies. Within the investigated volume range, *DM1* modification is the preferred structure at ambient conditions and under compression, whereas the *GS1* phase is the only low-density competitor that approaches stability on the expanded-volume side of the landscape. The orthorhombic *GS2*–*GS5* structures remain metastable throughout the volume range, and are best interpreted as higher-temperature metastable modifications.

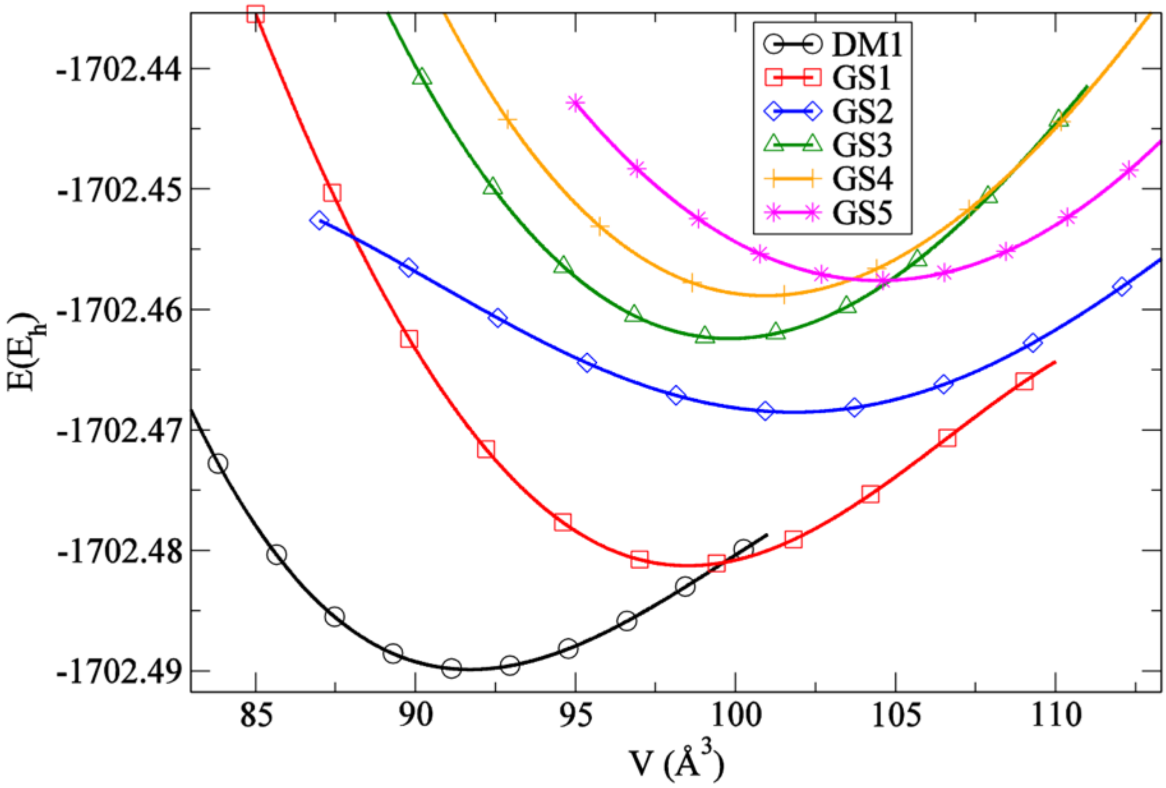

Figure 5.4. Energy-vs-volume curves for the six lowest-energy Ce_3O_3N modifications, calculated using LDA functional.

5.5 Magnetic Ordering in the Ce_3O_3N Ground State Structure

Cerium-containing mixed-anion compounds are natural candidates for magnetic calculations because the magnetically relevant local configuration is usually close to Ce^{3+} cation, for which the 4*f* shell remains partially occupied. The electronic configuration of neutral Ce can be written as $[Xe]4f^1\, 5d^1\, 6s^2$, while in the ideal trivalent ionic limit the two 6*s* electrons and one 5*d* electron are formally removed first, leaving Ce^{3+} in a $[Xe]4f^1$ configuration. Thus, forming a Ce^{3+} ion does not eliminate the 4*f* electron that carries the local spin moment. In a crystal, this notation should be interpreted as a useful local-configuration picture where the Ce 4*f* electron can remain largely localized, whereas the more spatially extended Ce 5*d* and 6*s* states participate more strongly in bonding [61, 62].

The calculations performed here are scalar-collinear VASP/PBE+U calculations without spin-orbit coupling for the ground-state Ce_3O_3N-*DM1* structure. They determine the exchange-driven energetic preference among ferromagnetic (FM), ferrimagnetic (FiM), and antiferromagnetic (AFM) collinear spin configurations. Because the tested collinear configurations are separated only by a few meV per relatively large unit cell, these calculations do not establish a robust magnetic ordering at this energy scale. More computationally demanding non-collinear DFT+U spin-orbit coupling (SOC) calculations, needed for calculation of orbital moments, anisotropy, and possible spin canting, were therefore not performed at this stage.

The Ce_3O_3N publication [10] is a non-spin-polarized CRYSTAL17 structure-prediction study using local basis set and LDA approximation, while the spin polarized VASP [53-55] PBE DFT+U [49, 66] results discussed below are additional calculations for the Ce_3O_3N-*DM1* ground-state modification. The non-spin-polarized LDA calculations are appropriate for exploring the structural energy landscape, whereas the correlated spin-polarized treatment is required to obtain a physically meaningful electronic structure for the system with localized Ce 4*f* states.

The magnetic configurations with collinear spins of Ce atoms were explored for the Ce_3O_3N-*DM1* ground-state phase using six formula units per unit cell, containing a total of eighteen Ce, eighteen O, and six N atoms. Three initial spin arrangements were tested under the same PBE+U scalar-collinear spin polarization setup: a compensated antiferromagnetic (AFM) arrangement, a ferrimagnetic (FiM) arrangement with twelve spin-up and six spin-down Ce sites, and the ferromagnetic (FM) state. The Dudarev correction was applied to the Ce 4*f* states with $U_{eff} = 5.0$ eV. The AFM configuration is the lowest-energy solution, while the FiM and FM states lie 2.76 and 3.38 meV per unit cell above it, respectively (Table 5.4). The FiM and FM arrangements are only 0.46 and 0.56 meV per formula unit higher than the AFM state, respectively. The AFM configuration is the lowest-energy state but the exchange-energy scale is very small, under one Kelvin per atom, so that the small thermal effects or even the spin-orbit coupling treatment could change the magnetic ordering.

The site-projected local magnetic moments on Ce are close to ± 0.983 μ_B in all three arrangements, which is consistent with a localized Ce $4f^1$ spin polarization. The total cell magnetization is 0 μ_B for the AFM state, 6 μ_B for the FiM state, and 18 μ_B for the FM state, corresponding to 0, 1, and 3 μ_B per Ce_3O_3N formula unit. The small difference between the total cell magnetization and the sum of the Ce PAW-sphere-projected local moments in the FM and FiM cases arises because not all spin density is contained inside the finite Ce PAW spheres; the remaining part is distributed in the interstitial region and as small induced polarization on the O and N atoms.

Table 5.4. Results for the collinear PBE+U magnetic configurations of the Ce_3O_3N-*DM1* ground-state modification. The ΔE is given relative to the lowest-energy AFM arrangement; the unit cell contains six Ce_3O_3N formula units. The initial Ce spin patterns are labeled by up/down arrows for spin-up/spin-down Ce states.

Initial Ce spin pattern	**E (eV/cell)**	**ΔE (meV/cell)**	**M_{tot} (μ_B/cell)**	**Ce moments (μ_B)**	**E_{gap} (eV)**
↑↑↑↓↓↓↑↑↑↓↓↓↑↑↑↓↓↓ (AFM)	-345.477485	0.00	0.00	+0.983 (×9), -0.983 (×9)	2.21
↑↑↑↑↑↑↓↓↓↓↓↓↑↑↑↑↑↑ (FiM)	-345.474722	2.76	6.00	+0.983 (×12), -0.983 (×6)	2.10
↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑↑ (FM)	-345.474104	3.38	18.00	+0.983 (×18)	2.02

The antiferromagnetic spin orientations shown in Figure 5.5 represent the lowest-energy scalar-collinear state used for the subsequent electronic structure calculation. The arrows represent scalar-collinear spin directions and are not fully relaxed magnetic-moment vectors with a determined easy axis, i.e. the crystallographic direction along which a magnetic moment aligns to achieve minimum energy. Determining whether the Ce_3O_3N-*DM1* phase has an easy axis, easy plane, or more complex anisotropy would require non-collinear DFT+U SOC calculations with many different initial moment directions, analogous to the HoFSe calculations presented in Chapter 7.

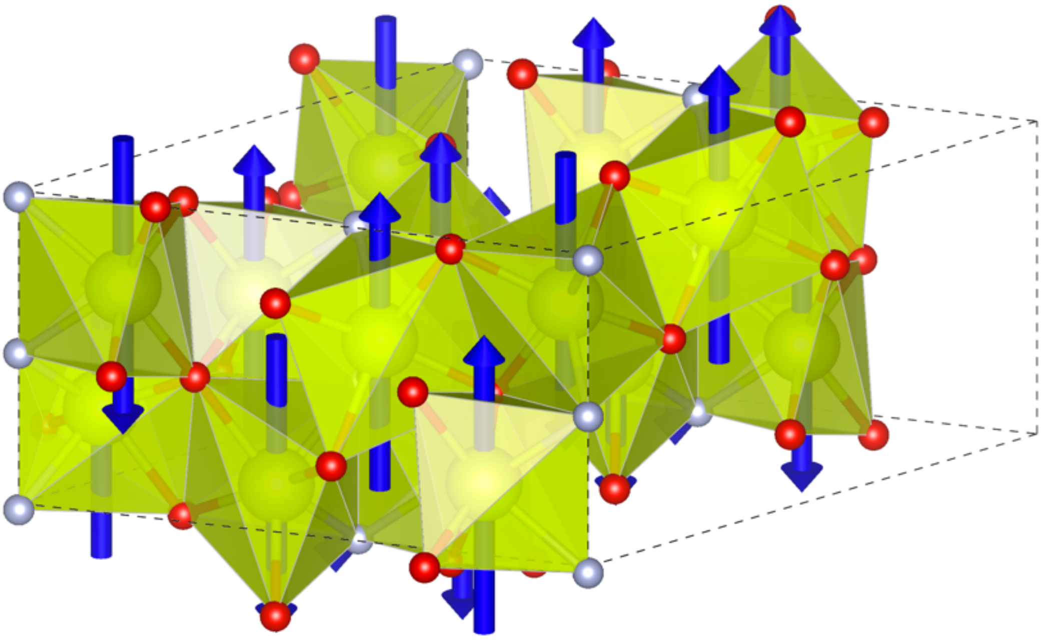

Figure 5.5. Magnetic structure of the minimum-energy collinear AFM configuration of Ce_3O_3N-*DM1* phase. The Ce spin directions are shown as blue arrows. The Ce, O, and N atoms are shown in green, red, and light grey, respectively.

5.6 Electronic Structure of the Ce_3O_3N Ground State Structure

Mixed-anion oxynitrides and nitrogen-containing oxide derivatives often show modified valence- and conduction-band edges relative to the corresponding oxides because nitrogen incorporation can raise the upper valence bands [6, 138, 139]. This trend follows from the fact that nitrogen 2*p* states are usually less strongly bound than oxygen 2*p* states, reflecting the lower electronegativity of nitrogen, although the exact quantitative effect in a particular compound depends on orbital mixing, coordination geometry, and the position of cation-derived states, i.e., bands with dominant contributions from Ce orbitals such as Ce 4*f* or Ce 5*d*. In Ce_3O_3N, the situation is more complex because correlated Ce 4*f* states can lie near the gap region and their placement is sensitive to the Ce local environment, spin polarization, and the choice of Hubbard correction [62, 66, 142].

The nonmagnetic calculations are suitable for exploring the structural landscape of the Ce_3O_3N crystal system. However, the electronic structure requires a correlated spin-polarized treatment of the Ce 4*f* states: the non-spin-polarized PBE calculation for the ground state Ce_3O_3N-*DM1* modification gives a metallic band structure with Ce 4*f* states crossing the Fermi level (Figure 5.7), whereas spin-polarized PBE+U calculations with U_{eff} = 5.0 eV open a semiconducting gap for all three tested collinear magnetic configurations (AFM, FiM, and FM), shown in Figures 5.6, 5.8, and 5.9. The transition from a metallic non-spin-polarized PBE result to semiconducting PBE+U results is primarily a consequence of applying the Hubbard-U correction to the Ce 4*f* states (and even the U_{eff} as small as 1 eV opens the band gap), while the magnetic arrangement controls the detailed spin splitting and the precise gap size.

The AFM solution is an indirect-gap semiconductor where the valence-band maximum lies at Γ, whereas the conduction-band minimum occurs at the L point on the plotted high-symmetry path, resulting in an indirect band gap of 2.21 eV (Table 5.4).

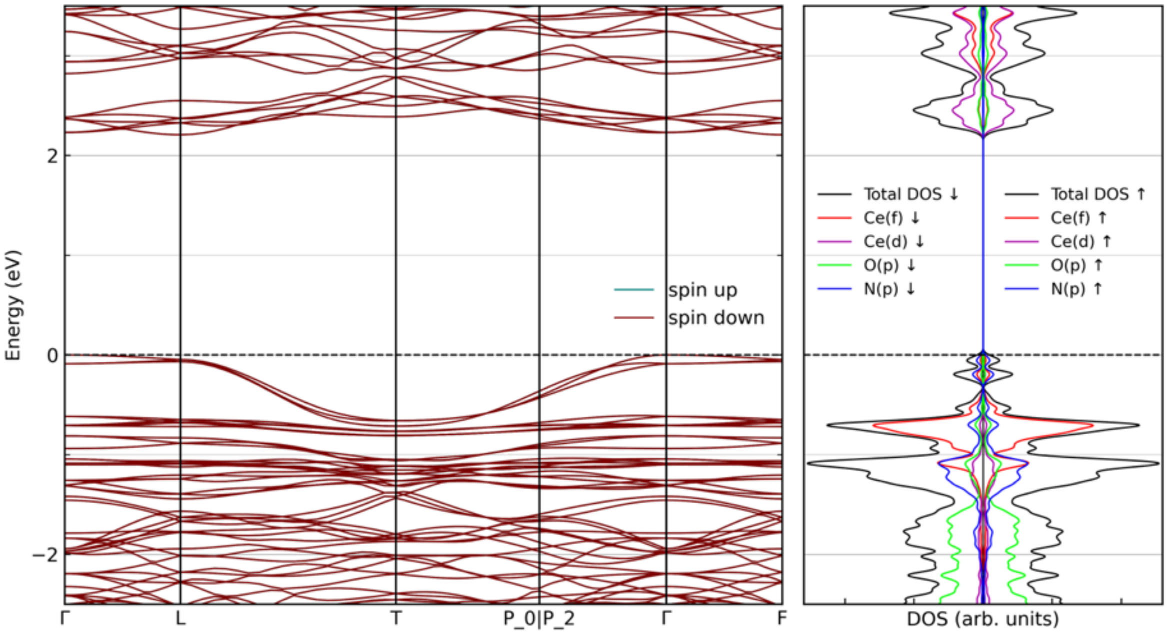

Figure 5.6. Band structure and projected DOS of the minimum-energy AFM configuration of the Ce_3O_3N-*DM1* structure, calculated using PBE+U with $U_{eff} = 5.0$ eV for the Ce 4*f* states.

The projected DOS indicates that the lowest conduction-band states have substantial Ce 5*d* character, while narrow unoccupied states with strong Ce 4*f* character occur higher in the conduction region. The upper valence states are dominated mainly by anion *p* contributions, with a smaller Ce 4*f* contribution near the edge of the gap. This is physically consistent with a DFT+U treatment of a Ce^{3+} $4f^1$ system: the on-site U localizes and shifts the Ce 4*f*-derived states, while the immediate band edges retain mixed Ce 5*d* and anion *p* character.

In contrast to the metallic non-spin-polarized PBE result, all three spin-polarized PBE+U solutions are semiconducting because of the Hubbard U treatment, which separates the Ce 4*f* states. Changing the magnetic arrangement from AFM to FiM and FM reduces the band gap from 2.21 to 2.10 and 2.02 eV, respectively, while the FiM and FM states remain only 2.76 and 3.38 meV per cell above the AFM state. The FiM and FM band structures in Figures 5.8 and 5.9 show the same qualitative near-edge features as the AFM solution, but with more pronounced shifts between spin-up and spin-down states as the net magnetization increases.

Figures 5.7–5.9 complete the comparison between the non-spin-polarized PBE calculation and the spin-polarized PBE+U magnetic configurations. The plain PBE result for the electronic band structure in Figure 5.7 shows metallic behavior because narrow Ce 4*f*-derived bands cross the Fermi level, confirming that a semilocal non-spin-polarized treatment is not sufficient for the Ce 4*f* band structure calculations. By contrast, the FiM and FM PBE+U band structures in Figures 5.8 and 5.9

are semiconducting, with the similar band-edge description as the AFM ground state: anion-p valence states, Ce-derived conduction states, and narrow Ce 4*f* bands shifted away from the Fermi level. The small but non-negligible narrowing of the band gap appears as the arrangement changes from the AFM to the FiM and finally to the FM configurations.

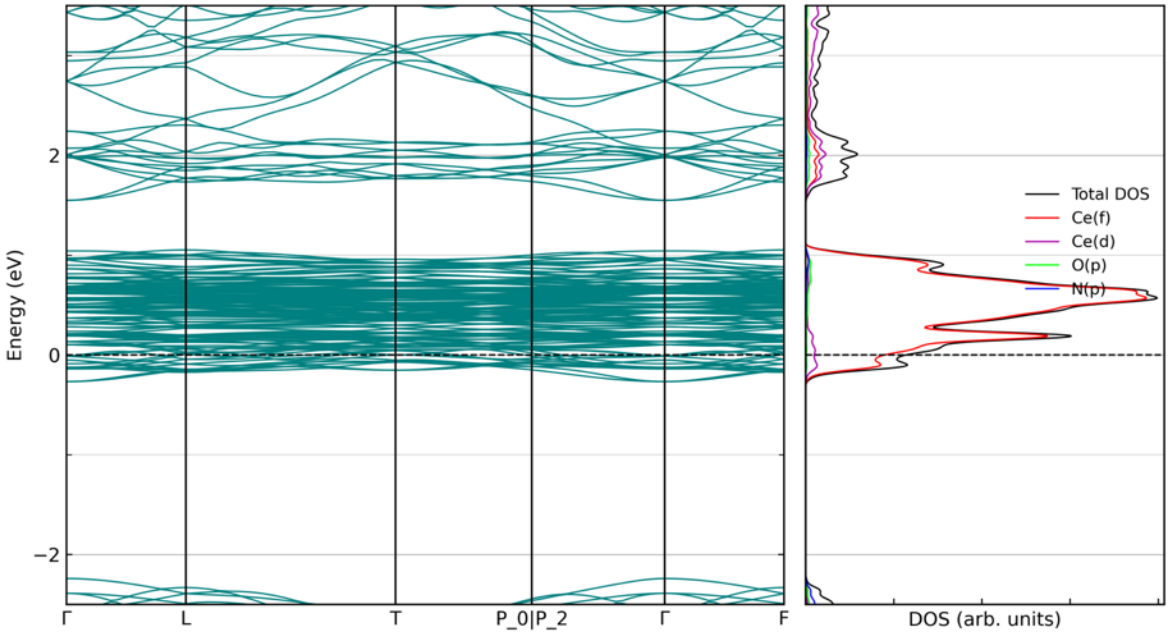

Figure 5.7. Band structure and projected DOS of the Ce_3O_3N-*DM1* structure, calculated using the non-spin-polarized DFT calculation with PBE functional.

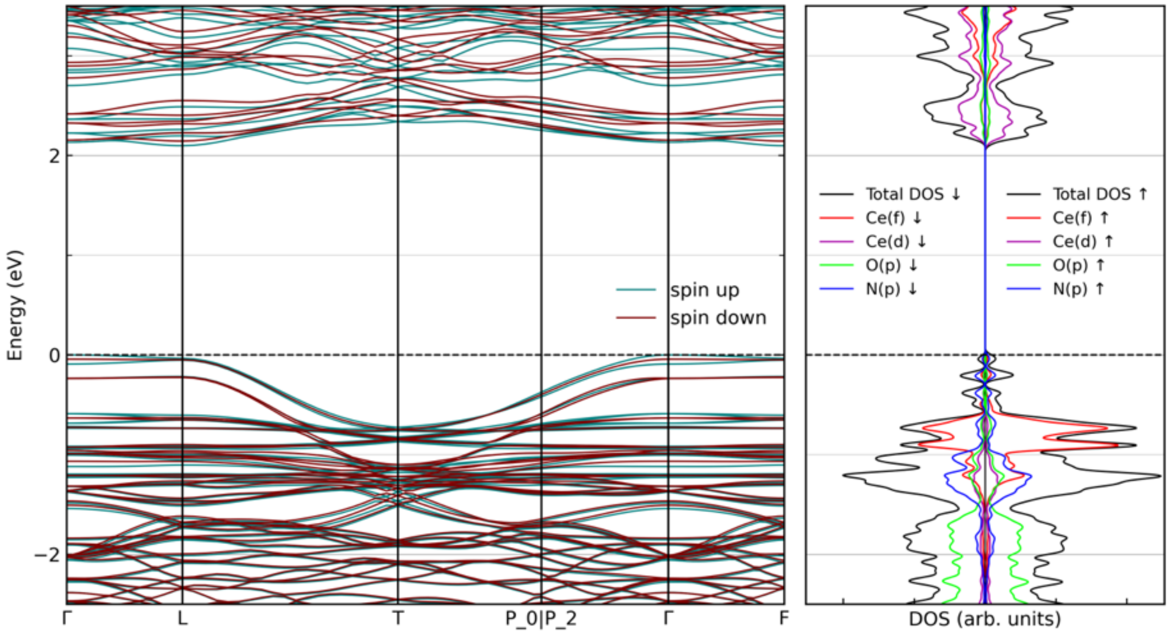

Figure 5.8. Band structure and projected DOS of the ferrimagnetic (FiM) configuration of the Ce_3O_3N-*DM1* structure, calculated using PBE+U with $U_{eff} = 5.0$ eV for the Ce 4*f* states.

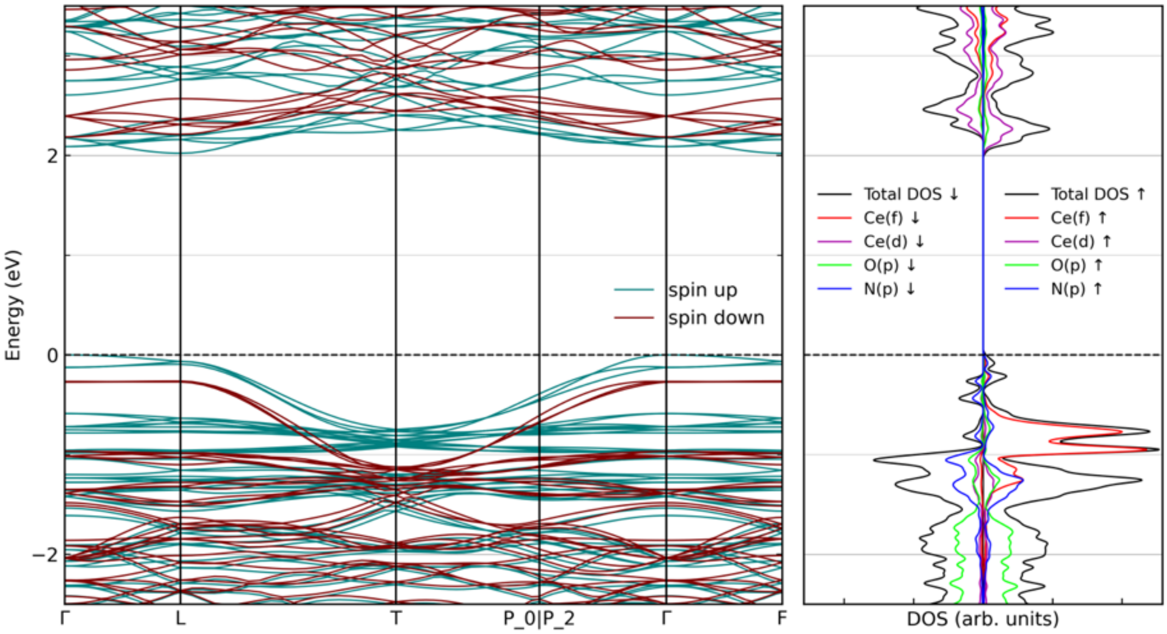

Figure 5.9. Band structure and projected DOS of the ferromagnetic (FM) configuration of the Ce_3O_3N-*DM1* structure, calculated using PBE+U with $U_{eff} = 5.0$ eV for the Ce 4*f* states.

5.7 Conclusion

The hierarchy of ground-state energies summarized in Table 5.2 shows the importance of complementary structure-search methods in Ce_3O_3N energy landscape exploration. The global minimum, Ce_3O_3N-*DM1*, did not appear among the global optimization structures, but it emerged after the *ab initio* relaxation of a data-mined proustite structure type. In the Ce_3O_3N system's structural exploration, global optimization efficiently sampled broad low-energy basins on the empirical energy landscape, while data mining introduced crystallographically diverse prototypes that can relax into structure types not easily reached by the global search alone, thus enabling a more complete search of the Ce_3O_3N energy landscape.

Taken together, Table 5.2 and Figures 5.1–5.3 show that even the small relative-energy interval separating the most relevant Ce_3O_3N candidates corresponds to clearly different Ce coordinations, local Ce oxygen–nitrogen distributions, and layer- and polyhedra-connectivity patterns. In particular, the contrast between the dense eightfold-coordinated *DM1* ground state, the low-density monoclinic *GS1* phase, and the orthorhombic *GS2–GS5* family makes it clear that the low-energy landscape in Ce_3O_3N is defined by several genuinely different structure types, rather than by small distortions of one structural arrangement.

The thermodynamic analysis shows that the ground state Ce_3O_3N-*DM1* is also the dominant modification at ambient conditions and under compression. On the expanded volume side of the landscape, Ce_3O_3N-*GS1* becomes the most relevant structure that can be interpreted as the effective-negative-pressure structure candidate, whereas the orthorhombic *GS2–GS5* family remains metastable throughout the investigated volume range.

The scalar-collinear magnetic calculations identify Ce_3O_3N-*DM1* as an antiferromagnetic semiconductor within the PBE+U description. Among the three tested spin arrangements, the AFM state with an indirect Γ–L band gap of 2.21 eV is slightly lower in energy than the FiM and FM configurations. The small energy separation between the explored magnetic states indicates that the magnetic energy landscape is shallow so thermal effects on this scale, additional spin patterns, changes in U_{eff}, or spin-orbit coupling could modify the detailed magnetic ordering. Even though the exploration of the Ce_3O_3N system would further benefit from a full treatment of magnetic ordering using additional spin patterns and non-collinear DFT+U SOC calculations, the present calculations clearly show that a correlated spin-polarized DFT+U treatment of the Ce 4*f* states is required to obtain a physically meaningful electronic structure for the ground-state modification.

6. Structure Prediction, Pressure-Dependent Stability, and Electronic Structure of Scandium Oxychloride (ScOCl)

6.1 Introduction

Scandium oxychloride (ScOCl) has recently attracted scientific and technological interest for several reasons. Scandium-containing materials are relevant as stabilizing components in zirconia-based solid oxide fuel-cell electrolytes [164], while layered transition-metal oxyhalides have been discussed because of their magnetic and electronic properties and their possible future use in photocatalytic and electronic applications [165, 166]. The only experimentally observed ScOCl modification crystallizes in the orthorhombic *Pmmn* (No. 59) space group [167-169]. This structure consists of layers of edge-connected ScO_4Cl_2 octahedra, with oxygen anions inside the layers and chlorine anions pointing outward, leading to van der Waals-type interactions between neighboring layers [168].

This chapter follows the published structure-prediction study of the ScOCl system [9] and extends it with additional electronic-structure calculations for the two most relevant low-energy phases. The discussion includes the computational setup and results of the energy-landscape exploration, the optimized geometry and structure-type analysis of the low-energy ScOCl modifications, the thermodynamic stability and pressure response derived from the calculated $E(V)$ and $H(p)$ curves, and electronic-structure calculations for the two most relevant phases.

6.2 Computational Setup and Energy Landscape Exploration Overview

The ScOCl energy landscape was explored with a multi-methodological approach that combined global optimization on the empirical level, data mining, and local optimization on the *ab initio* level [9]. In the first stage, simulated annealing with periodic stochastic quenches was carried out without symmetry restrictions in order to sample a large number of local minima [13, 93]. A fast two-body empirical potential, containing Lennard-Jones terms together with exponentially damped Coulomb contributions, was used so that extensive searches could be performed at reasonable computational cost [9]. In the second stage, data mining over the ICSD database was used to identify additional ABX-type structure candidates by analogy with known crystal structures [100, 140, 170].

After combining the global search and the data-mining step, more than 5000 ScOCl structure candidates had been collected [9]. Most candidates were low-symmetry variants or closely related representatives of a smaller number of underlying structure types rather than clearly distinct polymorphs. The α- and β-modifications were recovered repeatedly as local minima during the global optimization and were also found in the data-mining stage, whereas the γ- and δ-modifications were obtained only from global optimization, together with many lower-symmetry variants. The most promising candidates were then selected for full *ab initio* structural relaxation without symmetry restrictions in CRYSTAL17 [51, 52], using the LDA-PZ and hybrid PBE0 exchange-correlation functionals [9].

6.3 Optimized Geometry – Symmetry and Structure Type Analysis

After local *ab initio* optimization, four ScOCl modifications remained as the most relevant minima on the energy landscape: the experimentally known α-phase and the predicted β-, γ-, and δ-phases [9]. These four modifications define the main structure types that emerged from the ScOCl structural landscape exploration (Table 6.1).

Table 6.1. Summary of the four most relevant ScOCl polymorphs after local optimization with CRYSTAL17.

Modification	Structure type / space group	Erel, LDA (kJ/mol / eV)	Erel, PBE0 (kJ/mol / eV)
α-ScOCl	*FeOCl* type, Orthorhombic *Pmmn* (59)	0.0	0.0
β-ScOCl	*PbFCl* type, Tetragonal *P4/nmm* (129)	8.9 / 0.092	36.5 / 0.378
γ-ScOCl	Tetragonal $P4_2/m$ (84)	24.7 / 0.256	21.3 / 0.221
δ-ScOCl	Monoclinic *C*2/*m* (12)	42.8 / 0.444	32.8 / 0.340

The α-ScOCl phase is the ambient-pressure ground state. It crystallizes in the orthorhombic space group *Pmmn* (No. 59) and adopts the layered *FeOCl*-type arrangement [9, 167-169]. In the distorted ScO_4Cl_2 octahedra, each scandium atom is coordinated by four oxygen and two chlorine atoms. Two Sc-O distances are 2.0534 Å, two others are 2.1184 Å, and the Sc–Cl distance is 2.4874 Å, calculated at the LDA level. The slabs are formed by sharing O–O and O–Cl edges and are stacked along the c direction (Figure 6.1). This description is consistent with the experimental structural characterization of ScOCl and the more recent theoretical results for this phase [9, 167-169].

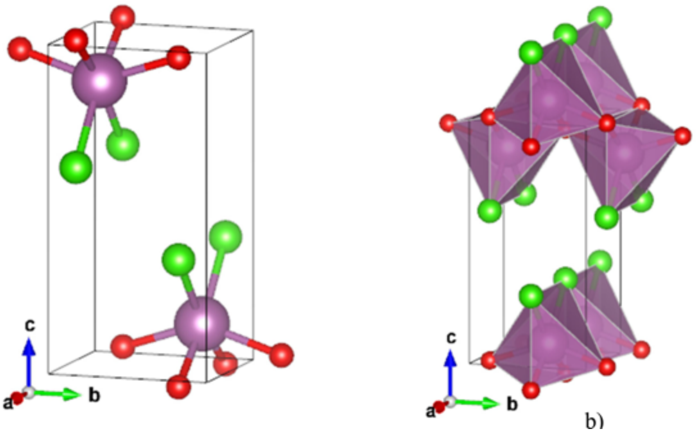

Figure 6.1. Optimized α-ScOCl structure (*Pmmn*) with (right) and without Sc-centered coordination polyhedra (left).

The *β*-ScOCl modification is the theoretically predicted high-pressure alternative (Figure 6.2). It crystallizes in the tetragonal space group *P*4/*nmm* (No. 129) and corresponds to the *PbFCl*, or *matlockite*, structure type [9, 171]. The *β*-phase should be thermodynamically preferred at elevated pressure and low temperature [9]. In this modification scandium is eightfold coordinated by four oxygen and four chlorine atoms, forming ScO4Cl4 polyhedra. The Sc–O distance is 2.06114 Å, whereas the Sc–Cl distance is 2.8741 Å. The oxygen-containing layers are located between chlorine-containing regions, while the coordination increase from six to eight marks the denser character expected for a pressure-stabilized modification. The same matlockite-like type is also known in related oxyhalide compounds, including LaOCl and *PbFCl*-type rare-earth oxyhalides [171, 172].

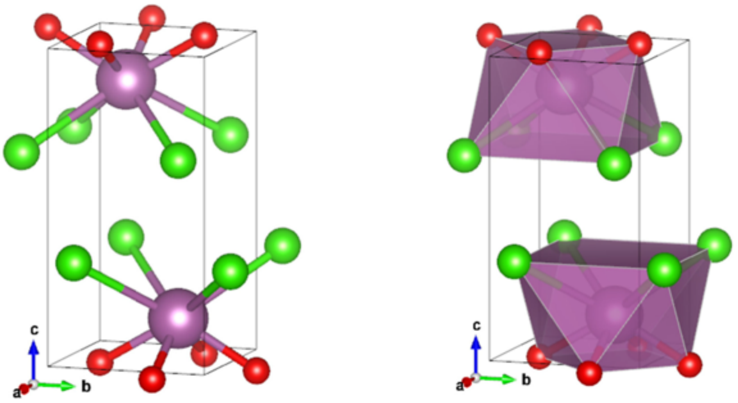

Figure 6.2. Optimized *β*-ScOCl structure (*P*4/*nmm*) with (right) and without Sc-centered coordination polyhedra (left).

The *γ*-ScOCl modification is also tetragonal, but with space group $P4_2/m$ (No. 84). In contrast to *β*-ScOCl, its scandium-centered coordination polyhedra are deformed octahedra connected by a combination of edges and corners, so that the arrangement resembles a distorted rutile or magnesium-fluoride-type framework [9, 173-175]. Each scandium atom is sixfold coordinated by three oxygen and three chlorine atoms. The structure is shown in Figure 6.3.

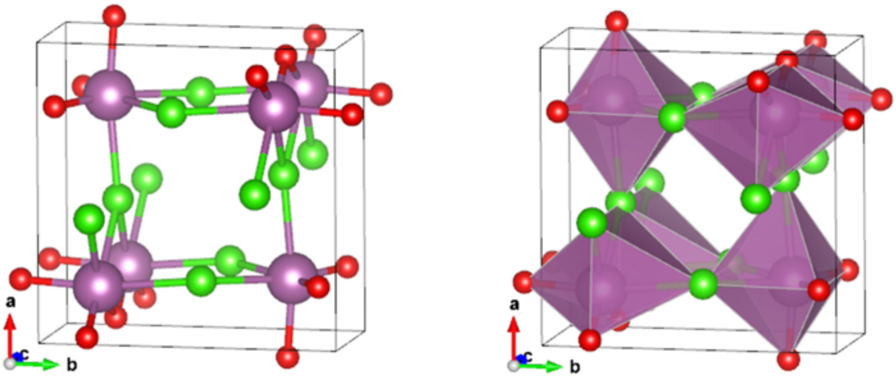

Figure 6.3. Optimized *γ*-ScOCl structure ($P4_2/m$) with (right) and without Sc-centered coordination polyhedra (left).

The *δ*-ScOCl modification is the lowest-symmetry member of the final set and crystallizes in the monoclinic space group *C*2/*m* (No. 12). Its structure contains two kinds of deformed coordination octahedra. In the first one scandium is coordinated by four oxygen and two chlorine atoms, while in the second one it is coordinated by three oxygen and three chlorine atoms [9]. These octahedra are linked by shared faces and corners. Because the predicted δ-ScOCl phase lies substantially above the ground-state α-phase in energy, it is best interpreted as a higher-temperature or non-equilibrium candidate rather than as the principal pressure-stabilized phase. The structure is shown in Figure 6.4.

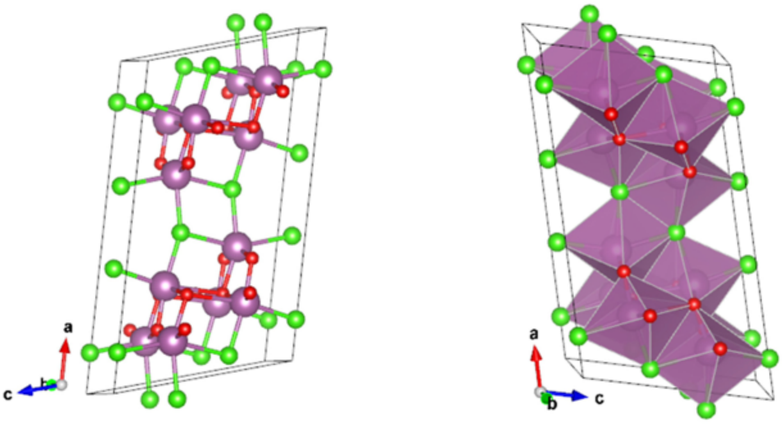

Figure 6.4. Optimized δ-ScOCl structure ($C2/m$) with (right) and without Sc-centered coordination polyhedra (left).

Table 6.2. Selected structural parameters and characteristic distances for the most relevant ScOCl modifications.

Phase	Lattice parameters	Characteristic coordination and Sc-anion distances	Average Sc-anion distance (Å)
α-ScOCl	LDA: a = 3.45, b = 3.90, c = 7.83 PBE0: a = 3.52, b = 3.94, c = 8.81	ScO_4Cl_2 octahedra; Sc-O = 2 × 2.0534 and 2 × 2.1184 Å; Sc-Cl = 2 × 2.4874 Å	2.2198
β-ScOCl	LDA: a = 3.55, c = 6.93 PBE0: a = 3.58, c = 8.08	ScO_4Cl_4 polyhedra; Sc-O = 4 × 2.06114 Å; Sc-Cl = 4 × 2.8741 Å	2.4678
γ-ScOCl	LDA: a = 7.60, c = 3.43 PBE0: a = 7.76, c = 3.47	Deformed ScO_3Cl_3 octahedra; Sc-O = 2 × 2.0014 and 1 × 1.9937 Å; Sc-Cl = 2 × 2.6082 and 1 × 2.6255 Å	2.3064
δ-ScOCl	LDA: a = 14.22, b = 3.92, c = 8.16, β = 106.23° PBE0: a = 15.50, b = 3.95, c = 8.12, β = 106.76°	Two distinct octahedral environments. first polyhedron (ScO_4Cl_2): Sc-O = 2 × 2.0323, 1 × 1.8892, 1 × 2.6287 Å; Sc-Cl = 1 × 2.3777, 1 × 2.4099 Å second polyhedron (ScO_3Cl_3): Sc-O = 2 × 2.0219, 1 × 2.0787 Å; Sc-Cl = 1 × 2.4592, 1 × 2.5327, 1 × 2.6062 Å	2.2283 2.2868

The optimized structures show that the compositionally compact ternary ScOCl system supports several distinct low-energy structure types: a layered orthorhombic ground state, a denser layered high-pressure form, a second tetragonal metastable alternative, and a monoclinic low-symmetry candidate [9]. Selected structural parameters and characteristic distances are summarized in Table 6.2.

6.4 Thermodynamic Stability and Pressure Response

The thermodynamic discussion is based on the relative energies of the optimized structure candidates and the *E(V)* and *H(p)* relations calculated at the LDA and PBE0 levels [9]. In both descriptions, α-ScOCl is the stable ambient-pressure phase. The β- and γ-modifications lie much closer in energy than the δ-phase and therefore dominate the pressure-related competition among the predicted polymorphs.

6.4.1 The *E(V)* Equations of State

The calculated *E(V)* curves show that α-ScOCl has the deepest energy minimum regardless of whether the LDA or hybrid PBE0 functional is used [9]. This confirms the orthorhombic *FeOCl*-type phase as the equilibrium structure at ambient conditions. The *E(V)* curves of β-, γ-, and δ-modifications also exhibit well-defined local minima, which supports their role as structurally meaningful local minima on the calculated energy landscape.

The layered α-phase remains the most stable configuration at standard pressure. From the shape of the *E(V)* curves (Figures 6.5 and 6.6), the role of density becomes clear. The β-phase has its equilibrium minimum in the higher-density (lower-volume) part of the landscape and is therefore favored as pressure is applied [9], which is consistent with the increase of the scandium coordination number from six in α-ScOCl to eight in β-ScOCl. The γ-phase represents a metastable high-pressure alternative with a distorted octahedral Sc environment: it is the closest metastable phase with respect to α-ScOCl at low pressure and with respect to β-ScOCl at sufficiently high pressure. Because the δ-phase lies much higher in energy, it is more likely to be stabilized only under high-temperature not easily obtainable under the usual synthesis conditions.

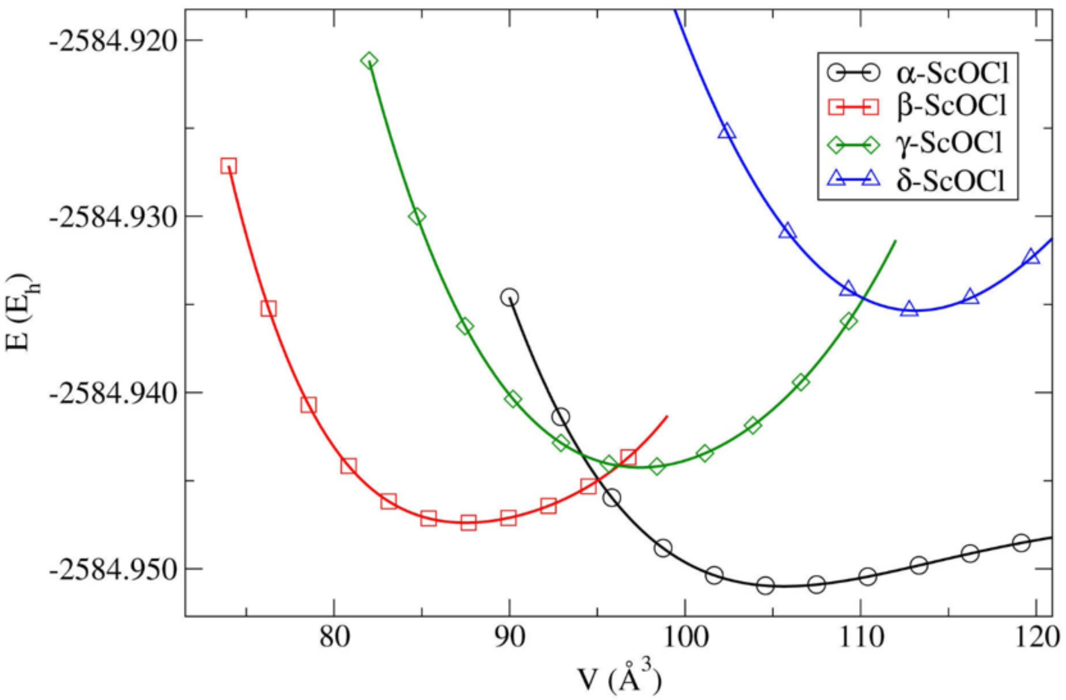

Figure 6.5. Calculated *E(V)* curves for *α*-, *β*-, *γ*-, and *δ*-ScOCl at the LDA level.

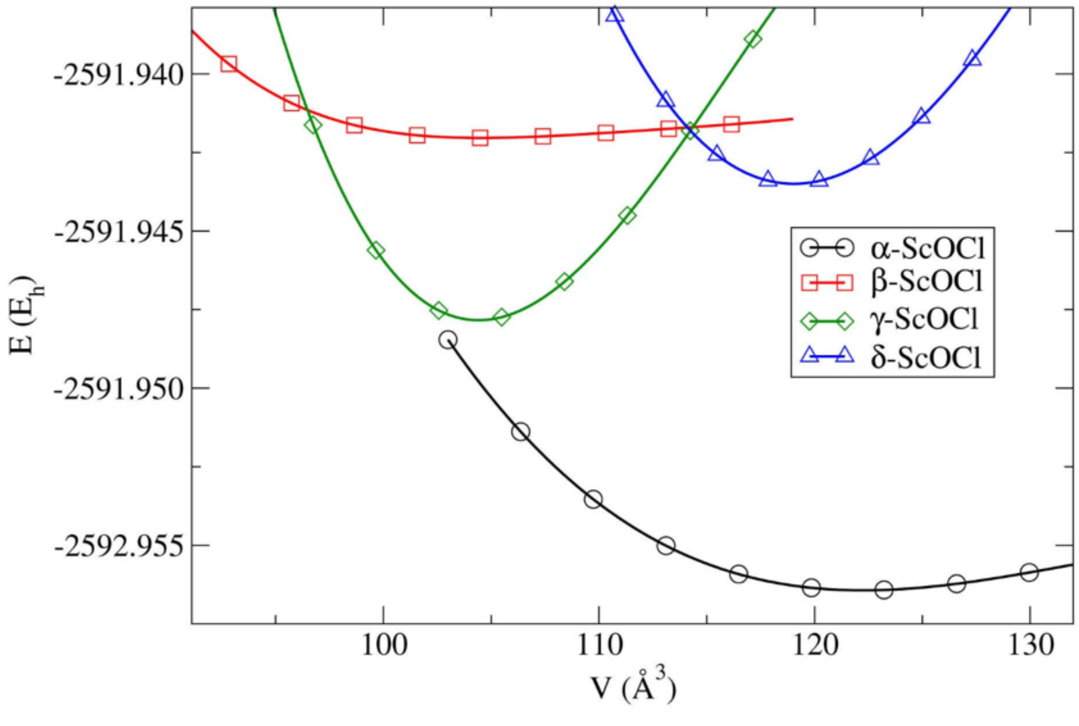

Figure 6.6. Calculated *E(V)* curves for *α*-, *β*-, *γ*-, and *δ*-ScOCl at the PBE0 level.

6.4.2 Enthalpy vs Pressure and Transition Pressures

The pressure dependence becomes more explicit in the calculated enthalpy curves $H(p) = E + pV$. At the LDA level, the $\alpha \rightarrow \beta$ phase transition is predicted at approximately 1 GPa with LDA functional calculations (Figure 6.7), whereas the corresponding transition pressure is approximately 3.7 GPa when the hybrid PBE0 functional is used (Figure 6.8) [9]. In both cases the qualitative result is the same: α-ScOCl is stable at ambient pressure, while β-ScOCl becomes the thermodynamically preferred phase under compression.

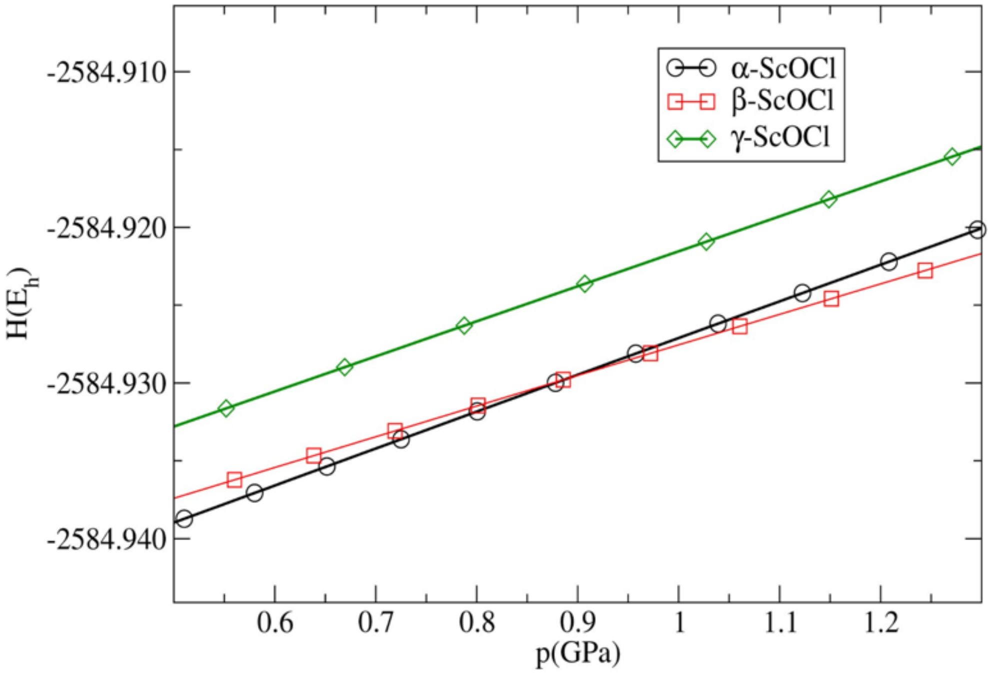

Figure 6.7. Calculated *H(p)* curves for α-, β-, and γ-ScOCl at the LDA level.

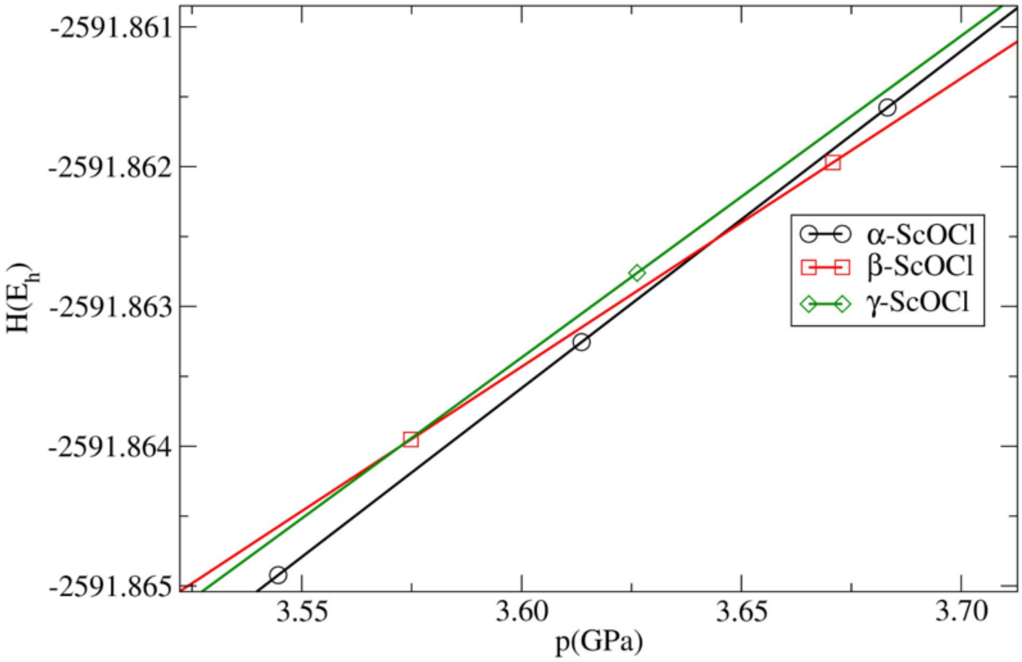

Figure 6.8. Calculated *H(p)* curves for α-, β-, and γ-ScOCl at the PBE0 level.

The γ-phase remains relevant because it stays close in enthalpy to the α- and β-phases and may act as an intermediate metastable modification during compression or decompression [9]. There is a possibility that *γ*-ScOCl could appear as an intermediary phase under pressure, leading to the pressure-induced pathway *α*-ScOCl (stable at standard conditions) → *γ*-ScOCl (metastable) → *β*-ScOCl (stable at high pressures) [9].

The energy difference between *γ*-ScOCl and *α*-ScOCl corresponds to a thermal k_BT scale of approximately 800 K, so moderately high temperatures may be sufficient to make the *γ*-phase accessible. For *δ*-ScOCl, the energy difference to the *α*-phase corresponds to approximately 1500 K, implying that substantially higher temperatures would be required for experimental stabilization [9].

6.5 Electronic Structure of the α-ScOCl and β-ScOCl Phases

The ScOCl paper [9] focuses on structure prediction and thermodynamic analysis, and it does not include electronic structure calculations. This subsection presents the additional band-structure and density-of-states plots of the *α*-ScOCl and *β*-ScOCl modifications, calculated using the VASP software package. These two phases were selected because *α*-ScOCl is the ground state structure at standard pressure and *β*-ScOCl is the principal high-pressure phase identified in the study.

The electronic-structure calculations for the *α*-ScOCl and *β*-ScOCl phases were performed using several exchange-correlation treatments. Table 6.3 shows the numerical band-gap comparison, while Figures 6.9-6.12 show the PBE and HSE06 band structures and projected densities of states emphasized in the discussion. The additional LDA, PBEsol, and PBE0 band structures and projected DOS are shown in Appendix A6 (Figures A6.1-A6.6), and their corresponding band-gap values are included in Table 6.3. Across all calculations, both *α*-ScOCl and *β*-ScOCl remain indirect-gap semiconductors.

Table 6.3. The band-gap comparison of *α*-ScOCl and *β*-ScOCl phases using VASP calculations with different functionals.

Functional	***α*-ScOCl E_{gap} (eV)**	***α*-ScOCl gap character**	***β*-ScOCl E_{gap} (eV)**	***β*-ScOCl gap character**
LDA	4.09	Indirect	3.86	Indirect
PBE	4.04	Indirect	3.90	Indirect
PBEsol	4.07	Indirect	3.87	Indirect
PBE0	6.53	Indirect	6.64	Indirect
HSE06	5.76	Indirect	5.56	Indirect

Figures 6.9 and 6.10 show that *α*-ScOCl is an indirect-gap semiconductor in both the PBE and HSE06 descriptions. The PBE gap is 4.04 eV, while HSE06 increases it to 5.76 eV. The highest occupied bands are relatively flat near the band edge and are dominated by anion *p* states, mostly Cl *p* states with additional O *p* contribution, whereas the lowest unoccupied bands are mainly Sc *d* states.

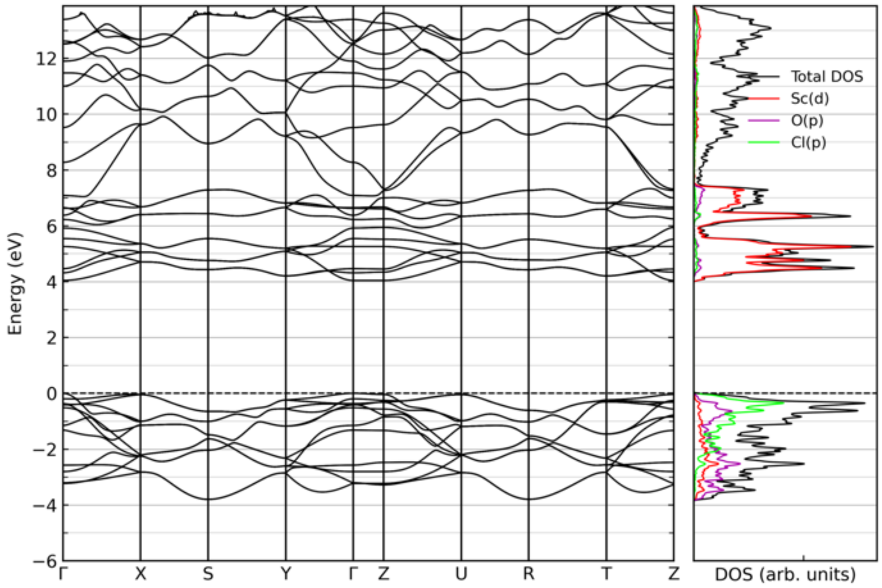

Figure 6.9. Electronic band structure and projected density of states of *α*-ScOCl calculated with PBE functional.

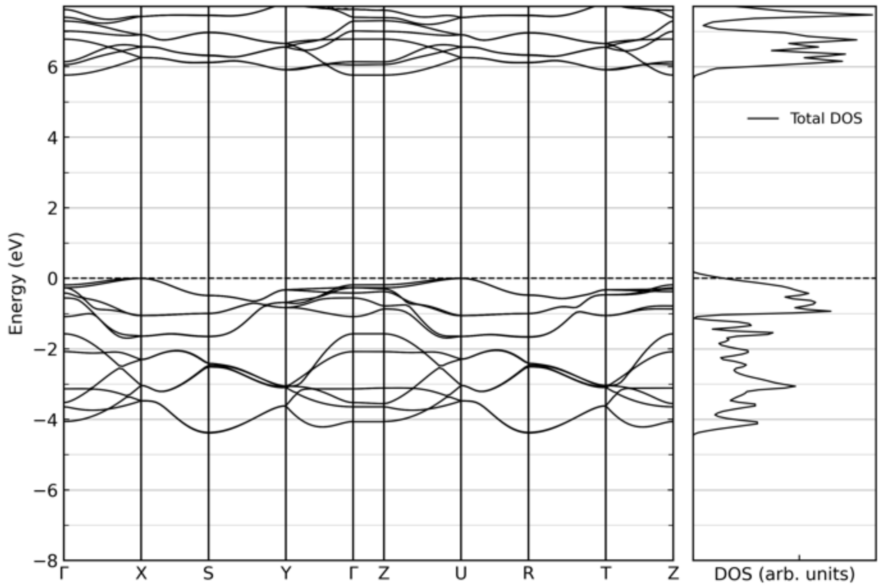

Figure 6.10. Electronic band structure and projected density of states of *α*-ScOCl calculated with HSE06 functional.

Figures 6.11 and 6.12 present the PBE and HSE06 band structures and projected DOS of *β*-ScOCl. The *β*-phase is also an indirect-gap semiconductor, with a PBE band gap of 3.90 eV and an HSE06 band gap of 5.56 eV, so the pressure-stabilized *P*4/*nmm* modification does not change the indirect nature of the gap. The projected DOS shows the same main near-gap description as in *α*-ScOCl: the valence-band edge is dominated by anion *p* states, with strong Cl *p* and additional O *p* contribution, whereas the conduction-band edge is dominated by Sc *d* states. Relative to *α*-ScOCl, the denser eightfold Sc coordination in *β*-ScOCl slightly reduces the gap in both the PBE and HSE06 calculations.

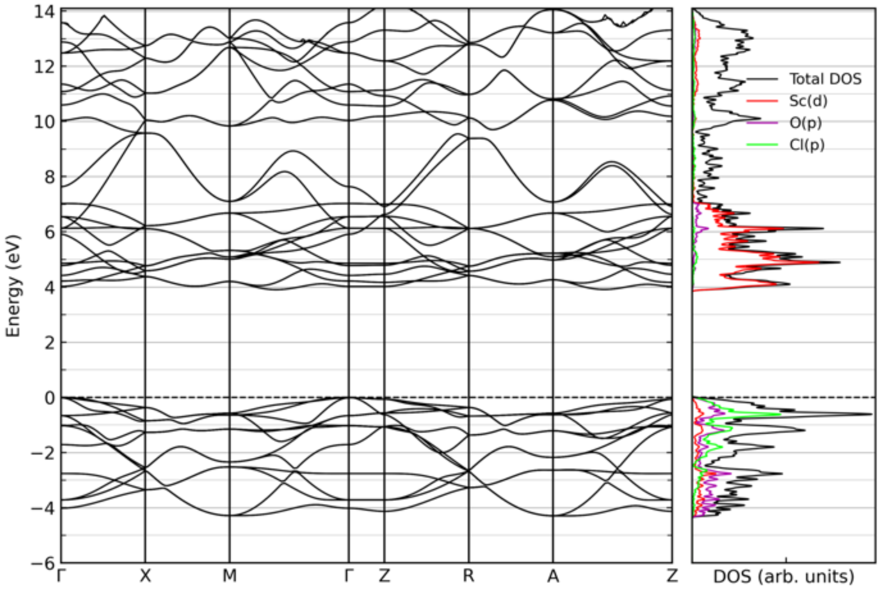

Figure 6.11. Electronic band structure and projected density of states of *β*-ScOCl calculated with PBE functional.

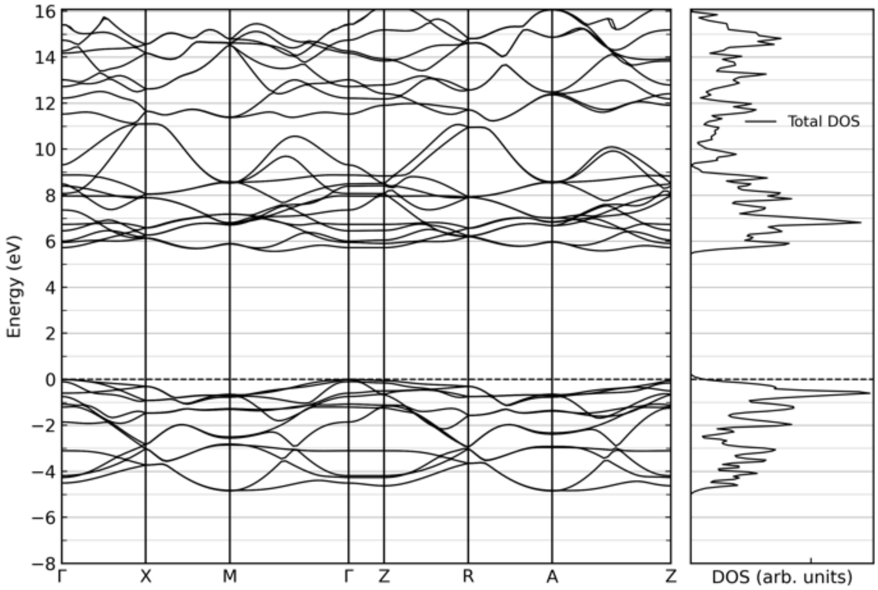

Figure 6.12. Electronic band structure and projected density of states of *β*-ScOCl calculated with HSE06 functional.

The comparison of electronic structure results obtained with different functionals shows that the qualitative electronic picture is robust. Both *α*-ScOCl and *β*-ScOCl remain indirect-gap semiconductors with anion-p-dominated valence-band edges and Sc d-dominated conduction-band near the edges. In the PBE calculations, *α*-ScOCl has a gap of approximately 4.04-4.09 eV, while *β*-ScOCl has a slightly smaller gap of approximately 3.86-3.92 eV. The HSE06 functional preserves this ordering and increases the gaps to 5.76 eV for α-ScOCl and 5.56 eV for β-ScOCl. The PBE0 functional gives the largest gaps and slightly reverses the ordering, with 6.53 eV for α-ScOCl and 6.64 eV for β-ScOCl. Hence, the indirect-gap character and the qualitative orbital composition picture of the bands are the same for all functionals, whereas the size of the band gaps is sensitive to the choice of exchange-correlation functional.

6.6 Conclusion

The energy-landscape exploration of ScOCl shows that the combined global-optimization and data-mining workflow successfully recovers the experimentally known α-ScOCl phase and identifies three additional low-energy structural candidates [9]. In total, more than 5000 candidate structures were generated from global optimization and ICSD-based data mining. After *ab initio* relaxation, four modifications remained as the most relevant minima: orthorhombic α-ScOCl in space group *Pmmn* (No. 59), tetragonal β-ScOCl in space group P4/nmm (No. 129), tetragonal γ-ScOCl in space group $P4_2/m$ (No. 84), and monoclinic δ-ScOCl in space group *C*2/*m* (No. 12). The α- and β-phases were obtained both from the global search and from data mining, whereas the γ- and δ-phases were found only through global optimization [9]. This result indicates that the ScOCl landscape is dominated by the *FeOCl*-type ground-state structure and it also contains plausible pressure- or temperature-stabilized (meta)stable polymorphs.

The thermodynamic analysis based on the calculated *E(V)* and *H(p)* curves establishes α-ScOCl as the stable ambient-pressure phase and β-ScOCl as the most probable high-pressure modification [9]. The enthalpy curves give the same qualitative pressure response for both used functionals, predicting an $\alpha \rightarrow \beta$ transition at approximately 1 GPa for LDA and approximately 3.7 GPa for PBE0 functional. The γ-phase remains thermodynamically metastable in the high-pressure region and may therefore appear as an intermediate or competing phase during pressure-induced transformation, while δ-ScOCl remains a higher-energy candidate more likely to be relevant under elevated-temperature or non-equilibrium synthesis conditions [9].

The electronic-structure calculations show that both α-ScOCl and β-ScOCl are wide-gap semiconductors, with an indirect band gap in the LDA, PBE, PBEsol, HSE06, and PBE0 descriptions, and the same qualitative band structure description retained with all functionals. The electronic structure near the band gap is characterized by a clear separation between anion-derived valence states and scandium-derived conduction states: the upper valence bands are dominated mainly by Cl *p* states with additional O *p* contribution, whereas the bottom of the conduction band is formed primarily by Sc *d* states. Hybrid functionals increase the calculated band gaps substantially relative to the LDA, PBE, PBEsol, and PBE descriptions, but they do not qualitatively change the dominant orbital character of the band-edge states.

Compared with much broader LaOI, HoFSe, and LaFS energy landscape explorations discussed elsewhere in this thesis, the ScOCl study yields a smaller final set of low-energy polymorphs. Nevertheless, its combined global-optimization and data-mining approach successfully recovers the experimentally known α-phase and identifies additional (meta)stable modifications [9]. The broader workflows developed after the ScOCl study, including larger-scale global searches, expanded data-mining explorations, and more extensive post-processing procedures, as applied in the later HoFSe and LaFS chapters, could in future be used to explore the ScOCl system in greater detail. Such an extended search would test whether additional metastable ScOCl modifications exist beyond the four phases identified in the original study.

7. Energy Landscape Exploration, Structure Prediction, Thermodynamic Stability, Magnetic Anisotropy, and Electronic Structure of Holmium Fluoroselenide (HoFSe)

7.1 Introduction

This chapter discusses the theoretical part of the HoFSe study [11], extended here by additional magnetic and electronic-structure calculations. Because the published study is a combined theoretical and experimental exploration of the HoFSe system, the experimental synthesis and characterization reported in that work, including single-crystal X-ray diffraction, scanning electron microscopy, and energy-dispersive X-ray spectroscopy, are not treated as thesis results and they are mentioned only where they provide an external consistency check for the theoretically predicted structure types and relative stability trends. The emphasis of this chapter is on crystal-structure prediction, symmetry and structure-type analysis, thermodynamic stability of selected low-energy polymorphs, and the magnetic and electronic structure of the ground-state modification.

HoFSe is a mixed-anion rare-earth fluoroselenide belonging to the broader family of rare-earth chalcohalide and chalcogenide-related compounds. Rare-earth chalcogenides have been discussed as an emerging class of optical and infrared-transmitting materials [176]. The broader rare-earth chalcogenide, oxide, and halide literature also includes examples relevant to rare-earth-ion-doped photovoltaic upconversion materials [177], laser spectroscopy and tunable solid-state laser materials [178], low-field magnetotransport [179], lanthanide-containing light-emitting organic–inorganic hybrids [180], bright upconversion luminescence in rare-earth-doped oxide nanocrystals [181], and oxygen vacancies at rare-earth oxide surfaces [182]. Related rare-earth fluoride, selenide, and chalcogenide-glass systems have also been studied because of their optical properties [183], superconductivity-related behavior in fluorine-containing rare-earth layered compounds [184], infrared optical-fiber applications [185], near- and mid-infrared active rare-earth-doped selenide glasses and fibers [186, 187], and upconversion-nanoparticle applications in nanophotonics and theranostics [188]. Within the narrower group of holmium-containing fluoride and fluoroselenide-related materials, reported examples include mid-infrared holmium-doped fluoride fiber lasers [189], Ho(III)-selective membrane sensors [190], fluoride-containing holmium hydroxide structural chemistry [191], and lanthanide-doped sodium holmium fluoride nanoparticles proposed for photothermal-therapy-related applications [192].

Within the group of ternary rare-earth compounds studied in this thesis, HoFSe is especially notable because several polymorphs were identified within the same combined theoretical and experimental study [11]. The theoretical energy-landscape search recovered the experimentally observed structures among the low-energy candidates and also predicted additional metastable, high-pressure, and low-density modifications [11]. Thus, HoFSe provides both a rich theoretical polymorphic landscape and an experimentally supported benchmark for validating the crystal-structure prediction workflow. The magnetic and electronic properties discussed in this chapter were not part of the HoFSe publication – they are additional thesis results and represent a first exploration of the magnetic ground state, collinear magnetic ordering, non-collinear spin-orbit-coupled magnetic configurations, and the electronic band structure of the HoFSe ground-state modification.

In the following, this chapter proceeds from the computational framework to the structural, thermodynamic, magnetic, and electronic results. First, the computational setup and energy-landscape exploration workflow are described, including global optimization, data-mining-based structure generation, DFT refinement, and the selection of the most relevant low-energy HoFSe candidates. The discussion then turns to global and local optimization results, followed by symmetry analysis and structure-type classification of the optimized polymorphs. Thermodynamic stability is analyzed through the calculated *E(V)* curves, pressure-dependent *H(p)* relations, and the resulting transition pressures. The magnetic part of the chapter examines the Ho magnetic ground state and anisotropy, beginning with collinear spin-arrangement calculations and then extending the analysis to non-collinear spin–orbit-coupled calculations. Finally, the electronic structure of the magnetic HoFSe ground-state modification is discussed, followed by a summary of the main conclusions.

7.2 Computational Setup and Energy Landscape Exploration Overview

The computational framework used for HoFSe energy-landscape exploration follows the workflow applied elsewhere in this thesis for LaOI, Ce_3O_3N, ScOCl, and LaFS. Crystal-structure candidates of HoFSe were first generated by global optimization on an empirical-potential landscape and by complementary prototype-based data mining, and were then refined on the *ab initio* level. Within this workflow, the aim of the exploratory stage is to generate the broadest possible pool of structurally plausible and potentially experimentally accessible candidates before more computationally intensive first-principles calculations are applied.

In the global-search part, simulated annealing [93] was carried out through stochastic Monte Carlo moves on the empirical-potential landscape, combined with periodic local quench optimizations to locate nearby minima along the search trajectory, as implemented in the G42+ framework [13]. No symmetry restrictions were imposed during the search, so both the lattice parameters and the atomic coordinates were allowed to vary freely. A computationally efficient two-body empirical potential containing damped Coulomb and Lennard-Jones terms was used in order to make the exploration of very large structure sets feasible. The global optimization was performed for cells containing four and six formula units, corresponding to 12 and 18 atoms, respectively. In total, approximately 1.82 million structure candidates were generated, from which approximately two hundred relevant candidates were selected for local optimization on the basis of empirical energy, symmetry, and frequency of occurrence on the structural landscape.

The post-processing stage converted the raw search output into a structurally interpretable dataset. Symmetry determination and idealization were carried out with the SFND and RGS algorithms [96, 119], duplicate structures were removed using CMPZ algorithm [97], and the corresponding analysis workflow was handled within the KPLOT software package [98]. This part of the workflow is methodologically important because the HoFSe landscape contains many near-related configurations and often multiple distinct structure types that share the same space-group symmetry. Without symmetry-aware clustering and comparison, these candidates could not be sorted into a compact set of genuinely distinct minima for subsequent first-principles analysis.

The complementary prototype-based data-mining stage explored the ICSD database for ternary ABX, ANX, and AXY prototypes compatible with HoFSe stoichiometry [100, 140]. The filtering followed a knowledge-discovery-in-databases logic [99, 100, 141] that consists of selection of ternary systems from the database, restriction to prototype families matching the required chemical formulas, elimination of redundant structures, and final conversion into a smaller list of distinct candidates [170, 193]. Starting from 91,739 ternary entries, the dataset was reduced to 639 structures matching the relevant stoichiometric patterns, and then to 30 unique prototype candidates that were subsequently subjected to *ab initio* structural relaxation.

For the local optimization part, the CRYSTAL17 code [51, 52] was used. Full structural relaxation was performed without symmetry constraints, using the analytical gradients and stress evaluation for both atom coordinates and lattice vectors [102, 194, 195]. The first-principles structural optimization calculations were carried out using the GGA-PBE functional [49], while calculations for selected low-energy candidates were repeated using the HSE06 hybrid functional [137]. An effective-core pseudopotential basis set was used for holmium [196], and all-electron basis sets were employed for fluorine and selenium [31, 197].

The additional magnetic and electronic-structure calculations were performed for the *exp1*-type ground-state modification using VASP within the GGA-PBE DFT+U approach, with $U_{\mathrm{eff}} = 5.3$ eV applied to the Ho 4*f* states in the Dudarev formulation [66]. Collinear calculations were carried out for fully relaxed primitive-cell Ho spin arrangements and for additional double-cell arrangements, while non-collinear spin-orbit-coupled calculations were initialized from the lowest-energy collinear antiferromagnetic geometry. Band-structure and projected-DOS calculations were then evaluated for the relevant magnetic ground-state configurations.

7.3 Global and Local Optimization. Symmetry and Structure Type Analysis of Optimized Structures.

A total of 1.82 million structures were generated in the unconstrained global optimization part of the HoFSe search, and their statistics is summarized in Table 7.1. As expected for such stochastic searches without imposed symmetry, the landscape is dominated by the triclinic *P*1 space group, which accounts for more than a million occurrences. The physically more important information, however, is the repeated appearance of higher-symmetry space groups. For the 4-formula-unit searches, space groups No. 62, 123, 19, and 84 recur with particularly high frequency, whereas the 6-formula-unit searches favor structures in space groups No. 11, 6, 59, 31, 129, and 123. About two hundred structure-type representatives from these and many other recurring high-symmetry space groups, together with representative amorphous *P*1 candidates, were selected for subsequent DFT local optimization.

Table 7.1. Frequency of occurrence of various space groups generated during global optimization, categorized by the number of formula units per simulation cell.

4 HoFSe		6 HoFSe		Total	
Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence
*P*1 (1)	585,680	*P*1 (1)	453,813	*P*1 (1)	1,039,493
Pnma (62)	154,079	$P2_1/m$ (11)	12,102	*Pnma* (62)	154,079
*P*4/*mmm* (123)	150,157	*Pm* (6)	6,748	*P*4/*mmm* (123)	151,939
$P2_12_12_1$ (19)	58,730	*P*-1 (2)	5,290	$P2_12_12_1$ (19)	58,730
$P4_2/m$ (84)	53,020	$P2_1$ (4)	4,379	$P4_2/m$ (84)	53,020
P-1 (2)	34,734	*Pmmn* (59)	3,876	*P-1* (2)	40,024
$P2_1$ (4)	30,438	$Pmn2_1$ (31)	2,133	$P2_1/m$ (11)	39,014
$P4_2/mmc$ (131)	29,851	*P*4/*nmm* (129)	2,046	$P2_1$ (4)	34,817
$P2_1/m$ (11)	26,912	*P*4/*mmm* (123)	1,782	$P4_2/mmc$ (131)	29,918
*P*4/*nmm* (129)	21,026	*P*2 (3)	949	*Pm* (6)	26,844
Pm (6)	20,096	*Pc* (7)	860	*P*4/*nmm* (129)	23,072
$Pmc2_1$ (26)	16,268	*Pmm*2 (25)	795	$Pmc2_1$ (26)	16,305
$P2_13$ (198)	15,722	$P2_12_12$ (18)	343	$P2_13$ (198)	15,722
Pbcm (57)	12,695	*P*422 (89)	305	$Pmn2_1$ (31)	12,837
Pc (7)	10,898	*Cmmm* (65)	298	*Pbcm* (57)	12,695
OTHER	99,694	OTHER	4,281	OTHER	111,491
TOTAL	1,320,000	TOTAL	500,000	TOTAL	1,820,000

Two complementary routes were used to explore the low-energy region of the HoFSe landscape. Approximately two hundred candidates were selected from the global search output for local optimization, while data mining contributed thirty prototype-derived candidates. After the DFT structural relaxation and total-energy evaluation, twenty promising HoFSe modifications remained in the relevant low-energy part of the landscape (Table 7.2). This set contains the three experimentally observed polymorphs, labelled exp1, exp2, and exp3, together with seventeen additional theoretically predicted phases. Among the final twenty structures, twelve have unique GO origin, four have unique DM origin, and four are found by both methods. The *exp1*-type structure is recovered by both GO and DM searches, whereas *exp2*-type and *exp3*-type structures are recovered through the data mining path [11]. The recovery of all experimentally observed structures among the low-energy candidates validates the theoretical workflow and also shows that data mining supplements the stochastic search in regions of the landscape associated with previously discovered prototypes. While Table 7.2 summarizes the low-energy structure-type set, their full GGA-PBE and HSE06 structural data are given in Tables A7.1-A7.3 of Appendix A7.

Table 7.2. Low-energy HoFSe structure candidates sorted according to the PBE energy-ranking index (E_{RI}).

E_{RI}	Modification	SG No.	Energy / FU [E_h]	GO	DM	Exp.	Volume / FU (Å³)	Cell parameters (Å / °)

1	*exp1*-type	62	-2538.90444	+	+	+	67.61	a = 10.091;
								b = 4.168;
								c = 6.430

2	*α*-type	11	-2538.90394	+			68.33	a = 6.464;
								b = 4.166;
								c = 15.396;
								β = 98.52

3	*β*-type	62	-2538.90388	+			58.65	a = 6.761;
								b = 3.878;
								c = 8.947

4	*γ*-type	63	-2538.90326	+			69.92	a = 5.656;
								b = 11.890;
								c = 4.159

5	*PbFCl*	129	-2538.90304	+	+		55.44	a = 3.945;
								b = 3.945;
								c = 7.123

6	*δ*-type	12	-2538.90302	+			67.53	a = 19.572;
								b = 4.165;
								c = 10.110;
								β = 100.53

7	*CoYC*	131	-2538.90291	+			74.81	a = 4.380;
								b = 4.380;
								c = 7.798

8	*ε*-type	87	-2538.90281	+			60.15	a = 11.138;
								b = 11.138;
								c = 3.879

9	*ZrOS*	198	-2538.90240	+	+		63.01	a = 6.317;
								b = 6.317;
								c = 6.317

10	*exp2*-type	11	-2538.90227		+	+	67.46	a = 10.106;
								b = 4.165;
								c = 13.284;
								β = 105.15

11	*exp3*-type	62	-2538.90167		+	+	67.19	a = 10.128;
								b = 4.158;
								c = 19.148

12	*ZrBeSi*	194	-2538.90013		+		56.82	a = 4.051;
								b = 4.051;
								c = 7.995;
								γ = 120.00

13	*ζ*-type	31	-2538.90011	+			67.19	a = 4.155;
								b = 15.309;
								c = 6.338

14	*η*-type	12	-2538.89962	+			61.93	a = 15.161;
								b = 4.105;
								c = 7.973;
								β = 93.12

15	*θ*-type	11	-2538.89962	+			69.79	a = 11.013;
								b = 4.158;
								c = 9.249;
								β = 98.64

16	*3R*-type	166	-2538.89930	+	+		66.55	a = 4.133;
								b = 4.133;
								c = 26.993;
								γ = 120.00

E_{RI}	Modification	SG No.	Energy / FU [E_h]	GO	DM	Exp.	Volume / FU (Å³)	Cell parameters (Å / °)

17	*2H*-type	194	-2538.89900		+		67.04	a = 4.153;
								b = 4.153;
								c = 17.956;
								γ = 120.00

18	*ι*-type	14	-2538.89871	+			63.73	a = 7.020;
								b = 5.293;
								c = 7.649;
								β = 116.23

19	*κ*-type	47	-2538.89870	+			74.91	a = 4.313;
								b = 4.391;
								c = 11.865

20	*λ*-type	12	-2538.89826	+			72.89	a = 15.486;
								b = 4.142;
								c = 11.103;
								β = 125.04

The predicted HoFSe structures are grouped according to the part of the thermodynamic landscape in which they are most relevant: standard-pressure and possible high-temperature candidates, high-pressure candidates, and low-density candidates on the effective negative-pressure side. The main structural models discussed in detail are the global-minimum *exp1*-type phase, the experimentally observed *exp2*- and *exp3*-type polymorphs, the most relevant high-pressure and negative-pressure/low-density candidates, and selected polytypic variants. Structural models for the remaining polymorphs are shown in Figures 7.7–7.11. The crystallographic summary that includes Ho-centered coordination numbers, Ho–anion distances, and Ho-centered polyhedra connectivity of the twenty low-energy structures is given in Table 7.3.

Table 7.3. Summary of the Ho-centered coordination environments of the low-energy HoFSe polymorphs.

Modification	Space group (No.)	The p/T region[*]	Volume / FU (Å³)	Ho-anion coordination numbers[**]	Ho-anion average distance (Å)	Ho–F distances (Å)	Ho–Se distances (Å)	Ho polyhedral connectivity
exp1-type	*Pnma* (62)	standard *p*	67.61	7	2.674	2.32–2.37	2.89–2.96	corner
α-type	$P2_1/m$ (11)	standard *p* / high *T*	68.33	7	2.672	2.32–2.37	2.86–2.98	edge; corner
β-type	*Pnma* (62)	high pressure	58.65	8	2.760	2.35–2.37	2.96–3.04	edge; corner
γ-type	*Cmcm* (63)	low density	69.92	7	2.669	2.32–2.36	2.86–2.98	edge
PbFCl	*P4/nmm* (129)	high pressure	55.44	9	2.767	2.50–2.50	2.96–3.05	edge
δ-type	*C2/m* (12)	standard *p* / high *T*	67.53	6, 7, 8	2.683	2.29–2.44	2.87–2.97	edge; corner
CoYC	$P4_2/mmc$ (131)	low density	74.81	6	2.685	2.19–2.19	2.93–2.93	corner
ε-type	*I4/m* (87)	high pressure	60.15	8	2.752	2.36–2.40	2.94–2.99	edge
ZrOS	$P2_13$ (198)	high pressure	63.01	7	2.679	2.34–2.34	2.93–2.94	edge; corner
exp2-type	$P2_1/m$ (11)	standard *p* / high *T*	67.46	6, 7, 8	2.687	2.29–2.48	2.86–2.99	edge; corner
exp3-type	*Pnma* (62)	standard *p* / high *T*	67.19	6, 7, 8	2.691	2.30–2.47	2.85–2.99	edge; corner
ZrBeSi	$P6_3/mmc$ (194)	high pressure	56.82	9	2.831	2.34–2.34	3.08–3.08	edge
ζ-type	$Pmn2_1$ (31)	standard *p* / high *T*	67.19	7	2.676	2.32–2.43	2.87–2.96	edge; corner
η-type	*C2/m* (12)	high pressure	61.93	7, 8	2.722	2.29–2.59	2.86–3.01	face; edge; corner
θ-type	$P2_1/m$ (11)	standard *p* / high *T*	69.79	6, 7	2.723	2.16–2.38	2.87–3.25	face; edge; corner
3R-type	*R*-3*m* (166)	standard *p* / high *T*	66.55	6, 8	2.695	2.40–2.40	2.86–2.92	corner
2H-type	$P6_3/mmc$ (194)	standard *p* / high *T*	67.04	6, 8	2.697	2.40–2.40	2.87–2.92	corner
ι-type	$P2_1/c$ (14)	standard *p* / high *T*	63.73	7	2.683	2.33–2.39	2.88–3.01	edge; corner
κ-type	*Pmmm* (47)	low density	74.91	6	2.685	2.16–2.20	2.91–2.97	corner
λ-type	*C2/m* (12)	standard *p* / high *T*	72.89	6, 7	2.740	2.18–2.21	2.85–3.18	edge; corner

[*] Regions refer to the thermodynamic grouping used in the text: standard pressure / high-T, high pressure, and low density.

[**] Holmium coordination numbers and Ho-anion average distances were determined by using the cutoffs of 2.8 Å for Ho–F and 3.4 Å for Ho–Se distances. Coordination numbers of all distinct Ho local environments (inequivalent Ho sites) are shown.

The standard-pressure subset contains *exp1*-, α-, δ-, *exp2*-, *exp3*-, ζ-, θ-, *2H*-, *3R*-, ι-, and λ-type HoFSe. The *exp1*-type ground state is the *LiCaN*-type structure in the *Pnma* (No. 62) space group; Ho is sevenfold coordinated in $[\mathrm{HoF_3Se_4}]$ polyhedra that form a dense network of linked coordination polyhedra. The α-type phase is a monoclinic $P2_1/m$ (No. 11) polymorph with sevenfold Ho coordination and it is structurally related to the experimentally observed monoclinic family. The δ-type phase is a $C2/m$ (No. 12) candidate with site-dependent 6-, 7-, and 8-fold Ho coordination, showing that the standard-pressure region of the landscape can accommodate several distorted Ho-centered polyhedra. The *exp2*-type phase in the $P2_1/m$ (No. 11) space group is the experimentally observed *YFSe*-type structure; its crystallographically distinct Ho sites show 6-, 7-, and 8-fold mixed coordinations, ranging from octahedral-like to pentagonal-bipyramidal and higher-coordinate distorted environments. The *exp3*-type phase in the *Pnma* (No. 62) space group is the experimentally observed *ErFSe*-type polymorph, whose elongated c axis and site-dependent Ho coordination point toward a polytypic extension of the exp1 structure type. The ζ-type phase in the $Pmn2_1$ (No. 31) space group has structural features similar to the *exp2*-type phase, whereas the θ-type phase in the $P2_1/m$ (No. 11) space group is related to the *exp2*-/*exp3*-type stacking motifs. The *2H* phase in the $P6_3/mmc$ (No. 194) space group and the *3R* phase in the R-3m (No. 166) space group are polytypic variants identified by the data mining method, while the ι-type phase in the $P2_1/c$ (No. 14) space group and the λ-type phase in the $C2/m$ (No. 12) space group are higher-energy monoclinic structures that are more likely to be relevant under more extreme conditions [11].

The high-pressure structure candidates are β-, *PbFCl-*, ε-, *ZrOS-*, *ZrBeSi-*, and η-type HoFSe modifications. The β-type phase in the *Pnma* (No. 62) space group has eightfold coordinated Ho (3F + 5Se) and consists of edge- and corner-connected Ho-centered mixed-anion polyhedra. The *PbFCl*-type HoFSe in the $P4/nmm$ (No. 129) space group has ninefold coordinated Ho (4F + 5Se) and is dominated by edge-connected Ho-centered polyhedra in a compact tetragonal structure. The ε-type phase in the $I4/m$ (No. 87) space group also contains eightfold coordinated Ho and may be viewed as a compressed tetragonal candidate with a density similar to the *PbFCl*-type structure. The *ZrBeSi*-type phase in the $P6_3/mmc$ (No. 194) space group has ninefold Ho coordination (3F + 6Se) in a layered arrangement, while *ZrOS*-type HoFSe in the $P2_13$ (No. 198) space group has sevenfold Ho coordination. The η-type phase in the $C2/m$ (No. 12) space group has site-dependent seven- to eightfold Ho coordination and the most complex connection pattern in this compressed group, with corner-, edge-, and face-sharing contacts between Ho-centered polyhedra. These structure descriptions show that compression can be accommodated either by an increase in Ho coordination or by denser packing at lower coordination [11]. The *PbFCl* structure type is common in related rare-earth mixed-anion compounds and in high-pressure binary analogues, such as high-pressure fluoride and sulfide systems, making it an experimentally plausible pressure-induced structure [198, 199].

The low-density region of the low-energy landscape is populated by γ-, *CoYC-*, and κ-type HoFSe structures. The γ-type phase in the *Cmcm* (No. 63) space group retains sevenfold Ho coordination (3F + 4Se), whereas *CoYC*-type in the $P4_2/mmc$ (No. 131) and κ-type in the *Pmmm* (No. 47) space groups both have sixfold Ho coordination (2F + 4Se) and the largest equilibrium volumes of the twenty low-energy structures. Taken together, the analysis shows a clear structural progression: compressed-region modifications generally favor eight- and ninefold Ho coordination or more tightly connected edge- or face-sharing polyhedra; the standard-pressure family is dominated by sevenfold and site-dependent six to eight coordination; and the lowest-density structures mostly favor reduced sixfold environments and more open packing arrangements.

The *exp1*-type ground state and the two experimentally observed *exp2*-type and *exp3*-type structures are presented in Figures 7.1 and 7.2, showing the excellent agreement between the theoretically predicted low-energy minima and the experimentally confirmed polymorphs [11].

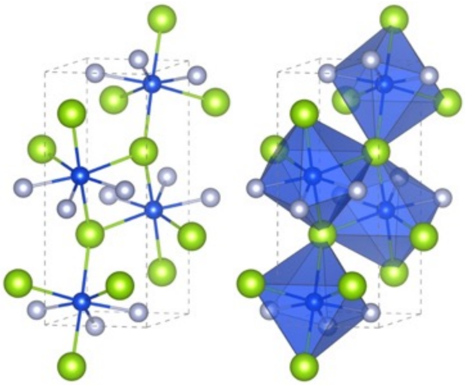

Figure 7.1. Structural model of the global-minimum *exp1*-type HoFSe modification.

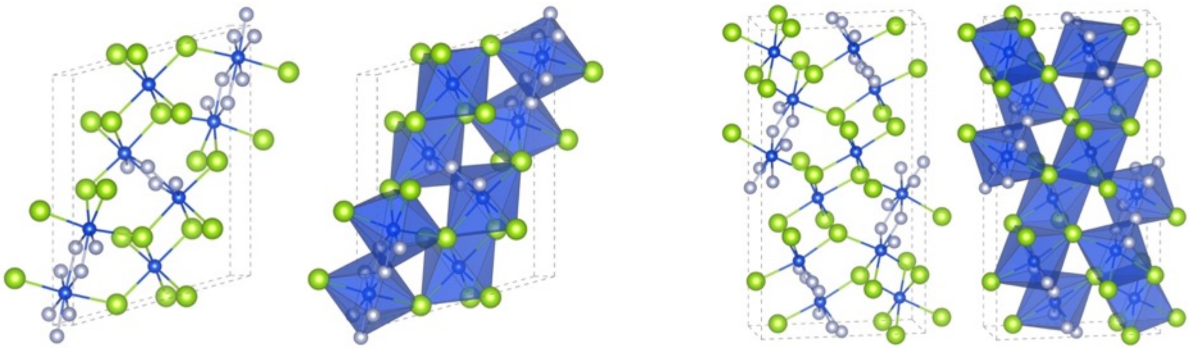

Figure 7.2. Structural models of the theoretically predicted and experimentally observed *exp2*-type (left) and *exp3*-type (right) HoFSe modifications.

The α-type and δ-type modifications, shown in Figure 7.3, represent low-energy standard-pressure metastable candidates. They are particularly relevant because they are predicted between the exp1 ground-state structure and the experimentally observed *exp2*- and *exp3*-type polymorphs in the energy ranking, suggesting that they may be accessible as metastable phases under suitable synthesis conditions.

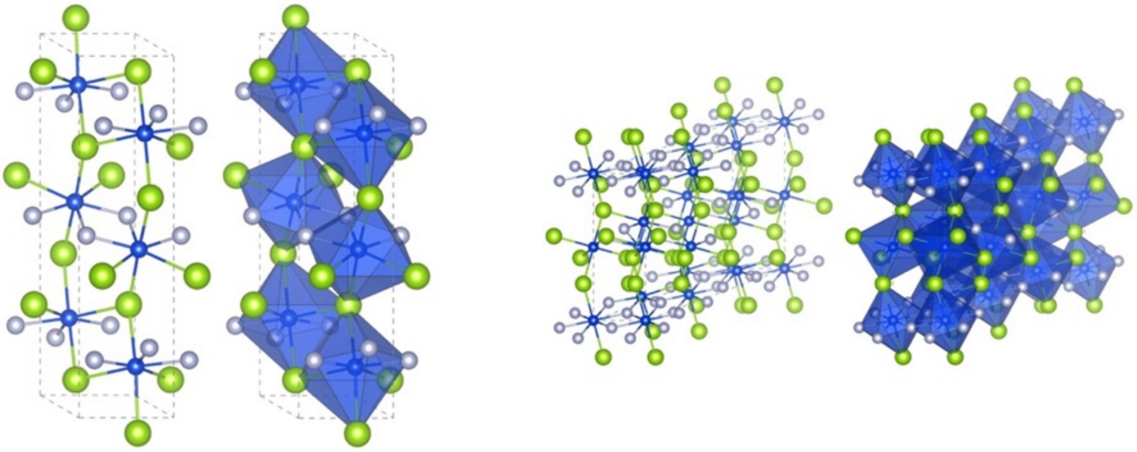

Figure 7.3. Structural model of the predicted α-type (left) and δ-type (right) HoFSe modifications.

The 2H and 3R polytypes, shown in Figure 7.4, further demonstrate that the HoFSe landscape includes stacking-derived variants with different repeat periods along the stacking direction. These polytypic candidates are related to the same broader polytypism expressed by the experimentally observed *exp3*-type structure.

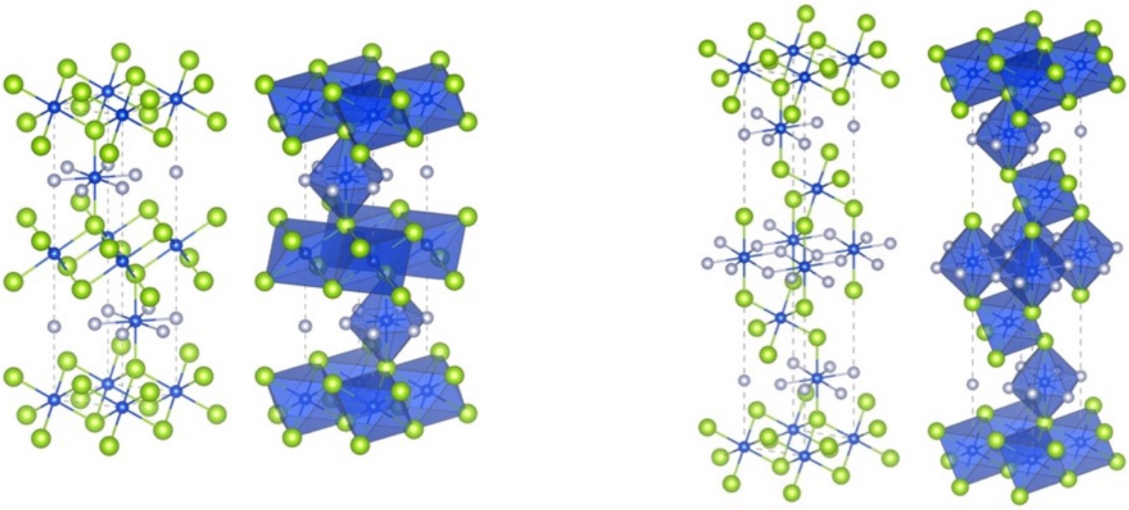

Figure 7.4. Structural models of the predicted 2H (left) and 3R (right) HoFSe polytypes.

Figure 7.5 shows the *β*-type and *PbFCl*-type structures, which are the main high-pressure competitors. The *β*-type phase represents the first pressure-induced competitor of the exp1 ground state, while the *PbFCl*-type phase becomes important at still smaller volumes / higher pressures. The *CoYC*-type low-density modification is shown in Figure 7.6, as it becomes relevant on the effective negative-pressure part of the landscape. The remaining structural models are shown in Figures 7.7–7.11.

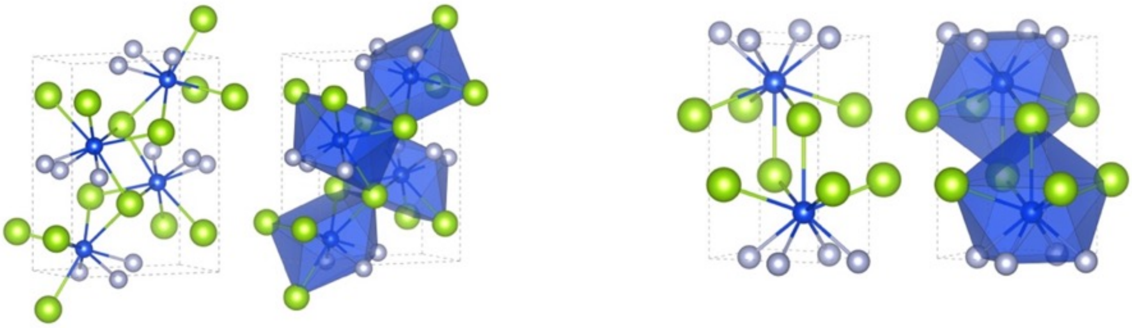

Figure 7.5. Structural models of the predicted *β*-type (left) and *PbFCl*-type (right) HoFSe modifications.

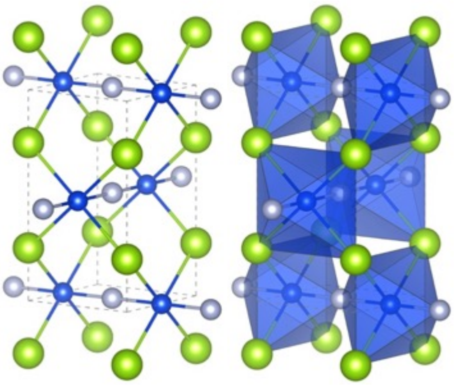

Figure 7.6. Structural model of the predicted *CoYC*-type modification.

Figure 7.7 shows the γ-type orthorhombic *Cmcm* (No. 63) low-density modification with sevenfold Ho coordination, together with the ε-type tetragonal *I*4/*m* (No. 87) modification, which belongs to the compressed/high-pressure group. The cubic *ZrOS*-type $P2_13$ (No. 198) modification, with sevenfold Ho coordination and the hexagonal *ZrBeSi*-type $P6_3/mmc$ (No. 194) structure, with ninefold Ho coordination in a layered arrangement are shown in Figure 7.8. The ζ-type $Pmn2_1$ (No. 31) modification, structurally related to the experimentally observed monoclinic family, and the η-type *C*2/*m* (No. 12) high-pressure candidate with site-dependent seven- to eightfold Ho coordination are presented in Figure 7.9. The θ-type $P2_1/m$ (No. 11) structure, related to the *exp2*- and *exp3*-type stacking arrangements, and the ι-type $P2_1/c$ (No. 14) higher-energy monoclinic structure are shown in Figure 7.10. And the Figure 7.11 presents the low-density κ-type *Pmmm* (No. 47) modification with sixfold Ho coordination and the λ-type *C*2/*m* (No. 12) higher-energy monoclinic structure [11].

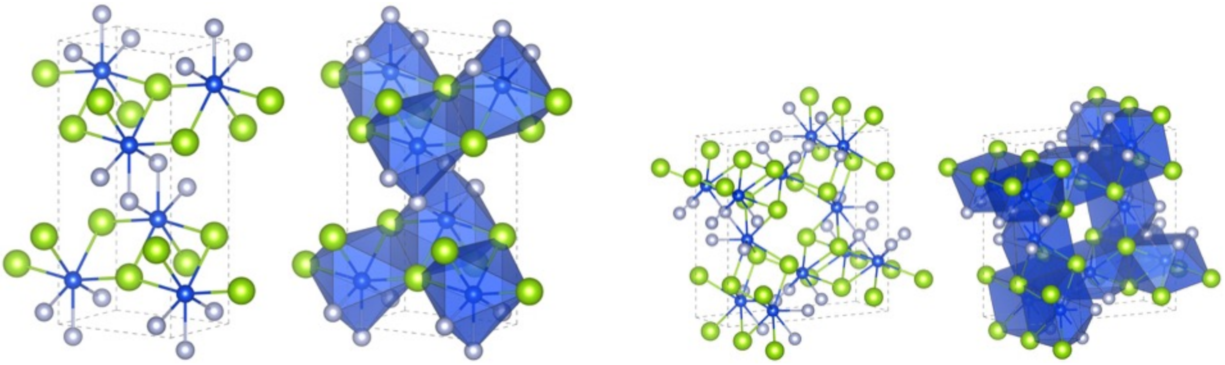

Figure 7.7. Structural model of the predicted γ-type (left) and ε-type (right) HoFSe modification.

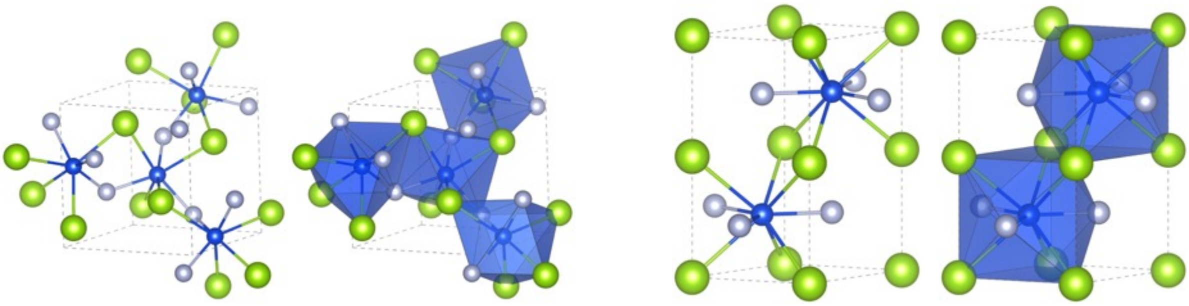

Figure 7.8. Structural model of the predicted *ZrOS*-type (left) and *ZrBeSi*-type (right) HoFSe modification.

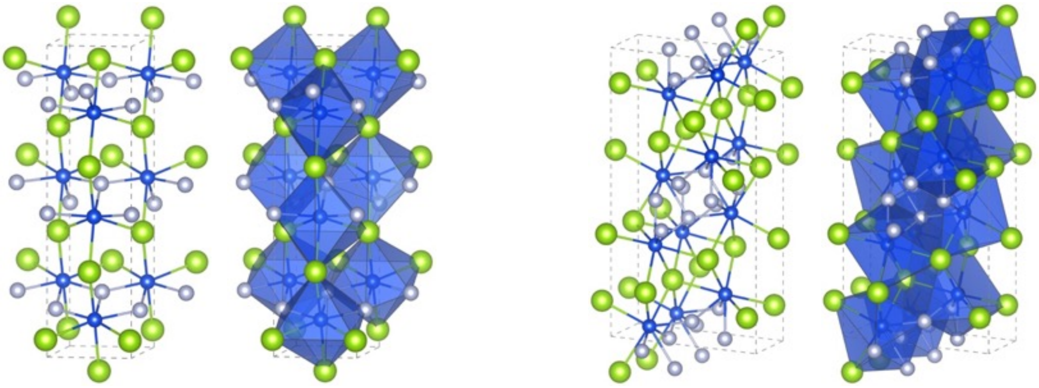

Figure 7.9. Structural model of the predicted ζ-type (left) and η-type (right) HoFSe modification.

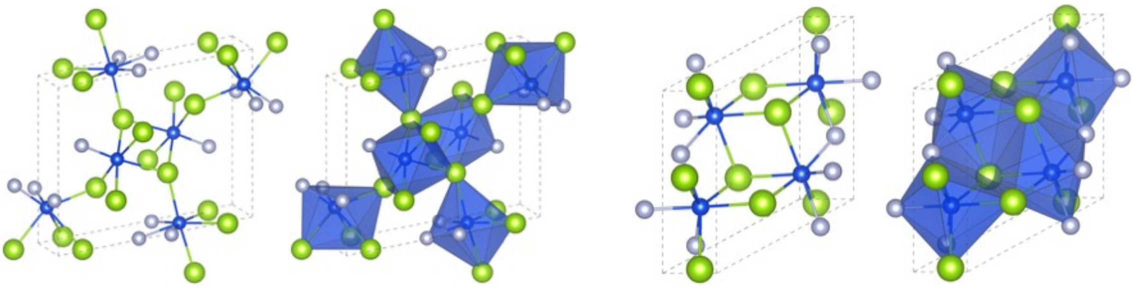

Figure 7.10. Structural model of the predicted θ-type (left) and ι-type (right) HoFSe modification.

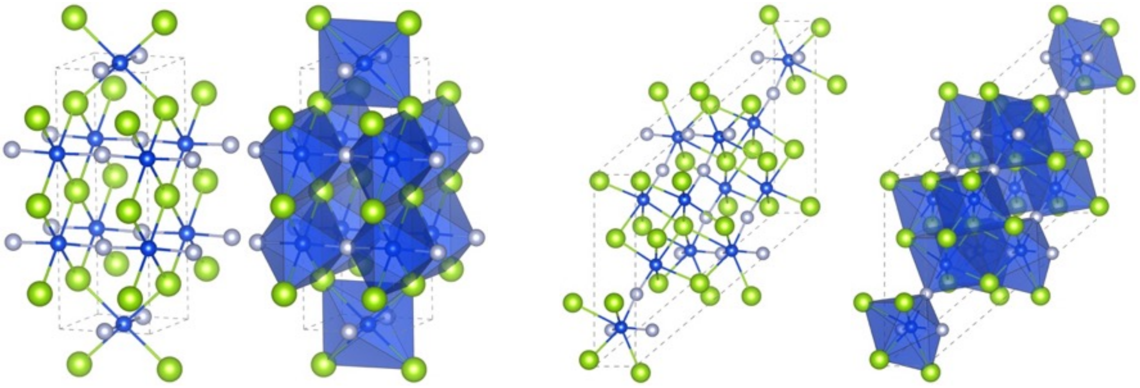

Figure 7.11. Structural model of the predicted κ-type (left) and λ-type (right) HoFSe modification.

7.4 Thermodynamic Stability and Pressure Response

The thermodynamic analysis of HoFSe is based on the calculated energy vs volume, *E(V)*, and enthalpy vs pressure, *H(p)*, relations of the candidate structures. Local optimization shows which structure types are present on the landscape as low-energy modifications, while the *E(V)* and *H(p)* curves determine which of these structure types become thermodynamically relevant at different pressure and temperature conditions.

7.4.1 The *E(V)* Equations of State

The *E(V)* curves for twenty low-energy HoFSe modifications provide a compact view of the HoFSe energy landscape. The experimentally realized *exp1*-type structure is the global minimum and therefore the ambient-pressure ground state, while a considerable number of nearby local minima show that HoFSe is a strongly polymorphic mixed-anion compound (Figures 7.12-7.15). The following discussion is based primarily on the GGA-PBE *E(V)* curves while the corresponding HSE06 results are shown in Figures 7.16–7.19.

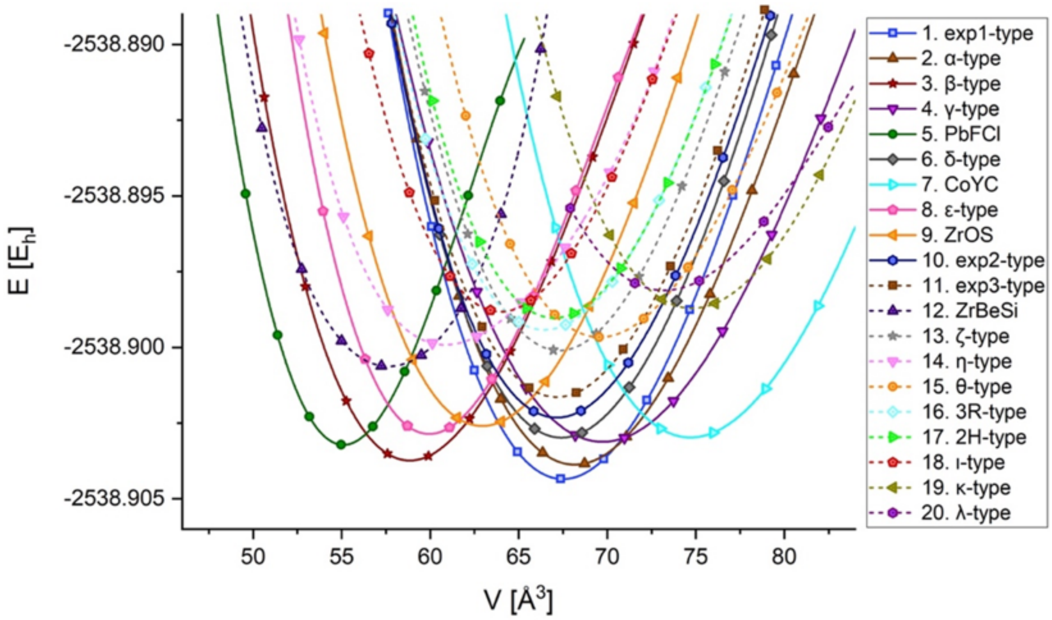

Figure 7.12. The *E(V)* curves for the twenty most relevant HoFSe modifications at the PBE level.

The GGA-PBE *E(V)* curves for the standard-pressure and possible high-temperature candidates are presented in Figure 7.13, while their HSE06 counterpart is shown in Figure 7.17. At approximately ambient pressure, HoFSe has many metastable modifications near the *exp1*-type ground state. The *α*-type and *δ*-type structures are the lowest-energy predicted metastable competitors in this subset, whereas the experimentally synthesized *exp2*- and *exp3*-type polymorphs lie higher in energy. The polytypes *2H* and *3R*, together with *ζ*-, *θ*-, *ι*-, and *λ*-type candidates, occupy the highest-energy part of this standard-pressure structure candidates' family.

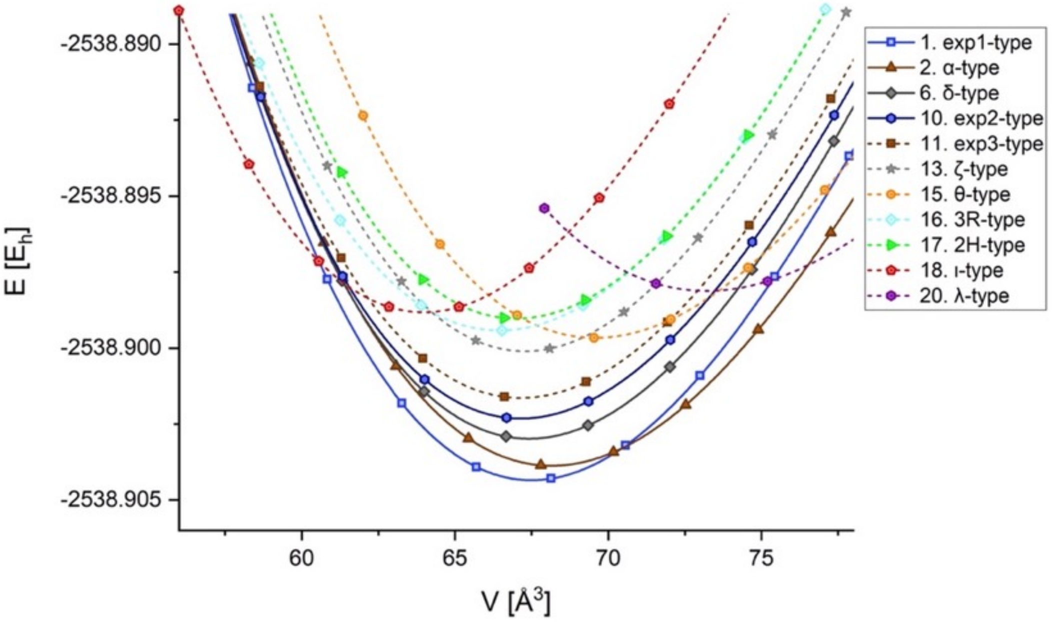

Figure 7.13. The *E(V)* curves for the ambient-pressure ground state and metastable standard-pressure HoFSe structures calculated using GGA-PBE functional.

The GGA-PBE *E(V)* curves for the high-pressure structure candidates are shown in Figure 7.14. The corresponding HSE06 curves are shown in Figure 7.18. In the high-pressure region, the *E(V)* curves identify *β*-type and *PbFCl*-type HoFSe phases as the main competitors. Their thermodynamic stabilization under pressure is established after calculating the *H(p)* dependence, but the *E(V)* curves already show that these two phases are the most relevant candidates for the high-pressure enthalpy analysis. Other high-pressure candidates, such as ε-, *ZrOS*-, *ZrBeSi*-, and *η*-type phases, remain as well-defined local minima but they lie higher in total energy in the low-volume range and do not become the principal high-pressure competitors.

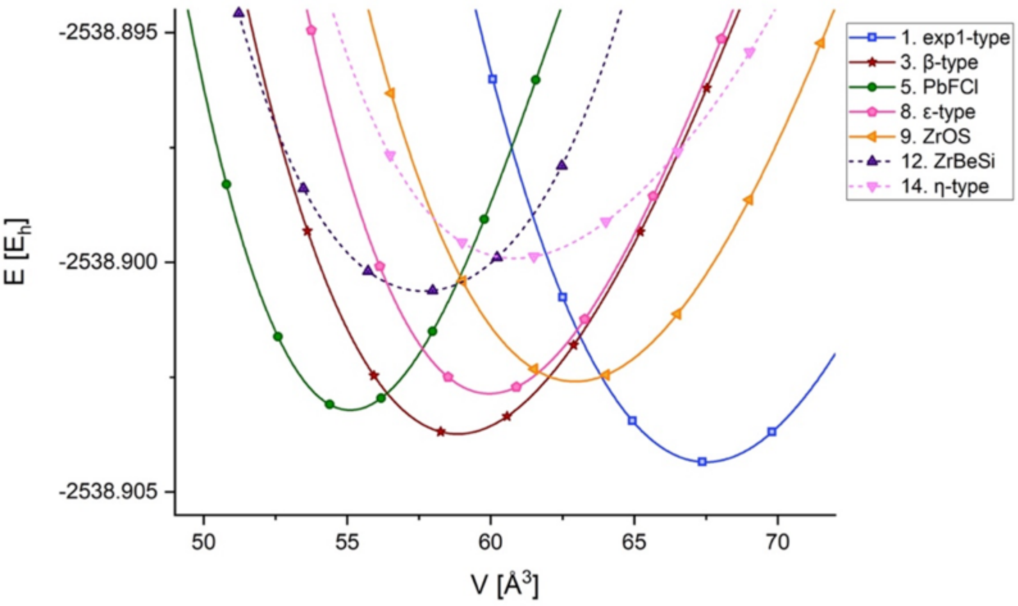

Figure 7.14. The *E(V)* curves for the most relevant high-pressure HoFSe modifications at the PBE level, reproduced from the HoFSe manuscript.

The *E(V)* curves of the negative-pressure structure candidates calculated using the GGA-PBE functional are shown in Figure 7.15, while their HSE06 counterparts are shown in Figure 7.19. The effective negative-pressure subset consists of *γ*-type, *CoYC*-type, and *κ*-type HoFSe modifications. Among them, *γ*-type and *CoYC*-type occupy the lowest part of the expanded *E(V)* region, whereas *κ*-type remains substantially higher in energy and should be regarded as metastable negative-pressure candidate structure.

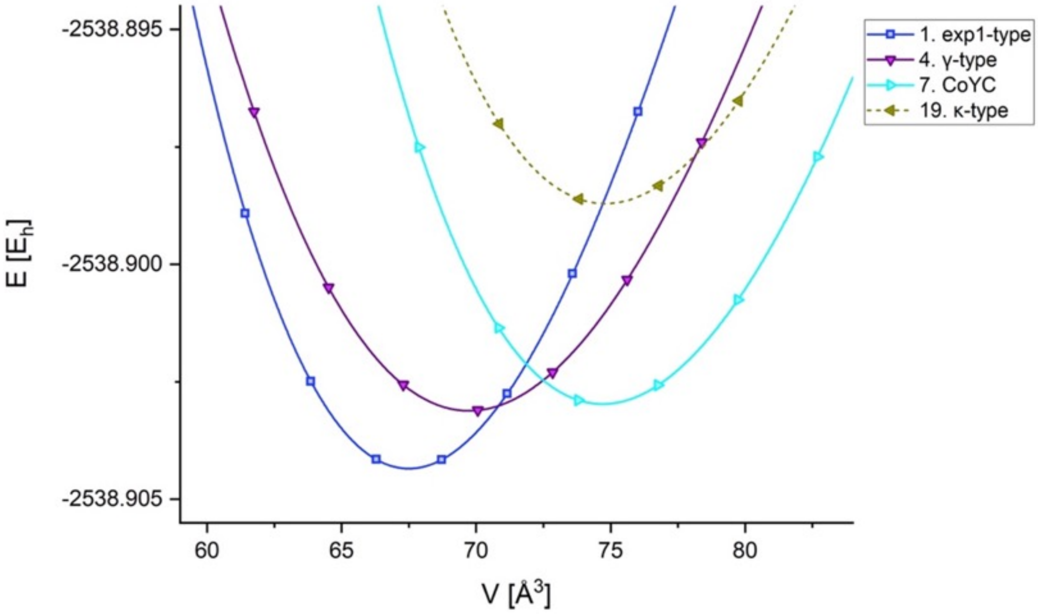

Figure 7.15. The *E(V)* curves for the negative-pressure HoFSe candidates at the GGA-PBE level.

The *E(V)* curves therefore partition the HoFSe landscape into three broad thermodynamic groups: the ambient-pressure family with possible high-temperature phases around *exp1*-type global minimum, the high-pressure group of structures with *β*- and *PbFCl*-type HoFSe as the most probable pressure-stabilized candidates, and the low-density negative-pressure group with *CoYC*-type and *γ*-type as their main representatives.

The HSE06 *E(V)* curves in Figures 7.16–7.19 confirm the main qualitative thermodynamic grouping obtained from GGA-PBE calculations. The *exp1*-type phase remains the global minimum at ambient pressure, and the HSE06 *E(V)* overview in Figure 7.16 preserves the separation between the standard-pressure family, the compressed high-pressure candidates, and the expanded low-density candidates, although the energy ranking is changed for some of the structures. In the standard-pressure/high-temperature subset (Figure 7.17), the *α*- and *δ*-type structures remain the closest predicted metastable competitors to *exp1*-type, while experimentally observed *exp2* and *exp3* remain higher-energy polymorphs. In the high-pressure subset (Figure 7.18), *β*-type and *PbFCl*-type remain the main high-pressure competitors to the *exp1*-type modification. In the expanded region (Figure 7.19), *γ*-type and *CoYC*-type again form the low-energy part of the negative-pressure landscape, while *κ*-type remains a higher-energy low-density structure. Although HSE06 changes quantitative energy separations, it does not alter the main structural interpretation of the HoFSe energy landscape.

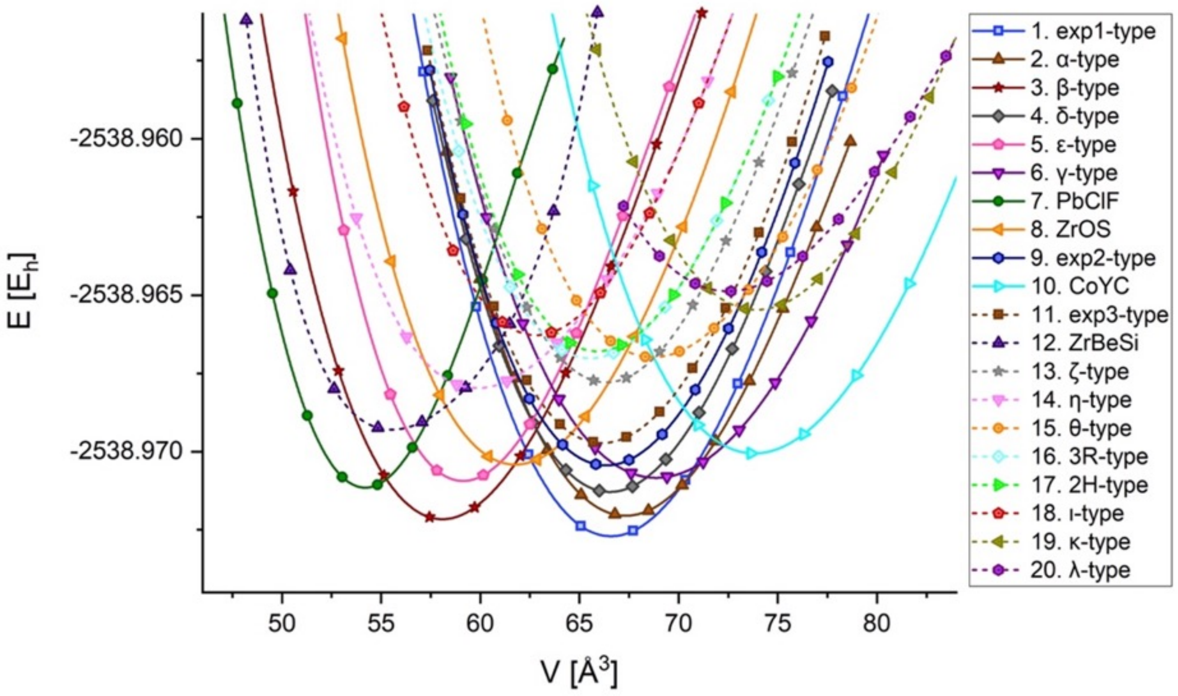

Figure 7.16. The HSE06 *E(V)* curves for the twenty low-energy HoFSe modifications.

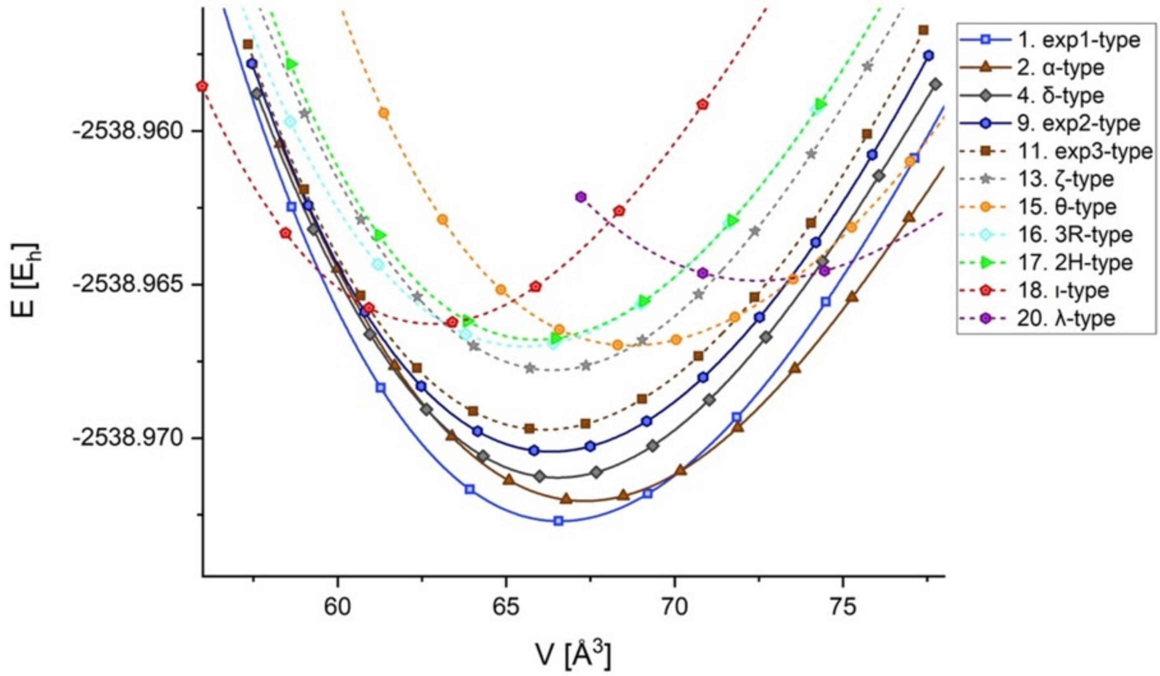

Figure 7.17. The HSE06 *E(V)* curves for the ambient-pressure and high-temperature feasible HoFSe modifications.

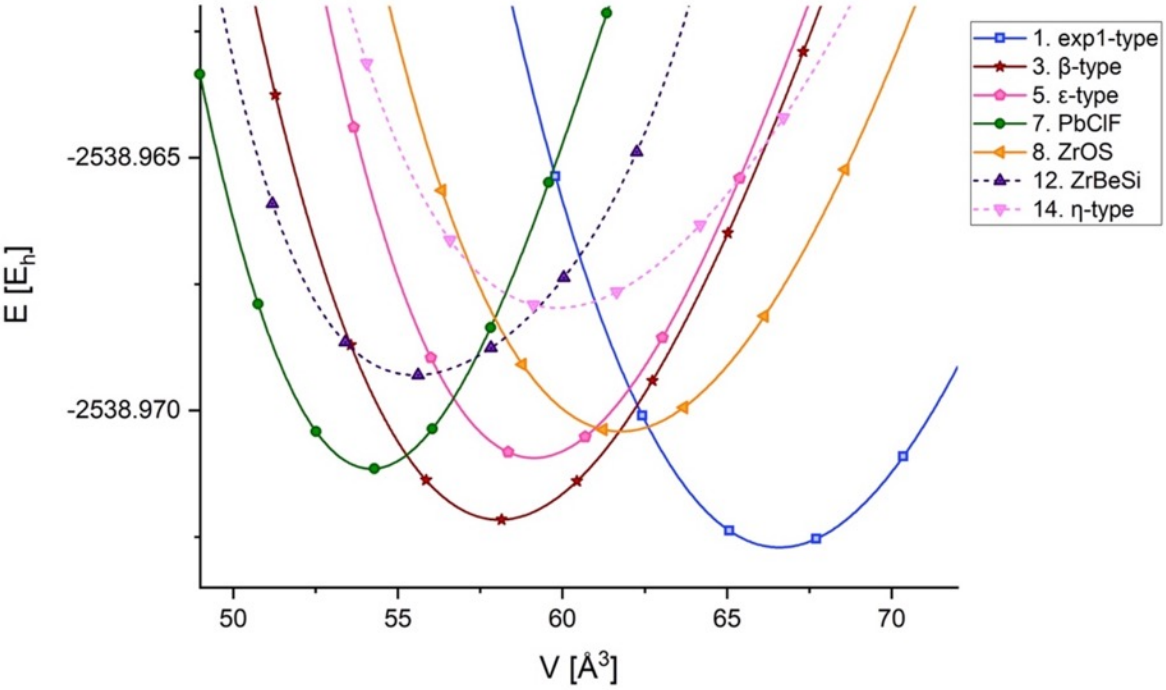

Figure 7.18. The HSE06 *E(V)* curves for the high-pressure feasible HoFSe structure candidates.

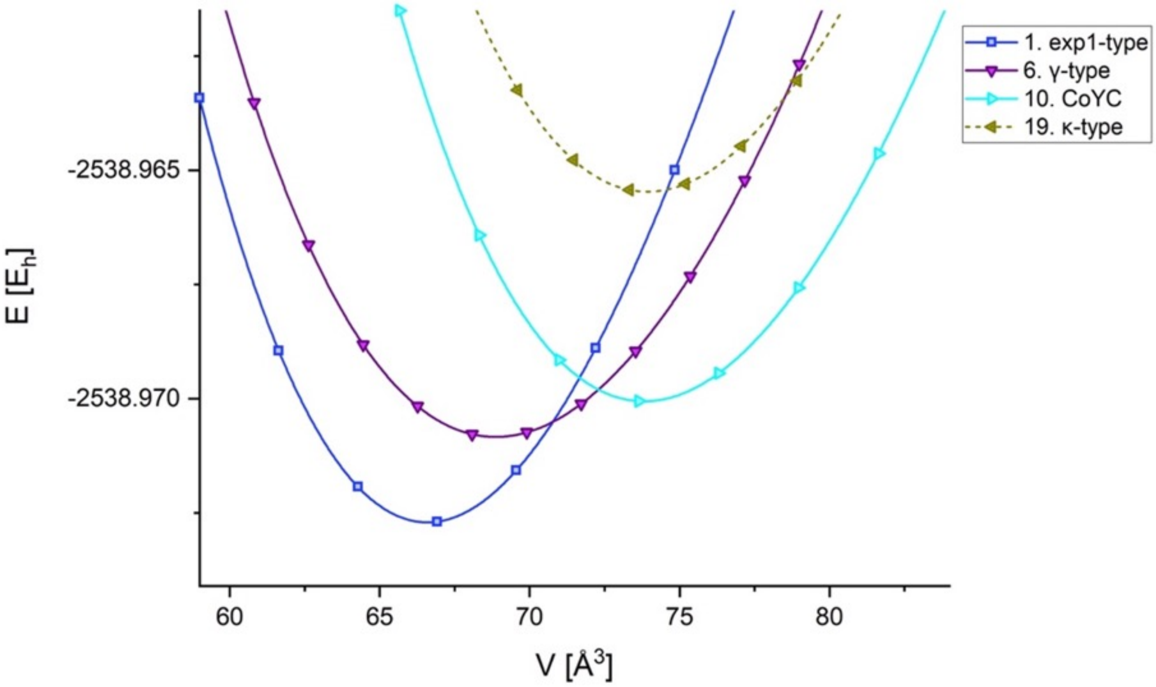

Figure 7.19. The HSE06 *E(V)* curves for the negative-pressure feasible HoFSe candidates.

7.4.2 Enthalpy vs Pressure and Transition Pressures

A clearer picture of phase competition under pressure is obtained after fitting the *E(V)* data by calculating the corresponding pressure-dependent *H(p)* curves. The enthalpy vs pressure curves for the high-pressure region calculated with the GGA-PBE functional are shown in Figure 7.20, while the corresponding HSE06 *H(p)* curves are shown in Figure 7.22.

In the positive-pressure range, the *β*-type phase becomes the lowest-enthalpy structure above 0.3 GPa within GGA-PBE, replacing *exp1*-type HoFSe as the thermodynamically most stable phase. With further compression, the *PbFCl*-type structure becomes lowest in enthalpy and the most stable at pressures above 0.6 GPa. From the HSE06 high-pressure *H(p)* curves, the same qualitative sequence is retained, but the transition pressures are approximately 0.3 GPa for *exp1* → *β* and approximately 1.15 GPa for the *β* → *PbFCl* phase transition. The other high-pressure phases identified in the *E(V)* analysis remain metastable in the investigated pressure range.

This sequence is physically consistent with the structural characteristics identified in Section 7.2. The *exp1*-type phase contains a sevenfold Ho coordination environment, and among the relevant high-pressure structures, the *β* phase contains an eightfold Ho-centered mixed-anion local environment, while the *PbFCl*-type phase shows a still denser ninefold Ho coordination. The corresponding HSE06 curves in Figure 7.22 agree with the general PBE picture, even though the transition pressures and the ordering of some metastable competitors change slightly.

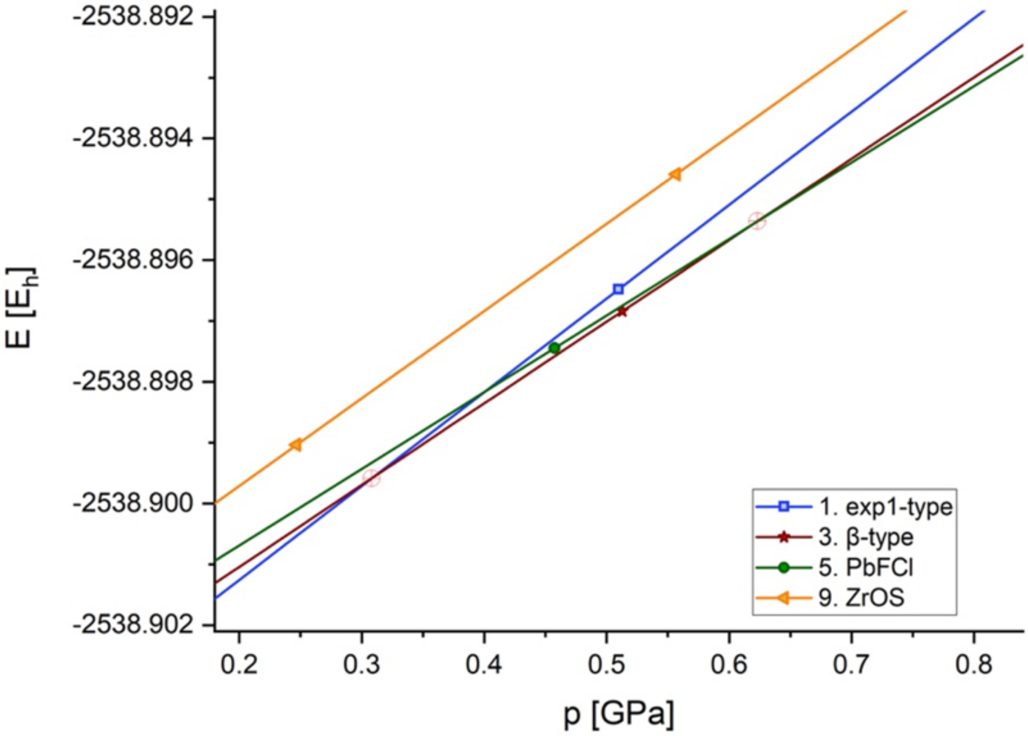

Figure 7.20. The *H(p)* curves for the four most relevant high-pressure HoFSe modifications calculated using the GGA-PBE functional.

The negative-pressure region is described by the enthalpy vs pressure, *H(p)*, curves shown in Figure 7.21. Here the *CoYC*-type modification becomes the lowest-enthalpy phase below approximately -0.8 GPa, while *γ*-type remains metastable throughout the explored pressure interval. In contrast to the positive-pressure side where two successive phase transitions are predicted, the expanded side of the HoFSe landscape shows only one low-density modification that becomes thermodynamically stable in the investigated pressure interval. The HSE06 negative-pressure *H(p)* curves shown in Figure 7.23 preserve the same qualitative behavior, but the *exp1* → *CoYC* transition is shifted to a lower negative pressure of approximately -1.55 GPa. Therefore, HSE06 calculations confirm the same qualitative phase sequence while shifting the phase transition pressures.

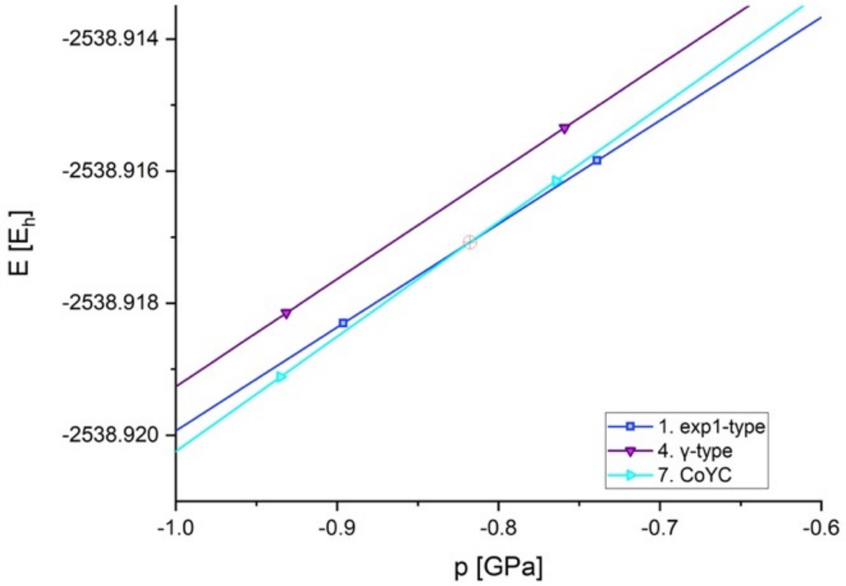

Figure 7.21. The *H(p)* curves for the relevant negative-pressure HoFSe modifications calculated using the GGA-PBE functional.

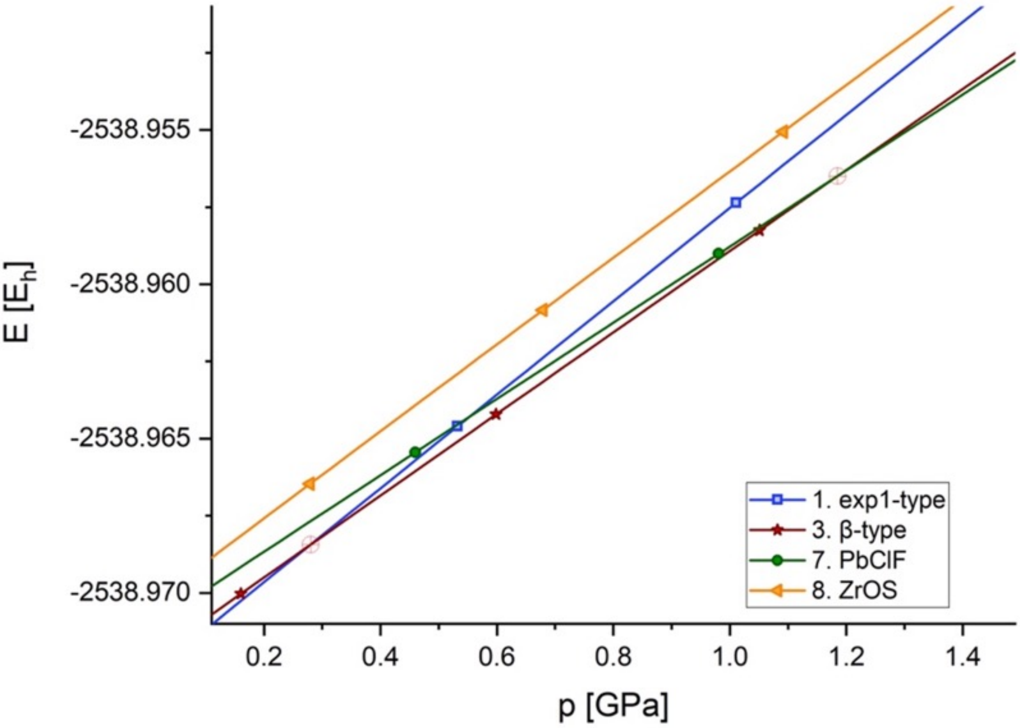

Figure 7.22. The HSE06 *H(p)* curves for the four most relevant high-pressure HoFSe modifications.

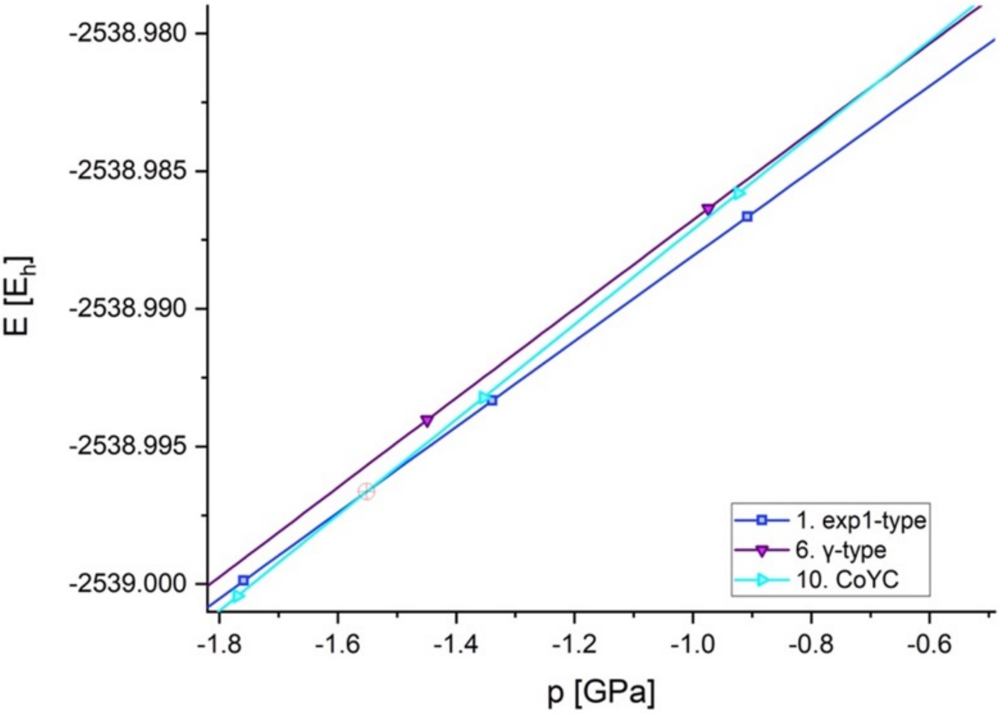

Figure 7.23. The HSE06 *H(p)* curves for the three most relevant negative-pressure HoFSe modifications.

7.5 Magnetic Ground State and Anisotropy

Spin-polarized DFT+U calculations were performed for the *exp1*-type ground-state modification using two approaches: scalar-collinear calculations for various inequivalent Ho spin arrangements, and non-collinear calculations that include spin-orbit coupling (SOC) for different initial directions of the local Ho magnetic moments. Throughout this chapter, the symbols ↑ and ↓ denote the collinear Ho spin-up and spin-down orientations, respectively.

All the structures with different collinear spin arrangements were fully optimized so that the final relaxed geometries could respond to the selected Ho spin pattern. The collinear calculations were initialized by assigning positive or negative local Ho magnetic moments, corresponding to the up (↑) and down (↓) spin-polarization patterns. The Ho 4*f* states were treated with the Dudarev [66] DFT+U approach using U_{eff} = 5.3 eV for Ho 4*f* states and the GGA-PBE functional. After the structural relaxations, band-structure and projected-DOS calculations were performed. Non-collinear spin calculations including spin-orbit coupling were then performed on the lowest-energy collinear ↑↓↑↓ AFM-1 relaxed geometry, using DFT+U with the same U_{eff} = 5.3 eV and GGA-PBE functional. Twelve different three-dimensional starting directions of the local Ho magnetic moments were used to explore the non-collinear spin/SOC energy landscape.

7.5.1 Collinear Ho Spin Arrangements

The results for all five symmetry-inequivalent primitive-cell collinear Ho-spin arrangements are shown in Table 7.4. The antiferromagnetic ↑↓↑↓ AFM-1 arrangement in the primitive cell of the *exp1*-HoFSe phase (Figure 7.24) is the lowest-energy magnetic state calculated using DFT+U with PBE functional and $U_{\mathrm{eff}} = 5.3$ eV for Ho $4f$ states. The second antiferromagnetic arrangement, ↑↓↓↑, lies 5.4 meV per cell higher, while the ferrimagnetic ↑↑↑↓ FiM-1 state and the antiferromagnetic ↑↑↓↓ AFM-3 state follow with approximately 6.4 and 8.6 meV above the ground state energy, respectively. The ferromagnetic ↑↑↑↑ state is the highest-energy primitive-cell configuration, 10.5 meV per cell above the ↑↓↑↓ AFM-1 ground state.

Table 7.4. The collinear PBE+U spin configurations for *exp1*-type HoFSe ground state. The ΔE is given relative to the lowest-energy ↑↓↑↓ AFM-1 arrangement; the cell contains four HoFSe formula units, i.e., four Ho atoms, and their spin up/down orderings are shown using up/down arrows.

Ho spin pattern	E (eV/cell)	ΔE (meV/cell)	M_{tot} (μ_B/cell)	Ho moments (μ_B)	E_{gap} (eV)
↑↓↑↓ (AFM-1)	-91.77983	0.00	0.00	+4.01, -4.01, +4.01, -4.01	1.597
↑↓↓↑ (AFM-2)	-91.77446	5.37	0.00	+4.01, -4.01, -4.01, +4.01	1.590
↑↑↑↓ (FiM-1)	-91.77343	6.40	8.00	+4.00, +4.01, +4.00, -4.01	1.581
↑↑↓↓ (AFM-3)	-91.77123	8.60	0.00	+4.01, +4.01, -4.01, -4.01	1.595
↑↑↑↑ (FM)	-91.76938	10.45	16.00	+4.00, +4.00, +4.00, +4.00	1.590

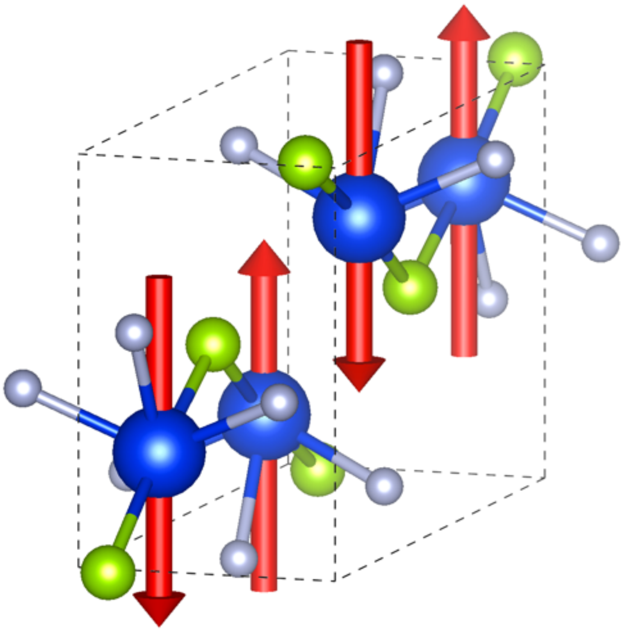

Figure 7.24. Magnetic structure of the minimum-energy primitive-cell collinear ↑↓↑↓ AFM-1 configuration. Ho moments are shown as red arrows; Ho, F, and Se atoms are shown in blue, light gray, and green, respectively.

The projected Ho spin moments remain well localized and close to 4 μ_B for all five primitive-cell collinear patterns, while the induced F and Se moments are negligible. These local projected moments cancel almost exactly only in the compensated antiferromagnetic configurations, giving zero total magnetization for ↑↓↑↓ AFM-1, ↑↓↓↑ AFM-2, and ↑↑↓↓ AFM-3 arrangements. The ferrimagnetic

↑↑↑↓ FiM-1 and ferromagnetic ↑↑↑↑ configurations retain finite total magnetization of approximately 8 and 16 μ_B per unit cell, respectively.

Summary of results for an additional set of 18 non-equivalent collinear configurations in a supercell containing two primitive cells is shown in Table 7.5. These results confirm the same magnetic hierarchy obtained in the primitive cell: compensated antiferromagnetic arrangements are the lowest-energy states, whereas ferrimagnetic and fully ferromagnetic configurations lie higher in energy. The three lowest antiferromagnetic supercell arrangements lie within approximately 4.4 meV per supercell, the first ferrimagnetic configuration appears approximately 12.5 meV above the ground state, and the fully ferromagnetic configuration lies 26.1 meV above it. Thus, increasing the cell size and allowing additional non-equivalent spin sequences does not change the main result that *exp1*-HoFSe modification favors antiferromagnetic ordering.

Table 7.5. Supercell collinear PBE+U spin configurations for *exp1*-type HoFSe. ΔE is given relative to the lowest-energy ↓↓↑↑↑↑↓↓ AFM-2 spin-arrangement. The equivalent primitive-cell orientation is given in parentheses where applicable.

Collinear-spin arrangement	E (eV/cell)	ΔE (meV/cell)	M_{tot} (μ_B/cell)	Ho moments (μ_B)
↓↓↑↑↑↑↓↓ (AFM-2)	-183.565877	0.00	0.00	-4.01, -4.01, +4.01, +4.01, +4.01, +4.01, -4.01, -4.01
↑↑↑↑↓↓↓↓ (AFM-3)	-183.564483	1.39	0.00	+4.01, +4.01, +4.01, +4.01, -4.01, -4.01, -4.01, -4.01
↓↓↑↑↓↓↑↑ (AFM-1)	-183.561480	4.40	0.00	-4.01, -4.01, +4.01, +4.01, -4.01, -4.01, +4.01, +4.01
↓↑↓↑↑↑↑↑ (FiM-2)	-183.553333	12.54	16.00	-4.01, +4.00, -4.01, +4.00, +4.00, +4.00, +4.00, +4.00
↑↑↑↑↓↓↑↑ (FiM-1)	-183.550234	15.64	16.00	+4.00, +4.01, +4.00, +4.00, -4.01, -4.01, +4.00, +4.00
↓↑↓↓↑↑↓↑ (AFM-4)	-183.548496	17.38	0.00	-4.01, +4.01, -4.01, -4.01, +4.01, +4.01, -4.00, +4.01
↑↑↓↑↑↓↑↑ (FiM-3)	-183.546046	19.83	16.00	+4.01, +4.00, -4.01, +4.00, +4.00, -4.01, +4.00, +4.00
↑↑↑↑↑↓↓↓ (FiM-4)	-183.545801	20.08	8.00	+4.01, +4.01, +4.00, +4.01, +4.01, -4.01, -4.01, -4.01
↑↑↓↑↓↑↑↑ (FiM-5)	-183.544814	21.06	16.00	+4.00, +4.01, -4.01, +4.00, -4.01, +4.00, +4.01, +4.01
↓↑↑↓↓↑↑↓ (AFM-5)	-183.544647	21.23	0.00	-4.01, +4.01, +4.01, -4.01, -4.01, +4.01, +4.01, -4.01
↓↑↑↑↓↓↑↑ (FiM-6)	-183.544441	21.44	8.00	-4.01, +4.01, +4.01, +4.01, -4.01, -4.01, +4.01, +4.00
↑↑↓↑↓↓↑↑ (FiM-7)	-183.543793	22.08	8.00	+4.01, +4.01, -4.01, +4.01, -4.01, -4.01, +4.01, +4.01
↓↑↓↑↓↑↓↑ (AFM-6)	-183.543535	22.34	0.00	-4.01, +4.01, -4.01, +4.01, -4.01, +4.01, -4.01, +4.01
↓↑↑↑↓↑↑↑ (FiM-8)	-183.540993	24.88	16.00	-4.01, +4.00, +4.01, +4.00, -4.01, +4.00, +4.00, +4.00
↑↓↑↓↑↓↓↑ (AFM-7)	-183.540268	25.61	0.00	+4.01, -4.01, +4.01, -4.01, +4.01, -4.01, -4.01, +4.01
↑↑↑↑↑↑↑↑ (FM)	-183.539759	26.12	32.00	+4.00, +4.00, +4.00, +4.00, +4.00, +4.00, +4.00, +4.00
↑↑↓↑↑↓↓↓ (AFM-8)	-183.539522	26.36	0.00	+4.01, +4.01, -4.01, +4.01, +4.01, -4.01, -4.01, -4.01
↑↑↑↑↓↑↑↑ (FiM-9)	-183.536922	28.96	24.00	+4.00, +4.00, +4.00, +4.00, -4.01, +4.00, +4.00, +4.00

7.5.2 Non-Collinear Spin-Orbit-Coupled Calculations

The non-collinear magnetic calculations were performed for the primitive cell of the *exp1*-type structure starting from the structurally optimized antiferromagnetic ↑↓↑↓ AFM-1 arrangement, but allowing the different arrangements of initial local Ho moments to freely rotate in three dimensions while including spin-orbit coupling. This is an essential extension for the HoFSe system because Ho is a rare-earth element with atomic configuration [Xe]$4f^{11}6s^2$, while the trivalent Ho^{3+} cation relevant for rare-earth fluoroselenides has a [Xe]$4f^{10}$ electronic configuration. For such localized 4*f* systems, spin-orbit coupling and the resulting magnetocrystalline anisotropy can strongly influence the preferred moment direction, with the total energy depending on the moment orientation relative to the crystal lattice [62, 200].

The energy ranking of the non-collinear spin–orbit-coupled calculations is summarized in Table 7.6. The two lowest-energy solutions, NCL-7 and NCL-1, are the lowest-energy magnetic arrangements of the *exp1*-type HoFSe structure. The NCL-7 is the absolute minimum, with a small residual total moment of $|M_{tot}| \approx 0.96$ μ_B per primitive cell, representing a weakly uncompensated, spin-orbit-coupled canted antiferromagnetic arrangement, while NCL-1 lies only 0.69 meV/cell higher and is an almost completely compensated antiferromagnetic state, with $|M_{tot}| \approx 0.02$ μ_B. The NCL-7 and NCL-1 spin arrangements are shown in Figures 7.25 and 7.26, respectively.

A second group is formed by NCL-6, NCL-4, NCL-8, and NCL-5 spin configurations. These states lie within approximately 1.1–1.7 meV/cell above NCL-7 in total energy and have small to moderate uncompensated total moments of approximately 1.56–1.92 μ_B per primitive cell. They can be described as low-energy weakly-uncompensated antiferromagnetic or ferrimagnetic-like solutions. Their near-degeneracy with the lowest-energy NCL-7/NCL-1 group shows that the low-energy magnetic landscape of *exp1*-HoFSe is shallow with respect to several spin-orbit-coupled Ho-moment orientations.

The next group consists of more strongly uncompensated ferrimagnetic-like arrangements. The NCL-10, NCL-9, and NCL-11 have substantially larger total magnetic moments, $|M_{tot}| \approx 3.20$, 4.76, and 6.82 μ_B per primitive cell, respectively, and are also significantly higher in energy, lying approximately 8.6–9.2 meV/cell above NCL-7. This indicates that increasing the uncompensated part of the total magnetic moment generally tends to destabilize the magnetic solution. However, the total moment is not the only parameter that determines the magnetic ground state.

The final group, NCL-12, NCL-2, and NCL-3, has mostly compensated or only weakly uncompensated total moments of about 0.19–1.55 μ_B per cell, yet these states are much higher in energy, from approximately 10 to 31 meV/cell above the ground state. The clearest example is NCL-3, which is the highest-energy solution in the set, despite being an almost fully compensated antiferromagnetic state, with $|M_{tot}| \approx 0.19$ μ_B. Conversely, NCL-1 is also almost fully compensated, with $|M_{tot}| \approx 0.02$ μ_B, but is the second lowest-energy state. Although magnetic compensation broadly correlates with the energy ranking, the dominant physical factor is the spin-orbit-coupled orientation of the Ho magnetic moments and the associated magnetocrystalline anisotropy, i.e., how the Ho magnetic moments are oriented relative to the crystal-field environment and crystallographic directions. Additionally, the magnitude of local Ho spin-projected moments remains close to 3.90 μ_B for all the non-collinear arrangements, again indicating that the main differences among these states lie in the orientations of the individual Ho magnetic moments and in the degree of their mutual compensation.

Table 7.6. The non-collinear / SOC PBE+U energy ranking for the primitive-cell *exp1*-type HoFSe structure, ordered by rising energy difference ΔE, given relative to the lowest-energy NCL-7 spin arrangement.

Spin arrangement	E (eV/cell)	ΔE (meV/cell)	$\lvert M_{tot}\rvert$ (μ_B/cell)	M_{tot} vector (μ_B)	|Ho moments| (μ_B)
NCL-7	-95.485679	0.00	0.959	(+0.009, -0.042, -0.958)	3.898–3.899
NCL-1	-95.484993	0.69	0.021	(+0.005, +0.020, -0.004)	3.897–3.897
NCL-6	-95.484614	1.07	1.916	(-0.971, -0.572, -1.550)	3.897–3.898
NCL-4	-95.484536	1.14	1.784	(+0.773, +0.581, -1.499)	3.897–3.899
NCL-8	-95.484517	1.16	1.832	(+0.214, -1.254, +1.317)	3.897–3.899
NCL-5	-95.483961	1.72	1.566	(+0.004, -0.001, -1.566)	3.897–3.899
NCL-10	-95.477097	8.58	3.200	(+1.185, +1.105, -2.760)	3.896–3.899
NCL-9	-95.476878	8.80	4.763	(+0.097, +4.488, +1.591)	3.896–3.901
NCL-11	-95.476514	9.17	6.816	(+3.630, +0.312, -5.760)	3.894–3.899
NCL-12	-95.475898	9.78	1.059	(+0.087, +0.168, -1.042)	3.896–3.900
NCL-2	-95.468203	17.48	1.548	(-0.219, -0.046, +1.532)	3.897–3.899
NCL-3	-95.454777	30.90	0.188	(-0.000, +0.188, -0.001)	3.898–3.899

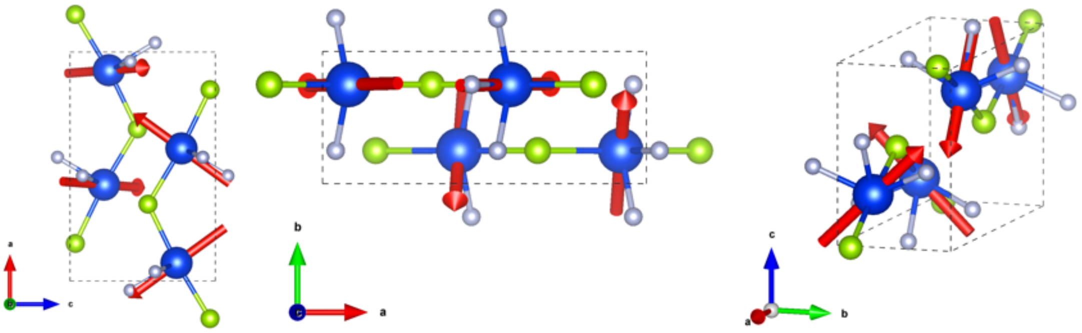

Figure 7.25. Magnetic structure of the lowest-energy non-collinear NCL-7 arrangement. The red arrows represent the Ho moment directions in the primitive-cell *exp1*-type structure. The projections are shown on the *xz* plane (left) and *xy* plane (middle), followed by an isometric view (right).

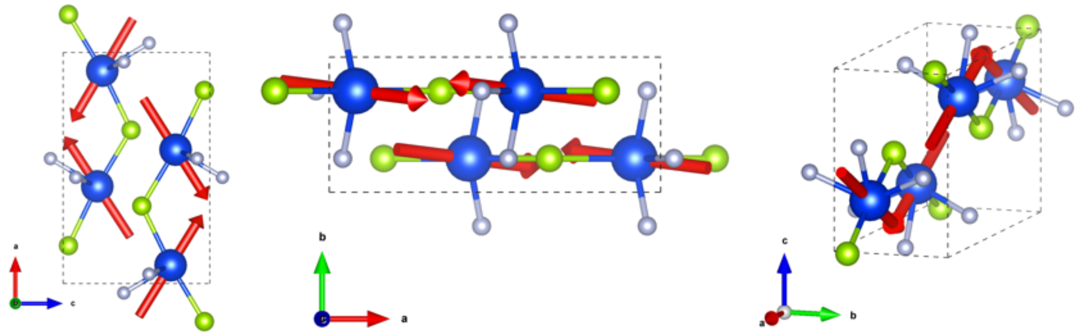

Figure 7.26. Magnetic structure of the fully compensated antiferromagnetic NCL-1, the second-lowest-energy non-collinear arrangement. The projections are shown on the *xz* plane (left) and *xy* plane (middle), followed by an isometric view (right).

7.6 Electronic Structure of the Magnetic Ground State

The electronic structure was first analyzed for the set of primitive-cell collinear magnetic arrangements. Since ↑↓↑↓ AFM-1 is the lowest-energy solution in the scalar-collinear primitive-cell calculations, its band structure and projected density of states are taken to represent the primitive-cell collinear-spin ground state in the discussion. The result is a semiconducting state with a band gap of approximately 1.60 eV. The valence-band maximum is dominated mainly by Se(*p*) states, while F(*p*) states and the occupied part of the Ho 4*f* states lie deeper in the valence region, approximately 5-6 eV below the Fermi level (Figure 7.27). The lowest conduction band is a very narrow Ho 4*f*-derived band approximately 1.6 eV above the valence-band edge, reflecting both the localized character of the Ho 4*f* states and the DFT+U splitting between occupied and unoccupied 4*f* states. The ↑↓↑↓ AFM-1 arrangement also shows the most visibly degenerate band pattern among the collinear configurations, suggesting that the enhanced degeneracy is associated with the different symmetry of the relaxed antiferromagnetic ground state compared to other AFM states. More dispersive conduction bands beginning approximately 1 eV above the lowest Ho 4*f*-derived conduction band are dominated mainly by Ho 5*d* states.

The other collinear spin arrangements give very similar band gaps of approximately 1.58-1.60 eV, (Table 7.4), showing that changing the Ho spin pattern mainly affects the total energy, while the gap size remains nearly unchanged.

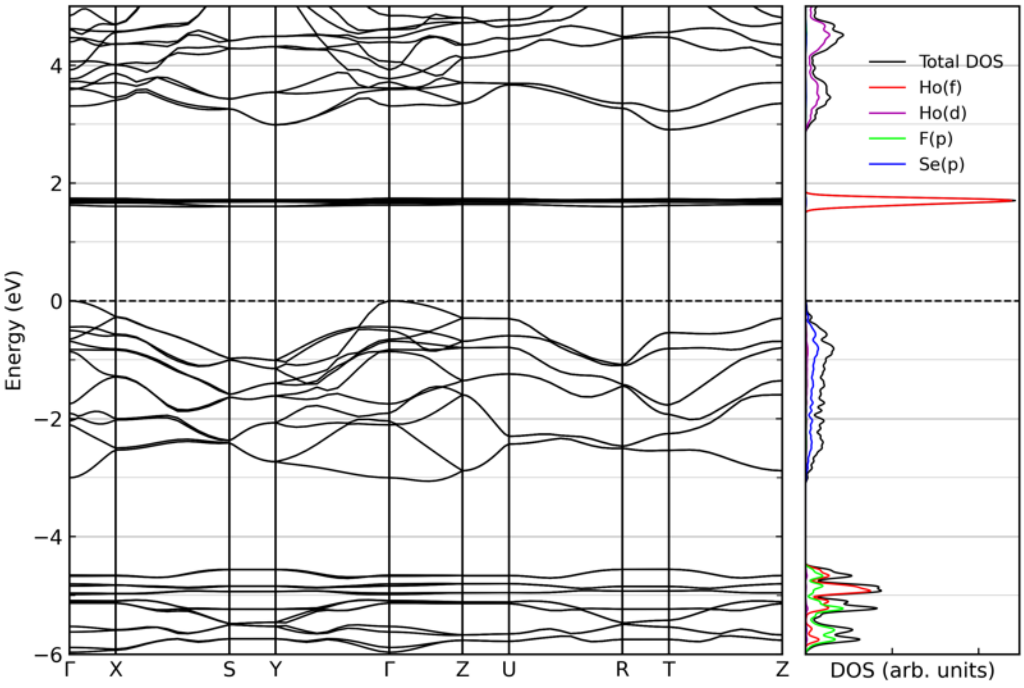

Figure 7.27. Electronic band structure and projected DOS of the ground-state collinear-spin configuration ↑↓↑↓ AFM-1 of *exp1*-HoFSe phase, calculated with PBE+U (U_{eff} = 5.3 eV for Ho 4*f* states).

The non-collinear NCL-7 band structure (Figure 7.28) remains semiconducting and is very similar to the NCL-1 band structure shown in Figure 7.33, indicating that the two lowest non-collinear/SOC states differ mainly in magnetic orientation and total energy. The most striking SOC-related feature is the splitting of the narrow Ho(*f*)-dominated conduction manifold into two nearly dispersionless localized bands approximately 1.6 and 1.8 eV above the Fermi level. The Se(*p*)-dominated valence-band region, the deeper valence-band F(*p*) states, and the Ho(*d*)-derived conduction bands remain essentially unchanged compared with the collinear calculation.

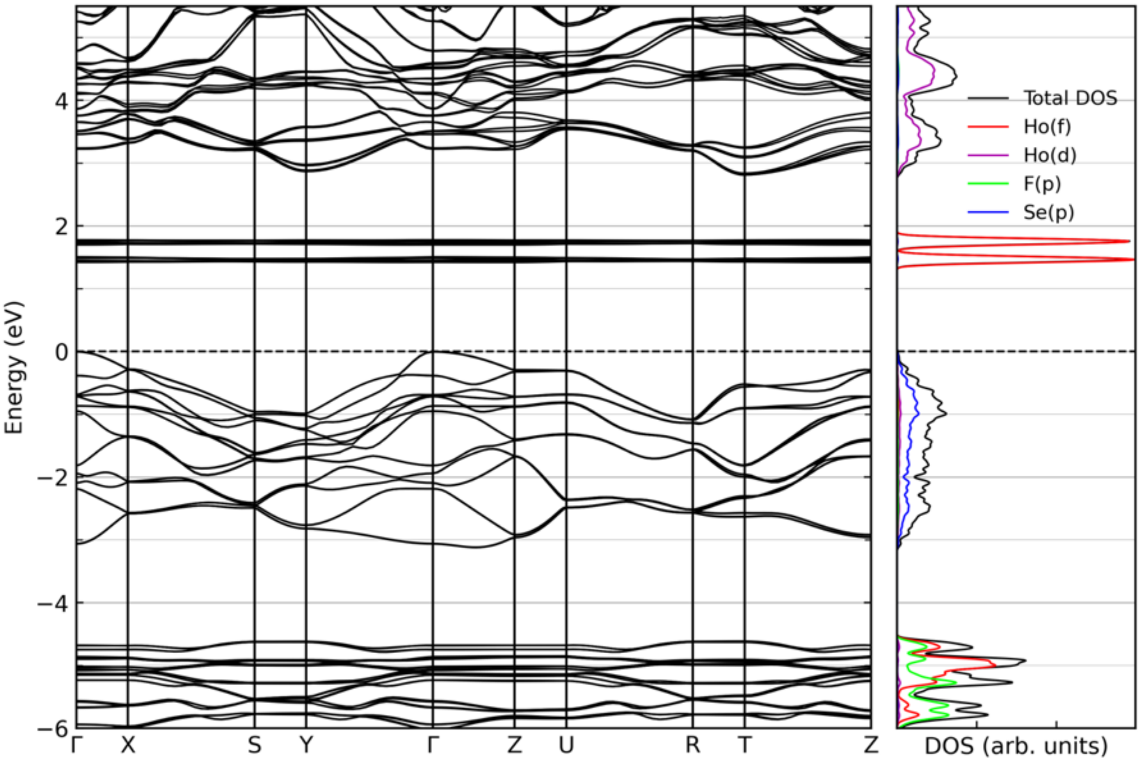

Figure 7.28. Band structure and projected DOS of the minimum-energy non-collinear/SOC NCL-7 configuration, calculated with PBE+U (U_{eff} = 5.3 eV for Ho 4*f* states).

The band structures in Figures 7.29–7.33 show that the semiconducting electronic structure is robust with respect to the tested magnetic arrangements. The collinear AFM-2, FiM-1, AFM-3, and FM configurations in Figures 7.29–7.32 retain the similar projected-DOS pattern as the AFM-1 ground state: the upper valence region is dominated mainly by Se(*p*) states, F(*p*) states lie deeper within the valence band, and the low-lying conduction region contains narrow Ho 4*f* states followed by more dispersive Ho 5*d* bands. The principal differences among the collinear configurations are small changes in spin splitting and band degeneracy. The NCL-1 band structure in Figure 7.33 is also very similar to the NCL-7 ground-state electronic band structure, confirming that the two lowest non-collinear/SOC states differ slightly by their magnetic orientation and total energy, while preserving the same semiconducting band structure.

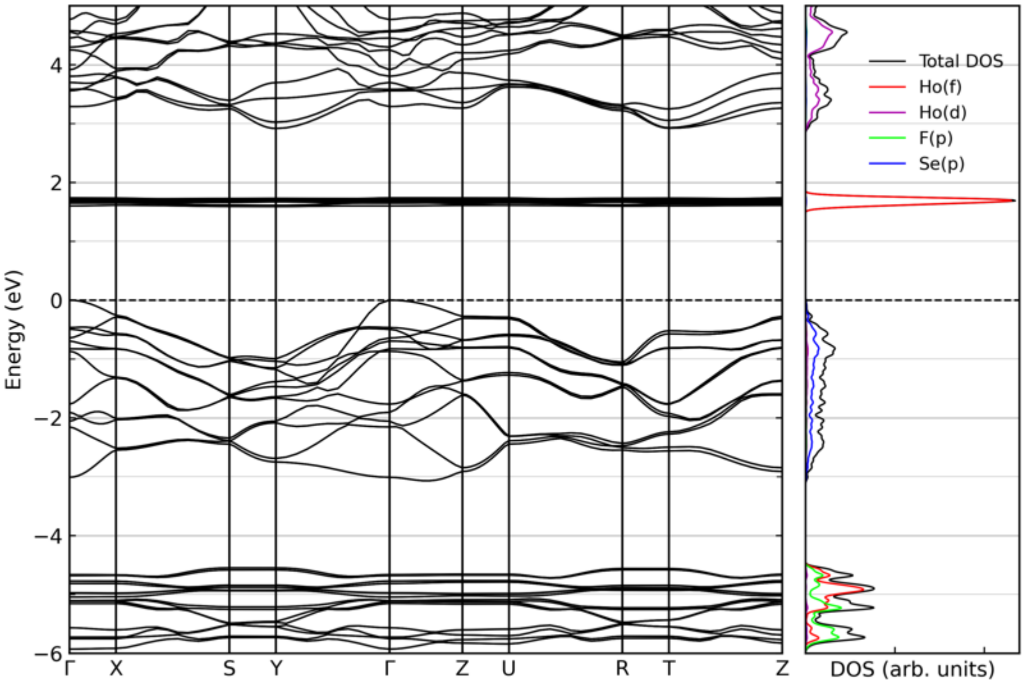

Figure 7.29. Electronic band structure and projected DOS of the AFM collinear-spin configuration ↑↓↓↑ AFM-2 of *exp1*-HoFSe phase, calculated with PBE+U ($U_{eff} = 5.3$ eV for Ho 4*f* states).

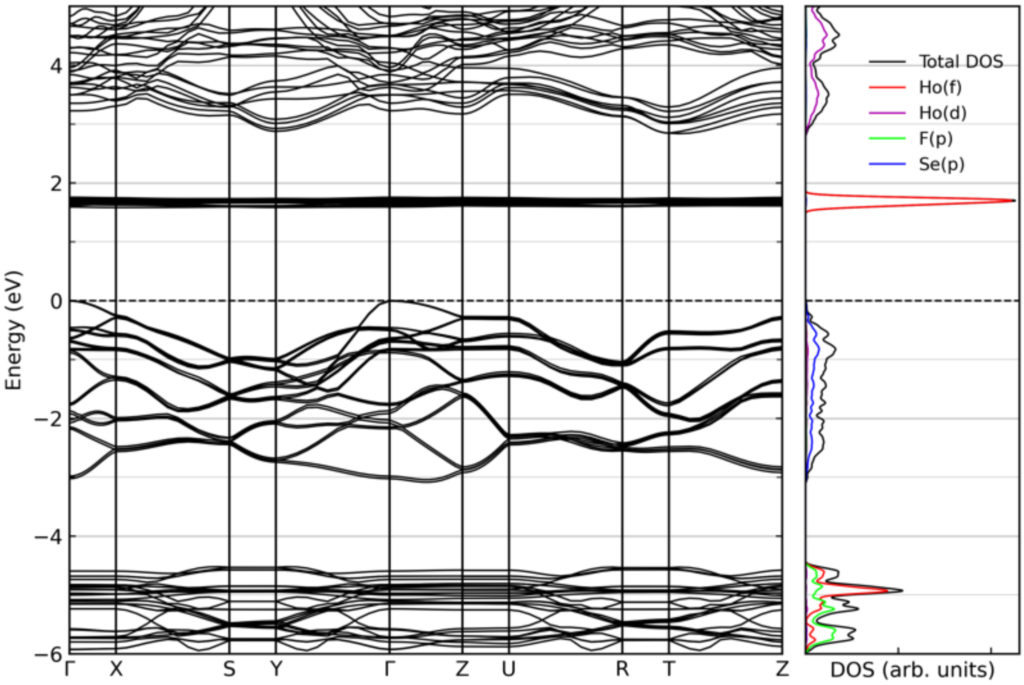

Figure 7.30. Electronic band structure and projected DOS of the FiM collinear-spin configuration ↑↑↑↓ FiM-1 of *exp1*-HoFSe phase, calculated with PBE+U ($U_{eff} = 5.3$ eV for Ho 4*f* states).

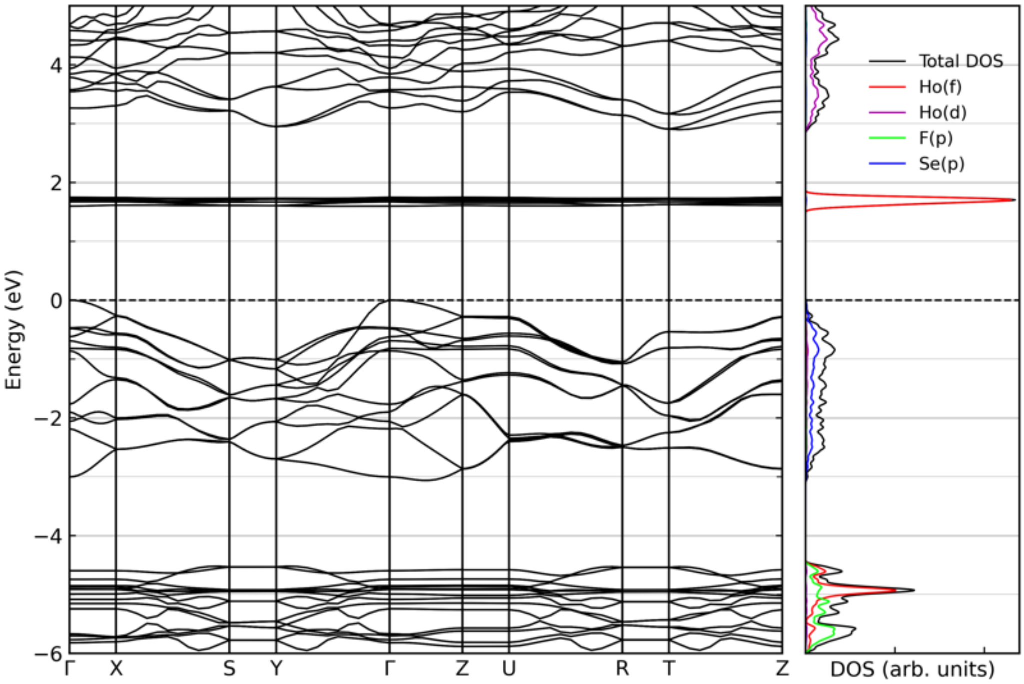

Figure 7.31. Electronic band structure and projected DOS of the AFM collinear-spin configuration ↑↑↓↓ AFM-3 of *exp1*-HoFSe phase, calculated with PBE+U (U_{eff} = 5.3 eV for Ho 4*f* states).

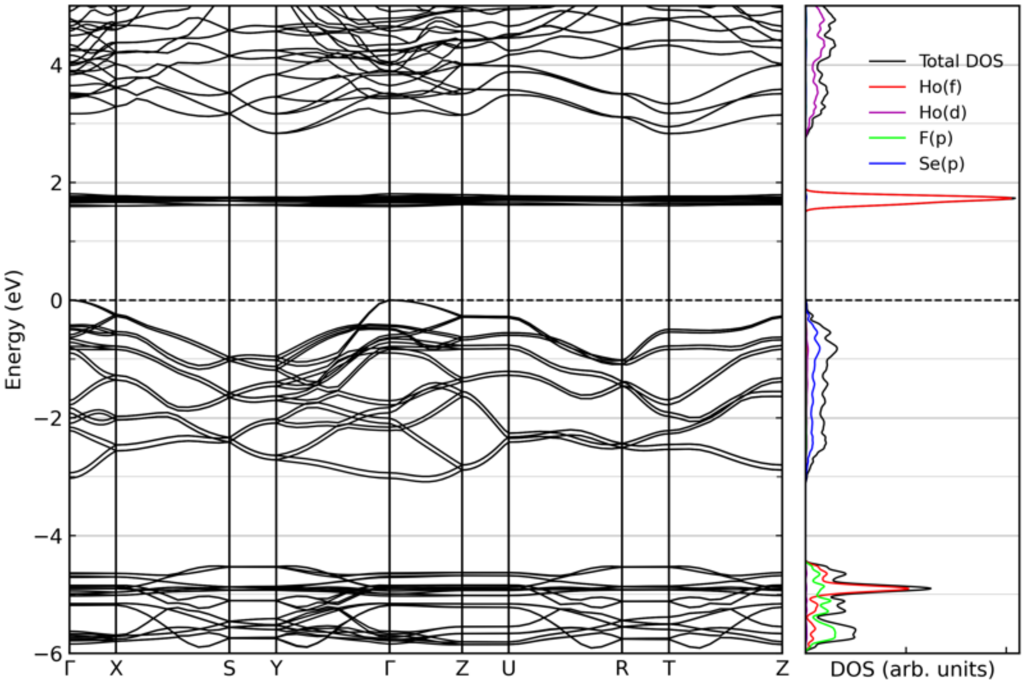

Figure 7.32. Electronic band structure and projected DOS of the FM collinear-spin configuration ↑↑↑↑ FM of *exp1*-HoFSe phase, calculated with PBE+U (U_{eff} = 5.3 eV for Ho 4*f* states).

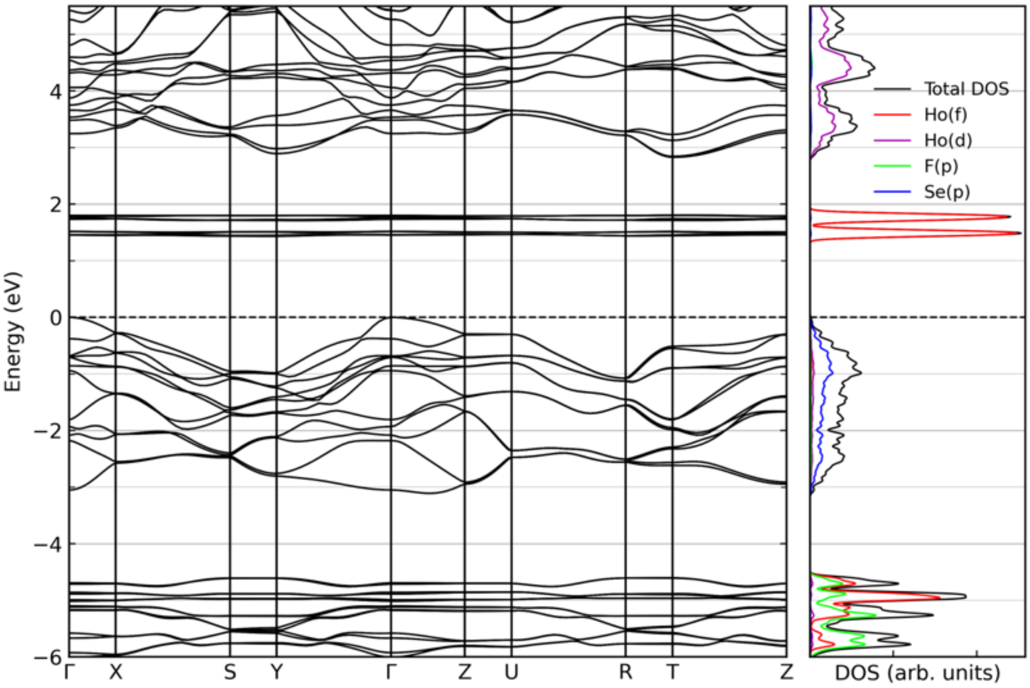

Figure 7.33. Band structure and projected DOS of the non-collinear/SOC NCL-1 configuration, calculated with PBE+U (U_{eff} = 5.3 eV for Ho 4*f* states)

7.7 Conclusion

The HoFSe study demonstrates how combining energy-landscape exploration with experimental crystal chemistry can provide a very broad structural picture of this mixed-anion rare-earth compound. Global optimization and data mining searches generated a large pool of candidates, and subsequent DFT refinement reduced this candidate group to twenty most relevant low-energy polymorphs. The recovery of the experimentally observed *exp1*-type, *exp2*-type, and *exp3*-type structures among the low-energy candidates confirms that the combined search strategy captures the physically important parts of the HoFSe landscape. At the same time, the seventeen additional predicted phases extend the experimentally known HoFSe structural landscape by identifying possible metastable, high-pressure, low-density, and high-temperature structure types.

The structural analysis shows that HoFSe polymorphism covers orthorhombic, monoclinic, tetragonal, hexagonal, rhombohedral, and cubic symmetry classes, with Ho coordination numbers ranging mostly from six to nine depending on the structure type. The *exp1*-type ground state is a sevenfold coordinated *LiCaN*-type structure, while compression favors denser eight- and ninefold Ho environments such as the *β*-type and *PbFCl*-type modifications. In contrast, the low-density side of the landscape is characterized by more open arrangements such as *CoYC*-type and *κ*-type structures with reduced Ho coordination.

The thermodynamic analysis identifies the *exp1*-type phase as the ambient-pressure ground state. Under pressure, GGA-PBE calculations predict a phase transition first to *β*-type HoFSe above approximately 0.3 GPa and then to *PbFCl*-type modification above approximately 0.6 GPa. The HSE06 *H(p)* curves confirm the same qualitative phase sequence, while shifting the $\beta \rightarrow PbFCl$ transition to approximately 1.15 GPa. On the expanded, effective negative-pressure side of the landscape, the *CoYC*-type modification becomes thermodynamically favored below approximately -0.8 GPa within GGA-PBE and below approximately -1.55 GPa within HSE06 calculations. The agreement between the GGA-PBE and HSE06 phase sequence shows that the predicted pressure response leads to robust phase transitions very likely to be achieved in the experiments.

The magnetic calculations provide an extension to the published HoFSe study. Collinear PBE+U calculations show that antiferromagnetic arrangements are lowest in energy in both primitive-cell and supercell calculations. The non-collinear spin-orbit-coupled calculations then reveal a relatively shallow magnetic-anisotropy landscape in which the weakly uncompensated canted antiferromagnetic NCL-7 arrangement is the absolute minimum, while the almost completely compensated antiferromagnetic NCL-1 arrangement lies only about 0.7 meV/cell higher. The comparison of low-energy and high-energy non-collinear states shows that magnetic compensation of the antiferromagnetic arrangement alone does not determine the ground state and that the orientation of the Ho magnetic moments relative to the crystal-field is also important.

The electronic-structure results show that all *exp1*-type magnetic states are semiconducting. The lowest-energy collinear ↑↓↑↓ AFM-1 configuration has a band gap of approximately 1.60 eV, and all the other collinear configurations show nearly the same gap size. The non-collinear NCL-7 ground state remains semiconducting and differs mainly through the spin-orbit-induced splitting of the narrow Ho 4*f*-derived conduction band into two almost dispersionless bands. In all magnetic configurations considered, the top of the valence-band is dominated mainly by Se(*p*) states, while the bottom of the conduction band contains localized Ho 4*f*-derived states followed by more dispersive Ho *d* states above. Changing the magnetic arrangement and including spin-orbit coupling in the calculations modify degeneracies and splittings of the bands near the band gap, but do not change the overall semiconducting character of *exp1*-HoFSe, with band gap remaining approximately the same for all the configurations.

8. Energy Landscape, Pressure-Induced Polymorphism, Electronic and Vibrational Properties, and AI-Guided High-Pressure Search in Lanthanum Fluorosulfide (LaFS)

8.1 Introduction

The structural, thermodynamic, electronic, and vibrational results for lanthanum fluorosulfide (LaFS) summarized in this chapter are based primarily on the LaFS study [12], except for the AI-guided exploration in the extreme high-pressure region discussed in Section 8.7. The LaFS system was investigated using a hierarchical multi-stage workflow that coupled broad crystal-structure exploration with first-principles refinement. Similar to earlier energy-landscape exploration based studies on LaOI [8] and HoFSe [11], and within the broader energy-landscape framework [13], the LaFS exploration consisted of (i) global optimization on an empirical-potential energy landscape over different cell sizes and pressures, (ii) complementary prototype-based data mining, (iii) post-processing by symmetry analysis, structural comparison, clustering, and screening, and (iv) final DFT relaxation and electronic and vibrational property evaluation. This established approach is supplemented by an AI-guided search aimed specifically at the extreme high-pressure region.

Lanthanum fluorosulfide belongs to the broader family of ternary rare-earth mixed-anion compounds, a group of materials in which the combination of rare-earth cations with different anion species offers a useful route for tuning structure and physical properties. Rare-earth elements and their compounds are important in modern technologies including electronics, magnetic materials, energy-storage systems, optical materials, and other functional applications [1, 201]. Rare-earth chalcogenides and related chalcohalide materials have also been discussed in connection with thermoelectric conversion, laser technology, low-field magnetotransport, lanthanide-containing light-emitting organic-inorganic hybrids, and photovoltaic concepts [176-180].

Rare-earth fluoride sulfides and fluoride selenides are particularly interesting because the combination of lanthanide cations with two chemically different anions can modify the local lanthanide coordination environment, the structural topology, and the electronic band structure. Earlier work on lanthanoid fluoride sulfides and selenides has pointed to their semiconducting and luminescence-related potential [202, 203]. Experimentally, rare-earth(III) fluoride sulfides are generally synthesized by high-temperature solid-state reactions involving rare-earth metals, sulfur, and rare-earth trifluorides, and the known LaFS phase adopts the tetragonal *PbFCl*-type, or *matlockite*-type, structure [29, 204].

Recent studies of chemically related rare-earth mixed-anion systems, including LaFSe and HoFSe, showed that such compounds can possess rich structural landscapes and strongly structure-dependent electronic properties [11, 30, 31]. The LaFS study continues this line of research by combining large-scale energy-landscape exploration, DFT structural refinement, thermodynamic stability analysis, band-structure and density-of-states calculations, phonon dispersions, and an additional AI-guided search for extreme high-pressure polymorphs [12].

8.2 Computational Setup and Energy Landscape Exploration Overview

Global structural searches were carried out by simulated annealing through stochastic Monte Carlo walks on an empirical-potential landscape, combined with periodic stochastic quenches as implemented in the G42+ framework [13]. No symmetry constraints were imposed, and both the cell parameters and all atomic coordinates were allowed to vary freely. To keep the exploration of millions of candidates computationally tractable, a fast two-body empirical potential composed of damped Coulomb and Lennard-Jones terms was used.

The global optimization covered periodic cells containing 1, 2, 3, 4, 6, 8, 12, and 16 formula units (up to 48 atoms), so that both simple and more complex ordering motifs could be sampled. Explicit external pressures of 0, 0.02, 1, 5, 10, 50, and 100 GPa were applied, and at finite pressure the objective function was the empirical enthalpy. To broaden the accessible range of local coordinations, the effective ionic radii in the empirical potential were additionally scaled by up to ±20% relative to their baseline values. This systematically modified the short-range repulsive length scale and complemented the explicit pressure scan by mimicking more compressive or more expansive conditions.

Because empirical searches cannot guarantee exhaustive prototype coverage within feasible computational cost, a complementary data-mining stage was also employed. In total, 484 ternary ABX, ANX, and AXY prototypes from the ICSD and 77 low-energy structures from chemically related rare-earth systems were mapped to LaFS by atomic substitution, generating 2,981 additional candidates [12].

Selected structure candidates were optimized by DFT calculations using VASP [54, 55] and the PAW method [53]. Structural relaxation and primary energetic ranking were carried out with GGA-PBE [49, 120], while selected low-energy polymorphs were re-evaluated with the screened hybrid HSE06 functional [121]. Thermodynamic stability under pressure was assessed by fitting $E(V)$ data to a third-order Birch–Murnaghan equation of state and by analyzing the corresponding $H(p)$ curves. For electronic properties, separate self-consistent calculations on the relaxed geometries were followed by non-self-consistent DOS and band-structure calculations from the converged charge density. Phonon dispersions were obtained within PBE by density-functional perturbation theory (DFPT) and post-processed with Phonopy [38]. A compact summary of the numerical VASP settings used in these calculations is collected in Appendix A8.

The electronic-structure part of the study was used to connect polymorphism with the structure-dependent electronic properties of LaFS. The calculated band structures and projected densities of states were analyzed together to characterize the electronic structures of the LaFS polymorphs, including their band-gap behavior, band dispersion, and dominant orbital contributions. The vibrational properties study served a complementary role: phonon dispersions were calculated for the ambient-pressure ground state and the two relevant high-pressure candidates in order to test whether the predicted phases correspond to local dynamical minima and whether the pressure-driven phase competition is associated with any soft-mode instability.

Machine learning entered the workflow at two distinct stages. First, rapid ML-assisted pre-relaxation and energy ranking using universal interatomic-potential tools implemented through MatGL [107, 108] were used during the standard screening stage within the STyX package. Second, for the targeted high-pressure search of the LaFS system, a custom AI-ELX graph-neural-network model was trained on more than 20,000 DFT self-consistent points spanning different structures, volumes, and pressure environments, and was then used to guide the selection of the most promising extreme-pressure candidates. This AI model represents crystal structures as periodic graphs in which atoms act as nodes and interatomic neighborhoods act as edges, an approach closely related to message-passing atomistic models and crystal-graph neural networks developed for materials applications [83, 88, 89].

Since the AI-ELX methodology is described in one of the previous chapters, mainly the aspects directly related to the LaFS study are addressed here. In the present chapter, the important point is its practical role as a computationally fast workflow that can screen candidate structures, pre-relax selected GO/DM configurations, and generate approximate $E(V)$ trends before computationally-expensive DFT calculations are performed. The principal hyperparameters used for the LaFS AI-

ELX model are shown in Appendix A3 (Table A3.3), while the actual AI-guided high-pressure screening results including the newly-found phases are presented in Section 8.7.

8.3 Optimized Geometry – Symmetry and Structure Type Analysis

The combined global-optimization and prototype-based data-mining procedures yielded 3,966,798 candidate structures (3,963,817 from GO and 2,981 from DM) [12]. Their statistical distribution over cell size and applied pressure is summarized in Tables 8.1 and 8.2.

Table 8.1. Frequency of occurrence of various space groups generated during global optimization, categorized by the number of formula units per simulation cell.

1 LaFS		2 LaFS		3 LaFS		4 LaFS	
Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence
*P*1 (1)	306,632	*P*4/*nmm* (129)	272,834	*P*1 (1)	282,803	*Pnma* (62)	383,407
Cm (8)	74,760	$P2_1/m$ (11)	66,663	*Cm* (8)	173,758	*P*1 (1)	231,489
*P*3*m*1 (156)	39,618	*P*4/*mmm* (123)	61,598	*C*2 (5)	31,968	*P*4/*nmm* (129)	173,011
Pm (6)	17,941	*C*2/*m* (12)	14,500	*Amm*2 (38)	23,254	$P2_12_12_1$ (19)	57,076
*R*3*m* (160)	13,497	*P*1 (1)	11,175	*Pm* (6)	21,903	$P6_3/mmc$ (194)	32,479
*C*2 (5)	2,232	$Pmn2_1$ (31)	9,489	*P*-6*m*2 (187)	12,218	$P2_1/m$ (11)	21,661
*Amm*2 (38)	400	*P*-6*m*2 (187)	8,989	*R*3*m* (160)	9,498	*Pm* (6)	17,323
*Imm*2 (44)	321	*P*-1 (2)	6,293	*P*-62*m* (189)	5,474	$P2_1$ (4)	15,710
*P*2 (3)	166	*Pm* (6)	3,098	*Imm*2 (44)	2,596	$P2_13$ (198)	13,003
*Fmm*2 (42)	27	$P2_1$ (4)	1,438	*P*3*m*1 (156)	2,238	*P*-6*m*2 (187)	11,228
P-1 (2)	20	*Cm* (8)	1,137	*Pmm*2 (25)	562	*Pmm*2 (25)	9,648
*Pmm*2 (25)	18	*Imm*2 (44)	925	*P*31*m* (157)	437	*P*-1 (2)	7,813
*Cmm*2 (35)	6	*Cc* (9)	847	*F*-43*m* (216)	273	$Pmn2_1$ (31)	7,017
F-43*m* (216)	6	$Pmc2_1$ (26)	320	*P*2 (3)	141	$P2_1/c$ (14)	5,102
*P*2/*m* (10)	1	*Amm*2 (38)	172	*I*-4*m*2 (119)	52	*Cm* (8)	4,436
OTHER	0	OTHER	1,032	OTHER	19	OTHER	36,932
TOTAL	455,645	TOTAL	460,510	TOTAL	567,194	TOTAL	1,027,335

6 LaFS		8 LaFS		12 LaFS		16 LaFS	
Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence	Space group	Freq. of occurrence
*P*1 (1)	327,547	*P*1 (1)	296,705	*P*1 (1)	319,069	*P*1 (1)	106,025
$P2_1/m$ (11)	82,473	*Pnma* (62)	47,169	*Pm* (6)	9,772	*Pnma* (62)	1,145
Pm (6)	52,042	*Pm* (6)	24,401	*Pnma* (62)	7,398	*Pm* (6)	277
*P*4/*nmm* (129)	20,820	*Cm* (8)	7,288	$P2_1/m$ (11)	4,225	*P*-1 (2)	153
F-43*m* (216)	17,829	*P*4/*nmm* (129)	6,021	*Cm* (8)	2,595	$P2_1$ (4)	53
P-1 (2)	15,447	$P2_1$ (4)	5,172	$P2_1$ (4)	1,714	$P2_1/m$ (11)	52
Cm (8)	12,584	$P2_1/m$ (11)	4,619	$Pmc2_1$ (26)	1,494	*Pmm*2 (25)	52
$Pmn2_1$ (31)	12,063	*P*-1 (2)	3,482	*P*-62*m* (189)	1,472	*Cm* (8)	38
$P2_1$ (4)	10,571	$P2_1/c$ (14)	3,212	*P*-1 (2)	1,103	*Pc* (7)	12
Pc (7)	5,065	*Pbam* (55)	1,902	*P*4/*nmm* (129)	455	*P*2 (3)	10
P-6*m*2 (187)	4,077	*Pc* (7)	1,687	*Pc* (7)	338	$Pca2_1$ (29)	10
*P*3*m*1 (156)	3,831	*Pmm*2 (25)	1,592	*F*-43*m* (216)	326	*Pbcn* (60)	9
*Amm*2 (38)	2,863	$Pmn2_1$ (31)	1,260	$P2_1/c$ (14)	179	$P2_1/c$ (14)	5
*C*2/*m* (12)	2,832	$Pmc2_1$ (26)	1,152	$Pmn2_1$ (31)	164	*P*-42*m* (111)	5
$Pmc2_1$ (26)	1,841	*F*-43*m* (216)	998	*Pmmn* (59)	112	*P*-4 (81)	2
OTHER	8,821	OTHER	7,166	OTHER	335	OTHER	2
TOTAL	580,706	TOTAL	413,826	TOTAL	350,751	TOTAL	107,850

Table 8.2. Frequency of occurrence of various space groups generated during global optimization, categorized by external applied pressure.

0 Pa	
Space group	Freq. of occurrence
$P1$ (1)	758,006
$Pnma$ (62)	170,887
$P4/nmm$ (129)	82,806
Pm (6)	48,451
Cm (8)	45,762
$P2_1/m$ (11)	43,369
$P2_12_12_1$ (19)	29,879
$P4/mmm$ (123)	27,603
$P2_1$ (4)	15,232
$P\text{-}1$ (2)	14,407
$Pmn2_1$ (31)	14,110
$F\text{-}43m$ (216)	11,910
$C2$ (5)	10,275
$P6_3/mmc$ (194)	8,277
$R3m$ (160)	7,911
OTHER	41,253
TOTAL	1,330,138

20 MPa	
Space group	Freq. of occurrence
$P1$ (1)	352,031
$P4/nmm$ (129)	60,941
$Pnma$ (62)	53,939
Cm (8)	24,498
$P4/mmm$ (123)	24,206
Pm (6)	23,919
$P2_1/m$ (11)	22,787
$P2_12_12_1$ (19)	8,125
$Pmn2_1$ (31)	6,812
$C2$ (5)	6,671
$P\text{-}1$ (2)	5,927
$P2_1$ (4)	5,467
$F\text{-}43m$ (216)	4,810
$C2/m$ (12)	3,451
$R3m$ (160)	3,227
OTHER	17,537
TOTAL	624,348

1 GPa	
Space group	Freq. of occurrence
$P1$ (1)	182,373
$Pnma$ (62)	44,121
$P4/nmm$ (129)	38,385
Cm (8)	21,604
$P2_1/m$ (11)	14,606
Pm (6)	13,727
$P4/mmm$ (123)	10,784
$P2_12_12_1$ (19)	7,642
$C2$ (5)	7,124
$P2_1$ (4)	3,575
$P\text{-}1$ (2)	3,524
$F\text{-}43m$ (216)	3,179
$Pmn2_1$ (31)	2,945
$P2_13$ (198)	2,788
$R3m$ (160)	2,637
OTHER	12,059
TOTAL	371,073

5 GPa	
Space group	Freq. of occurrence
$P1$ (1)	193,976
$Pnma$ (62)	46,295
$P4/nmm$ (129)	43,550
Cm (8)	28,435
$P2_1/m$ (11)	20,987
Pm (6)	14,259
$C2$ (5)	6,585
$P2_12_12_1$ (19)	6,458
$R3m$ (160)	4,992
$C2/m$ (12)	4,550
$P\text{-}1$ (2)	3,753
$Pmn2_1$ (31)	3,425
$P2_1$ (4)	3,218
$P4/mmm$ (123)	3,207
$P\text{-}6m2$ (187)	2,680
OTHER	15,963
TOTAL	402,333

10 GPa	
Space group	Freq. of occurrence
$P1$ (1)	168,727
$P4/nmm$ (129)	55,248
$Pnma$ (62)	44,565
Cm (8)	36,155
$P2_1/m$ (11)	26,992
Pm (6)	14,918
$P2_12_12_1$ (19)	5,008
$P6_3/mmc$ (194)	4,328
$C2/m$ (12)	4,311
$C2$ (5)	4,086
$P\text{-}6m2$ (187)	4,011
$R3m$ (160)	3,436
$P2_1$ (4)	3,150
$P\text{-}1$ (2)	3,075
$Pmn2_1$ (31)	1,560
OTHER	12,098
TOTAL	391,668

50 GPa	
Space group	Freq. of occurrence
$P1$ (1)	120,704
$P4/nmm$ (129)	88,486
Cm (8)	62,266
$Pnma$ (62)	41,285
$P2_1/m$ (11)	27,386
Pm (6)	15,412
$P\text{-}6m2$ (187)	10,596
$P6_3/mmc$ (194)	7,724
$Amm2$ (38)	3,611
$Pmm2$ (25)	3,498
$P3m1$ (156)	3,210
$P\text{-}62m$ (189)	2,383
$P\text{-}1$ (2)	1,967
$P2_1$ (4)	1,933
$Pmc2_1$ (26)	1,292
OTHER	8,360
TOTAL	400,113

100 GPa	
Space group	Freq. of occurrence
$P1$ (1)	105,628
$P4/nmm$ (129)	103,725
Cm (8)	57,876
$Pnma$ (62)	38,027
$P3m1$ (156)	37,631
$P2_1/m$ (11)	23,566
$Amm2$ (38)	19,427
Pm (6)	16,071
$P\text{-}6m2$ (187)	13,489
$P6_3/mmc$ (194)	6,925
$Pmm2$ (25)	4,475
$P\text{-}62m$ (189)	2,281
$P2_1$ (4)	2,083
$P\text{-}1$ (2)	1,663
$Cmcm$ (63)	1,398
OTHER	9,879
TOTAL	444,144

All pressures	
Space group	Freq. of occurrence
$P1$ (1)	1,881,445
$P4/nmm$ (129)	473,141
$Pnma$ (62)	439,119
Cm (8)	276,596
$P2_1/m$ (11)	179,693
Pm (6)	146,757
$P4/mmm$ (123)	65,955
$P2_12_12_1$ (19)	58,118
$P3m1$ (156)	45,725
$P\text{-}6m2$ (187)	36,730
$C2$ (5)	35,913
$P2_1$ (4)	34,658
$P\text{-}1$ (2)	34,316
$P6_3/mmc$ (194)	33,418
$Pmn2_1$ (31)	29,994
OTHER	192,239
TOTAL	3,963,817

Low-symmetry $P1$ structures dominate the landscape under most search conditions. This is a natural consequence of unconstrained stochastic exploration, especially at high effective temperatures in simulated annealing, where many trial configurations visit low-symmetry regions of configuration space before quenching. At the same time, the repeated appearance of higher-symmetry prototypes, especially $Pnma$ and $P4/nmm$, indicates that these arrangements correspond to robust minima on the empirical-potential landscape. At 100 GPa the share of $P1$ structures decreases, while dense high-symmetry phases become comparatively more prominent [12].

All GO and DM candidates were then processed with Structure Type eXplorer (STyX), which uses the pymatgen library for symmetry determination and structural comparison [105], together with rapid ML-assisted screening based on universal interatomic potentials implemented through the MatGL library [107, 108]. The workflow performed symmetry determination, structural comparison, clustering of non-$P1$ representatives, ranking, and finally the selection for subsequent refinement, while storing the results in an SQL database for reproducible retrieval and statistics. After filtering and clustering, the dataset contained 21,275 unique non-$P1$ representative structures and about 1.88 million $P1$ structures; because of their large number and diversity, the $P1$ structures were treated as unique and were not compared among themselves, which was not computationally feasible for such a large and heterogeneous group. About 300 structurally distinct candidates were finally selected for DFT refinement on the basis of empirical energy, recurrence on the landscape, symmetry, and rapid ML-assisted energy predictions.

Following DFT refinement and *E(V)* analysis, seventeen low-energy LaFS polymorphs were selected as the most relevant part of the *ab initio* landscape. Their space groups, energies, volumes per formula unit, and optimized lattice parameters at the PBE and HSE06 levels are summarized in Table 8.3.

Table 8.3. Energy-based ranking and geometric parameters of the 17 most promising LaFS polymorphs, evaluated at the PBE and HSE06 levels.

Modification and structure type	Space group No.	PBE				HSE06				Unit cell params. (Å) - other results
		E_{RI}	Energy / FU (eV)	Vol. / FU (Å^3)	Unit cell params. (Å)	E_{RI}	Energy / FU (eV)	Vol. / FU (Å^3)	Unit cell params. (Å)	
PbFCl-type	129	1	-20.160	57.63	a = 4.051 c = 7.023	1	-23.523	56.84	a = 4.034 c = 6.987	a=4.042[a] c=6.960[a] a=4.09[b] c=7.03[b] a=4.13[c] c=7.14[c]
TiNiSi-type	62	2	-20.138	58.21	a = 6.980 b = 4.032 c = 8.274	2	-23.495	57.47	a = 6.950 b = 4.016 c = 8.237	
A-LaFS	62	3	-20.123	58.01	a = 6.998 b = 4.033 c = 16.445	3	-23.477	57.27	a = 6.964 b = 4.015 c = 16.386	
B-LaFS	25	4	-20.109	57.84	a = 4.040 b = 16.307 c = 7.023	4	-23.461	57.09	a = 4.024 b = 16.233 c = 6.992	
C-LaFS	38	5	-20.095	58.28	a = 4.048 b = 16.344 c = 14.094	5	-23.449	57.45	a = 4.029 b = 10.737 α = 98.39	
ZrNiAl-type	189	6	-20.084	57.86	a = 7.040 c = 4.044	7	-23.432	57.14	a = 7.016 c = 4.022	
D-LaFS	87	7	-20.083	62.86	a = 11.264 c = 3.963	6	-23.432	62.09	a = 11.216 c = 3.949	
E-LaFS	109	8	-20.071	58.69	a = 4.056 c = 42.808	9	-23.430	57.85	a = 4.039 c = 42.545	
F-LaFS	59	9	-20.062	58.81	a = 4.034 b = 4.069 c = 21.496	10	-23.418	58.00	a = 4.015 b = 4.053 c = 21.381	
ZrBeSi-type	194	10	-20.059	59.24	a = 4.185 c = 7.811	8	-23.430	58.40	a = 4.159 c = 7.796	a=4.23[b] c=7.77[b] a=4.27[c] c=7.92[c]
G-LaFS	25	11	-20.034	59.11	a = 4.057 b = 4.066 c = 14.334	11	-23.392	58.29	a = 4.042 b = 4.046 c = 14.257	
H-LaFS	26	12	-20.025	61.27	a = 4.188 b = 11.255 c = 7.799	12	-23.369	60.54	a = 4.170 b = 11.204 c = 7.775	
I-LaFS	6	13	-20.020	60.43	a = 8.320 b = 4.060 c = 21.533 β = 94.44	13	-23.364	59.64	a = 8.287 b = 4.041 c = 21.435 β = 94.44	
J-LaFS	62	14	-20.010	74.79	a = 12.573 b = 4.281 c = 5.559	15	-23.328	73.98	a = 12.495 b = 4.266 c = 5.552	
LiMgN-type	62	15	-20.009	70.53	a = 10.091 b = 4.305 c = 6.494	14	-23.340	69.82	a = 10.089 b = 4.284 c = 6.462	
3R-type	166	16	-19.976	70.01	a = 9.093 α = 27.349	16	-23.305	69.32	a = 9.104 α = 27.15	
LiCuO-type	82	17	-19.847	88.32	a = 11.131 c = 5.702	17	-23.133	87.43	a = 11.089 c = 5.689	

a Experimental results [31].
b Theoretical HSE06 results [31].
c Theoretical B3LYP results [31].

Detailed information containing unit-cell parameters and fractional coordinates for all seventeen polymorphs at the GGA-PBE and HSE06 levels is presented in Appendix A8 (Tables A8.1 and A8.2).

The DFT calculations confirm the experimentally documented *PbFCl*-type (*matlockite*-type) phase [29, 204] as the ambient-pressure ground state of LaFS. This result is consistent at both the PBE and HSE06 levels. The phase crystallizes in the tetragonal *P*4/*nmm* space group (No. 129), and each La atom is coordinated by four F and five S atoms, forming edge-sharing ninefold mixed-anion coordination polyhedra that resemble mono-capped square antiprisms. The average La–anion distance is 2.81 Å, with four La–F bonds of 2.60 Å and five La–S bonds between 2.93 and 2.98 Å. Related fluorosulfide structure types have also been discussed in LnSF systems, as well as for earlier rare-earth chalcogenide fluorides [205, 206]. The corresponding structural model is shown in Figure 8.1.

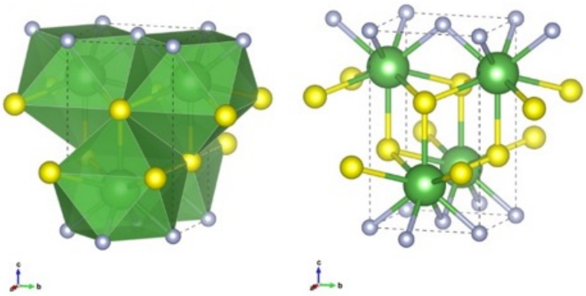

Figure 8.1. Structural model of the global minimum *PbFCl*-type LaFS modification.

Besides the ground state, a number of low-lying local minima were identified. At ambient pressure, the *C*-LaFS and *ZrBeSi*-type modifications are particularly noteworthy because they lie relatively close in energy to the *PbFCl*-type phase and illustrate distinct ways in which mixed-anion coordination can be arranged at similar volumes.

The *C*-LaFS modification crystallizes in the orthorhombic *Amm*2 space group (No. 38) and is the fifth lowest-energy phase at ambient pressure. Lanthanum occupies five inequivalent crystallographic sites, all with distorted ninefold mixed-anion coordination. Depending on the site, the nearest-neighbor environment contains four F and five S atoms, three F and six S atoms, or six F and three S atoms; the corresponding La–F distances span 2.35–2.93 Å and the La–S distances 2.88–3.23 Å (Figure 8.2, left).

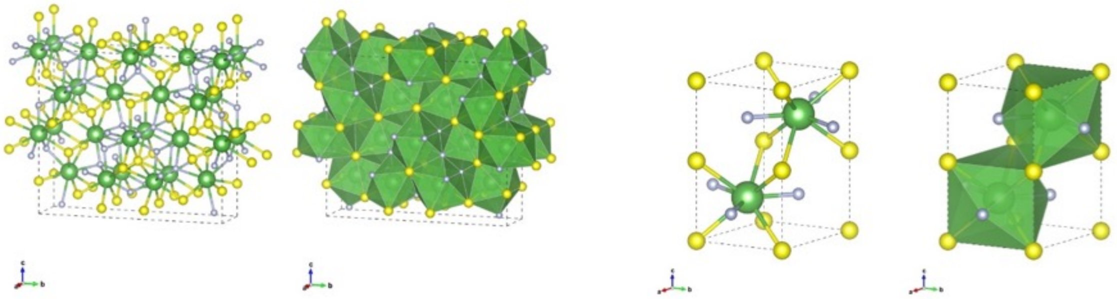

Figure 8.2. Structural models of the *C*-LaFS (left) and *ZrBeSi*-type (right) modifications.

The *ZrBeSi*-type modification crystallizes in the hexagonal $P6_3/mmc$ space group (No. 194) and is ranked tenth in the GGA-PBE energy hierarchy. It adopts the layered arrangement characteristic of the ZrBeSi structure type. There is one crystallographically inequivalent La site with ninefold coordination: three F atoms lie approximately coplanar with the central La atom at 2.42 Å, while six S atoms occur at the distance of 3.11 Å. Related *ZrBeSi*-type phases have been reported previously in chemically related HoFSe and LaFSe systems [11, 30].

In the high-pressure region, the *TiNiSi*-type and *A*-LaFS modifications emerge as the two most relevant candidates. The *TiNiSi*-type phase crystallizes in orthorhombic *Pnma* (No. 62) and contains ninefold La coordination, with an average La–anion distance of 2.82 Å, similar to that in the *PbFCl*-type phase. Its coordination polyhedra contain four La–F neighbor pairs with distances of 2.52–2.95 Å and five La–S pairs with distances between 2.93 and 3.01 Å. The competing *A*-LaFS phase also adopts *Pnma* (No. 62), but has two symmetry-inequivalent La environments, both corresponding to ninefold mixed-anion coordination with 3F+6S and 5F+4S first-neighbor environments, respectively. The average La–anion distance is 2.82 Å, with La–F distances of 2.53–2.99 Å and La–S distances of 2.89–3.03 Å. The *TiNiSi*-type phase is also known in HoFSe, where it occurs in the high-pressure region [11]. The structures are shown in Figure 8.3.

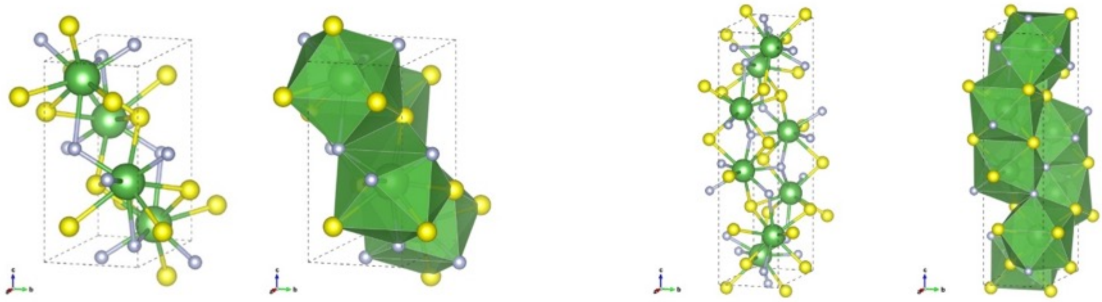

Figure 8.3. Structural models of the high-pressure *TiNiSi*-type (left) and *A*-LaFS (right) modifications.

In the effective negative-pressure region, where low-density structures may in principle be favored by synthesis routes involving low-density amorphous precursor compounds [163, 207], several low-density modifications were identified. Among them, *J*-LaFS, *LiMgN*-type, and *LiCuO*-type have the lowest energies at large volumes. *J*-LaFS and *LiMgN*-type both crystallize in orthorhombic *Pnma* (No. 62), but they represent distinct structure types, whereas *LiCuO*-type adopts the tetragonal *I*-4 space group (No. 82) (Figure 8.4).

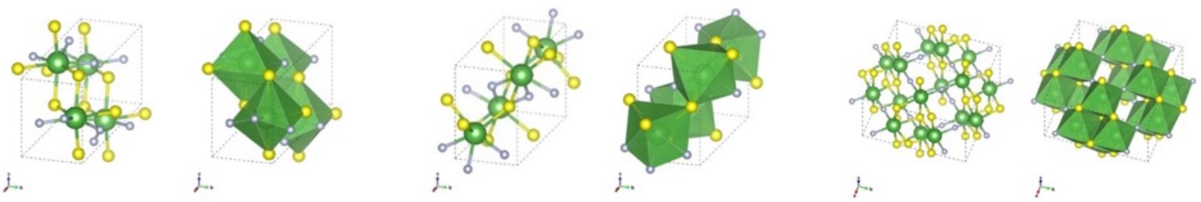

Figure 8.4. Structural models of the negative-pressure *J*-LaFS (left), *LiMgN*-type (middle), and *LiCuO*-type (right) modifications.

Polytypism is another notable feature of the LaFS landscape. The *E*-LaFS and *F*-LaFS modifications, ranked eighth and ninth in the GGA-PBE energy hierarchy, are closely related polytypes distinguished by long repeat distances along the stacking direction: *E*-LaFS crystallizes in tetragonal $I4_1md$ (No. 109) with c = 42.81 Å, whereas *F*-LaFS adopts orthorhombic *Pmmn* (No. 59) with c = 21.50 Å. Both structures contain three inequivalent sites for La, F, and S, consistent with long-period layered stacking sequences. The *3R*-type phase, ranked sixteenth, represents a more symmetric three-layer stacking variant in rhombohedral *R*-3*m* (No. 166), with a considerably simpler asymmetric unit containing two inequivalent La sites and one inequivalent site for each anion species. An analogous 3R polytype has also been reported for HoFSe [11]. These polytypic structures are shown in Figure 8.5.

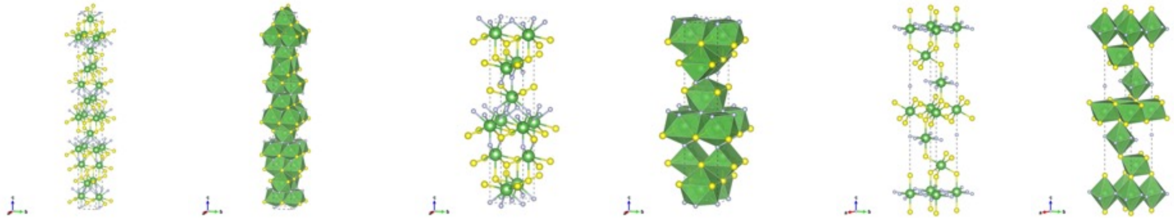

Figure 8.5. Structural models of the polytypic *E*-LaFS (left), *F*-LaFS (middle), and *3R*-type (right) modifications.

The *B*-LaFS and *ZrNiAl*-type modifications (Figure 8.6) are compact low-energy structures with volumes close to that of the *PbFCl*-type ground state. *B*-LaFS crystallizes in the orthorhombic *Pmm*2 space group (No. 25) and contains five inequivalent ninefold-coordinated La sites, whereas the *ZrNiAl*-type phase is a hexagonal *P*-62*m* (No. 189) modification with two inequivalent ninefold La environments. The *D*-LaFS is a tetragonal *I*4/*m* (No. 87) phase with a single eightfold La environment, while *G*-LaFS is an orthorhombic *Pmm*2 (No. 25) phase with four inequivalent ninefold-coordinated La sites (Figure 8.7). The *H*-LaFS and *I*-LaFS (Figure 8.8) illustrate lower-symmetry structures with diverse La coordination. The *H*-LaFS contains three inequivalent La sites with coordination numbers 8, 8, and 9, whereas *I*-LaFS has a monoclinic *Pm* (No. 6) space group with a broad range of inequivalent La environments, including 6-, 7-, 8-, and 9-fold coordination.

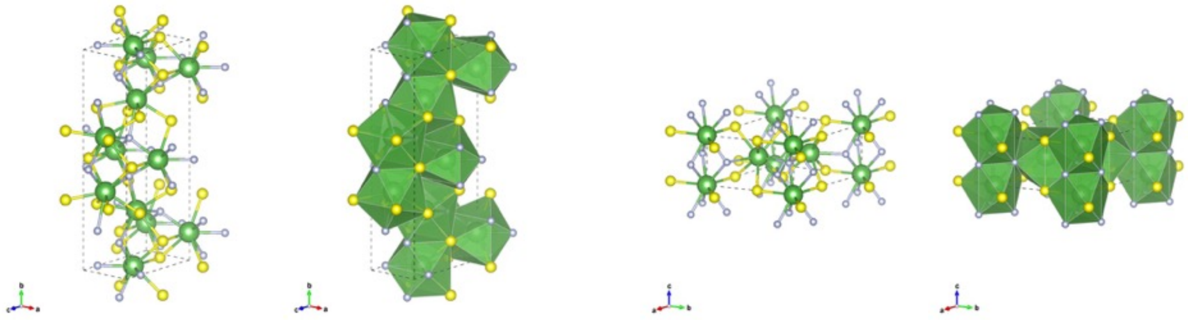

Figure 8.6. Structural models of the *B*-LaFS (left) and *ZrNiAl*-type (right) modifications.

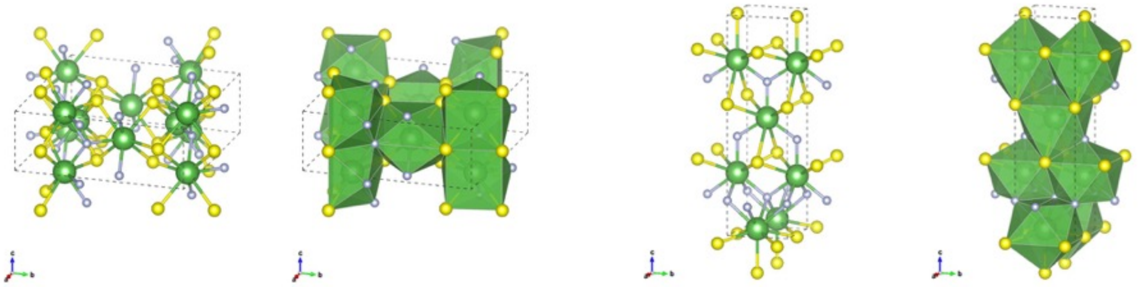

Figure 8.7. Structural models of the *D*-LaFS (left) and *G*-LaFS (right) modifications.

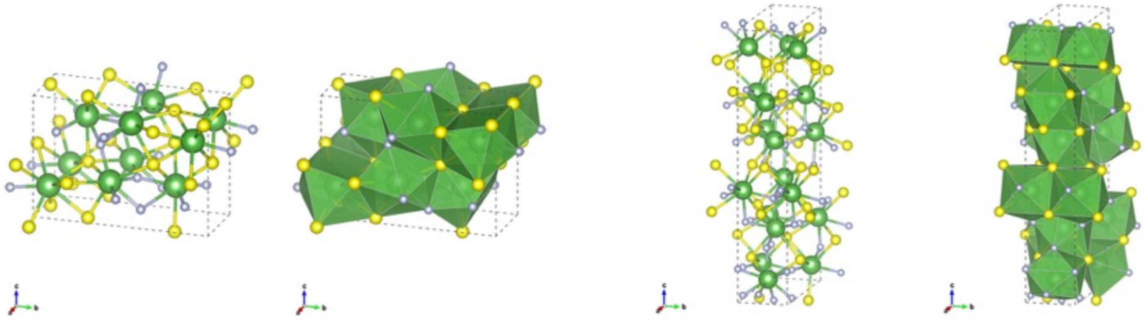

Figure 8.8. Structural models of the *H*-LaFS (left) and *I*-LaFS (right) modifications.

8.4 Thermodynamic Stability and Pressure Response

This section uses the calculated energy vs volume *E(V)* curves and the corresponding enthalpy vs pressure *H(p)* relations to determine the relative stability of the LaFS polymorphs under compression and expansion. The aim is to establish the stability ranges of the relevant low-energy phases, distinguish thermodynamically stable from metastable structures, and predict the pressures at which pressure-induced phase transitions occur.

8.4.1 The *E(V)* Equations of State

Thermodynamic stability as a function of pressure was analyzed from high-resolution energy–volume, *E(V)*, data computed for the seventeen low-energy polymorphs. For the most relevant structures, fixed-volume optimizations spanning approximately ±20% around equilibrium were fitted with a third-order Birch–Murnaghan equation of state.

The resulting GGA-PBE *E(V)* relations are shown in Figure 8.9 for the full low-energy set and in Figures 8.10, 8.11, and 8.12 for the subsets most relevant to the ambient pressure high-temperature region, high-pressure region, and effective negative-pressure regions, respectively.

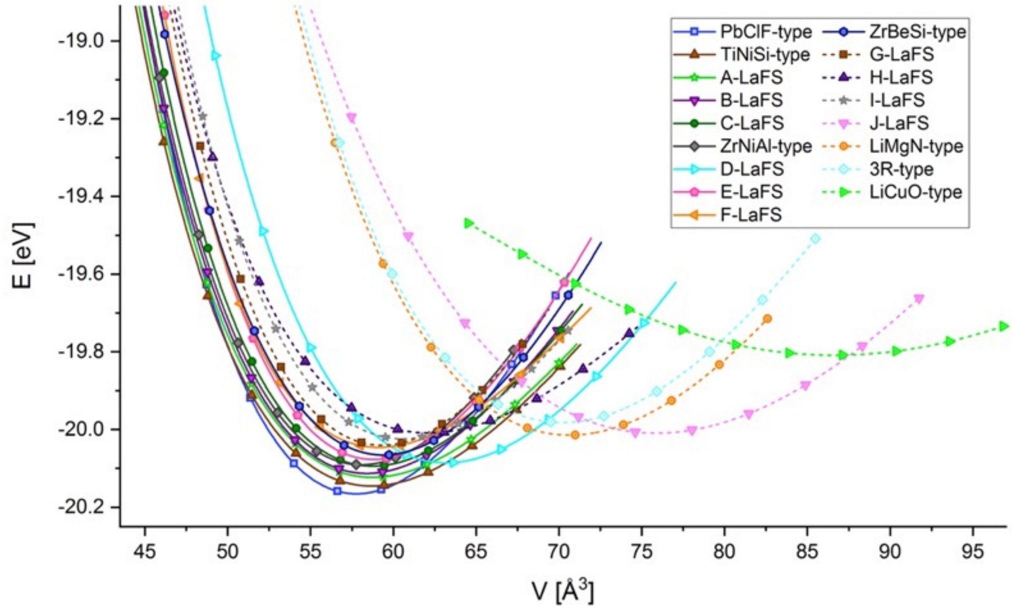

Figure 8.9. The *E(V)* curves for the seventeen predicted low-energy LaFS polymorphs.

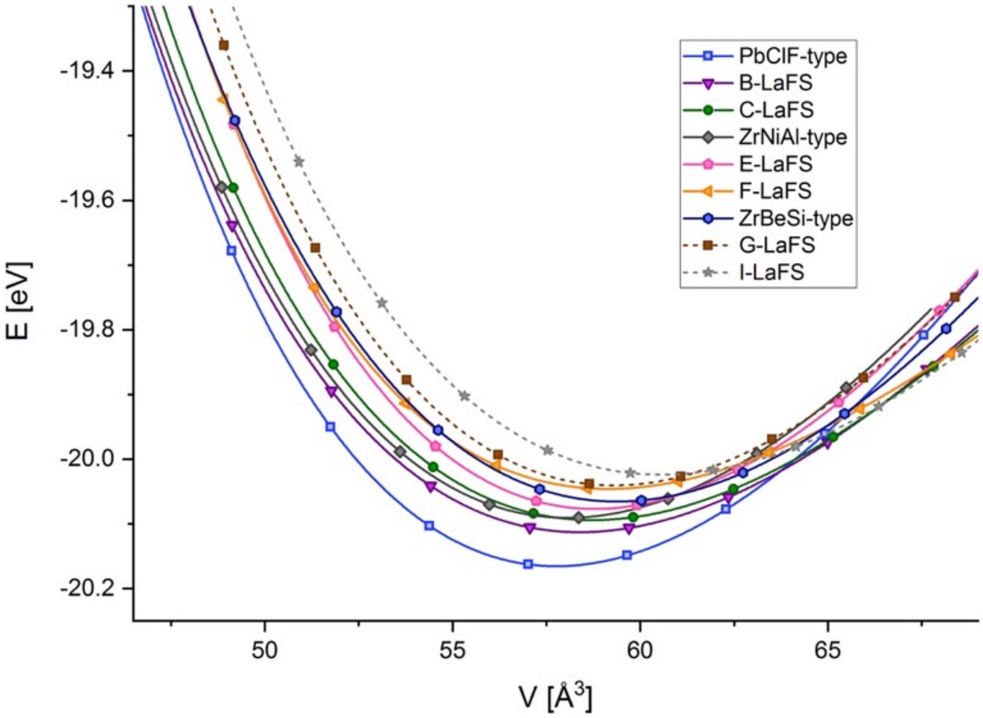

Figure 8.10. The *E(V)* curves of the predicted high-temperature LaFS phases.

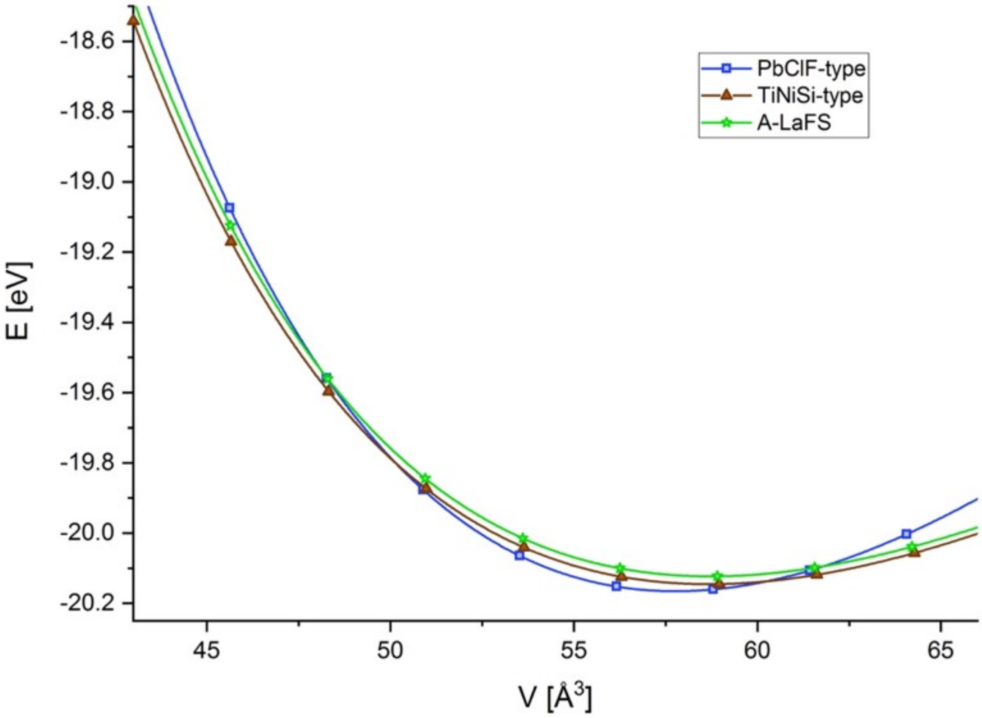

Figure 8.11. The *E(V)* curves of the predicted high-pressure candidates.

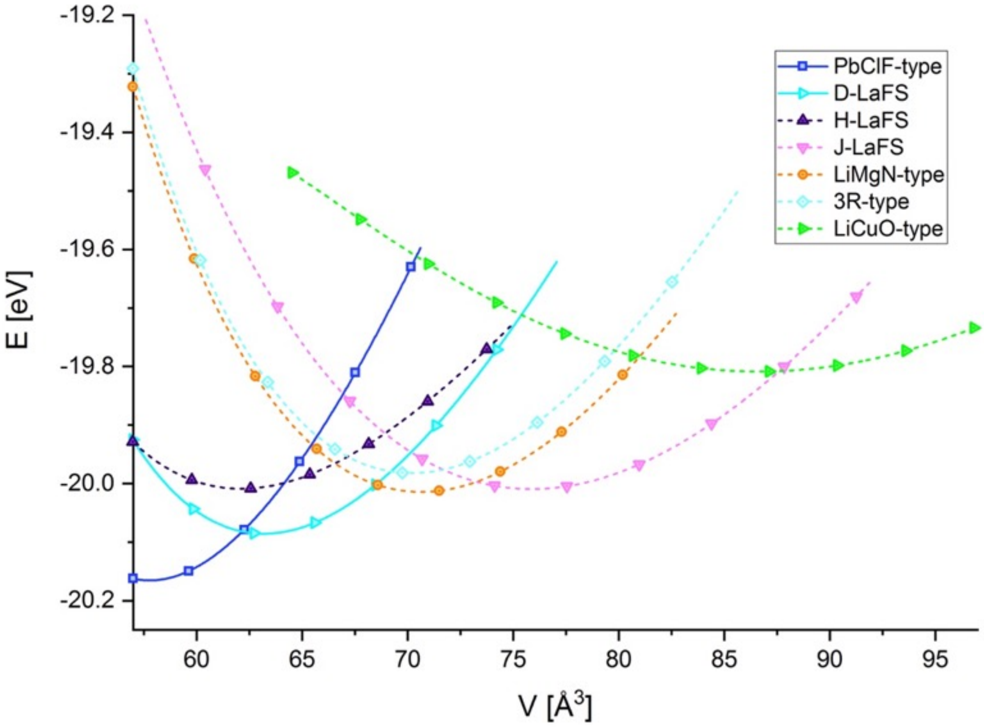

Figure 8.12. The *E(V)* curves of the predicted low-density (effective negative-pressure) modifications.

The *E(V)* curves confirm that the *PbFCl*-type phase occupies the deepest minimum at ambient pressure, establishing it as the thermodynamic ground state [12]. They also separate the landscape into three practically useful regions: low-energy ambient-pressure structures, compressed high-pressure candidates, and expanded low-density modifications. The curves in Figure 8.10 show that several ambient-pressure metastable phases remain relatively close in energy to the ground state over a narrow volume window, which is why they remain structurally relevant despite not becoming stable at zero pressure. In Figure 8.11 the *TiNiSi*-type phase approaches and eventually overtakes the *PbFCl*-type phase under compression, whereas the *A*-LaFS phase remains nearby in energy but does not become the thermodynamic minimum. By contrast, the expanded-volume curves in Figure 8.12 show that the *J*-LaFS, *LiMgN*-type, and *LiCuO*-type modifications become competitive only after substantial lattice expansion, consistent with their interpretation as effective negative-pressure structures.

8.4.2 Enthalpy vs Pressure and Transition Pressures

Pressure-dependent enthalpies were obtained from the fitted *E(V)* curves through *H(p)* = E + pV in order to identify phase competition and transition pressures. Figure 8.13 shows *H(p)* curves for the three most relevant structures in the high-pressure region: *PbFCl*-type, *TiNiSi*-type, and *A*-LaFS [12].

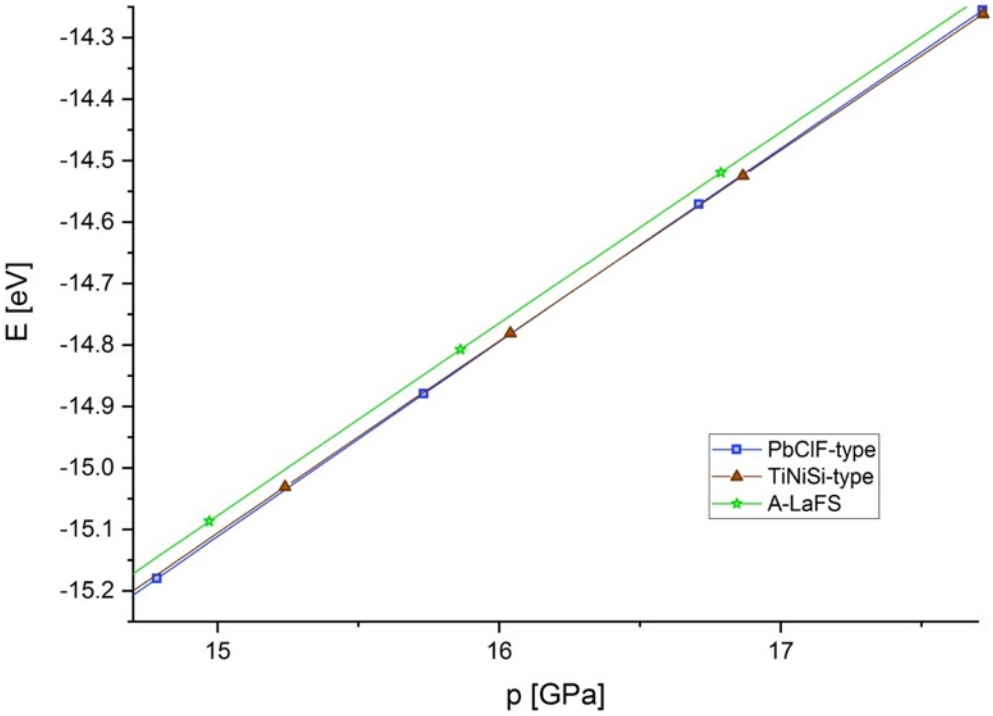

Figure 8.13. Enthalpy vs pressure, *H(p)*, curves of the three most relevant high-pressure LaFS modifications calculated from GGA-PBE *E(V)* data.

The calculated enthalpy crossings predict a pressure-induced transition from the tetragonal *PbFCl*-type ground state to the orthorhombic *TiNiSi*-type phase at approximately 16.3 GPa within GGA-PBE. The *A*-LaFS enthalpy remains higher than those of the competing phases throughout the investigated pressure range, so this phase stays metastable.

In the effective negative-pressure range, the *H(p)* curves in Figure 8.14 identify the sequence of low-density candidates. *J*-LaFS becomes the lowest-enthalpy modification below approximately -1.3 GPa, while *LiCuO*-type becomes the lowest-enthalpy phase below approximately -2.4 GPa. *LiMgN*-type does not cross into a region of absolute stability and remains metastable throughout the inspected negative-pressure interval.

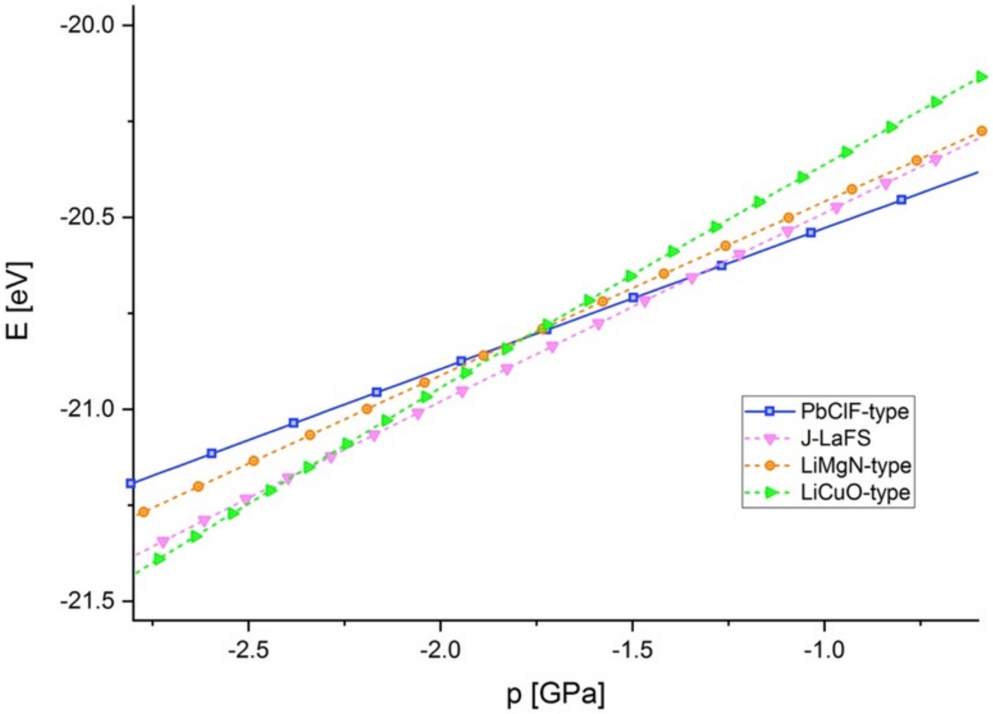

Figure 8.14. Enthalpy vs pressure curves for the effective negative-pressure phases calculated using GGA-PBE *E(V)* data.

8.5 Electronic Structure

Table 8.4 summarizes the calculated band gaps and their direct or indirect character for the low-energy LaFS polymorphs, while Table 8.5 explores the relationship between the GGA-PBE band gaps and the volume per formula unit, average La–anion distances, and La coordination numbers. Across the low-energy modifications, the GGA-PBE band gaps are in the range of 1.38–3.40 eV, while HSE06 calculated band gaps are significantly wider, in the range of 2.23–4.58 eV for different polymorphs.

Table 8.4. Electronic band gap parameters and transition characteristics across the investigated LaFS polymorphs.

Structure type	Space group No.	GGA-PBE			HSE06			Other band gap results (eV)
		E_{RI}	Band gap (eV)	Band-gap character	E_{RI}	Band gap (eV)	Band-gap character	
PbFCl-type	129	1	1.49	Direct	1	2.31	Direct	2.71(a) 2.37(b) 2.90(c)
TiNiSi-type	62	2	2.16	Direct	2	3.15	Direct	
A-LaFS	62	3	1.94	Direct	3	2.84	Direct	
B-LaFS	25	4	1.72	Indirect	4	2.62	Indirect	
C-LaFS	38	5	1.50	Direct	5	2.34	Direct	
ZrNiAl-type	189	6	2.03	Indirect	7	2.99	Direct	
D-LaFS	87	7	2.00	Direct	6	3.00	Direct	
E-LaFS	109	8	1.56	Indirect	9	2.28	Direct	
F-LaFS	59	9	1.49	Indirect	10	2.31	Direct	
ZrBeSi-type	194	10	1.96	Direct	8	2.96	Direct	2.95(b) 3.48(c)
G-LaFS	25	11	1.46	Indirect	11	2.31	Direct	
H-LaFS	26	12	1.92	Indirect	12	2.85	Indirect	
I-LaFS	6	13	1.38	Direct	13	2.23	Direct	
J-LaFS	62	14	2.85	Indirect	15	4.04	Indirect	
LiMgN-type	62	15	3.40	Indirect	14	4.58	Indirect	
3R-type	166	16	3.03	Indirect	16	3.82	Indirect	
LiCuO-type	82	17	2.53	Direct	17	3.45	Direct	

(a) Experimental data [208].
(b) Theoretical HSE06 data [31].
(c) Theoretical B3LYP data [31].

Table 8.5. The electronic band gaps, band-gap character, volume per formula unit, average La–anion distances, and La coordination numbers for the investigated LaFS polymorphs calculated using GGA-PBE functional.

Structure type	**Space group No.**	**ERI**	**Band gap (eV)**	**Band-gap character**	**Vol. / FU (Å³)**	**Average La–anion distances (Å)**	**La-anion coordination numbers**
I-LaFS	6	13	1.38	Direct	60.43	2.81 (La-all); 2.97 (La-S); 2.59 (La-F)	6, 7, 8, 9, 9, 9, 9, 9, 9, 9, 9, 9
G-LaFS	25	11	1.46	Indirect	59.11	2.85 (La-all); 3.07 (La-S); 2.51 (La-F)	9, 9, 9, 9
PbFCl-type	129	1	1.49	Direct	57.63	2.81 (La-all); 2.97 (La-S); 2.60 (La-F)	9
F-LaFS	59	9	1.49	Indirect	58.81	2.79 (La-all); 2.99 (La-S); 2.55 (La-F)	7, 9, 9
C-LaFS	38	5	1.50	Direct	58.28	2.83 (La-all); 2.99 (La-S); 2.61 (La-F)	9, 9, 9, 9, 9
E-LaFS	109	8	1.56	Indirect	58.69	2.84 (La-all); 3.04 (La-S); 2.55 (La-F)	9, 9, 9
B-LaFS	25	4	1.72	Indirect	57.84	2.82 (La-all); 2.94 (La-S); 2.66 (La-F)	9, 9, 9, 9, 9
H-LaFS	26	12	1.92	Indirect	61.27	2.82 (La-all); 3.01 (La-S); 2.53 (La-F)	8, 8, 9

A-LaFS	62	3	1.94	Direct	58.01	2.82 (La-all); 2.95 (La-S); 2.66 (La-F)	9, 9
ZrBeSi-type	194	10	1.96	Direct	59.24	2.88 (La-all); 3.11 (La-S); 2.42 (La-F)	9
D-LaFS	87	7	2.00	Direct	62.86	2.80 (La-all); 2.99 (La-S); 2.47 (La-F)	8
ZrNiAl-type	189	6	2.03	Indirect	57.86	2.82 (La-all); 2.93 (La-S); 2.69 (La-F)	9, 9
TiNiSi-type	62	2	2.16	Direct	58.21	2.82 (La-all); 2.97 (La-S); 2.64 (La-F)	9
LiCuO-type	82	17	2.53	Direct	88.32	2.72 (La-all); 2.91 (La-S); 2.34 (La-F)	6
J-LaFS	62	14	2.85	Indirect	74.79	2.72 (La-all); 2.90 (La-S); 2.48 (La-F)	7
3R-type	166	16	3.03	Indirect	70.01	2.73 (La-all); 2.89 (La-S); 2.51 (La-F)	6, 8
LiMgN-type	62	15	3.40	Indirect	70.53	2.72 (La-all); 2.92 (La-S); 2.47 (La-F)	7

* Distances and coordination numbers were taken from the PBE-optimized structures using La–F nearest-neighbor distances up to 3.00 Å and La–S nearest-neighbor distances up to 3.40 Å. Coordination numbers of all distinct La local environments (inequivalent La sites) are shown.

For the ambient-pressure *PbFCl*-type ground state, the calculations reveal a direct band gap of 1.49 eV at Γ-point with the GGA-PBE functional, and 2.31 eV with HSE06 functional [12]. Both values underestimate the experimental optical gap of 2.71 eV [208], but HSE06 is much closer and agrees well with earlier HSE06 results [31]. The *ZrBeSi*-type phase also shows excellent agreement with previous hybrid-functional calculations: the present HSE06 gap is 2.96 eV, compared with 2.95 eV reported previously [31]. For the remaining polymorphs, the LaFS study provides the first systematic set of predicted band-gap data [12].

Most predicted modifications exhibit larger band gaps than the *PbFCl*-type ground state under both functionals. The main exceptions are *G*-LaFS and *I*-LaFS, while *F*-LaFS is notable because its gap remains essentially identical to that of the ground state at both the GGA-PBE and HSE06 levels, making it an interesting polytypic candidate for future layered or heterostructure concepts with similar electronic properties.

Polymorphism changes both the magnitude of the gap and its direct/indirect character. The *PbFCl*-type, *TiNiSi*-type, *A*-LaFS, *C*-LaFS, *D*-LaFS, *ZrBeSi*-type, *I*-LaFS, and *LiCuO*-type are direct-gap semiconductors at both DFT levels. Conversely, *B*-LaFS, *H*-LaFS, *J*-LaFS, *LiMgN*-type, and *3R*-type band gaps show indirect character. The *PbFCl*-type and *3R*-type band-structure and DOS plots are shown in Figures 8.15 and 8.16, respectively. However, *ZrNiAl*-type, *E*-LaFS, *F*-LaFS, and *G*-LaFS are indirect within GGA-PBE but direct within HSE06, which can be explained by nearly degenerate band extrema that are sensitive to the exchange-correlation treatment. The band structures in Figures 8.17–8.20 illustrate this behavior explicitly. In the *ZrNiAl*-type phase, the top of the valence band is nearly degenerate at M and Γ, while the conduction-band minimum remains at Γ. In *E*-LaFS phase, the relevant extrema at Γ and X are close in energy for both the valence- and conduction-band edges, whereas in F-LaFS, the analogous near-degeneracy occurs between Γ and Z for both band edges. In G-LaFS modification, the valence-band maximum is nearly degenerate between Γ and Z, while the conduction-band minimum lies at Γ. These small energy differences of the valence and/or conduction band extrema fully explain why the direct/indirect character changes between the GGA-PBE and HSE06 descriptions for these four polymorphs.

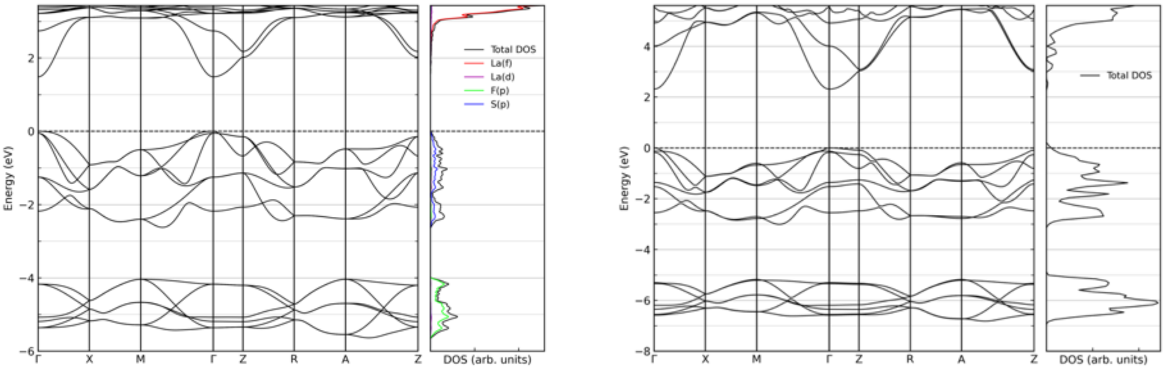

Figure 8.15. Electronic band structure and projected density of states for the *PbFCl*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

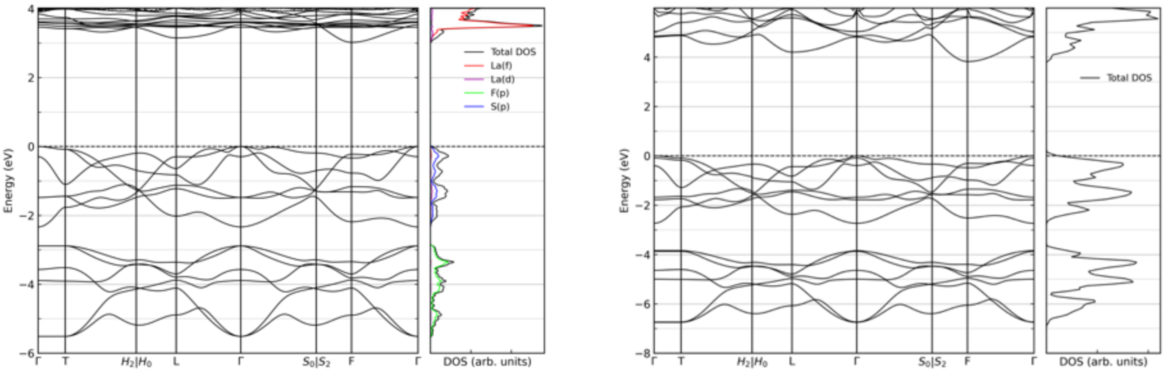

Figure 8.16. Electronic band structure and projected density of states for the *3R*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

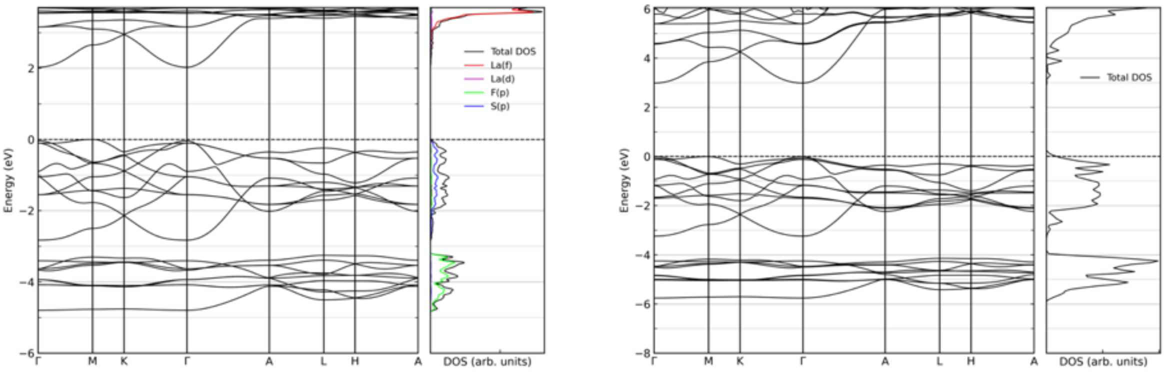

Figure 8.17. Electronic band structure and projected density of states for the *ZrNiAl*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

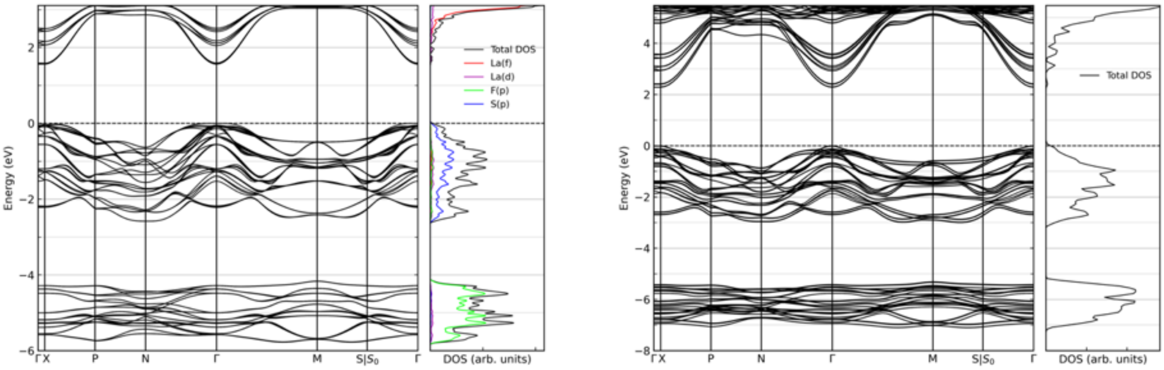

Figure 8.18. Electronic band structure and projected density of states for the *E*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

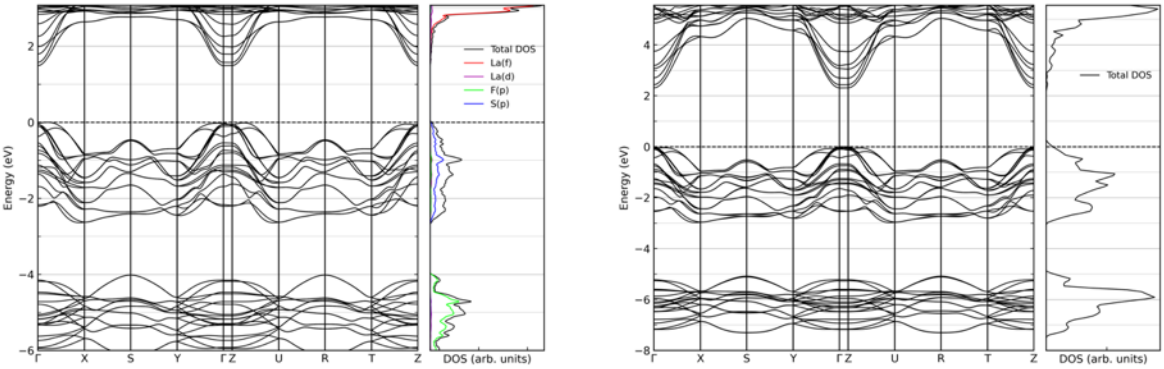

Figure 8.19. Electronic band structure and projected density of states for the *F*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

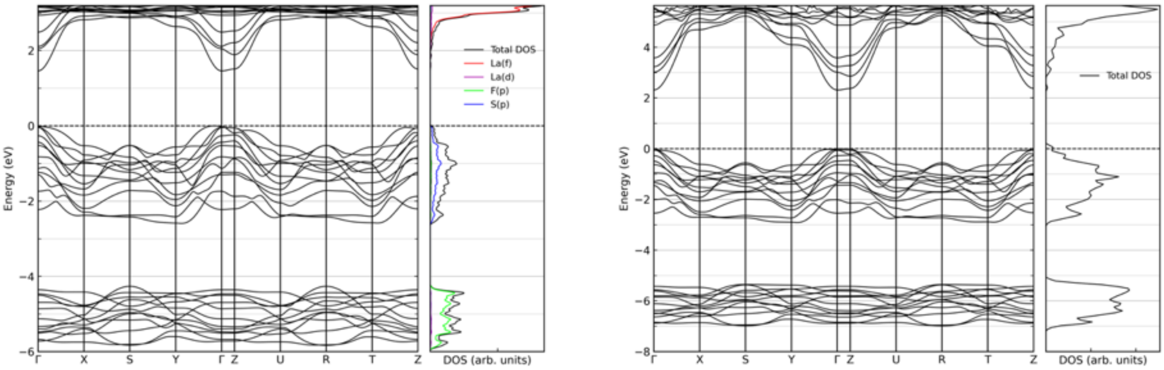

Figure 8.20. Electronic band structure and projected density of states for the *G*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

The band-gap trend is governed not only by volume, but also by the structural topology and effective coordination of the La-centered mixed-anion polyhedra. Table 8.5 shows that the largest GGA-PBE band gaps occur in the lower-density negative-pressure structures: *LiMgN*-type, *3R*-type, *J*-LaFS, and *LiCuO*-type. These phases combine large volumes per formula unit with lower La coordination numbers (6–7, except for the second eightfold La site in *3R*-type) than most of the denser low-energy structures, which are mostly eight- or ninefold coordinated. Their average La–anion distances are also shorter, approximately 2.72–2.73 Å, compared with approximately 2.79–2.88 Å for the other thirteen structures. This indicates that band-gap widening is associated with lower structural density, reduced La-centered coordination, and weaker connectivity between neighboring La-centered coordination polyhedra. In more open structures with fewer short La–anion–La connection pathways, orbital overlap between neighboring polyhedra is reduced, which can narrow the near-edge bands and increase the separation between the S(*p*)-dominated valence states and La(*f*)-dominated conduction states. This behavior is not controlled by a single parameter, however: *LiCuO*-type has by far the largest volume and a very short average La–F distance, but its gap is smaller than those of *LiMgN*-type and *3R*-type. Thus, the gap reflects the combined influence of volume, local La coordination, and the arrangement and connectivity of La-centered polyhedra.

The projected density of states shows a consistent near-gap orbital character across the polymorphs. The states forming the valence-band maximum are dominated by S(*p*) states, whereas the states at the bottom of the conduction band are dominated primarily by La(*f*) states. In the *PbFCl*-type direct-gap phase and in the *3R*-type indirect-gap phase used here as a contrast case, this separation between sulfur-derived valence states and lanthanum-derived conduction states is especially clear, and the same near-gap orbital partitioning is retained across essentially all orbital-projected DOS plots analyzed for the LaFS polymorphs. The remaining GGA-PBE and HSE06 band structures and DOS plots are shown in Appendix A8 Figures A8.1–A8.11.

8.6 Vibrational Properties – Phonon Dispersion and Dynamical Stability

Phonon dispersions were calculated to test local dynamical stability and to examine whether the predicted phase competition is accompanied by soft-mode instabilities [12]. Density-functional perturbation theory (DFPT) calculations within GGA-PBE were carried out for the *PbFCl*-type ground state and for the *TiNiSi*-type and *A*-LaFS high-pressure candidates at 0, 12, 18, and 24 GPa, and the resulting force constants were post-processed with Phonopy [38].

At 0 GPa, the *PbFCl*-type ground state shows no imaginary frequencies along the selected high-symmetry path (Figure 8.21), confirming its local dynamical stability, which is consistent with earlier theoretical and experimental assessments of the ambient-pressure phase [31, 208]. The *TiNiSi*-type and *A*-LaFS phases are also free of imaginary branches at 0 GPa (Figures 8.22 and 8.23), showing that they are locally stable even though they are thermodynamically metastable at ambient pressure. The corresponding high-pressure phonon dispersions at 12, 18, and 24 GPa likewise show no imaginary modes for any of the three phases considered. This suggests that pressure-induced or pressure-quenched retention of such phases is plausible in principle, but phonons alone do not prove easy recoverability at 0 GPa, because the actual decompression pathway and the relevant kinetic barriers were not calculated here. With increasing pressure, the spectra shift to higher frequencies and the acoustic branches become slightly stiffer, as expected for lattice compression.

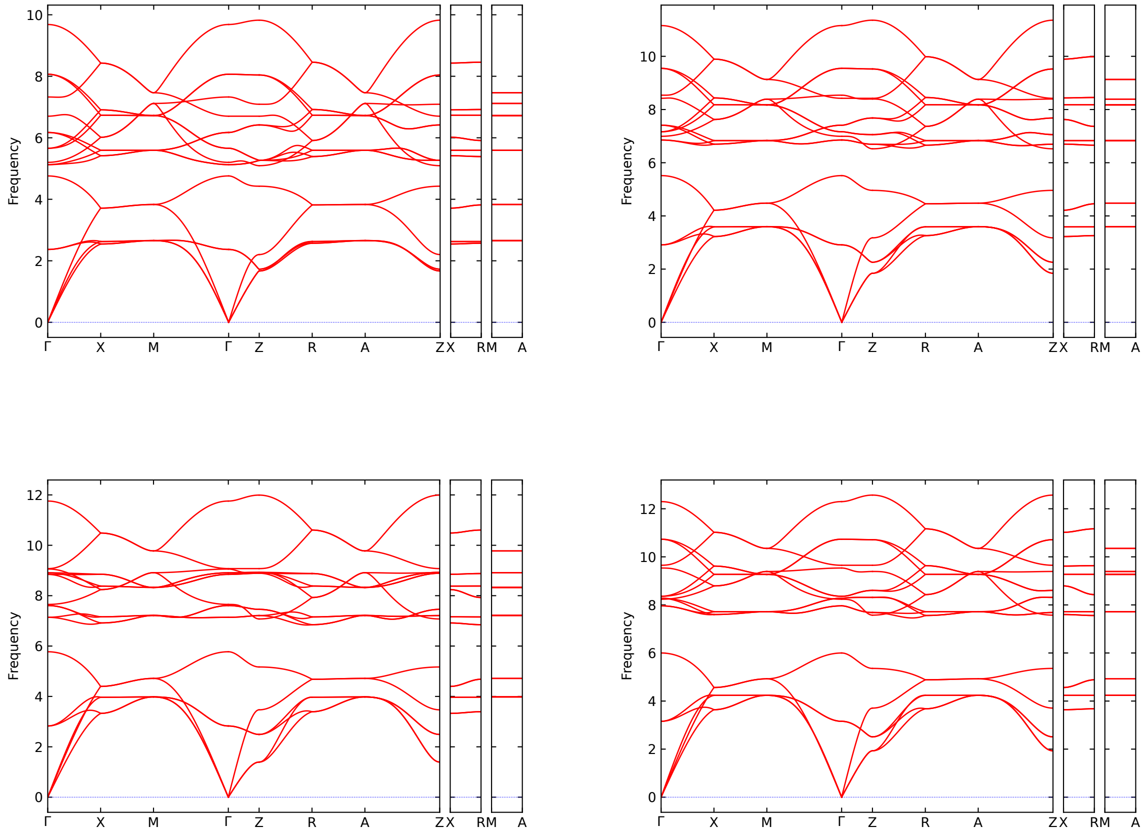

Figure 8.21. Phonon dispersion spectra for *PbFCl*-type LaFS phase calculated at 0 GPa (top-left), 12 GPa (top-right), 18 GPa (bottom-left), and 24 GPa (bottom-right) using the GGA-PBE functional.

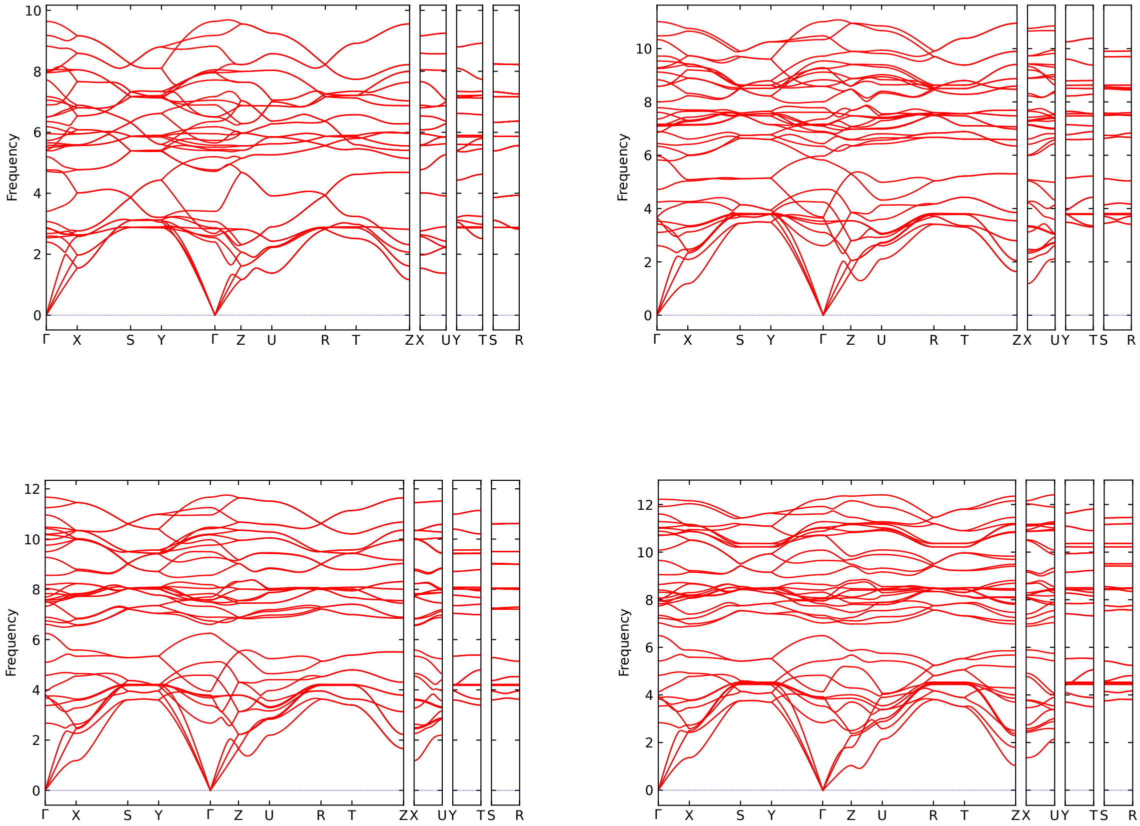

Figure 8.22. Phonon dispersion spectra for *TiNiSi*-type LaFS phase calculated at 0 GPa (top-left), 12 GPa (top-right), 18 GPa (bottom-left), and 24 GPa (bottom-right) using the GGA-PBE functional.

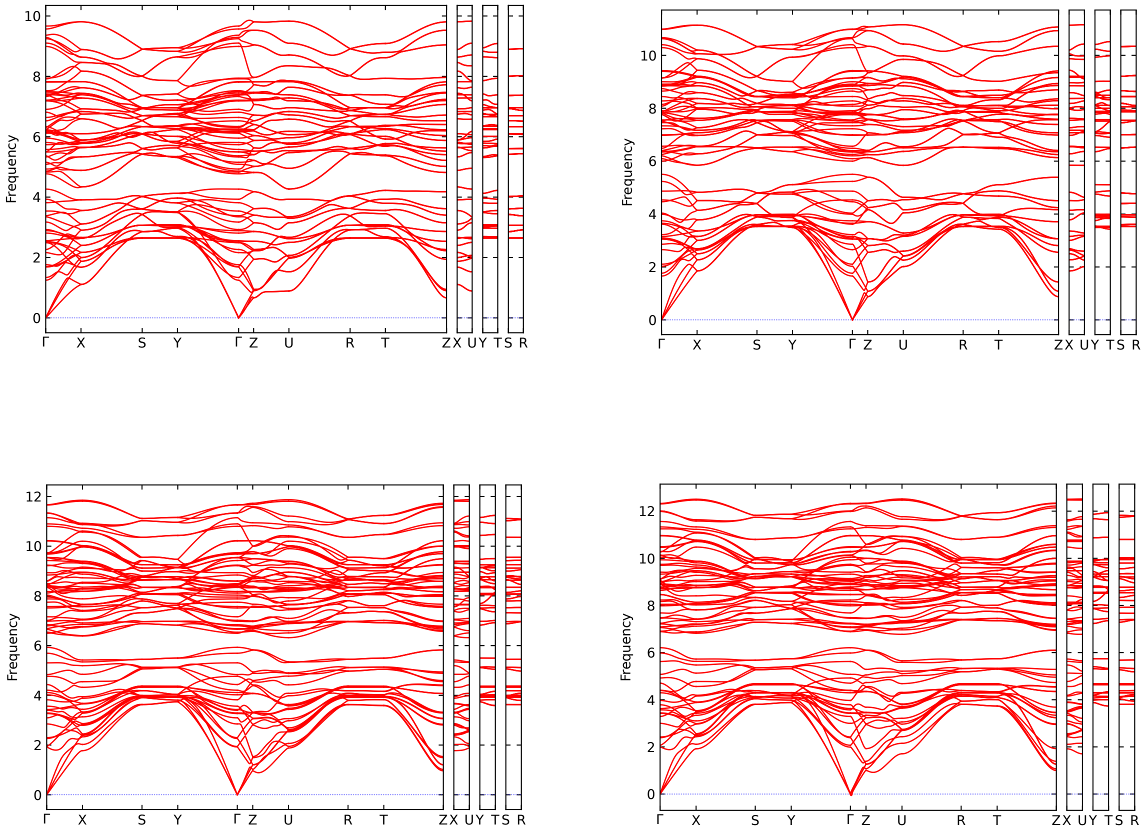

Figure 8.23. Phonon dispersion spectra for *A*-LaFS phase calculated at 0 GPa (top-left), 12 GPa (top-right), 18 GPa (bottom-left), and 24 GPa (bottom-right) using the GGA-PBE functional.

Because no soft mode appears in the inspected structures around the transition region, the calculated phase boundary is consistent with enthalpy competition between dynamically stable phases rather than with a phonon-driven instability of one phase into another. Within the sampled pressure range each of the inspected structures remains a well-defined local minimum, and the change in the stable phase is traced to crossings of their *H(p)* curves instead of to the loss of local stability of one structure through a soft vibrational mode.

8.7 Results of the AI-Driven Search for High-Pressure Phases

The search for previously unidentified extreme high-pressure polymorphs within the large GO/DM data set required a screening strategy that was much faster and more computational-resource-efficient than performing non-selective DFT calculations on all candidates. The DFT is sufficiently accurate for final ranking, but it is too expensive for exhaustive high-pressure exploration, whereas universal ML potentials may not resolve subtle energy differences reliably enough inside one specific rare-earth mixed-anion system. To address this, a system-specific AI-ELX graph-neural-network model was trained on approximately 20,000 self-consistent LaFS DFT data points spanning different structures, volumes, and pressure environments. The training-data distribution is shown in Figure 8.24. The role of AI-ELX here was to guide DFT calculations toward the most relevant (high-pressure) region of phase space.

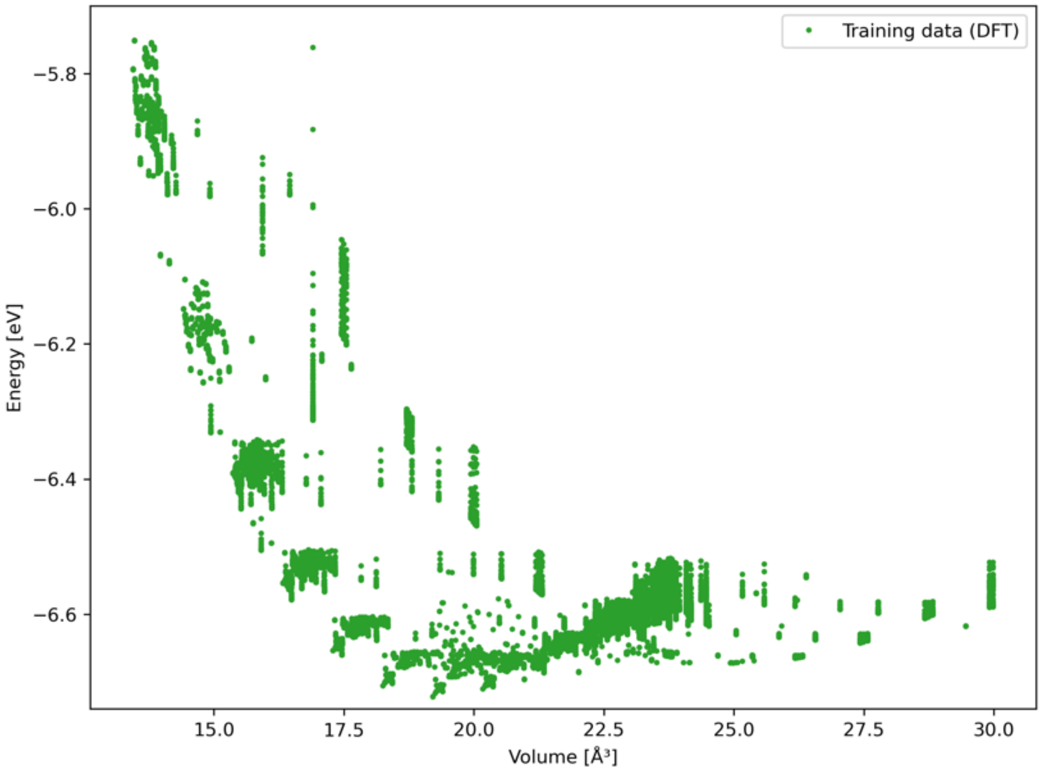

Figure 8.24. Distribution of the DFT training data used for the AI-ELX model applied to LaFS.

The trained model was then used to pre-relax selected GO and DM candidates and to evaluate their energies over a range of volumes, producing AI-predicted *E(V)* curves before DFT refinement. This step reduced the number of candidates that had to be explored explicitly at the first-principles level and made the subsequent DFT stage more focused on the target structures.

The AI-generated *E(V)* curves in Figure 8.25 made it possible to discard large numbers of unpromising structural candidates whose minima or curvatures were inconsistent with competitive high-pressure behavior. Candidates whose AI-calculated *E(V)* curves indicated plausible compressed-region minima were retained for DFT follow-up, so the screening acted as a step that concentrated computational effort on the most relevant part of the landscape.

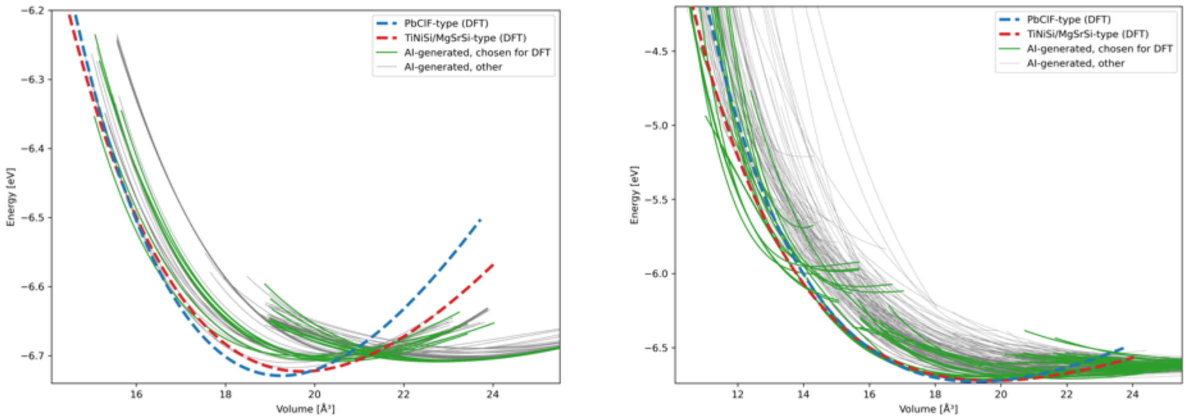

Figure 8.25. The *E(V)* Equation of state curves generated by the AI-ELX model for GO (left) and DM (right) structures, highlighting the separation between competitive and unpromising high-pressure structures.

After DFT refinement of the AI-selected structures, two additional candidates were identified in the targeted extreme high-pressure region. In Figure 8.26 these are denoted α-LaFS and β-LaFS, and are compared with the *E(V)* curves of the previously found ground-state *PbFCl*-type and high-pressure *TiNiSi*-type phases.

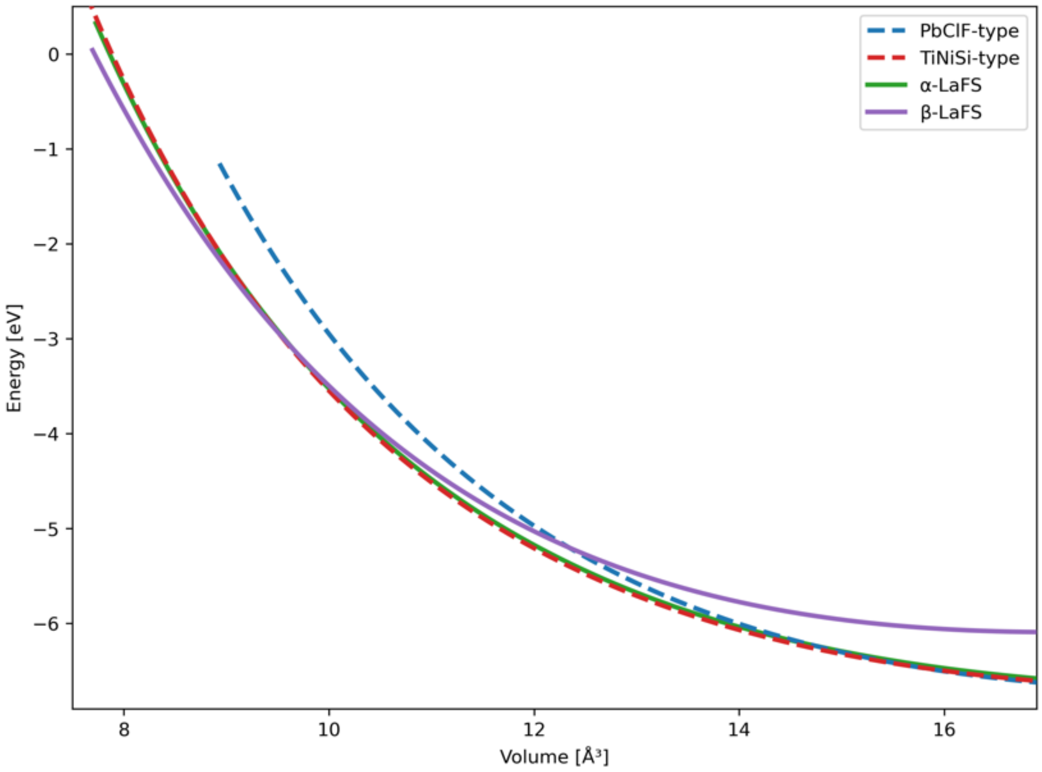

Figure 8.26. Energy vs volume, *E(V)*, curves for the AI-discovered α and β phases in the extreme high-pressure region.

The two found extreme-pressure candidate structure types are shown in Figure 8.27. The α-LaFS candidate crystallizes in the orthorhombic *Pbcn* space group (No. 60) and uses a comparatively large unit cell with several symmetry-inequivalent La, S, and F sites, indicating a more complex ordering distortion than the simpler reference high-pressure structures. By contrast, β-LaFS crystallizes in tetragonal $P4/nmm$ (No. 129) and contains only one symmetry-inequivalent site for each chemical species, so its geometry is much simpler and has the same symmetry as the ambient pressure ground-state *PbFCl*-type modification.

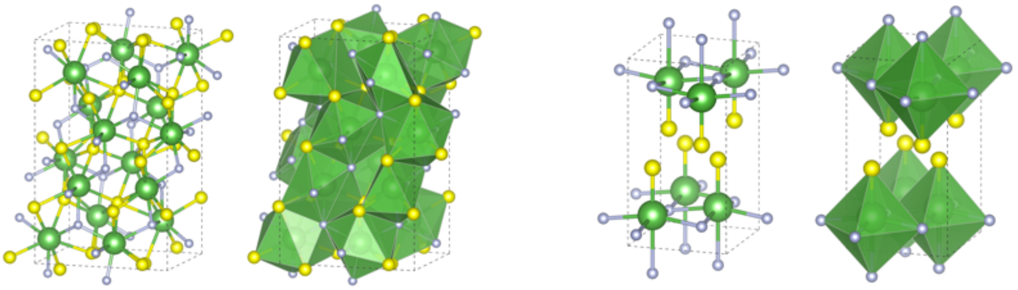

Figure 8.27. Crystallographic representations of the AI-discovered extreme-pressure α-LaFS (left) and β-LaFS (right) structure types.

Figure 8.28 shows the enthalpy vs pressure, *H(p)*, relations for the *TiNiSi*-type phase and the two AI-discovered extreme high-pressure candidates, *α*-LaFS and *β*-LaFS. The *TiNiSi*-type modification remains the lowest-enthalpy phase up to 206 GPa, confirming that it is not only the first high-pressure phase predicted from the established energy-landscape search workflow, but also the most favorable phase over a wide pressure interval. Below 195 GPa, the enthalpy ordering is *TiNiSi*-type < *α*-LaFS < *β*-LaFS. At 195 GPa, the *α*-LaFS and *β*-LaFS curves cross, so that *β*-LaFS becomes lower in enthalpy than *α*-LaFS, although *TiNiSi*-type still remains the most stable modification. The second and most important crossing occurs at 206 GPa, where *β*-LaFS becomes lower in enthalpy than the *TiNiSi*-type phase and therefore emerges as the predicted thermodynamically favored extreme high-pressure modification. Between 206 and 256 GPa, the enthalpy ordering is *β*-LaFS < *TiNiSi*-type < *α*-LaFS. At 256 GPa, the *α*-LaFS and *TiNiSi*-type curves cross, after which *β*-LaFS still remains the lowest-enthalpy phase, but now followed by *α*-LaFS and then *TiNiSi*-type as metastable modifications. Therefore, *β*-LaFS modification is the only AI-discovered phase that becomes thermodynamically favored, while *α*-LaFS remains a competitive metastable phase in the investigated extreme high-pressure region.

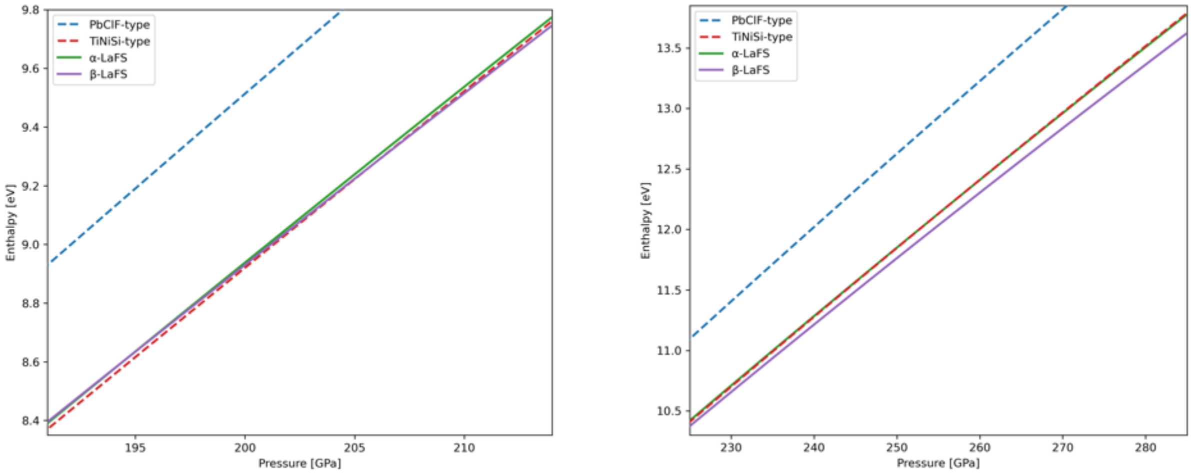

Figure 8.28. Enthalpy vs pressure, *H(p)*, curves for the *TiNiSi*-type phase and the AI-discovered *α*-LaFS and *β*-LaFS phases in the extreme high-pressure region. The curves show the *α*-LaFS/*β*-LaFS and *TiNiSi*-type/*β*-LaFS crossings at 195 and 206 GPa, respectively (left) and the crossings between *TiNiSi*-type and *α*-LaFS at 256 GPa (right).

The AI-guided search shows that a system-specific GNN model can function as an efficient screening layer within energy-landscape exploration. It is fast enough to evaluate many candidates and also accurate enough to guide first-principles calculations toward the most promising structures. For the LaFS system, this strategy enabled a deeper and more targeted search for extreme high-pressure polymorphs, expanding the knowledge of its relevant structural landscape modifications.

8.8 Conclusion

To obtain a broad overview of the LaFS energy landscape, global optimization over different cell sizes and pressures was supplemented by prototype-based data mining, generating almost four million candidate structures. Symmetry analysis, structural comparison, clustering, recurrence statistics, energy ranking, and ML-assisted pre-screening reduced this very large dataset to a tractable group of candidates for subsequent DFT calculations. This workflow provides a systematic overview of the low-energy LaFS landscape.

After the DFT refinement, the experimentally known *PbFCl*-type structure was confirmed as the ambient-pressure ground state, while sixteen additional low-energy polymorphs were identified. These include tetragonal, orthorhombic, hexagonal, rhombohedral, and monoclinic modifications, as well as several layered or polytypic variants. The most important high-pressure-related candidates are the orthorhombic *TiNiSi*-type and *A*-LaFS phases, whereas the effective negative-pressure region contains low-density *J*-LaFS, *LiMgN*-type, and *LiCuO*-type modifications. Together, these phases show that LaFS has a rich polymorphic landscape with several structurally distinct local minima under different thermodynamic conditions.

The *E(V)* and *H(p)* curve analyses establish the pressure-dependent stability relations among the main phases. Within GGA-PBE, the *PbFCl*-type ground state is predicted to transform to the orthorhombic *TiNiSi*-type phase at approximately 16.3 GPa, while *A*-LaFS remains metastable throughout the investigated positive-pressure range. In the effective negative-pressure regime, *J*-LaFS becomes the lowest-enthalpy modification below approximately -1.3 GPa and then the *LiCuO*-type below approximately -2.4 GPa, whereas *LiMgN*-type remains metastable, identifying both compression-driven and expansion-driven regions of the energy landscape.

The electronic-structure calculations show that all investigated LaFS polymorphs are semiconducting. The GGA-PBE-calculated band gaps range from 1.38 to 3.40 eV, while HSE06 calculations increase the gaps to the range of 2.23 to 4.58 eV. Several phases are direct-gap semiconductors at both DFT levels, including *PbFCl*-type, *TiNiSi*-type, *A*-LaFS, *C*-LaFS, *D*-LaFS, *ZrBeSi*-type, *I*-LaFS, and *LiCuO*-type, whereas *B*-LaFS, *H*-LaFS, *J*-LaFS, *LiMgN*-type, and *3R*-type remain indirect. The comparison with various structural descriptors shows that the gap depends on the unit cell volume, La coordination, La-anion distances, and the connectivity of La-centered polyhedra. The orbital character of the band-edge states is consistent across the polymorphs: the valence-band maximum is dominated mainly by S(*p*) states, while the conduction-band minimum is dominated mostly by La(*f*) states.

The phonon calculations support the local dynamical stability of the ground-state *PbFCl*-type phase and high-pressure *TiNiSi*-type and *A*-LaFS modifications, showing no imaginary phonon branches at 0 GPa, while their corresponding phonon dispersions at 12, 18, and 24 GPa also remain free of imaginary modes. Thus, the predicted pressure-induced phase competition is not associated with a dynamical instability of the investigated phases, but rather with phase transitions between dynamically stable local minima.

The AI-guided high-pressure search extended the structural landscape beyond the standard GO and DM workflow by focusing DFT effort on candidates most likely to be relevant in the extreme high-pressure region. With relatively low additional computational effort, the AI-ELX screening and subsequent DFT calculations identified two additional high-pressure candidates, α-LaFS in the orthorhombic *Pbcn* space group and β-LaFS in the tetragonal *P*4/*nmm* space group. The predicted phase transition between *TiNiSi*-type and β-LaFS occurs at a pressure of 206 GPa, above which β-LaFS becomes the extreme-pressure-stabilized modification, while α-LaFS becomes the most favorable metastable modification above 256 GPa. These results show that the AI-assisted search can reveal additional competitive regions of the LaFS energy landscape and provide useful starting points for future structural explorations under high-pressure conditions, as well as under any other targeted conditions explored within the AI-ELX-driven energy landscape exploration workflow.

9. Summary

The aim of this thesis was to investigate how broad energy-landscape exploration, crystallographic analysis, and first-principles calculations can be combined to discover, classify, and analyze new structure candidates in ternary rare-earth mixed-anion compounds. The systems explored in the results chapters, LaOI, Ce_3O_3N, ScOCl, HoFSe, and LaFS, were chosen because they combine relative simplicity of chemical composition with considerable structural, thermodynamic, electronic, and in some cases magnetic complexity. In each case, the goal was to find the global minimum and explore and understand its surrounding landscape, which contains stable and metastable polymorphs, and to further explore that landscape by identifying pressure-induced phase transitions and the physical properties of the predicted modifications.

The computational workflow used throughout the thesis was hierarchical, with broad global optimization and data-mining steps generating large sets of candidate structures, while local first-principles calculations were used for the final structural relaxation and property analysis. This separation was necessary because direct *ab initio* exploration of all possible configurations would be computationally impractical. At the same time, the thesis shows that, for very large structure datasets generated by empirical searches and prototype mapping (containing millions of structures), the manual structural analysis is not practical: the important structural refinements are obtained only after extensive automated symmetry analysis, structure-type classification, followed by DFT refinement, thermodynamic analysis, and in selected cases calculations of electronic, magnetic, and vibrational properties.

An important methodological contribution of this thesis is the development and use of STyX, the Structure Type eXplorer software package. The STyX workflow was introduced as a post-processing layer between large-scale structure generation and final DFT calculations. Its role is to turn a very large and redundant set of candidate structures into a much smaller, reproducible, and queryable set of representative structure types. The application imports structure files and their metadata, stores the results in an SQLite database, determines and idealizes symmetry, groups related structures into structure types, keeps track of parent-child relations between representatives and duplicates, and enables final candidate selection through explicit customizable database queries. In most respects, STyX automates the structure-prediction post-processing work that is often performed manually, but on datasets so large that reliable manual analysis would be impractical.

The second methodological development is AI-ELX, an artificial-intelligence-assisted energy landscape exploration workflow. This GNN model is trained on system-specific DFT data and can, in principle, predict structure energies with DFT-level accuracy in the region of the landscape covered by the training data. The model represents periodic crystals as graphs, predicts DFT-level energies from structural input, and can be coupled to differentiable optimization of atomic coordinates and cell parameters. Its most important use in this thesis is to reduce the number of computationally expensive DFT calculations by focusing them on the most promising candidates in the selected region of the energy landscape. Through AI-predicted $E(V)$ curves and subsequent enthalpy analysis, the method enables targeted searches for high-pressure phases or any other selected regions of the landscape. Its application to the LaFS system demonstrates how such an AI-guided workflow can extend a conventional global search to new structures with targeted properties.

The LaOI study provided the first example of this thesis energy landscape exploration applied to a rare-earth oxyiodide. The experimentally known tetragonal α-LaOI phase was confirmed as the deepest minimum of the calculated landscape, while six additional low-energy polymorphs, β-, γ-, δ-, ε-, ζ-, and η-LaOI, were identified as metastable candidates. The global exploration generated more than one million local minima, making the large-scale structure post-processing step of the search crucial. Several predicted structures, especially β-LaOI and η-LaOI, occur rarely on the landscape and could have been missed by a less detailed global search. The resulting polymorph set spans rhombohedral, monoclinic, tetragonal, orthorhombic, and cubic symmetry classes, and illustrates the structural flexibility of the LaOI composition.

The LaOI thermodynamic analysis shows that *α*-LaOI remains the stable ground-state modification in the inspected volume and pressure range at both the GGA-PBE and HSE06 levels, while the other six structures remain metastable. The electronic-structure calculations show that all optimized LaOI polymorphs are semiconducting, with the calculated band-gap size and direct or indirect character depending strongly on the structure type. The HSE06 band gap of *α*-LaOI agrees very well with the available experimental optical gap, supporting the reliability of the electronic-structure description. The projected density of states shows a consistent separation between anion valence states and La conduction states for all polymorphs.

The Ce_3O_3N study extended the work to a hypothetical (not-yet-synthesized) ternary oxynitride, where the central question was primarily the identification of a plausible ground-state structure. The most important result is that the global minimum, Ce_3O_3N-*DM1*, emerged after relaxing a data-mined proustite-derived initial model, showing that global optimization and data mining should be used as complementary methods for landscape exploration. While global optimization efficiently samples broad basins of the empirical energy landscape, data mining introduces crystallographically diverse prototypes that can relax into structure types not easily found by the global search alone.

The Ce_3O_3N landscape contains several distinct structure types in a relatively narrow energy range. The dense, eightfold-coordinated *DM1* phase is the ambient-pressure ground state and remains stable under compression. In contrast, the monoclinic *GS1* phase is a lower-density candidate that becomes relevant on the expanded-volume side of the landscape, while the orthorhombic *GS2–GS5* phases remain metastable. Because cerium has localized 4*f* electronic states, the Ce_3O_3N study was extended beyond the nonmagnetic structure prediction work by using scalar-collinear-spin PBE+U calculations. Within the explored set of spin arrangements, the AFM solution is lowest in energy and it is semiconducting, with an indirect Γ–L gap of approximately 2.21 eV. The FiM and FM states remain close in energy and also semiconducting, but with only slightly smaller gaps. Because the ordinary (non-spin-polarized) DFT calculations with the same GGA-PBE functional result in a metallic Ce_3O_3N state, these results show that Ce_3O_3N and similar systems with occupied localized 4*f* states require a spin-polarized treatment to obtain a physically meaningful electronic structure of the ground-state phase.

The ScOCl study addressed a simpler yet instructive oxyhalide system. The combined global-optimization and data-mining workflow recovered the experimentally known *α*-ScOCl phase and identified three additional low-energy modifications: *β*-, *γ*-, and *δ*-ScOCl. The *α*-phase is the stable orthorhombic *FeOCl*-type structure at ambient conditions, while *β*-ScOCl is the main high-pressure candidate. The *γ*-phase remains close enough in enthalpy to be considered a possible intermediate or competing metastable phase during pressure-induced transformation, and *δ*-ScOCl is a higher-energy candidate more likely to be relevant under elevated-temperature or non-equilibrium synthesis conditions.

The ScOCl thermodynamic analysis based on both *E(V)* and *H(p)* curves predicts an $\alpha \rightarrow \beta$ transition at relatively low pressure, with the transition pressure depending on the choice of exchange-correlation functional: approximately 1 GPa with LDA and approximately 3.7 GPa with PBE0, but nevertheless with both functionals giving the same qualitative pressure response. The electronic-structure calculations show that both *α*- and *β*-ScOCl are wide-gap indirect semiconductors. Across LDA, PBE, PBEsol, HSE06, and PBE0 functionals, the same qualitative band structure picture near the band gap is retained, where the upper valence region is dominated mainly by anion *p* states, especially Cl(*p*) with additional O(*p*) contribution, while the conduction-band bottom has primarily Sc(*d*) character.

The exploration of the HoFSe system is significantly broader because it combines theoretical structure prediction with experimental data and magnetic calculations. Energy landscape exploration and data mining generated a large group of 1.82 million candidates, and subsequent STyX workflow post-processing and DFT refinement identified the twenty most relevant low-energy polymorphs. The experimentally observed *exp1*-, *exp2*-, and *exp3*-type structures are all recovered among the low-energy candidates, confirming that the search strategy covered physically important regions of the HoFSe landscape. At the same time, the theoretical work extends the experimental study by predicting additional high-pressure, low-density, high-temperature, and metastable modifications.

The structural landscape of HoFSe encompasses orthorhombic, monoclinic, tetragonal, hexagonal, rhombohedral, and cubic symmetry classes, with Ho coordination numbers ranging from six to nine and strongly correlating with the equilibrium volumes of the phases. The *exp1*-type *LiCaN*-related structure is the ambient-pressure ground state, which under compression undergoes a phase transition to the *β*-type HoFSe modification, and with further increase in pressure, to the *PbFCl*-type HoFSe phase. The transition pressures, calculated with the GGA-PBE functional are approximately 0.3 GPa and 0.6 GPa, respectively, while the HSE06 calculations confirm the same qualitative phase sequence as GGA-PBE with the first transition remaining at 0.3 GPa, and the second transition increasing to approximately 1.15 GPa. On the effective negative-pressure side, the *CoYC*-type structure is favored below -0.8 GPa with the GGA-PBE functional and below -1.55 GPa with the HSE06 functional.

The HoFSe study also introduced magnetic calculations as a major extension of the structural analysis. Collinear PBE+U calculations show that antiferromagnetic arrangements are favored in both primitive-cell and supercell calculations. The non-collinear spin-orbit-coupled calculations then show that the magnetic-anisotropy landscape is relatively shallow: the NCL-7 state, a weakly uncompensated canted antiferromagnetic arrangement, is the lowest-energy solution, while the almost fully compensated NCL-1 antiferromagnetic state lies only about 0.7 meV per cell higher. Although the AFM arrangement is generally favored, with FiM states rising significantly in energy compared to the NCL-7 ground state, the magnetic ordering is not controlled only by the magnitude of the net magnetic moment, but also by the orientation of the Ho moments relative to the crystal field, as some other almost compensated antiferromagnetic states lie much higher in energy than the NCL-7 ground state. The HoFSe electronic-structure calculations show that all considered magnetic configurations remain semiconducting, with Se p-dominated valence-band maxima and localized Ho 4*f*-derived conduction states near the gap.

The LaFS study represents the most extensive crystal-structure prediction study in the thesis. The standard workflow combined global optimization over many different cell sizes and pressures with prototype-based data mining, generating almost four million candidate structures. Symmetry analysis, structure comparison, clustering, recurrence statistics, energy ranking, and ML-assisted pre-screening, performed using the STyX workflow, then reduced this very large set to a manageable group of candidates for subsequent DFT local optimization. The experimentally known *PbFCl*-type structure was confirmed as the ambient-pressure ground state, and sixteen additional low-energy polymorphs were identified. These include tetragonal, orthorhombic, hexagonal, rhombohedral, monoclinic, layered, and polytypic modifications.

Thermodynamically, LaFS shows polymorphism stabilized at both positive and effective negative pressures. Within GGA-PBE calculations, the *PbFCl*-type ground state is predicted to transform to the *TiNiSi*-type structure at about 16.3 GPa, while *A*-LaFS, another high-pressure candidate, remains metastable over the inspected pressure range. In the effective negative-pressure region, *J*-LaFS and then *LiCuO*-type modifications become the favored low-density phases under -1.3 GPa and -2.4 GPa, respectively. The electronic-structure calculations show that all investigated LaFS polymorphs are semiconducting, with GGA-PBE band gaps ranging from 1.38 to 3.40 eV and HSE06 band gaps ranging from 2.23 to 4.58 eV. The band-gap size and direct or indirect character depend on volume, La coordination, La-anion distances, and the connectivity of La-centered polyhedra. The valence-band maximum is consistently dominated mainly by S(*p*) states, while the conduction-band minimum is dominated mostly by La(*f*) states. Phonon calculations confirm the local dynamical stability of the *PbFCl*-type, *TiNiSi*-type, and *A*-LaFS structures across the tested pressure interval from 0 to 24 GPa, indicating that the predicted pressure-induced phase competition is enthalpy-controlled rather than driven by soft-mode instabilities.

The additional AI-guided LaFS high-pressure search provides a demonstration of the AI-ELX methodology as a thesis result. After the conventional global optimization and data-mining workflow had established the ambient- and moderate-pressure landscape, AI-ELX was used to focus the search on extreme-pressure candidates. The AI screening and subsequent DFT calculations identified additional high-pressure structures, including *α*-LaFS in the orthorhombic *Pbcn* space group and *β*-LaFS in the tetragonal *P*4/*nmm* space group. The predicted transition from *TiNiSi*-type LaFS to *β*-LaFS occurs above 200 GPa, and *α*-LaFS becomes the most favorable metastable candidate in the even higher-pressure region. More generally, this result shows that the AI-ELX model can act as a targeted extension of the standard workflow: it is not limited to LaFS or to high-pressure searches used here, but can be trained on a sufficiently large DFT dataset for a chosen system and directed toward any selected region of its energy landscape.

The methodological and structural exploration results are therefore closely connected: STyX made it possible to manage very large structure sets much more efficiently, while AI-ELX then extended the search workflow by adding a system-specific graph-neural-network model that can screen targeted regions of the landscape before final DFT validation. The results in Chapters 4–8 show why such developments are useful: the important structures are often rare, metastable, or located in pressure regions where exhaustive first-principles exploration would be impractical. By combining broad search, large-scale crystallographic analysis, DFT calculations, and machine-learning-assisted screening, this thesis provides a practical route toward discovering and interpreting new polymorphs in complex mixed-anion compounds.

The extensions of the energy landscape exploration to the study of pressure-induced phase transitions, magnetic ground-state arrangements, electronic band structures, phonon-stability results, and AI-augmented first-principles calculations all arise from the same underlying idea: to extensively explore the properties of a crystal system by combining modern computational tools and algorithms with increasingly powerful computational resources.

Appendix A1. List of Abbreviations in Alphabetical Order

AFM	Antiferromagnetic
AI-ELX	Artificial Intelligence Energy Landscape eXplorer
ANN	Artificial neural network
ASE	Atomic Simulation Environment
CE	Cross-entropy
CIF	Crystallographic Information File
CMPZ	CoMPare Zellen (compare cells)
DFPT	Density functional perturbation theory
DFT	Density functional theory
DM	Data mining
DOS	Density of states
EOS	Equation of state
$\mathbf{E_{RI}}$	Energy ranking index
FiM	Ferrimagnetic
FM	Ferromagnetic
GGA	Generalized-gradient approximation
GNN	Graph neural network
GO	Global optimization
GS	Global search
HK	Hohenberg–Kohn
ICSD	Inorganic Crystal Structure Database
JSON	JavaScript Object Notation
KDD	Knowledge discovery in databases
KS	Kohn–Sham
LDA	Local density approximation
LL	Levy–Lieb
LSDA	Local spin-density approximation
MAE	Magnetocrystalline anisotropy energy
ML	Machine learning
MLP	Multilayer perceptron
MSE	Mean-squared error
NCL	Non-collinear
PAW	Projector augmented wave
QMC	Quantum Monte Carlo
REEs (REE)	Rare earth elements
RGS	Raum Gruppen Sucher (space group finder)
SFND	Symmetrie FiNDer (symmetry finder)
SOC	Spin–orbit coupling
STyX	Structure Type eXplorer
xc	Exchange-correlation

Appendix A2. Space Group Classification Summary

Table A2.1 presents the standard crystallographic space groups, their corresponding point groups, their symmorphic or nonsymmorphic character, and a compact note on the screw axes and glide planes indicated by the Hermann-Mauguin symbol.

Table A2.1. Summary of crystallographic space groups, related point groups, and symmorphic or nonsymmorphic character including screw-axis and glide-plane content.

No.	Space group	Point group	Symmorphic/nonsymmorphic note
1	$P1$	C_1 (1)	symmorphic
2	P-1	C_i (-1)	symmorphic
3	$P2$	C_2 (2)	symmorphic
4	$P2_1$	C_2 (2)	nonsymmorphic: screw axis 2_1
5	$C2$	C_2 (2)	symmorphic
6	Pm	C_s (m)	symmorphic
7	Pc	C_s (m)	nonsymmorphic: glide plane c
8	Cm	C_s (m)	symmorphic
9	Cc	C_s (m)	nonsymmorphic: glide plane c
10	$P2/m$	C_{2h} ($2/m$)	symmorphic
11	$P2_1/m$	C_{2h} ($2/m$)	nonsymmorphic: screw axis 2_1
12	$C2/m$	C_{2h} ($2/m$)	symmorphic
13	$P2/c$	C_{2h} ($2/m$)	nonsymmorphic: glide plane c
14	$P2_1/c$	C_{2h} ($2/m$)	nonsymmorphic: screw axis 2_1; glide plane c
15	$C2/c$	C_{2h} ($2/m$)	nonsymmorphic: glide plane c
16	$P222$	D_2 (222)	symmorphic
17	$P222_1$	D_2 (222)	nonsymmorphic: screw axis 2_1
18	$P2_12_12$	D_2 (222)	nonsymmorphic: screw axis 2_1×2
19	$P2_12_12_1$	D_2 (222)	nonsymmorphic: screw axis 2_1×3
20	$C222_1$	D_2 (222)	nonsymmorphic: screw axis 2_1
21	$C222$	D_2 (222)	symmorphic
22	$F222$	D_2 (222)	symmorphic
23	$I222$	D_2 (222)	symmorphic
24	$I2_12_12_1$	D_2 (222)	nonsymmorphic: screw axis 2_1×3
25	$Pmm2$	C_{2v} ($mm2$)	symmorphic
26	$Pmc2_1$	C_{2v} ($mm2$)	nonsymmorphic: screw axis 2_1; glide plane c
27	$Pcc2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane c×2
28	$Pma2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane a
29	$Pca2_1$	C_{2v} ($mm2$)	nonsymmorphic: screw axis 2_1; glide plane c, a
30	$Pnc2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane n, c
31	$Pmn2_1$	C_{2v} ($mm2$)	nonsymmorphic: screw axis 2_1; glide plane n
32	$Pba2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane b, a
33	$Pna2_1$	C_{2v} ($mm2$)	nonsymmorphic: screw axis 2_1; glide plane n, a
34	$Pnn2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane n×2
35	$Cmm2$	C_{2v} ($mm2$)	symmorphic
36	$Cmc2_1$	C_{2v} ($mm2$)	nonsymmorphic: screw axis 2_1; glide plane c
37	$Ccc2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane c×2
38	$Amm2$	C_{2v} ($mm2$)	symmorphic
39	$Abm2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane b
40	$Ama2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane a
41	$Aba2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane b, a
42	$Fmm2$	C_{2v} ($mm2$)	symmorphic
43	$Fdd2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane d×2
44	$Imm2$	C_{2v} ($mm2$)	symmorphic
45	$Iba2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane b, a
46	$Ima2$	C_{2v} ($mm2$)	nonsymmorphic: glide plane a
47	$Pmmm$	D_{2h} (mmm)	symmorphic
48	$Pnnn$	D_{2h} (mmm)	nonsymmorphic: glide plane n×3
49	$Pccm$	D_{2h} (mmm)	nonsymmorphic: glide plane c×2
50	$Pban$	D_{2h} (mmm)	nonsymmorphic: glide plane b, a, n
51	$Pmma$	D_{2h} (mmm)	nonsymmorphic: glide plane a
52	$Pnna$	D_{2h} (mmm)	nonsymmorphic: glide plane n×2, a

No.	Space group	Point group	Symmorphic/nonsymmorphic note
53	*Pmna*	D_{2h} (*mmm*)	nonsymmorphic: glide plane n, a
54	*Pcca*	D_{2h} (*mmm*)	nonsymmorphic: glide plane c×2, a
55	*Pbam*	D_{2h} (*mmm*)	nonsymmorphic: glide plane b, a
56	*Pccn*	D_{2h} (*mmm*)	nonsymmorphic: glide plane c×2, n
57	*Pbcm*	D_{2h} (*mmm*)	nonsymmorphic: glide plane b, c
58	*Pnnm*	D_{2h} (*mmm*)	nonsymmorphic: glide plane n×2
59	*Pmmn*	D_{2h} (*mmm*)	nonsymmorphic: glide plane n
60	*Pbcn*	D_{2h} (*mmm*)	nonsymmorphic: glide plane b, c, n
61	*Pbca*	D_{2h} (*mmm*)	nonsymmorphic: glide plane b, c, a
62	*Pnma*	D_{2h} (*mmm*)	nonsymmorphic: glide plane n, a
63	*Cmcm*	D_{2h} (*mmm*)	nonsymmorphic: glide plane c
64	*Cmca*	D_{2h} (*mmm*)	nonsymmorphic: glide plane c, a
65	*Cmmm*	D_{2h} (*mmm*)	symmorphic
66	*Cccm*	D_{2h} (*mmm*)	nonsymmorphic: glide plane c×2
67	*Cmma*	D_{2h} (*mmm*)	nonsymmorphic: glide plane a
68	*Ccca*	D_{2h} (*mmm*)	nonsymmorphic: glide plane c×2, a
69	*Fmmm*	D_{2h} (*mmm*)	symmorphic
70	*Fddd*	D_{2h} (*mmm*)	nonsymmorphic: glide plane d×3
71	*Immm*	D_{2h} (*mmm*)	symmorphic
72	*Ibam*	D_{2h} (*mmm*)	nonsymmorphic: glide plane b, a
73	*Ibca*	D_{2h} (*mmm*)	nonsymmorphic: glide plane b, c, a
74	*Imma*	D_{2h} (*mmm*)	nonsymmorphic: glide plane a
75	*P*4	C_4 (4)	symmorphic
76	$P4_1$	C_4 (4)	nonsymmorphic: screw axis 4_1
77	$P4_2$	C_4 (4)	nonsymmorphic: screw axis 4_2
78	$P4_3$	C_4 (4)	nonsymmorphic: screw axis 4_3
79	*I*4	C_4 (4)	symmorphic
80	$I4_1$	C_4 (4)	nonsymmorphic: screw axis 4_1
81	*P*-4	S_4 (-4)	symmorphic
82	*I*-4	S_4 (-4)	symmorphic
83	*P*4/*m*	C_{4h} (4/*m*)	symmorphic
84	$P4_2/m$	C_{4h} (4/*m*)	nonsymmorphic: screw axis 4_2
85	*P*4/*n*	C_{4h} (4/*m*)	nonsymmorphic: glide plane n
86	$P4_2/n$	C_{4h} (4/*m*)	nonsymmorphic: screw axis 4_2; glide plane n
87	*I*4/*m*	C_{4h} (4/*m*)	symmorphic
88	$I4_1/a$	C_{4h} (4/*m*)	nonsymmorphic: screw axis 4_1; glide plane a
89	*P*422	D_4 (422)	symmorphic
90	$P4_212$	D_4 (422)	nonsymmorphic: screw axis 4_2
91	$P4_122$	D_4 (422)	nonsymmorphic: screw axis 4_1
92	$P4_12_12$	D_4 (422)	nonsymmorphic: screw axis 4_1, 2_1
93	$P4_222$	D_4 (422)	nonsymmorphic: screw axis 4_2
94	$P4_22_12$	D_4 (422)	nonsymmorphic: screw axis 4_2, 2_1
95	$P4_322$	D_4 (422)	nonsymmorphic: screw axis 4_3
96	$P4_32_12$	D_4 (422)	nonsymmorphic: screw axis 4_3, 2_1
97	*I*422	D_4 (422)	symmorphic
98	$I4_122$	D_4 (422)	nonsymmorphic: screw axis 4_1
99	*P*4*mm*	C_{4v} (4*mm*)	symmorphic
100	*P*4*bm*	C_{4v} (4*mm*)	nonsymmorphic: glide plane b
101	$P4_2cm$	C_{4v} (4*mm*)	nonsymmorphic: screw axis 4_2; glide plane c
102	$P4_2nm$	C_{4v} (4*mm*)	nonsymmorphic: screw axis 4_2; glide plane n
103	*P*4*cc*	C_{4v} (4*mm*)	nonsymmorphic: glide plane c×2
104	*P*4*nc*	C_{4v} (4*mm*)	nonsymmorphic: glide plane n, c
105	$P4_2mc$	C_{4v} (4*mm*)	nonsymmorphic: screw axis 4_2; glide plane c
106	$P4_2bc$	C_{4v} (4*mm*)	nonsymmorphic: screw axis 4_2; glide plane b, c
107	*I*4*mm*	C_{4v} (4*mm*)	symmorphic
108	*I*4*cm*	C_{4v} (4*mm*)	nonsymmorphic: glide plane c
109	$I4_1md$	C_{4v} (4*mm*)	nonsymmorphic: screw axis 4_1; glide plane d
110	$I4_1cd$	C_{4v} (4*mm*)	nonsymmorphic: screw axis 4_1; glide plane c, d
111	*P*-42*m*	D_{2d} (-42*m*)	symmorphic
112	*P*-42*c*	D_{2d} (-42*m*)	nonsymmorphic: glide plane c
113	$P\text{-}42_1m$	D_{2d} (-42*m*)	nonsymmorphic: screw axis 2_1
114	$P\text{-}42_1c$	D_{2d} (-42*m*)	nonsymmorphic: screw axis 2_1; glide plane c
115	*P*-4*m*2	D_{2d} (-42*m*)	symmorphic
116	*P*-4*c*2	D_{2d} (-42*m*)	nonsymmorphic: glide plane c
117	*P*-4*b*2	D_{2d} (-42*m*)	nonsymmorphic: glide plane b

No.	Space group	Point group	Symmorphic/nonsymmorphic note
118	*P*-4*n*2	D_{2d} (-42*m*)	nonsymmorphic: glide plane n
119	*I*-4*m*2	D_{2d} (-42*m*)	symmorphic
120	*I*-4*c*2	D_{2d} (-42*m*)	nonsymmorphic: glide plane c
121	*I*-42*m*	D_{2d} (-42*m*)	symmorphic
122	*I*-42*d*	D_{2d} (-42*m*)	nonsymmorphic: glide plane d
123	*P*4/*mmm*	D_{4h} (4/*mmm*)	symmorphic
124	*P*4/*mcc*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane c×2
125	*P*4/*nbm*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane n, b
126	*P*4/*nnc*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane n×2, c
127	*P*4/*mbm*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane b
128	*P*4/*mnc*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane n, c
129	*P*4/*nmm*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane n
130	*P*4/*ncc*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane n, c×2
131	$P4_2/mmc$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane c
132	$P4_2/mcm$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane c
133	$P4_2/nbc$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane n, b, c
134	$P4_2/nnm$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane n×2
135	$P4_2/mbc$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane b, c
136	$P4_2/mnm$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane n
137	$P4_2/nmc$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane n, c
138	$P4_2/ncm$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_2; glide plane n, c
139	*I*4/*mmm*	D_{4h} (4/*mmm*)	symmorphic
140	*I*4/*mcm*	D_{4h} (4/*mmm*)	nonsymmorphic: glide plane c
141	$I4_1/amd$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_1; glide plane a, d
142	$I4_1/acd$	D_{4h} (4/*mmm*)	nonsymmorphic: screw axis 4_1; glide plane a, c, d
143	*P*3	C_3 (3)	symmorphic
144	$P3_1$	C_3 (3)	nonsymmorphic: screw axis 3_1
145	$P3_2$	C_3 (3)	nonsymmorphic: screw axis 3_2
146	*R*3	C_3 (3)	symmorphic
147	*P*-3	C_{3i} (-3)	symmorphic
148	*R*-3	C_{3i} (-3)	symmorphic
149	*P*312	D_3 (32)	symmorphic
150	*P*321	D_3 (32)	symmorphic
151	$P3_112$	D_3 (32)	nonsymmorphic: screw axis 3_1
152	$P3_121$	D_3 (32)	nonsymmorphic: screw axis 3_1
153	$P3_212$	D_3 (32)	nonsymmorphic: screw axis 3_2
154	$P3_221$	D_3 (32)	nonsymmorphic: screw axis 3_2
155	*R*32	D_3 (32)	symmorphic
156	*P*3*m*1	C_{3v} (3*m*)	symmorphic
157	*P*31*m*	C_{3v} (3*m*)	symmorphic
158	*P*3*c*1	C_{3v} (3*m*)	nonsymmorphic: glide plane c
159	*P*31*c*	C_{3v} (3*m*)	nonsymmorphic: glide plane c
160	*R*3*m*	C_{3v} (3*m*)	symmorphic
161	*R*3*c*	C_{3v} (3*m*)	nonsymmorphic: glide plane c
162	*P*-31*m*	D_{3d} (-3*m*)	symmorphic
163	*P*-31*c*	D_{3d} (-3*m*)	nonsymmorphic: glide plane c
164	*P*-3*m*1	D_{3d} (-3*m*)	symmorphic
165	*P*-3*c*1	D_{3d} (-3*m*)	nonsymmorphic: glide plane c
166	*R*-3*m*	D_{3d} (-3*m*)	symmorphic
167	*R*-3*c*	D_{3d} (-3*m*)	nonsymmorphic: glide plane c
168	*P*6	C_6 (6)	symmorphic
169	$P6_1$	C_6 (6)	nonsymmorphic: screw axis 6_1
170	$P6_5$	C_6 (6)	nonsymmorphic: screw axis 6_5
171	$P6_2$	C_6 (6)	nonsymmorphic: screw axis 6_2
172	$P6_4$	C_6 (6)	nonsymmorphic: screw axis 6_4
173	$P6_3$	C_6 (6)	nonsymmorphic: screw axis 6_3
174	*P*-6	C_{3h} (-6)	symmorphic
175	*P*6/*m*	C_{6h} (6/m)	symmorphic
176	$P6_3/m$	C_{6h} (6/m)	nonsymmorphic: screw axis 6_3
177	*P*622	D_6 (622)	symmorphic
178	$P6_122$	D_6 (622)	nonsymmorphic: screw axis 6_1
179	$P6_522$	D_6 (622)	nonsymmorphic: screw axis 6_5
180	$P6_222$	D_6 (622)	nonsymmorphic: screw axis 6_2
181	$P6_422$	D_6 (622)	nonsymmorphic: screw axis 6_4
182	$P6_322$	D_6 (622)	nonsymmorphic: screw axis 6_3

No.	Space group	Point group	Symmorphic/nonsymmorphic note
183	*P6mm*	C_{6v} (*6mm*)	symmorphic
184	*P6cc*	C_{6v} (*6mm*)	nonsymmorphic: glide plane c×2
185	$P6_3cm$	C_{6v} (*6mm*)	nonsymmorphic: screw axis 6_3; glide plane c
186	$P6_3mc$	C_{6v} (*6mm*)	nonsymmorphic: screw axis 6_3; glide plane c
187	*P*-6*m*2	D_{3h} (-62*m*)	symmorphic
188	*P*-6*c*2	D_{3h} (-62*m*)	nonsymmorphic: glide plane c
189	*P*-62*m*	D_{3h} (-62*m*)	symmorphic
190	*P*-62*c*	D_{3h} (-62*m*)	nonsymmorphic: glide plane c
191	*P6/mmm*	D_{6h} (6/*mmm*)	symmorphic
192	*P6/mcc*	D_{6h} (6/*mmm*)	nonsymmorphic: glide plane c×2
193	$P6_3/mcm$	D_{6h} (6/*mmm*)	nonsymmorphic: screw axis 6_3; glide plane c
194	$P6_3/mmc$	D_{6h} (6/*mmm*)	nonsymmorphic: screw axis 6_3; glide plane c
195	*P*23	*T* (23)	symmorphic
196	*F*23	*T* (23)	symmorphic
197	*I*23	*T* (23)	symmorphic
198	$P2_13$	*T* (23)	nonsymmorphic: screw axis 2_1
199	$I2_13$	*T* (23)	nonsymmorphic: screw axis 2_1
200	*Pm*-3	T_h (*m*-3)	symmorphic
201	*Pn*-3	T_h (*m*-3)	nonsymmorphic: glide plane n
202	*Fm*-3	T_h (*m*-3)	symmorphic
203	*Fd*-3	T_h (*m*-3)	nonsymmorphic: glide plane d
204	*Im*-3	T_h (*m*-3)	symmorphic
205	*Pa*-3	T_h (*m*-3)	nonsymmorphic: glide plane a
206	*Ia*-3	T_h (*m*-3)	nonsymmorphic: glide plane a
207	*P*432	*O* (432)	symmorphic
208	$P4_232$	*O* (432)	nonsymmorphic: screw axis 4_2
209	*F*432	*O* (432)	symmorphic
210	$F4_132$	*O* (432)	nonsymmorphic: screw axis 4_1
211	*I*432	*O* (432)	symmorphic
212	$P4_332$	*O* (432)	nonsymmorphic: screw axis 4_3
213	$P4_132$	*O* (432)	nonsymmorphic: screw axis 4_1
214	$I4_132$	*O* (432)	nonsymmorphic: screw axis 4_1
215	*P*-43*m*	T_d (-43*m*)	symmorphic
216	*F*-43*m*	T_d (-43*m*)	symmorphic
217	*I*-43*m*	T_d (-43*m*)	symmorphic
218	*P*-43*n*	T_d (-43*m*)	nonsymmorphic: glide plane n
219	*F*-43*c*	T_d (-43*m*)	nonsymmorphic: glide plane c
220	*I*-43*d*	T_d (-43*m*)	nonsymmorphic: glide plane d
221	*Pm*-3*m*	O_h (*m*-3*m*)	symmorphic
222	*Pn*-3*n*	O_h (*m*-3*m*)	nonsymmorphic: glide plane n×2
223	*Pm*-3*n*	O_h (*m*-3*m*)	nonsymmorphic: glide plane n
224	*Pn*-3*m*	O_h (*m*-3*m*)	nonsymmorphic: glide plane n
225	*Fm*-3*m*	O_h (*m*-3*m*)	symmorphic
226	*Fm*-3*c*	O_h (*m*-3*m*)	nonsymmorphic: glide plane c
227	*Fd*-3*m*	O_h (*m*-3*m*)	nonsymmorphic: glide plane d
228	*Fd*-3*c*	O_h (*m*-3*m*)	nonsymmorphic: glide plane d, c
229	*Im*-3*m*	O_h (*m*-3*m*)	symmorphic
230	*Ia*-3*d*	O_h (*m*-3*m*)	nonsymmorphic: glide plane a, d

Appendix A3. Additional Energy Landscape Exploration Information

The Table A3.1 summarizes the roles of the pre-existing methods used before the new methods STyX and AI-ELX developments.

Table A3.1. Comparison of existing energy exploration methods and structure analysis methods used in the thesis workflow.

Method	**Function in the workflow**	**Strengths and limitations**
Global optimization using simulated annealing in G42+ software package	Stochastic global exploration of an empirical energy or enthalpy landscape using Metropolis moves and periodic quenches.	Efficient and symmetry-unbiased; can sample unexpected structure types and very large candidate sets. Empirical ranking is approximate and duplicates are numerous.
ICSD prototype data mining	Selection, filtering, and chemical remapping of known crystallographic prototypes to the target composition.	Adds chemically plausible and often high-symmetry starting structures. Limited by database coverage.
Structural refinement using SFND, RGS, and CMPZ algorithms in KPLOT or in Lisp scripts	Symmetry determination, space-group assignment, and finding duplicates and structure types after GO or DM phase.	Transforms the GO/DM output into interpretable structure families. Tolerance choices influence grouping. Limited to tens of thousands of structure candidates.
CRYSTAL local optimization	*Ab initio* geometry relaxation using localized Gaussian basis functions and analytical gradients.	Robust and accurate for structural relaxation and property calculations. Good for modeling finite or low-dimensional systems (molecules, polymers, surfaces). Requires careful basis-set selection to avoid basis set superposition error.
VASP local optimization	*Ab initio* geometry relaxation using plane-wave basis.	Robust and accurate for structural relaxation and property calculations. A complete, orthogonal basis set that scales systematically with a single energy cutoff parameter. Highly efficient and robust for large supercells, bulk materials, and periodic defects, but computationally expensive for 1D or 2D systems.

Tables A3.2 and A3.3 describe the present implementation of the AI-ELX model.

Table A3.2. Principal scripts used in the current AI-ELX implementation.

Script / component	**Primary role**	**Main input / output**
train_gnn.py	Training of the volume-aware graph neural network	VASP structure files and matching energy values → trained model checkpoint
optimizer.py	Differentiable relaxation of fractional atomic coordinates and cell parameters	Input structure + model checkpoint → relaxed structure and predicted *E(V)* points
fit_birch_murnaghan_ev.py	Fitting of predicted *E(V)* points to a Birch-Murnaghan equation of state	Predicted *E(V)* points → fitted *E(V)* curve, residuals, and metadata
calculate_enthalpy_from_ev.py	Calculation of pressure and enthalpy from a fitted *E(V)* curve	Fitted *E(V)* curve → *H(p)* curve

Table A3.3. Settings for the current AI-ELX implementation.

Setting	Representative value	Meaning
k	32	number of periodic nearest neighbors per atom in the representative training run
images	2 2 2	translated-image search range used for periodic neighbor construction
emb-dim	128	dimension of the learned elemental embedding
msg-hidden	256	hidden width of message, update, and readout MLP blocks
layers	4	number of message-passing / residual update stages
n-rbf	96	number of Gaussian radial basis functions for reduced distances
cutoff-red	7.0	support of the reduced-distance radial basis expansion
batch-size	32	mini-batch size used during supervised training
lr	2×10^{-4}	initial learning rate in the representative training run
epochs	300	maximum number of training epochs
rebuild-every	5	example graph-rebuilding interval used during surrogate relaxation

Appendix A4. Additional Figures for LaOI

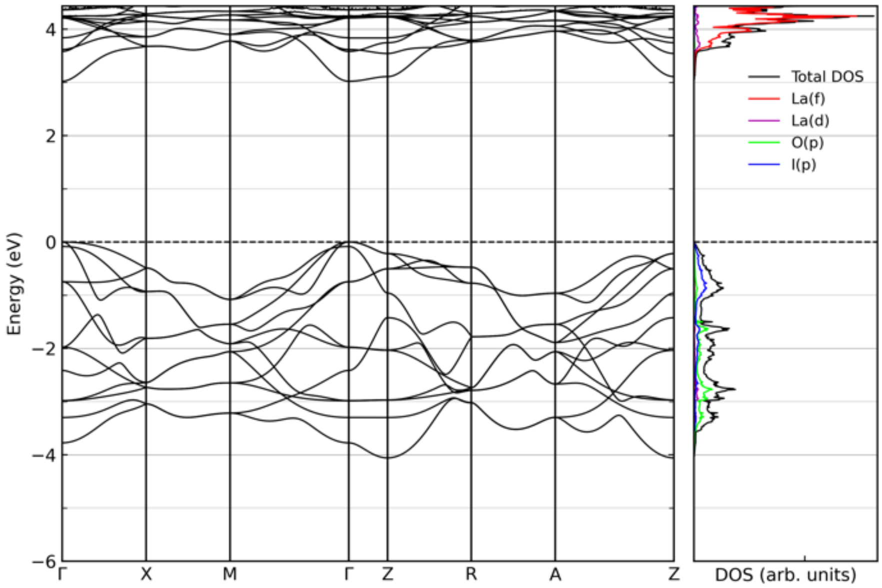

Figure A4.1. Band structure and DOS for α-LaOI calculated using LDA functional.

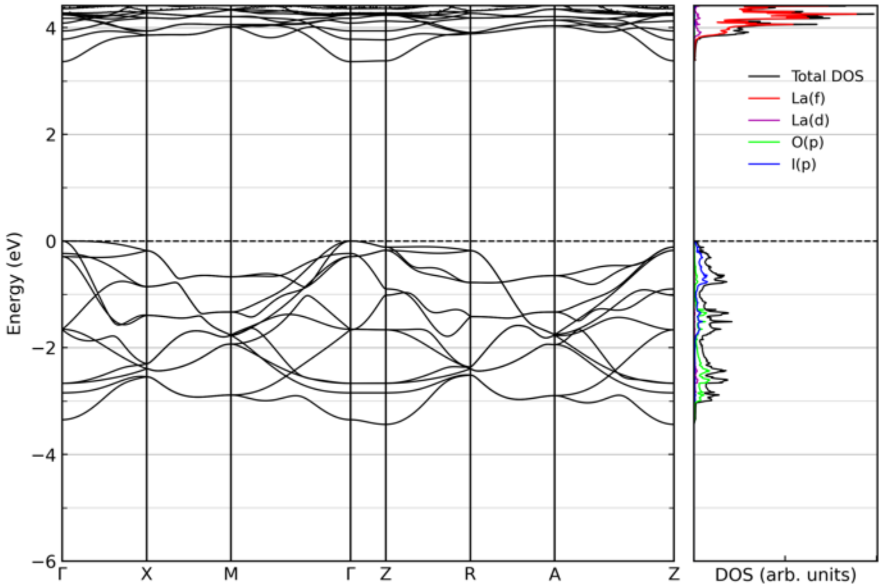

Figure A4.2. Band structure and DOS for α-LaOI calculated using PBE functional.

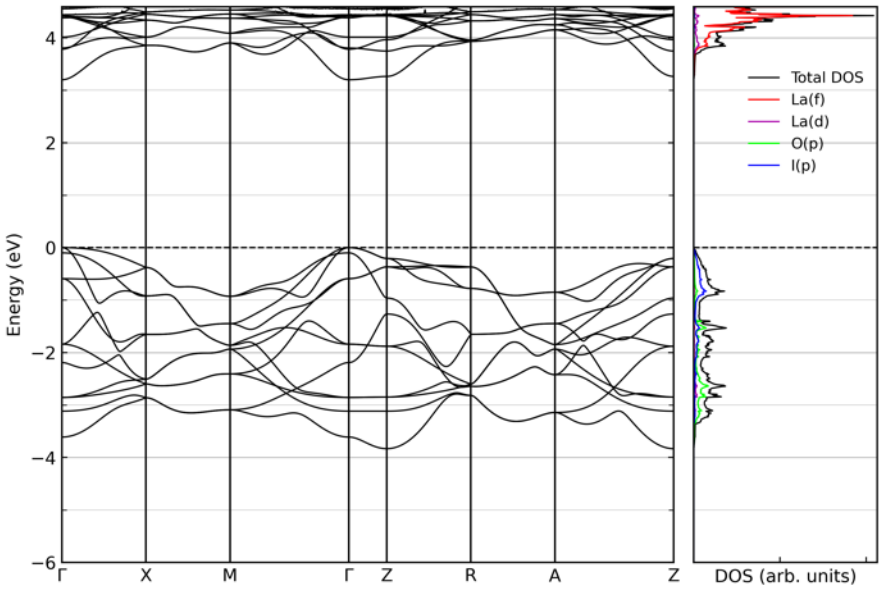

Figure A4.3. Band structure and DOS for α-LaOI calculated using PBEsol functional.

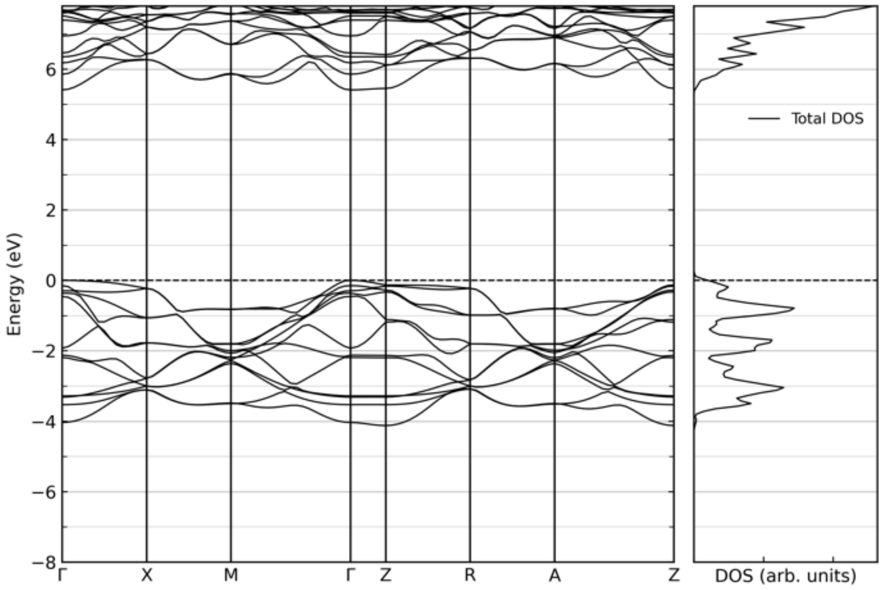

Figure A4.4. Band structure and DOS for α-LaOI calculated using PBE0 functional.

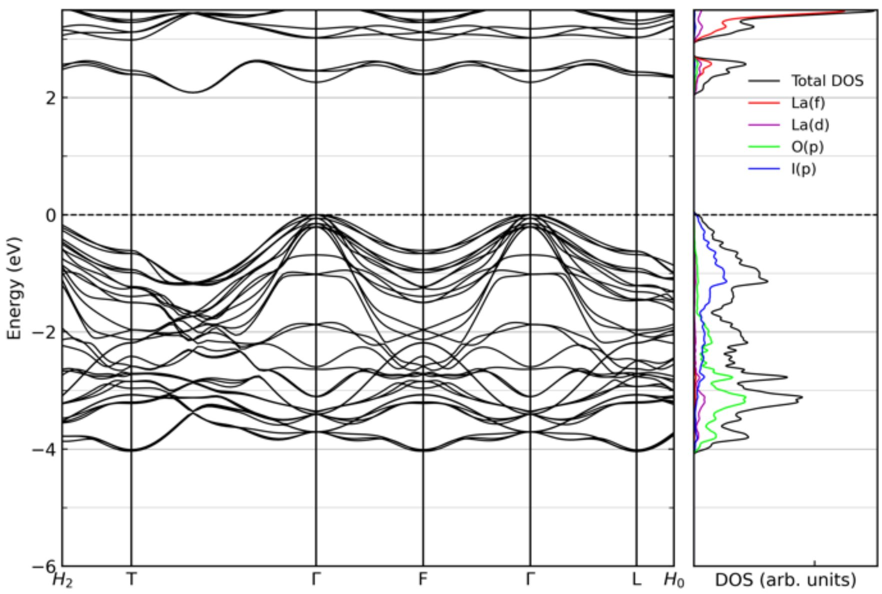

Figure A4.5. Band structure and DOS for *β*-LaOI calculated using LDA functional.

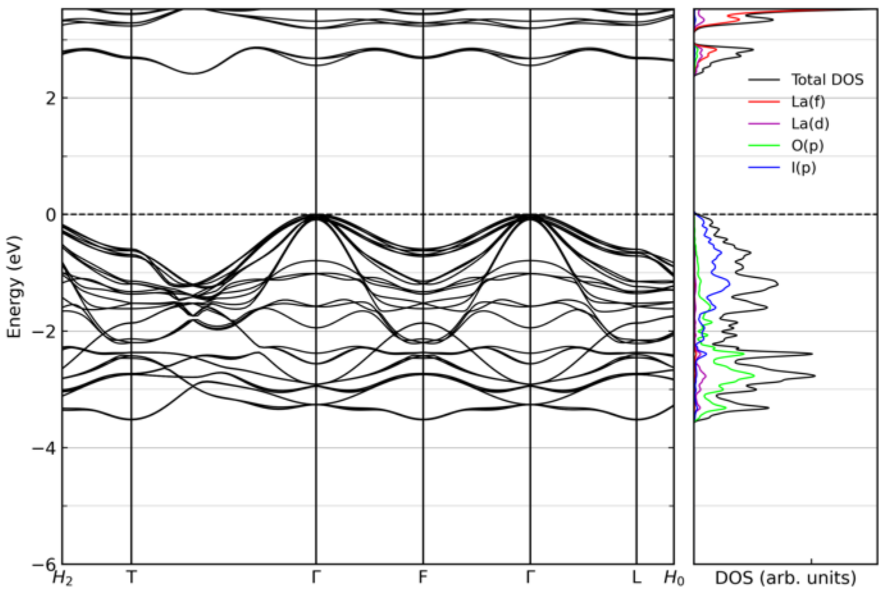

Figure A4.6. Band structure and DOS for *β*-LaOI calculated using PBE functional.

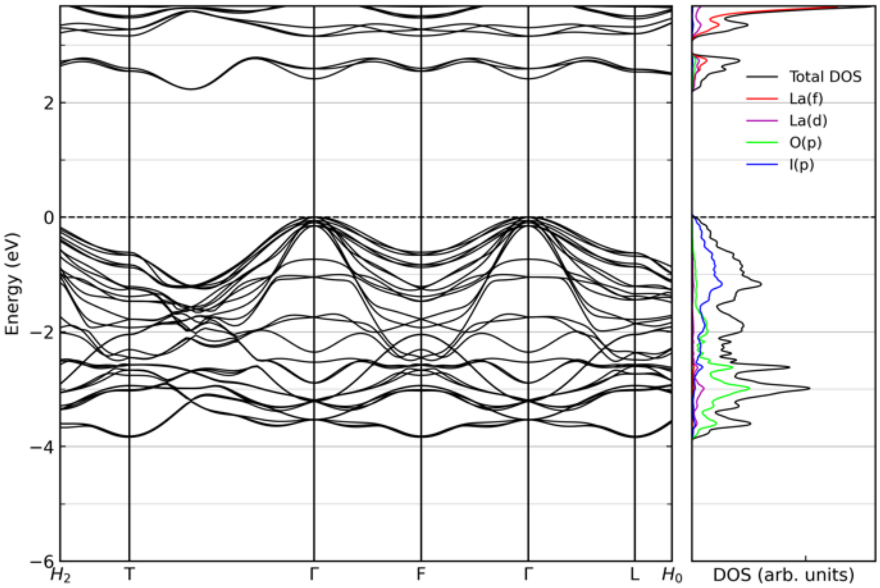

Figure A4.7. Band structure and DOS for *β*-LaOI calculated using PBEsol functional.

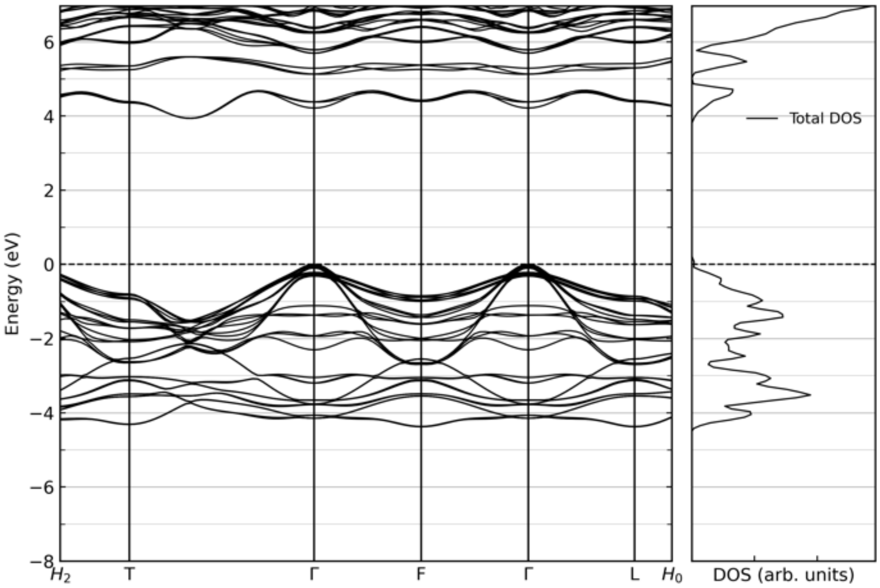

Figure A4.8. Band structure and DOS for *β*-LaOI calculated using PBE0 functional.

Appendix A5. Additional Information for Ce_3O_3N

Table A5.1. Space group, unit-cell parameters, atomic positions, and total-energy values for the optimized Ce_3O_3N modifications obtained by global search.

Modification	Space group (No.)	Cell parameters (Å) and fractional coordinates	Total energy (E_h)
Ce_3O_3N-GS1	*P*2/*m* (10)	a = 5.89; b = 3.62; c = 5.04; β = 113.2° Ce (0 0 0) Ce (0.3017 1/2 0.6238) O (0.2702 1/2 0.0652) O (1/2 0 1/2) N (0 0 1/2)	-1702.4813
Ce_3O_3N-GS2	*Amm*2 (38)	a = 3.59; b = 9.83; c = 5.78 Ce (1/2 0.1729 0.8911) Ce (0 0 0.3964) Ce (1/2 0 0.8038) O (0.5 0 0.1402) O (0 0.7712 0.1681) N (1/2 0 0.6528)	-1702.4685
Ce_3O_3N-GS3	*Imm*2 (44)	a = 3.38; b = 3.39; c = 17.39 Ce (0 0 0.7735) Ce (0 0 0.4276) Ce (0 0 0.0814) O (0 0 0.5589) O (0 0 0.2961) O (0 1/2 0.1775) N (0 0 0.9277)	-1702.4627
Ce_3O_3N-GS4	*Pmmm* (47)	a = 6.86; b = 3.53; c = 4.62 Ce (1/2 0 0) Ce (0.7607 1/2 1/2) O (0.2869 1/2 0) O (0 0 1/2) N (1/2 0 1/2)	-1702.4588
Ce_3O_3N-GS5	*Amm*2 (38)	a = 3.49; b = 3.32; c = 17.93 Ce (1/2 0 0.0970) Ce (0 0 0.4475) Ce (1/2 0 0.8038) O (0 0 0.5730) O (0 0 0.3143) O (1/2 0 0.6850) N (1/2 0 0.9444)	-1702.4576
Ce_3O_3N-GS6	*Pmmm* (47)	a = 3.53; b = 4.63; c = 6.85 Ce (0 0 0) Ce (1/2 1/2 0.7394) O (1/2 0 0.2132) O (0 1/2 1/2) N (0 1/2 0)	-1702.4509
Ce_3O_3N-GS7	*Pm*-3*m* (221)	a = 4.75 Ce (1/2 0 1/2) O (0 1/2 0) N (1/2 1/2 1/2)	-1702.4394

Table A5.2. Space group, unit-cell parameters, atomic positions, and total-energy values for the optimized Ce_3O_3N modifications obtained by data mining.

Modification	**Space group (No.)**	**Cell parameters (Å) and fractional coordinates**	**Total energy (E_h)**
Ce_3O_3N-DM1	*R*3*c* (161)	a = 10.17; c = 6.15 Ce (-0.1996 -0.0378 0.2416) O (-0.0656 0.2200 0.3762) N (0 0 0.0066)	-1702.4875
Ce_3O_3N-DM2	*P*6_3/*m* (176)	a = 7.21; c = 4.75 Ce (0.3715 0.3398 1/4) O (0.2747 0.9691 1/4) N (2/3 1/3 1/4)	-1702.4007
Ce_3O_3N-DM3	*I*-43*m* (217)	a = 10.49 Ce (1/2 0 1/4) Ce (0 0 0.2430) O (0.8876 0.6093 0.1124) N (0.1159 0.1159 0.8841)	-1702.2837

Appendix A6. Additional Figures for ScOCl

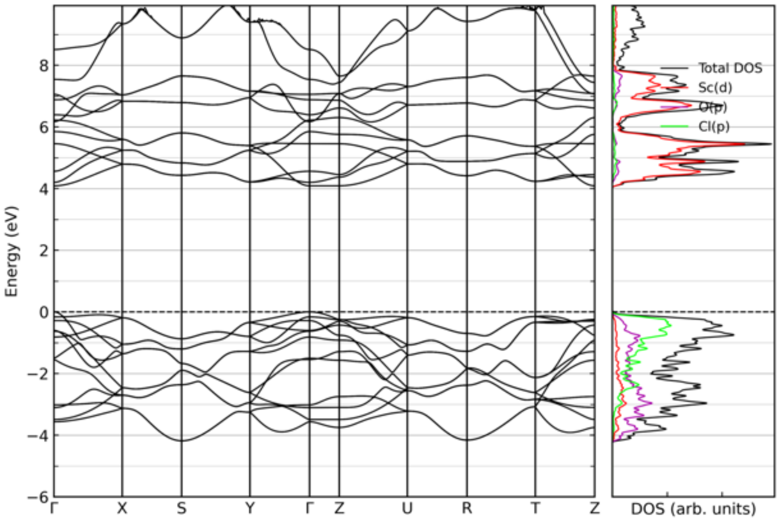

Figure A6.1. Electronic band structure and projected density of states of α-ScOCl calculated with LDA functional.

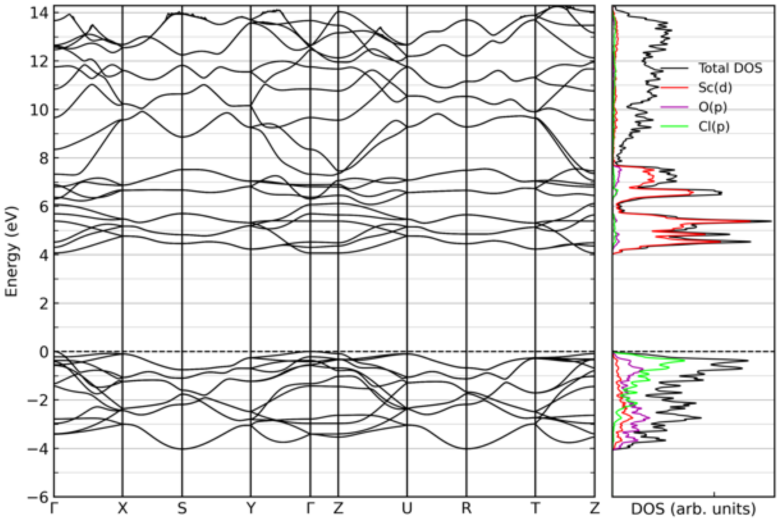

Figure A6.2. Electronic band structure and projected density of states of α-ScOCl calculated with PBEsol functional.

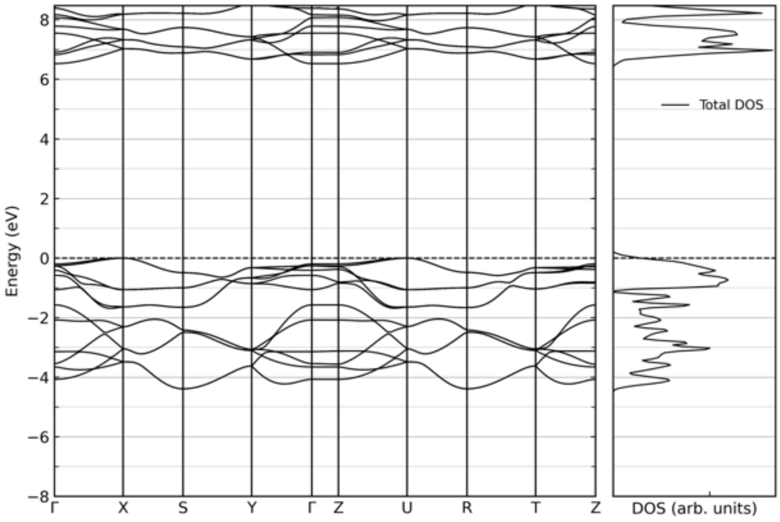

Figure A6.3. Electronic band structure and projected density of states of *α*-ScOCl calculated with PBE0 functional.

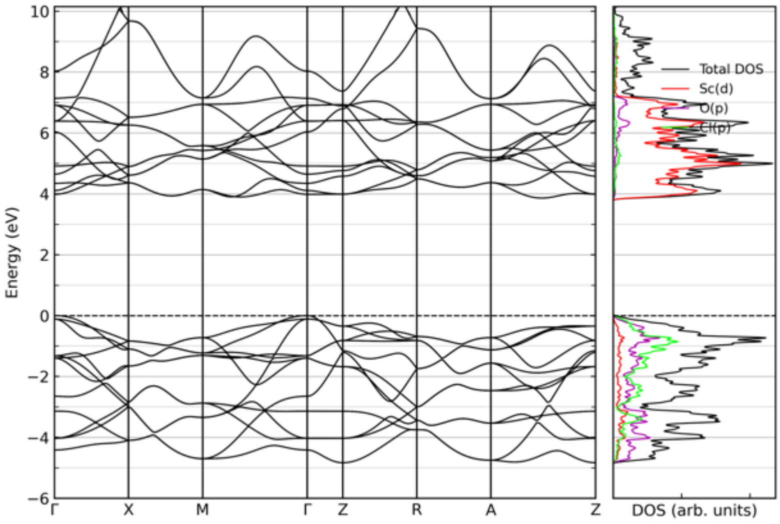

Figure A6.4. Electronic band structure and projected density of states of *β*-ScOCl calculated with LDA functional.

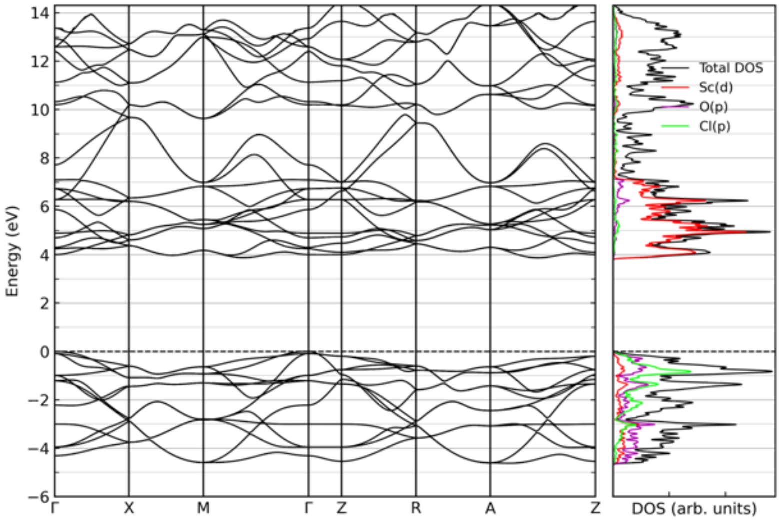

Figure A6.5. Electronic band structure and projected density of states of *β*-ScOCl calculated with PBEsol functional.

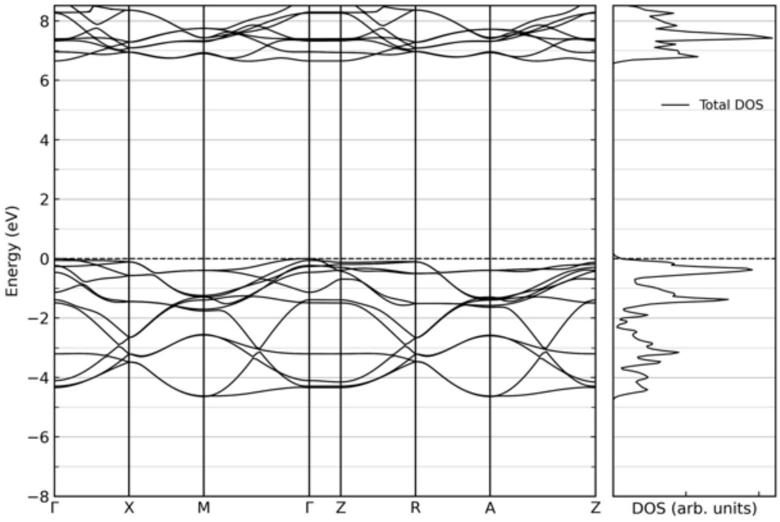

Figure A6.6. Electronic band structure and projected density of states of *β*-ScOCl calculated with PBE0 functional.

Appendix A7. Additional Information for HoFSe

Table A7.1. Full structural data for the low-energy HoFSe polymorphs calculated using GGA-PBE functional.

E_{RI}	**Modification and Space Group**	**Cell Parameters (pm / °)**	**Atomic Positions**
1	*exp1*-type; No. 62	a=1009.1 b=416.8 c=643.0	Ho 0.4203 0.2500 0.2323 F 0.0342 0.2500 -0.4106 Se 0.1638 0.2500 0.0324
2	*α*-type; No. 11	a=646.4 b=416.6 c=1539.6 β=98.52	Ho 0.1344 0.2500 0.2219 Ho 0.2863 0.2500 -0.4479 Ho -0.3668 -0.2500 0.1186 F -0.2107 0.2500 0.1463 F 0.4189 -0.2500 -0.4775 F -0.0190 -0.2500 0.1920 Se -0.0079 0.2500 0.3914 Se 0.4521 -0.2500 0.2786 Se -0.3117 -0.2500 -0.0622
3	*β*-type; No. 62	a=676.1 b=387.8 c=894.7	Ho -0.2818 0.2500 0.1013 F -0.3825 -0.2500 -0.0265 Se -0.4608 -0.2500 0.3255
4	*γ*-type; No. 63	a=565.6 b=1189.0 c=415.9	Ho 0.0000 0.1430 -0.2500 F 0.0000 0.0555 0.2500 Se 0.5000 -0.1774 0.2500
5	*PbFCl*; No. 129	a=394.5 c=712.3	Ho -0.2500 -0.2500 0.2165 F 0.2500 -0.2500 0.0000 Se -0.2500 -0.2500 -0.3552
6	*δ*-type; No. 12	a=1957.2 b=416.5 c=1011.0 β=100.53	Ho -0.1585 0.5000 0.1066 Ho -0.1738 0.5000 -0.3889 Ho 0.0000 0.0000 0.0000 Ho 0.0000 0.0000 -0.5000 F -0.0555 0.5000 0.0304 F -0.1099 -0.0000 0.0826 F -0.2207 -0.0000 -0.4537 Se 0.2415 0.5000 0.1599 Se -0.0722 0.0000 -0.2753 Se 0.0960 0.5000 -0.3919
7	*CoYC*; No. 131	a=438.0 c=779.8	Ho 0.0000 -0.5000 -0.5000 F -0.5000 -0.5000 -0.5000 Se 0.0000 0.0000 -0.2500
8	*ε*-type; No. 87	a=1113.8 c=387.9	Ho -0.0647 0.2796 -0.0000 F -0.1013 -0.4175 0.0000 Se 0.1200 0.1911 0.5000
9	*ZrOS*; No. 198	a=631.7	Ho -0.4413 -0.0587 0.0587 F -0.3419 0.1581 0.3419 Se 0.3274 -0.1726 -0.3274
10	*exp2*-type; No. 11	a=1010.6 b=416.5 c=1328.4 β=105.15	Ho 0.4547 0.2500 0.1272 Ho -0.0406 0.2500 0.1205 Ho 0.1201 0.2500 -0.3607 Ho -0.3756 0.2500 -0.3869 F -0.0999 0.2500 0.2906 F 0.4492 0.2500 0.4560 F 0.0733 0.2500 -0.2010 F 0.0059 0.2500 -0.0427 Se 0.2436 0.2500 0.2355 Se -0.3321 0.2500 0.0175 Se 0.4179 0.2500 -0.2690 Se -0.1585 0.2500 -0.4884
11	*exp3*-type; No. 62	a=1012.8 b=415.8 c=1914.8	Ho 0.0967 0.2500 0.0833 Ho 0.2486 0.2500 -0.2382 Ho 0.0817 0.2500 0.4194 F 0.2100 0.2500 0.3094 F 0.3713 0.2500 -0.1371 F 0.4511 0.2500 -0.0194 Se 0.3385 0.2500 0.1598 Se 0.3398 0.2500 0.4854 Se 0.4865 0.2500 -0.3225

E_{RI}	Modification and Space Group	Cell Parameters (pm / °)	Atomic Positions
12	*ZrBeSi*; No. 194	a=405.1 c=799.5	Ho 0.3333 -0.3333 -0.2500 F 0.3333 -0.3333 0.2500 Se 0.0000 0.0000 0.0000
13	*ζ*-type; No. 31	a=415.5 b=1530.9 c=633.8	Ho 0.5000 0.4446 -0.2133 Ho 0.0000 -0.1126 0.2563 Ho 0.5000 -0.2204 -0.2098 F 0.0000 0.4599 -0.0192 F 0.0000 0.1407 -0.3903 F 0.5000 0.1915 0.4376 Se 0.5000 0.2750 -0.0122 Se 0.5000 -0.3875 0.0022 Se 0.5000 -0.0583 -0.4381
14	*η*-type; No. 12	a=1516.1 b=410.5 c=797.3 β=93.12	Ho 0.1651 -0.5000 -0.3054 Ho -0.0509 0.0000 -0.2648 F -0.0103 -0.5000 0.3082 F 0.0852 -0.0000 -0.4412 Se 0.1936 0.5000 0.3320 Se -0.1197 0.0000 0.0616
15	*θ*-type; No. 11	a=1101.3 b=415.8 c=924.9 β=98.64	Ho -0.3300 -0.2500 0.4800 Ho -0.0177 0.2500 0.2276 Ho -0.3508 0.2500 -0.0261 F -0.4421 -0.2500 -0.0177 F 0.0067 0.2500 -0.3085 F -0.1474 -0.2500 -0.3959 Se 0.2730 -0.2500 -0.2864 Se 0.4110 -0.2500 0.3393 Se 0.1523 0.2500 0.0161
16	*3R*-type; No. 166	a=413.3 c=2699.3	Ho 0.0000 0.0000 -0.5000 Ho 0.0000 0.0000 0.0000 F -0.0000 -0.0000 0.1752 Se -0.3333 0.3333 -0.0586
17	*2H*-type; No. 194	a=415.3 c=1795.6	Ho 0.0000 0.0000 0.0000 Ho 0.3333 -0.3333 0.2500 F 0.0000 0.0000 0.2500 F 0.3333 -0.3333 -0.2500 Se 0.3333 -0.3333 0.0874
18	*ι*-type; No. 14	a=702.0 b=529.3 c=764.9 β=116.23	Ho -0.2295 -0.1184 0.2696 F -0.0882 -0.3894 0.1163 Se -0.3139 0.1088 -0.1048
19	*κ*-type; No. 47	a=431.3 b=439.1 c=1186.5	Ho 0.0000 0.0000 0.3284 Ho -0.5000 -0.5000 0.0000 F 0.0000 -0.5000 -0.3229 F 0.0000 -0.5000 0.0000 Se -0.5000 0.0000 -0.1637 Se -0.5000 0.0000 -0.5000
20	*λ*-type; No. 12	a=1548.6 b=414.2 c=1110.3 β=125.04	Ho 0.2727 -0.0000 0.3482 Ho 0.0344 0.5000 -0.1764 F -0.3795 -0.0000 -0.2803 F -0.0936 -0.0000 0.1139 Se 0.1190 0.0000 0.4272 Se 0.1434 0.5000 0.1481

Table A7.2. Full structural data for the low-energy HoFSe polymorphs calculated using HSE06.

E_{RI}	Modification and Space Group	Cell Parameters (pm / °)	Atomic Positions
1	*exp1*-type; No. 62	a=1006.0 b=414.1 c=640.2	Ho 0.4202 0.2500 0.7306 F 0.0344 0.2500 0.0883 Se 0.1634 0.2500 0.5323
2	*α*-type; No. 11	a=644.2 b=413.7 c=1534.2 β=98.46	Ho 0.1331 0.2500 0.2220 Ho 0.2876 0.2500 -0.4480 Ho -0.3651 -0.2500 0.1189 F -0.2099 0.2500 0.1468 F 0.4193 -0.2500 -0.4777 F -0.0194 -0.2500 0.1920 Se -0.0080 0.2500 0.3916 Se 0.4514 -0.2500 0.2787 Se -0.3130 -0.2500 -0.0625
3	*β*-type; No. 62	a=673.4 b=386.2 c=891.0	Ho -0.2825 0.2500 0.1002 F -0.3837 -0.2500 -0.0251 Se -0.4612 -0.2500 0.3250
4	*δ*-type; No. 12	a=1948.6 b=413.9 c= 1007.9 β=100.492	Ho -0.1578 0.5000 0.1065 Ho -0.1744 0.5000 -0.3888 Ho 0.0000 0.0000 0.0000 Ho 0.0000 0.0000 -0.5000 F -0.0552 0.5000 0.0304 F -0.1093. 0.0000 0.0824 F -0.2210 0.0000 -0.4537 Se 0.2418 0.5000 0.1595 Se -0.0724 0.0000 -0.2754 Se 0.0963 -0.5000 -0.3919
5	*ε*-type; No. 87	a=1107.8 c=386.1	Ho -0.0639 0.2802 -0.0000 F -0.1013 -0.4181 0.0000 Se 0.1200 0.1916 0.5000
6	*γ*-type; No. 63	a=564.1 b=1184.4 c=413.0	Ho 0.0000 0.1422 -0.2500 F 0.0000 0.0552 0.2500 Se 0.5000 -0.1777 0.2500
7	*PbFCl*; No. 129	a=393.1 c=705.3	Ho 0.2500 0.2500 0.2160 F 0.7500 0.2500 0.0000 Se 0.2500 0.2500 0.6457
8	*ZrOS*; No. 198	a=627.8	Ho -0.4405 -0.0595 0.0595 F -0.3434 0.1566 0.3434 Se 0.3293 -0.1707 -0.3293
9	*exp2*-type; No. 11	a=1009.1 b=413.5 c=1319.9 β=105.22	Ho 0.4527 0.2500 0.1274 Ho -0.0421 0.2500 0.1194 Ho 0.1206 0.2500 -0.3587 Ho -0.3743 0.2500 -0.3883 F -0.1079 0.2500 0.2852 F 0.4482 0.2500 0.4571 F 0.0585 0.2500 -0.2045 F 0.0370 0.2500 -0.0298 Se 0.2425 0.2500 0.2360 Se -0.3333 0.2500 0.0189 Se 0.4188 0.2500 -0.2702 Se -0.1570 0.2500 -0.4894
10	*CoYC*; No. 131	a=435.8 c=778.4	Ho 0.0000 -0.5000 0.5000 F -0.5000 -0.5000 0.5000 Se 0.0000 0.0000 -0.2500
11	*exp3*-type; No. 62	a=1009.7 b=413.3 c=1906.3	Ho 0.0964 0.2500 0.0834 Ho 0.2491 0.2500 -0.2372 Ho 0.0809 0.2500 0.4194 F 0.2088 0.2500 0.3101 F 0.3711 0.2500 -0.1366 F 0.4529 0.2500 -0.0199 Se 0.3379 0.2500 0.1603 Se 0.3395 0.2500 0.4853 Se 0.4866 0.2500 -0.3222
12	*ZrBeSi*; No. 194	a=402.6 c=796.4	Ho 0.3333 -0.3333 -0.2500 F 0.3333 -0.3333 0.2500 Se 0.0000 0.0000 0.0000

E_{RI}	Modification and Space Group	Cell Parameters (pm / °)	Atomic Positions
13	*ζ*-type; No. 31	a=413.2 b=1525.0 c=631.4	Ho -0.5000 0.4446 -0.2126 Ho 0.0000 -0.1127 0.2549 Ho -0.5000 -0.2201 -0.2083 F 0.0000 0.4607 -0.0195 F 0.0000 0.1415 -0.3901 F -0.5000 0.1907 0.4372 Se -0.5000 0.2750 -0.0115 Se -0.5000 -0.3872 0.0023 Se -0.5000 -0.0582 -0.4393
14	*η*-type; No. 12	a=1508.8 b=408.0 c=793.5 β=93.17	Ho 0.1650 -0.5000 -0.3048 Ho -0.0506 0.0000 -0.2662 F -0.0107 -0.5000 0.3096 F 0.0847 -0.0000 -0.4408 Se 0.1933 -0.5000 0.3323 Se -0.1194 0.0000 0.0614
15	*θ*-type; No. 11	a=1097.1 b=413.8 c=921.9 β=98.70	Ho -0.3294 -0.2500 0.4805 Ho -0.0173 0.2500 0.2277 Ho -0.3516 0.2500 -0.0260 F -0.4425 -0.2500 -0.0173 F 0.0063 0.2500 -0.3076 F -0.1472 -0.2500 -0.3954 Se 0.2728 -0.2500 -0.2867 Se 0.4113 -0.2500 0.3392 Se 0.1525 0.2500 0.0154
16	*3R*-type; No. 166	a=411.0 c=2693.9	Ho 0.0000 0.0000 -0.5000 Ho 0.0000 0.0000 0.0000 F -0.0000 -0.0000 0.1746 Se -0.3333 0.3333 -0.0586
17	*2H*-type; No. 194	a=412.7 c=1791.9	Ho 0.0000 0.0000 0.0000 Ho 0.3333 -0.3333 0.2500 F 0.0000 0.0000 0.2500 F 0.3333 -0.3333 -0.2500 Se 0.3333 -0.3333 0.0875
18	*ι*-type; No. 14	a=694.8 b=527.8 c=763.2 β=116.46	Ho -0.2275 -0.1219 0.2710 F -0.0889 -0.3895 0.1154 Se -0.3111 0.1114 -0.1017
19	*κ*-type; No. 47	a=429.3 b=436.4 c=1184.9	Ho 0.0000 0.0000 0.3283 Ho -0.5000 0.5000 0.0000 F 0.0000 0.5000 -0.3228 F 0.0000 0.5000 0.0000 Se -0.5000 0.0000 -0.1637 Se -0.5000 0.0000 -0.5000
20	*λ*-type; No. 12	a=1542.3 b=412.6 c=1106.0 β=124.90	Ho 0.2736 -0.0000 0.3480 Ho 0.0342 0.5000 -0.1765 F -0.3805 -0.0000 -0.2808 F -0.0919 0.0000 0.1159 Se 0.1196 0.0000 0.4270 Se 0.1440 0.5000. 0.1478

Table A7.3. Low-energy HoFSe structure candidates sorted according to the HSE06 energy-ranking index (E_{RI}).

E_{RI}	Modification	SG No.	Energy / FU [E_h]	GO	DM	Exp.
1	*exp1*-type	62	-2538.97274	+	+	+
2	*α*-type	11	-2538.97203	+		
3	*β*-type	62	-2538.97199	+		
4	*δ*-type	12	-2538.97127	+		
5	*ε*-type	87	-2538.97112	+		
6	*γ*-type	63	-2538.97095	+		
7	*PbFCl*	129	-2538.97090	+	+	
8	*ZrOS*	198	-2538.97048	+	+	
9	*exp2*-type	11	-2538.97046		+	+
10	*CoYC*	131	-2538.96998	+		
11	*exp3*-type	62	-2538.96973		+	+
12	*ZrBeSi*	194	-2538.96928		+	
13	*ζ*-type	31	-2538.96772	+		
14	*η*-type	12	-2538.96760	+		
15	*θ*-type	11	-2538.96703	+		
16	*3R*-type	166	-2538.96696	+	+	
17	*2H*-type	194	-2538.96669		+	
18	*ι*-type	14	-2538.96634	+		
19	*κ*-type	47	-2538.96548	+		
20	*λ*-type	12	-2538.96498	+		

Appendix A8. Additional Information for LaFS

The most relevant numerical settings used in the LaFS VASP calculations are summarized below.

- Plane-wave basis: A kinetic energy cutoff in the range of 520-700 eV was employed, depending on the purpose of the calculation.
- Electronic convergence: The electronic self-consistency threshold was typically set in the range of 10^{-6} eV for structural relaxations, about 10^{-7} eV for the self-consistent calculations used to generate converged charge densities, DOS, and band structure calculations, and tighter threshold of about 10^{-8} eV was used for the phonon calculations.
- Ionic convergence: Structural relaxations were considered converged when the residual forces reached values on the order of 10^{-2} eV/Å.
- Brillouin-zone sampling: Monkhorst-Pack-type meshes were used with a consistent reciprocal-space resolution across cells of different size and shape. In practice, minimum sampling corresponded approximately to a 6×6×6 mesh for small primitive cells. This was implemented either through explicit k-point meshes or through an equivalent reciprocal-space spacing parameter set to approximately 0.2 $Å^{-1}$.
- Smearing: Structural relaxations and self-consistent calculations for semiconducting LaFS polymorphs were performed using Gaussian smearing. The DOS calculations were likewise evaluated from the converged charge density using Gaussian smearing and a dense energy grid, whereas the band structures were obtained in a separate non-self-consistent step along high-symmetry paths in the Brillouin zone using the same converged charge density; in this step the tetrahedron scheme was employed for the electronic occupations.
- Spin treatment: All investigated LaFS polymorphs were treated as nonmagnetic, i.e., using non-spin-polarized calculations.

Table A8.1. Calculated unit-cell parameters and atomic positions for all polymorphs of LaFS computed using the GGA-PBE functional.

E_{RI}	Modification and Space Group	SG No.	Cell Parameters (Å / °)	Atom Positions
1	*PbFCl*-type	129	a = 4.051 c = 7.023	La 0.0000 0.5000 0.2317 S 0.0000 0.5000 0.6492 F 0.0000 0.0000 0.0000
2	*TiNiSi*-type	62	a = 6.980 b = 4.032 c = 8.274	La 0.2296 0.2500 0.6108 S 0.0293 0.7500 0.8208 F 0.1305 0.7500 0.4467
3	*A*-LaFS	62	a = 6.998 b = 4.033 c = 16.445	La 0.2147 0.7500 0.3174 La 0.2243 0.2500 0.0661 S 0.0086 0.2500 0.2208 S 0.1458 0.2500 0.4399 F 0.0039 0.2500 0.9396 F 0.1581 0.2500 0.6438

4	*B*-LaFS	25	a = 4.040 b = 16.307 c = 7.023	La 0.0000 0.1259 0.0059 La 0.5000 0.2483 0.4970 La 0.5000 0.3756 0.9964 La 0.0000 0.5000 0.4979 La 0.5000 0.0000 0.5488 S 0.0000 0.1230 0.4226 S 0.0000 0.2815 0.7902 S 0.0000 0.3429 0.2889 S 0.0000 0.5000 0.9164 S 0.5000 0.0000 0.1258 F 0.5000 0.1237 0.7726 F 0.5000 0.2089 0.1202 F 0.5000 0.4137 0.6157 F 0.0000 0.0000 0.7755 F 0.5000 0.5000 0.2673
5	*C*-LaFS	38	a = 4.048 b = 16.344 c = 14.094	La 0.0000 0.1297 0.7331 La 0.5000 0.1250 0.2410 La 0.5000 0.2316 0.9769 La 0.0000 0.0000 0.4872 La 0.0000 0.0000 0.9886 S 0.0000 0.1621 0.3824 S 0.0000 0.2129 0.1301 S 0.5000 0.1155 0.5788 S 0.0000 0.0000 0.1982 S 0.5000 0.0000 0.8289 F 0.0000 0.1427 0.9166 F 0.5000 0.0844 0.0387 F 0.5000 0.2062 0.7944 F 0.0000 0.0000 0.6609 F 0.5000 0.0000 0.3748
6	*ZrNiAl*-type	189	a = 7.040 c = 4.044	La 0.3333 0.6667 0.5000 La 0.0000 0.0000 0.0000 S 0.0000 0.5826 0.0000 F 0.0000 0.2316 0.5000
7	*D*-LaFS	87	a = 11.264 c = 3.963	La 0.0673 0.7286 0.0000 S 0.1187 0.1914 0.5000 F 0.0839 0.6029 0.5000
8	*E*-LaFS	109	a = 4.056 c = 42.808	La 0.0000 0.0000 0.3062 La 0.0000 0.0000 0.4697 La 0.0000 0.0000 0.8819 S 0.0000 0.0000 0.3999 S 0.0000 0.0000 0.7814 S 0.0000 0.0000 0.9493 F 0.0000 0.0000 0.0924 F 0.0000 0.0000 0.2515 F 0.0000 0.0000 0.5921
9	*F*-LaFS	59	a = 4.034 b = 4.069 c = 21.496	La 0.0000 0.0000 0.4250 La 0.0000 0.5000 0.0758 La 0.0000 0.5000 0.7494 S 0.0000 0.0000 0.6206 S 0.0000 0.5000 0.2124 S 0.0000 0.5000 0.8855 F 0.0000 0.0000 0.3169 F 0.0000 0.0000 0.9996 F 0.0000 0.5000 0.5008
10	*ZrBeSi*-type	194	a = 4.185 c = 7.811	La 0.3333 0.6667 0.7500 S 0.0000 0.0000 0.0000 F 0.3333 0.6667 0.2500

11	*G*-LaFS	25	a = 4.057 b = 4.066 c = 14.334	La 0.0000 0.5000 0.0370 La 0.5000 0.0000 0.2629 La 0.5000 0.0000 0.7753 La 0.5000 0.5000 0.5189 S 0.0000 0.0000 0.5950 S 0.0000 0.5000 0.3392 S 0.0000 0.5000 0.8378 S 0.5000 0.0000 0.9859 F 0.0000 0.0000 0.1576 F 0.5000 0.0000 0.4279 F 0.5000 0.5000 0.1584 F 0.5000 0.5000 0.6826
12	*H*-LaFS	26	a = 4.188 b = 11.255 c = 7.799	La 0.0000 0.0879 0.2717 La 0.0000 0.4386 0.3617 La 0.5000 0.2456 0.7881 S 0.0000 0.3132 0.0436 S 0.5000 0.0376 0.0109 S 0.5000 0.2533 0.4080 F 0.0000 0.1326 0.7219 F 0.0000 0.3848 0.6582 F 0.5000 0.4665 0.8191
13	*I*-LaFS	6	a = 8.320 b = 4.060 c = 21.533 β = 94.44	La 0.0697 0.0000 0.2372 La 0.1610 0.0000 0.5592 La 0.2171 0.5000 0.8807 La 0.2746 0.0000 0.0805 La 0.3537 0.5000 0.4051 La 0.4435 0.0000 0.7210 La 0.6018 0.5000 0.2415 La 0.6374 0.5000 0.5670 La 0.6793 0.0000 0.9087 La 0.8319 0.5000 0.0745 La 0.8515 0.0000 0.3976 La 0.9051 0.5000 0.7304 S 0.0475 0.5000 0.3317 S 0.1441 0.0000 0.7858 S 0.2121 0.5000 0.6587 S 0.2764 0.5000 0.1780 S 0.3596 0.0000 0.9588 S 0.4004 0.0000 0.3044 S 0.5047 0.5000 0.8218 S 0.5746 0.0000 0.4670 S 0.6077 0.0000 0.1306 S 0.8611 0.0000 0.6241 S 0.8901 0.5000 0.8737 S 0.9343 0.5000 0.4933 F 0.0021 0.0000 0.1216 F 0.1192 0.5000 0.0435 F 0.1964 0.0000 0.4386 F 0.3253 0.5000 0.5209 F 0.4681 0.0000 0.6031 F 0.6028 0.5000 0.6818 F 0.6607 0.5000 0.9817 F 0.6831 0.5000 0.3575 F 0.7462 0.0000 0.7566 F 0.7754 0.0000 0.2739 F 0.8537 0.0000 0.0121 F 0.8841 0.5000 0.1937
14	*J*-LaFS	62	a = 12.573 b = 4.281 c = 5.559	La 0.1435 0.2500 0.5477 S 0.1964 0.2500 0.0476 F 0.0472 0.7500 0.5852
15	*LiMgN*-type	62	a = 10.091 b = 4.305 c = 6.494	La 0.0925 0.7500 0.7602 S 0.1534 0.2500 0.4664 F 0.0581 0.2500 0.9338

16	*3R*-type	166	a = 4.299 c = 26.243	La -0.0000 -0.0000 0.5000 La 0.0000 0.0000 0.0000 S 0.0000 0.0000 0.1107 F 0.0000 0.0000 0.3471
17	*LiCuO*-type	82	a = 11.132 c = 5.702	La 0.0000 0.3011 0.5000 S 0.0000 0.3318 0.0000 F 0.1536 0.1536 0.5000

Table A8.2. Calculated unit-cell parameters and atomic positions for all polymorphs of LaFS computed using the HSE06 (hybrid) functional.

E_{RI}	Modification and Space Group	SG No.	Cell Parameters (Å / °)	Atom Positions
1	*PbFCl*-type	129	a = 4.034 c = 6.987	La 0.0000 0.5000 0.2311 S -0.0000 0.5000 0.6496 F -0.0000 -0.0000 0.0000
2	*TiNiSi*-type	62	a = 6.950 b = 4.016 c = 8.237	La 0.2288 0.7500 0.3895 S 0.0297 0.2500 0.1797 F 0.1299 0.2500 0.5526
3	*A*-LaFS	62	a = 6.964 b = 4.015 c = 16.386	La 0.2147 0.7500 0.6825 La 0.2239 0.2500 0.9339 S 0.0088 0.2500 0.7793 S 0.1463 0.2500 0.5601 F 0.0031 0.2500 0.0602 F 0.1582 0.2500 0.3563
4	*B*-LaFS	25	a = 4.024 b = 16.233 c = 6.992	La -0.0000 0.1261 0.0054 La 0.5000 0.2485 0.4971 La 0.5000 0.3756 0.9966 La -0.0000 0.5000 0.4985 La 0.5000 0.0000 0.5494 S 0.0000 0.1232 0.4227 S 0.0000 0.2817 0.7901 S 0.0000 0.3430 0.2887 S 0.0000 0.5000 0.9166 S 0.5000 0.0000 0.1252 F 0.5000 0.1238 0.7725 F 0.5000 0.2089 0.1196 F 0.5000 0.4135 0.6161 F 0.0000 0.0000 0.7750 F 0.5000 0.5000 0.2685
5	*C*-LaFS	38	a = 4.029 b = 10.737 α = 98.39	La -0.0000 0.6036 0.8626 La 0.5000 0.1159 0.3659 La 0.5000 0.2077 0.7453 La -0.0000 0.4874 0.4874 La -0.0000 0.9890 0.9890 S 0.0000 0.2204 0.5446 S 0.0000 0.3429 0.9171 S 0.5000 0.4634 0.6942 S -0.0000 0.1983 0.1983 S 0.5000 0.8287 0.8287 F 0.0000 0.0594 0.7742 F 0.5000 0.0007 0.5886 F 0.5000 0.1229 0.9543 F -0.0000 0.6609 0.6609 F 0.5000 0.3749 0.3749

7	*ZrNiAl*-type	189	a = 7.016 c = 4.022	La 0.3333 0.6667 0.5000 La -0.0000 -0.0000 -0.0000 S 0.0000 0.5821 -0.0000 F 0.0000 0.2311 0.5000
6	*D*-LaFS	87	a = 11.216 c = 3.949	La 0.0672 0.2717 -0.0000 S 0.1188 0.8082 0.5000 F 0.0839 0.3970 0.5000
9	*E*-LaFS	109	a = 4.039 c = 42.545	La 0.0000 0.0000 0.2803 La 0.0000 -0.0000 0.4438 La -0.0000 -0.0000 0.8682 S 0.0000 0.0000 0.3502 S -0.0000 -0.0000 0.8008 S -0.0000 0.0000 0.9685 F 0.0000 0.0000 0.1579 F -0.0000 0.0000 0.4985 F 0.0000 0.0000 0.6576
10	*F*-LaFS	59	a = 4.015 b = 4.053 c = 21.381	La 0.0000 0.0000 0.4251 La -0.0000 0.5000 0.0756 La -0.0000 0.5000 0.7492 S 0.0000 0.0000 0.6206 S 0.0000 0.5000 0.2125 S 0.0000 0.5000 0.8856 F -0.0000 -0.0000 0.3169 F -0.0000 -0.0000 0.9996 F -0.0000 0.5000 0.5009
8	*ZrBeSi*-type	194	a = 4.159 c = 7.796	La 0.3333 0.6667 0.7500 S -0.0000 -0.0000 0.0000 F 0.3333 0.6667 0.2500
11	*G*-LaFS	25	a = 4.042 b = 4.046 c = 14.257	La -0.0000 0.5000 0.0372 La 0.5000 0.0000 0.2629 La 0.5000 0.0000 0.7751 La 0.5000 0.5000 0.5188 S 0.0000 0.0000 0.5950 S 0.0000 0.5000 0.3391 S -0.0000 0.5000 0.8378 S 0.5000 -0.0000 0.9860 F -0.0000 -0.0000 0.1576 F 0.5000 0.0000 0.4278 F 0.5000 0.5000 0.1585 F 0.5000 0.5000 0.6825
12	*H*-LaFS	26	a = 4.170 b = 11.204 c = 7.775	La 0.0000 0.0875 0.7278 La 0.0000 0.4396 0.6378 La 0.5000 0.2461 0.2125 S 0.0000 0.3137 0.9568 S 0.5000 0.0372 0.9900 S 0.5000 0.2540 0.5915 F 0.0000 0.1327 0.2775 F 0.0000 0.3850 0.3412 F 0.5000 0.4660 0.1816
13	*I*-LaFS	6	a = 8.287 b = 4.041 c = 21.435 β = 94.44	La 0.0983 0.5000 0.2691 La 0.1478 -0.0000 0.6025 La 0.1644 0.5000 0.9261 La 0.3221 -0.0000 0.0911 La 0.3654 0.5000 0.4326 La 0.3954 0.5000 0.7585 La 0.5585 -0.0000 0.2790 La 0.6466 0.5000 0.5948 La 0.7225 -0.0000 0.9193 La 0.7847 0.5000 0.1193 La 0.8403 0.0000 0.4407 La 0.9294 -0.0000 0.7633 S 0.0664 0.5000 0.5064

				S 0.1116 0.5000 0.1264 S 0.1418 0.0000 0.3754 S 0.3886 -0.0000 0.8694 S 0.4260 0.0000 0.5331 S 0.4970 0.5000 0.1780 S 0.5979 0.0000 0.6957 S 0.6407 0.0000 0.0412 S 0.7218 0.5000 0.8223 S 0.7898 0.5000 0.3411 S 0.8584 0.0000 0.2140 S 0.9521 0.5000 0.6685 F 0.1123 0.5000 0.8063 F 0.1450 0.0000 0.9882 F 0.2222 0.0000 0.7263 F 0.2581 0.0000 0.2435 F 0.3138 0.5000 0.6431 F 0.3373 0.5000 0.0183 F 0.4002 0.5000 0.3181 F 0.5341 0.0000 0.3969 F 0.6750 0.5000 0.4788 F 0.8033 0.0000 0.5609 F 0.8790 0.5000 0.9562 F 0.9955 -0.0000 0.8791
15	*J*-LaFS	62	a = 12.495 b = 4.266 c = 5.552	La 0.1432 0.2500 0.5478 S 0.1961 0.2500 0.0476 F 0.0467 0.7500 0.5835
14	*LiMgN*-type	62	a = 10.089 b = 4.284 c = 6.462	La 0.0922 0.7500 0.7613 S 0.1537 0.2500 0.4665 F 0.0574 0.2500 0.9341
16	*3R*-type	166	a = 9.104 α = 27.15	La 0.5000 0.5000 0.5000 La 1.0000 0.0000 0.0000 S 0.1104 0.1104 0.1104 F 0.3469 0.3469 0.3469
17	*LiCuO*-type	82	a = 11.089 c = 5.689	La 0.0061 0.7962 0.2501 S 0.0045 0.8314 0.7500 F 0.1482 0.3455 0.2480

The GGA-PBE and HSE06 band structures in Figures A8.1–A8.11 show the remaining LaFS polymorphs whose full electronic-structure plots are not presented in the main text. These figures confirm that all predicted LaFS polymorphs are semiconducting and that HSE06 systematically widens the band gaps while preserving the overall semiconducting character. In all cases, the upper valence region remains dominated mainly by sulfur states and the lower conduction region by lanthanum states, whereas polymorphism changes the size of the gap and the *k*-space positions of the valence-band maximum and conduction-band minimum.

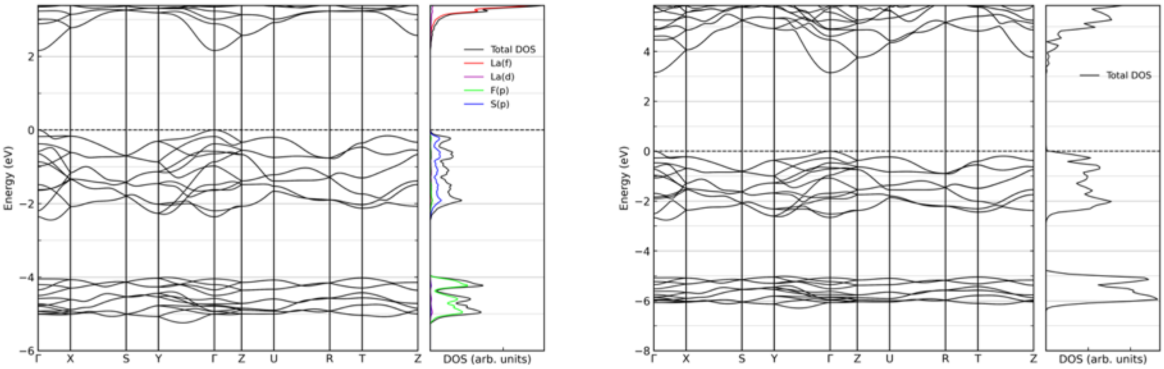

Figure A8.1. Electronic band structure and projected density of states for the *TiNiSi*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

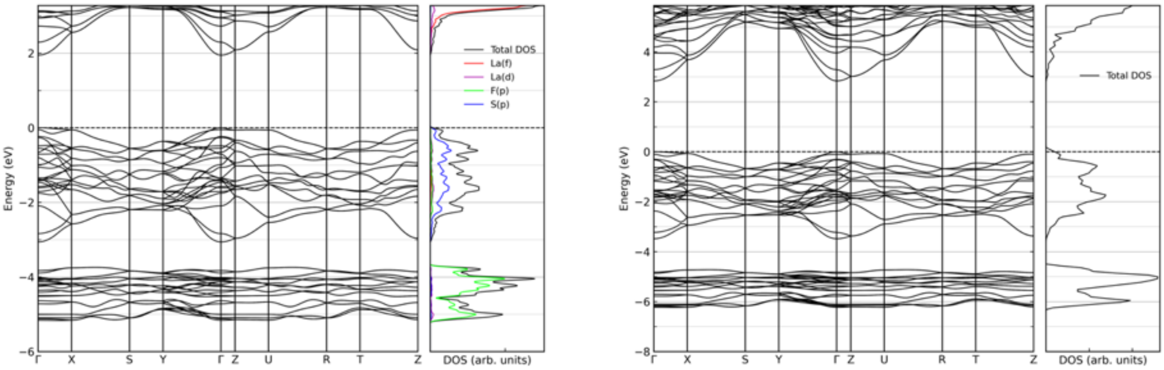

Figure A8.2. Electronic band structure and projected density of states for the *A*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

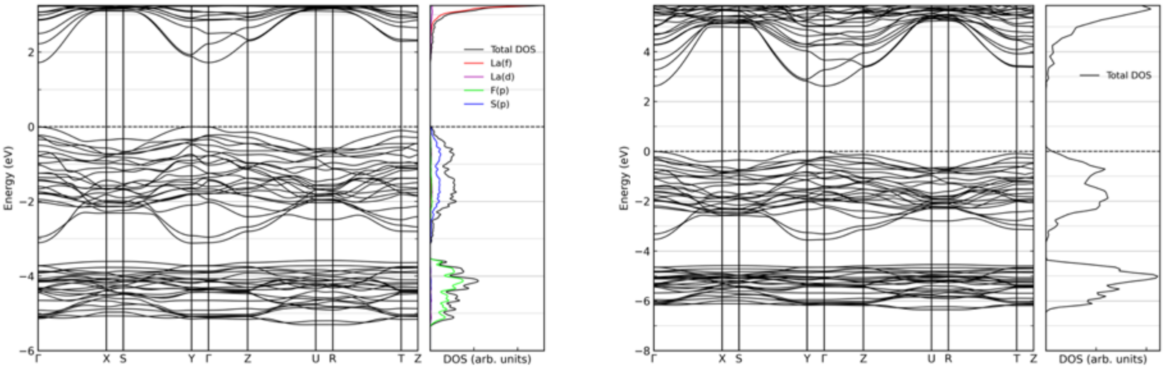

Figure A8.3. Electronic band structure and projected density of states for the *B*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

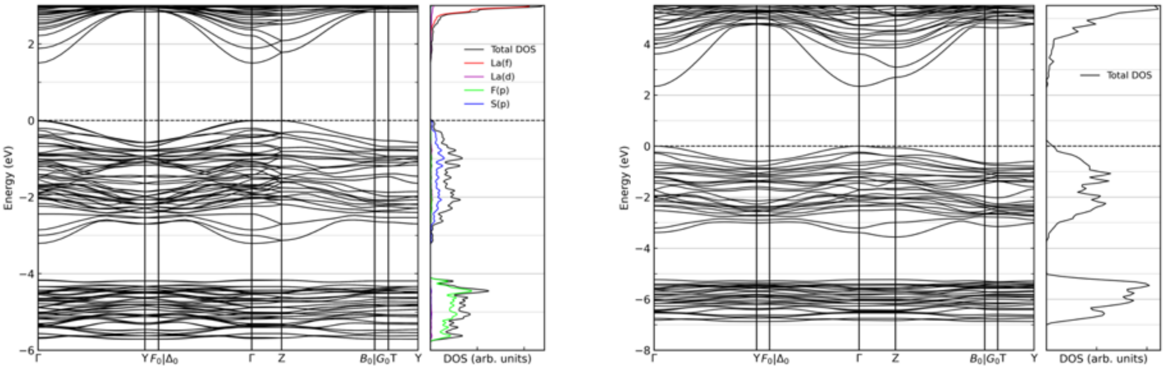

Figure A8.4. Electronic band structure and projected density of states for the *C*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

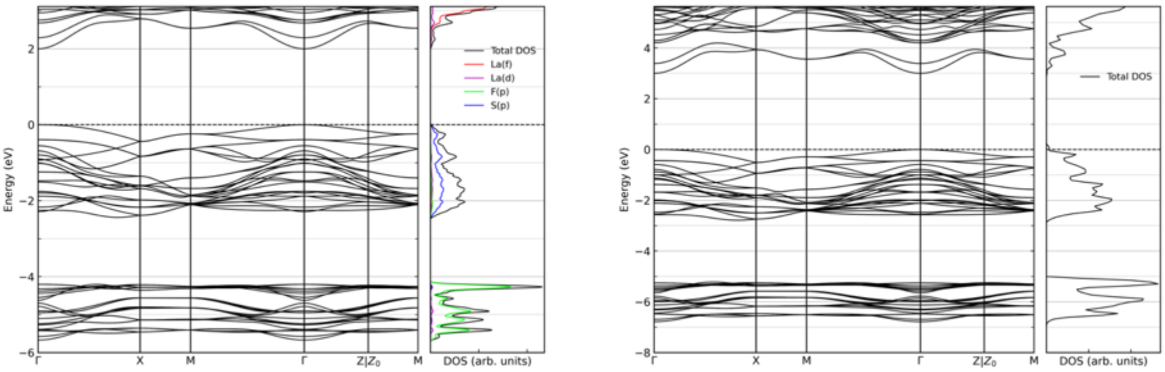

Figure A8.5. Electronic band structure and projected density of states for the *D*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

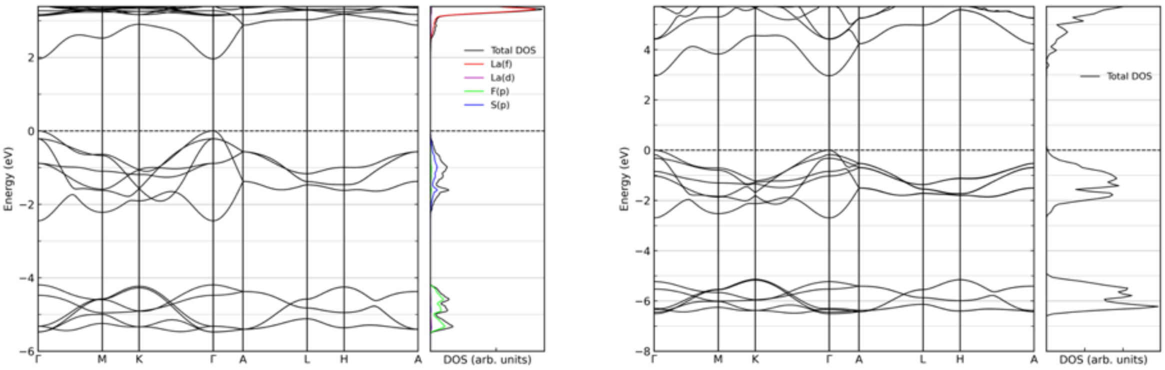

Figure A8.6. Electronic band structure and projected density of states for the *ZrBeSi*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

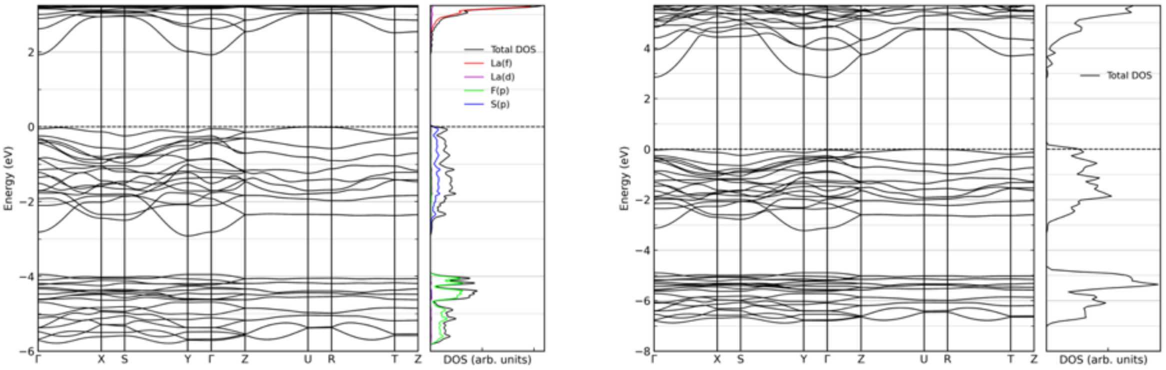

Figure A8.7. Electronic band structure and projected density of states for the *H*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

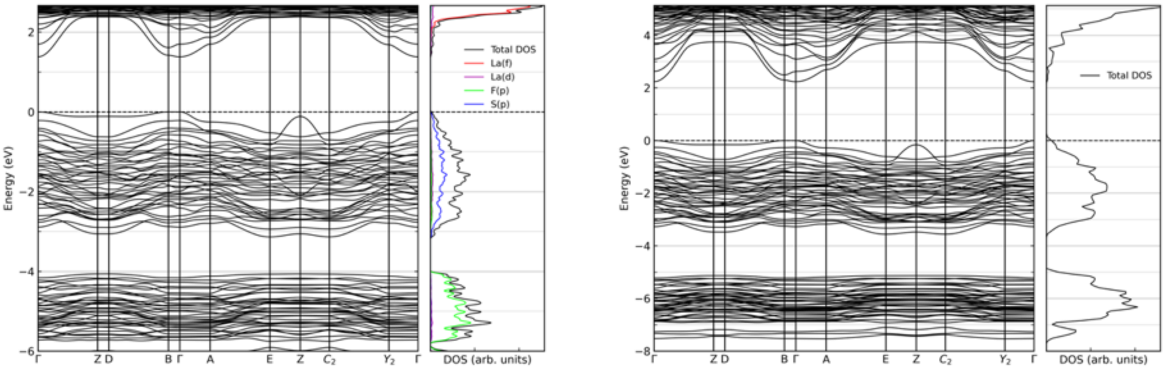

Figure A8.8. Electronic band structure and projected density of states for the *I*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

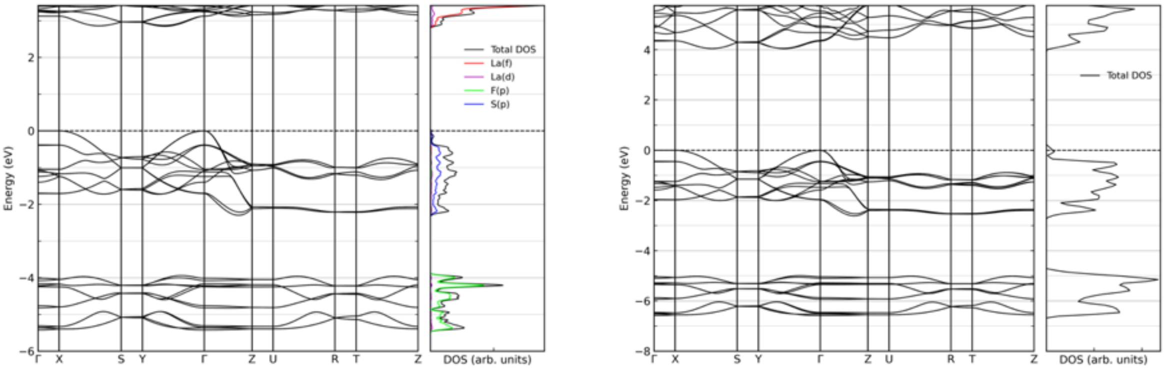

Figure A8.9. Electronic band structure and projected density of states for the *J*-LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

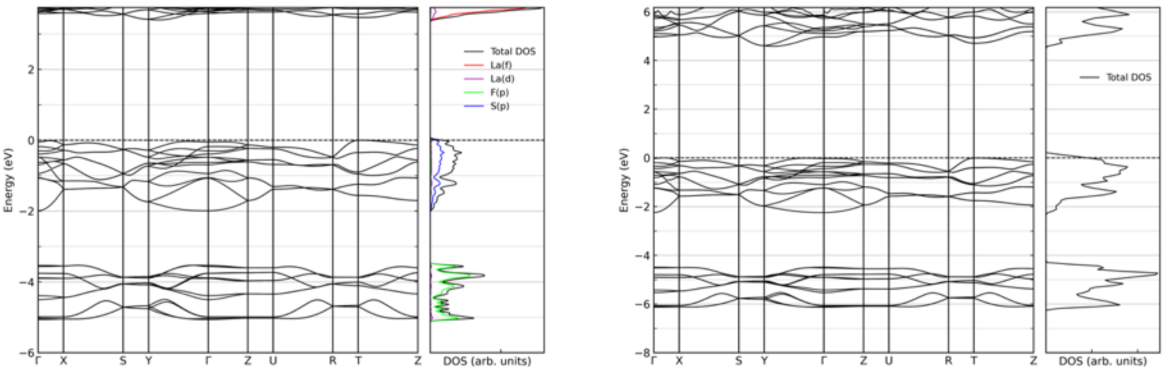

Figure A8.10. Electronic band structure and projected density of states for the *LiMgN*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

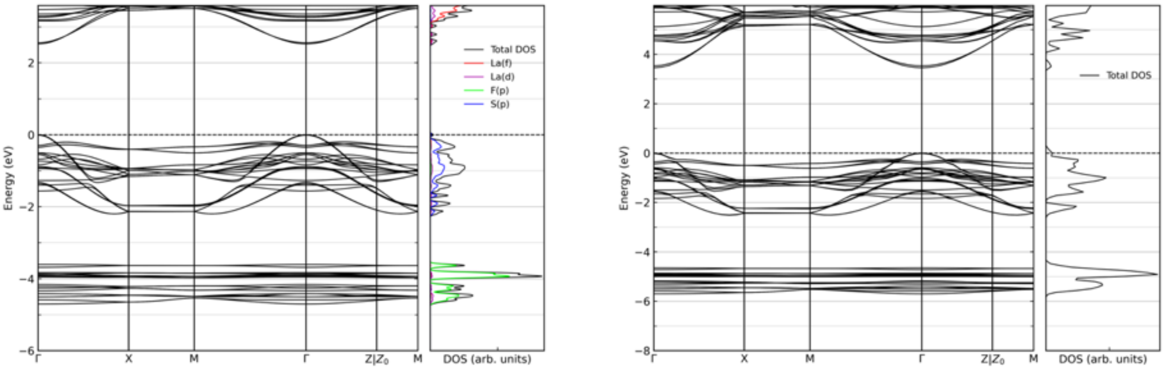

Figure A8.11. Electronic band structure and projected density of states for the *LiCuO*-type LaFS phase calculated using the GGA-PBE functional (left) and HSE06 functional (right).

References

[1] V. Balaram, Rare earth elements: A review of applications, occurrence, exploration, analysis, recycling, and environmental impact, Geoscience Frontiers 10(4) (2019) 1285-1303.

[2] B.J. Beaudry, K.A. Gschneidner, Preparation and basic properties of the rare earth metals in: K.A. Gschneidner, Jr., L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam, 1978, pp. 173–232.

[3] P. Thyssen, K. Binnemans, Accommodation of the Rare Earths in the Periodic Table: A Historical Analysis, Handbook on the Physics and Chemistry of Rare Earths, Elsevier2011, pp. 1-93.

[4] J.E. Powell, Separation chemistry, in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths1979, pp. 81-109.

[5] S. Cotton, Lanthanide and Actinide Chemistry, Wiley, Hoboken, 2024.

[6] H. Kageyama, K. Hayashi, K. Maeda, J.P. Attfield, Z. Hiroi, J.M. Rondinelli, K.R. Poeppelmeier, Expanding frontiers in materials chemistry and physics with multiple anions, Nat Commun 9(1) (2018) 772.

[7] M. Orr, G.R. Hebberd, E.E. McCabe, R.T. Macaluso, Structural Diversity of Rare-Earth Oxychalcogenides, ACS Omega 7(10) (2022) 8209-8218.

[8] M. Pejić, D. Zagorac, J. Zagorac, B. Matović, J.C. Schön, Structure prediction via global energy landscape exploration of the ternary rare-earth compound LaOI, Zeitschrift für anorganische und allgemeine Chemie 648(24) (2022) e202200308.

[9] D. Zagorac, J. Zagorac, M. Fonović, M. Pejić, J.C. Schön, Computational discovery of new modifications in scandium oxychloride (ScOCl) using a multi-methodological approach, Zeitschrift für anorganische und allgemeine Chemie 648(23) (2022) e202200198.

[10] J. Zagorac, J.C. Schön, B. Matović, M. Pejić, M. Prekajski Đorđević, D. Zagorac, Computational Discovery of New Feasible Crystal Structures in Ce3O3N, Crystals 13(5) (2023) 774.

[11] M. Pejić, D.D. Zimmermann, D. Zagorac, M. Fonović, J. Zagorac, J.C. Schön, T. Schleid, Structural exploration of holmium fluoride selenide (HoFSe): theory and experiment, Journal of Physics and Chemistry of Solids 208 (2026) 113000.

[12] M. Pejić, D. Zagorac, M. Fonović, J. Zagorac, J.C. Schön, Energy Landscape, Crystal Structures, and Electronic and Vibrational Properties of Lanthanum Fluoride Sulfide (LaFS) [Manuscript submitted for publication to Computational Condensed Matter], 2026.

[13] J. Schön, Nanomaterials - what energy landscapes can tell us, Processing and Application of Ceramics 9(3) (2015) 157-168.

[14] D.J. Wales, T.V. Bogdan, Potential energy and free energy landscapes, J Phys Chem B 110(42) (2006) 20765-76.

[15] Z.B. Goldschmidt, Atomic properties (free atom), in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam 1978, pp. 1-171.

[16] J.J. Rhyne, Highlights from the exotic phenomena of lanthanide magnetism, in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, Elsevier, Amsterdam 1988, pp. 293-322.

[17] K.A. McEwen, Magnetic and transport properties of the rare earths, in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam 1978, pp. 411-488.

[18] Hans R. Kirchmayr, Carl A. Poldy, R. Groessinger, R. Haferl, G. Hilscher, W. Steiner, G. Wiesinger, Magnetic properties of intermetallic compounds of rare earth metals, in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam, 1979, pp. 55–230.

[19] L. Eyring, The Binary Rare Earth Oxides, in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam 1979, pp. 337-399.

[20] G. Meyer, M.S. Wickleder, Simple and complex halides, in: K.A. Gschneidner, Jr., L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, Elsevier, Amsterdam 2000, pp. 53-129.

[21] A.D. Chervonnyi, Thermodynamic Properties of Lanthanide Fluorides and Chlorides in the Gaseous and Condensed States, in: J.-C.G. Bünzli, V.K. Pecharsky (Eds.), Handbook on the Physics and Chemistry of Rare Earths, Elsevier, Amsterdam 2012, pp. 165–484.

[22] F. Hulliger, Rare Earth Pnictides, in: K.A.G. Jr., L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam, 1979, pp. 153-236.

[23] J. Flahaut, Sulfides, selenides and tellurides, in: K.A. Gschneidner, Jr., L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam 1979, pp. 1-88.

[24] P. Wachter, Europium chalcogenides: EuO, EuS, EuSe and EuTe, in: J. Karl A. Gschneidner, L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam, 1979, pp. 507-574.

[25] A.A. Eliseev, G.M. Kuzmichyeva, Phase Equilibrium and Crystal Chemistry in Rare Earth Ternary Systems with Chalcogenide Elements, in: K.A. Gschneidner, Jr., L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, Elsevier, Amsterdam 1990, pp. 191–281.

[26] L.D. Gulay, M. Daszkiewicz, Ternary and Quaternary Chalcogenides of Si, Ge, Sn, Pb, and In, in: K.A. Gschneidner, Jr., J.-C.G. Bünzli, V.K. Pecharsky (Eds.), Handbook on The Physics and Chemistry of Rare Earths, Elsevier, Amsterdam 2011, pp. 157-273.

[27] P.O. Andreev, O.G. Mikhalkina, O.V. Andreev, A.V. Elyshev, Enthalpies of melting of LnSF compounds (Ln = La, Ce, Pr, Nd, Sm), Russian Journal of Physical Chemistry A 89(5) (2015) 731-736.

[28] F. Goubin, X. Rocquefelte, D. Pauwels, A. Tressaud, A. Demourgues, S. Jobic, Y. Montardi, The dielectric function of LnSF rare-earth fluorosulfides (Ln=La, Ce): experiment and theory, Journal of Solid State Chemistry 177(8) (2004) 2833-2840.

[29] T. Schleid, Drei Formen von Selten-Erd(III)-Fluoridsulfiden: A-LaFS, B-YFS und C-LuFS, Zeitschrift für anorganische und allgemeine Chemie 625(10) (1999) 1700-1706.

[30] C. Buyer, H. Grossholz, S. Wolf, D. Zagorac, J. Zagorac, J.C. Schön, T. Schleid, Crystal-Structure Prediction and Experimental Investigation of the Polymorphic Lanthanum Fluoride Selenides LaFSe and La2F4Se, Crystal Growth & Design 22(12) (2022) 7133-7142.

[31] D. Zagorac, C. Buyer, J. Zagorac, T. Škundrić, J.C. Schön, T. Schleid, Band-Gap Engineering and Unusual Behavior of Electronic Properties during Anion Substitution of Sulfur in LaFSe, Crystal Growth & Design 24(4) (2024) 1648-1657.

[32] I.U. Y. Uwatoko, M. Ohashi, T. Nakano, G. Oomi,, Thermal and Electronic Properties of Rare Earth Compounds at High Pressure, in: Jean-Claude G. Bünzli, Vitalij K. Pecharsky (Eds.), Handbook on the Physics and Chemistry of Rare Earths, Elsevier, Amsterdam, 2012, pp. 1-164.

[33] J. Elliott, P.G. Dawber, Symmetry in Physics Vol. 1, Macmillan Press, London, 1979.

[34] J. Elliott, P.G. Dawber, Symmetry in Physics Vol. 2, Macmillan Press, London, 1979.

[35] M. Damnjanović, Symmetry in Quantum Nonrelativistic Physics Faculty of Physics, University of Belgrade, Belgrade, 2014.

[36] W. Ludwig, C. Falter, Symmetries in Physics: Group Theory Applied to Physical Problems, 2nd ed., Springer, Berlin, 1996.

[37] R.M. Martin, Electronic Structure: Basic Theory and Practical Methods, 2nd ed., Cambridge University Press, Cambridge, 2020.

[38] A. Togo, L. Chaput, T. Tadano, I. Tanaka, Implementation strategies in phonopy and phono3py, J Phys Condens Matter 35(35) (2023) 353001.

[39] E. Kaxiras, Atomic and Electronic Structure of Solids, Cambridge University Press, Cambridge, 2003.

[40] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Physical Review 136(3B) (1964) B864-B871.

[41] N.D. Mermin, Thermal Properties of the Inhomogeneous Electron Gas, Physical Review 137(5A) (1965) A1441-A1443.

[42] M. Levy, Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem, Proc Natl Acad Sci U S A 76(12) (1979) 6062-5.

[43] E.H. Lieb, Density functionals for coulomb systems, International Journal of Quantum Chemistry 24(3) (1983) 243-277.

[44] W. Kohn, L.J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, Physical Review 140(4A) (1965) A1133-A1138.

[45] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, J.D. Joannopoulos, Iterative minimization techniques forab initiototal-energy calculations: molecular dynamics and conjugate gradients, Reviews of Modern Physics 64(4) (1992) 1045-1097.

[46] R.O. Jones, O. Gunnarsson, The density functional formalism, its applications and prospects, Reviews of Modern Physics 61(3) (1989) 689-746.

[47] D.M. Ceperley, B.J. Alder, Ground State of the Electron Gas by a Stochastic Method, Physical Review Letters 45(7) (1980) 566-569.

[48] J.P. Perdew, A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Physical Review B 23(10) (1981) 5048-5079.

[49] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys Rev Lett 77(18) (1996) 3865-3868.

[50] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys Rev A Gen Phys 38(6) (1988) 3098-3100.

[51] R. Dovesi, A. Erba, R. Orlando, C.M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro, B. Kirtman, Quantum-mechanical condensed matter simulations with CRYSTAL, WIREs Computational Molecular Science 8(4) (2018).

[52] R. Dovesi, F. Pascale, B. Civalleri, K. Doll, N.M. Harrison, I. Bush, P. D'Arco, Y. Noel, M. Rerat, P. Carbonniere, M. Causa, S. Salustro, V. Lacivita, B. Kirtman, A.M. Ferrari, F.S. Gentile, J. Baima, M. Ferrero, R. Demichelis, M. De La Pierre, The CRYSTAL code, 1976-2020 and beyond, a long story, J Chem Phys 152(20) (2020) 204111.

[53] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Physical Review B 59(3) (1999) 1758-1775.

[54] G. Kresse, J. Furthmuller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys Rev B Condens Matter 54(16) (1996) 11169-11186.

[55] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Computational Materials Science 6(1) (1996) 15-50.

[56] D.R. Hamann, M. Schlüter, C. Chiang, Norm-Conserving Pseudopotentials, Physical Review Letters 43(20) (1979) 1494-1497.

[57] N. Troullier, J.L. Martins, Efficient pseudopotentials for plane-wave calculations, Phys Rev B Condens Matter 43(3) (1991) 1993-2006.

[58] D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, Phys Rev B Condens Matter 41(11) (1990) 7892-7895.

[59] P.E. Blochl, Projector augmented-wave method, Phys Rev B Condens Matter 50(24) (1994) 17953-17979.

[60] L. Kleinman, D.M. Bylander, Efficacious Form for Model Pseudopotentials, Physical Review Letters 48(20) (1982) 1425-1428.

[61] R.J. Elliott, Magnetic Properties of Rare Earth Metals Springer New York, NY, 1972.

[62] P. Mohn, Magnetism in the Solid State: An Introduction, Springer, Berlin Heidelberg, 2006.

[63] R. Winkler, Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems, Springer, New York, 2003.

[64] P. Fulde, Crystal Fields, in: K.A.G. Jr., L. Eyring (Eds.), Handbook on the Physics and Chemistry of Rare Earths, North-Holland, Amsterdam, 1979, pp. 295-386.

[65] D. Hobbs, G. Kresse, J. Hafner, Fully unconstrained noncollinear magnetism within the projector augmented-wave method, Physical Review B 62(17) (2000) 11556-11570.

[66] S.L. Dudarev, G.A. Botton, S.Y. Savrasov, C.J. Humphreys, A.P. Sutton, Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study, Physical Review B 57(3) (1998) 1505-1509.

[67] B.R. Cooper, Phenomenological Theory of Magnetic Ordering: Importance of Interactions with the Crystal Lattice, in: R.J. Elliott (Ed.), Magnetic Properties of Rare Earth Metals, Springer, New York, 1972, pp. 17-80.

[68] S.Z.D. Cheng, Phase Transitions in Polymers: The Role of Metastable States, Elsevier2008.

[69] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, Cambridge, MA, 2016.

[70] F. Scarselli, M. Gori, A.C. Tsoi, M. Hagenbuchner, G. Monfardini, The graph neural network model, IEEE Trans Neural Netw 20(1) (2009) 61-80.

[71] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, P.S. Yu, A Comprehensive Survey on Graph Neural Networks, IEEE Trans Neural Netw Learn Syst 32(1) (2021) 4-24.

[72] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, The Bulletin of Mathematical Biophysics 5(4) (1943) 115-133.

[73] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, Psychol Rev 65(6) (1958) 386-408.

[74] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323(6088) (1986) 533-536.

[75] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86(11) (1998) 2278-2324.

[76] G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of Control, Signals, and Systems 2(4) (1989) 303-314.

[77] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2(5) (1989) 359-366.

[78] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Networks 4(2) (1991) 251-257.

[79] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, CoRR abs/1412.6980 (2014).

[80] J. Behler, M. Parrinello, Generalized neural-network representation of high-dimensional potential-energy surfaces, Phys Rev Lett 98(14) (2007) 146401.

[81] A.P. Bartok, M.C. Payne, R. Kondor, G. Csanyi, Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons, Phys Rev Lett 104(13) (2010) 136403.

[82] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, A.J. Smola, Deep Sets, Advances in Neural Information Processing Systems 30 (NIPS 2017) (2017).

[83] J. Gilmer, S.S. Schoenholz, P.F. Riley, O. Vinyals, G.E. Dahl, Neural Message Passing for Quantum Chemistry, in: D. Precup, Y.W. Teh (Eds.) Proceedings of the 34 th International Conference on Machine Learning, PMLR, Sydney, Australia, 2017, pp. 1263-1272.

[84] T. Kipf, M. Welling, Semi-Supervised Classification with Graph Convolutional Networks, International Conference on Learning Representations (2017).

[85] W. Hamilton, Z. Ying, J. Leskovec, Inductive Representation Learning on Large Graphs, in: I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems2017.

[86] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, Graph Attention Networks, International Conference on Learning Representations (2018).

[87] K. Xu, W. Hu, J. Leskovec, S. Jegelka, How powerful are graph neural networks?, International Conference on Learning Representations (2019).

[88] K. Schütt, P.-J. Kindermans, H.E. Sauceda, S. Chmiela, A. Tkatchenko, K.-R. Müller, SchNet: A continuous-filter convolutional neural network for modeling quantum interactions, Advances in Neural Information Processing Systems, 2017.

[89] T. Xie, J.C. Grossman, Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties, Phys Rev Lett 120(14) (2018) 145301.

[90] I.E. Lagaris, A. Likas, D.I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, IEEE Trans Neural Netw 9(5) (1998) 987-1000.

[91] V.I. Gorikhovskii, T.O. Evdokimova, V.A. Poletansky, Neural networks in solving differential equations, Journal of Physics: Conference Series 2308(1) (2022).

[92] D. Zagorac, J.C. Schön, J. Zagorac, M. Jansen, Theoretical investigations of novel zinc oxide polytypes and in-depth study of their electronic properties, RSC Advances 5(33) (2015) 25929-25935.

[93] S. Kirkpatrick, C.D. Gelatt, Jr., M.P. Vecchi, Optimization by simulated annealing, Science 220(4598) (1983) 671-80.

[94] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equation of State Calculations by Fast Computing Machines, The Journal of Chemical Physics 21(6) (1953) 1087-1092.

[95] J.C. Schön, G42+ manual, Max Planck Institute for Solid State Research, Stuttgart, 2015.

[96] R. Hundt, J.C. Scho¨n, A. Hannemann, M. Jansen, Determination of symmetries and idealized cell parameters for simulated structures, Journal of Applied Crystallography 32(3) (1999) 413-416.

[97] R. Hundt, J.C. Schön, M. Jansen, CMPZ– an algorithm for the efficient comparison of periodic structures, Journal of Applied Crystallography 39(1) (2006) 6-16.

[98] R. Hundt, KPLOT: A Program for Plotting and Analysing Crystal Structures, in: P.D.J.C. Schön, D.r.n.D. Zagorac (Eds.) Technicum Scientific Publishing, Stuttgart, Germany, 2016.

[99] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, From Data Mining to Knowledge Discovery in Databases, AI Magazine 17(3) (1996) 37-54.

[100] D. Zagorac, H. Muller, S. Ruehl, J. Zagorac, S. Rehme, Recent developments in the Inorganic Crystal Structure Database: theoretical crystal structure data and related features, J Appl Crystallogr 52(Pt 5) (2019) 918-925.

[101] T. Škundrić, D. Zagorac, J. Schön, M. Pejić, B. Matović, Crystal Structure Prediction of the Novel Cr2SiN4 Compound via Global Optimization, Data Mining, and the PCAE Method, Crystals 11(8) (2021).

[102] K. Doll, V.R. Saunders, N.M. Harrison, Analytical Hartree-Fock gradients for periodic systems, International Journal of Quantum Chemistry 82(1) (2001) 1-13.

[103] K. Doll, R. Dovesi, R. Orlando, Analytical Hartree Fock gradients with respect to the cell parameter for systems periodic in three dimensions, Theoretical Chemistry Accounts 112(5-6) (2004) 394-402.

[104] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V.R. Saunders, C.M. Zicovich-Wilson, CRYSTAL: a computational tool for the ab initio study of the electronic properties of crystals, Zeitschrift für Kristallographie - Crystalline Materials 220(5-6) (2005) 571-573.

[105] S.P. Ong, W.D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V.L. Chevrier, K.A. Persson, G. Ceder, Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis, Computational Materials Science 68 (2013) 314-319.

[106] A. Togo, K. Shinohara, I. Tanaka, Spglib: a software library for crystal symmetry search, Science and Technology of Advanced Materials: Methods 4(1) (2024).

[107] C. Chen, S.P. Ong, A universal graph deep learning interatomic potential for the periodic table, Nat Comput Sci 2(11) (2022) 718-728.

[108] T.W. Ko, B. Deng, M. Nassar, L. Barroso-Luque, R. Liu, J. Qi, A.C. Thakur, A.R. Mishra, E. Liu, G. Ceder, S. Miret, S.P. Ong, Materials Graph Library (MatGL), an open-source graph deep learning library for materials science and chemistry, npj Computational Materials 11(1) (2025).

[109] A. Hjorth Larsen, J. Jorgen Mortensen, J. Blomqvist, I.E. Castelli, R. Christensen, M. Dulak, J. Friis, M.N. Groves, B. Hammer, C. Hargus, E.D. Hermes, P.C. Jennings, P. Bjerre Jensen, J. Kermode, J.R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiotz, O. Schutt, M. Strange, K.S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K.W. Jacobsen, The atomic simulation environment-a Python library for working with atoms, J Phys Condens Matter 29(27) (2017) 273002.

[110] F. Birch, Finite Elastic Strain of Cubic Crystals, Physical Review 71(11) (1947) 809-824.

[111] F.D. Murnaghan, The Compressibility of Media under Extreme Pressures, Proc Natl Acad Sci U S A 30(9) (1944) 244-7.

[112] D. Zagorac, Energy landscape investigations of chemical matter and structure prediction of binary inorganic solids, University of Stuttgart, Stuttgart, Germany, 2012, p. 226.

[113] M. Udayakantha, P. Schofield, G.R. Waetzig, S. Banerjee, A full palette: Crystal chemistry, polymorphism, synthetic strategies, and functional applications of lanthanide oxyhalides, Journal of Solid State Chemistry 270 (2019) 569-592.

[114] L. Xu, Y. Xue, J. Xia, X. Qu, B. Lei, T. Yang, X. Zhang, N. Li, H. Zhao, M. Wang, M. Luo, C. Zhang, Y. Du, C. Yan, Construction of high quality ultrathin lanthanide oxyiodide nanosheets for enhanced CT imaging and anticancer drug delivery to efficient cancer theranostics, Biomaterials 230 (2020) 119670.

[115] N. Imanaka, M. Misran, N. Nunotani, Evidence for enormous iodide anion migration in lanthanum oxyiodide-based solid, Sci Adv 7(43) (2021) eabh0812.

[116] M. Udayakantha, J.V. Handy, R.D. Davidson, J. Kaur, G. Villalpando, L. Zuin, S. Chakraborty, S. Banerjee, Halide Replacement with Complete Preservation of Crystal Lattice in Mixed-Anion Lanthanide Oxyhalides, Angew Chem Int Ed Engl 60(28) (2021) 15582-15589.

[117] L.G. Sillen, A.L. Nylander, The crystal structure of La O Cl, La O Br, and La O I, Svensk Kemisk Tidskrift 53 (1941) 367-372.

[118] J. Flahaut, Les structures type PbFCl (EOI) et type anti-Fe2As (C38) des composés ternaires à deux anions MXY, Journal of Solid State Chemistry 9(2) (1974) 124-131.

[119] A. Hannemann, R. Hundt, J.C. Schön, M. Jansen, A New Algorithm for Space-Group Determination, Journal of Applied Crystallography 31(6) (1998) 922-928.

[120] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)], Physical Review Letters 78(7) (1997) 1396-1396.

[121] A.V. Krukau, O.A. Vydrov, A.F. Izmaylov, G.E. Scuseria, Influence of the exchange screening parameter on the performance of screened hybrid functionals, J Chem Phys 125(22) (2006) 224106.

[122] A. Chaudhry, A. Canning, R. Boutchko, Y.D. Porter-Chapman, E. Bourret-Courchesne, S.E. Derenzo, N. GrØnbech-Jensen, First Principles Calculations for Scintillation in Ce-Doped Y and La Oxyhalides, IEEE Transactions on Nuclear Science 56(3) (2009) 949-954.

[123] D. Kim, S. Park, S. Kim, S.G. Kang, J.C. Park, Blue-emitting Eu2+-activated LaOX (X = Cl, Br, and I) materials: crystal field effect, Inorg Chem 53(22) (2014) 11966-73.

[124] G. Shwetha, V. Kanchana, N. Yedukondalu, G. Vaitheeswaran, Ab initiostudy of scintillating lanthanide oxyhalide host materials, Materials Research Express 2(10) (2015).

[125] D. Zagorac, J. Zagorac, J.C. Schön, N. Stojanović, B. Matović, ZnO/ZnS (hetero)structures: ab initio investigations of polytypic behavior of mixed ZnO and ZnS compounds, Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials 74(6) (2018) 628-642.

[126] D.A. Odink, K. Song, S.M. Kauzlarich, Intercalation of pyridine into the layered samarium sulfide iodide (SmSI) structure of ytterbium oxide chloride (YbOCl), Chemistry of Materials 4(4) (1992) 906-911.

[127] J. Dabachi, M. Body, J. Dittmer, F. Fayon, C. Legein, Structural refinement of the RT LaOF phases by coupling powder X-Ray diffraction, (19)F and (139)La solid state NMR and DFT calculations of the NMR parameters, Dalton Trans 44(47) (2015) 20675-84.

[128] J. Hölsä, B. Piriou, M. Räsänen, IR- and Raman-active normal vibrations of rare earth oxyfluorides, REOF; RE=Y, La, and Gd, Spectrochimica Acta Part A: Molecular Spectroscopy 49(4) (1993) 465-470.

[129] M. Glätzle, M. Schauperl, C. Hejny, M. Tribus, K.R. Liedl, H. Huppertz, Orthorhombic HP-REOF (RE = Pr, Nd, Sm – Gd) – High-Pressure Syntheses and Single-Crystal Structures (RE = Nd, Sm, Eu), Zeitschrift für anorganische und allgemeine Chemie 642(20) (2016) 1134-1142.

[130] S. Wang, R. Deng, H. Guo, S. Song, F. Cao, X. Li, S. Su, H. Zhang, Lanthanide doped Y6O5F8/YF3 microcrystals: phase-tunable synthesis and bright white upconversion photoluminescence properties, Dalton Trans 39(38) (2010) 9153-8.

[131] T. Wen, W. Luo, Y. Wang, M. Zhang, Y. Guo, J. Yuan, J. Ju, Y. Wang, F. Liao, B. Yang, Multicolour and up-conversion fluorescence of lanthanide doped Vernier phase yttrium oxyfluoride nanocrystals, Journal of Materials Chemistry C 1(10) (2013) 1995-2001.

[132] S.E. Dutton, D. Hirai, R.J. Cava, Low temperature synthesis of LnOF rare-earth oxyfluorides through reaction of the oxides with PTFE, Materials Research Bulletin 47(3) (2012) 714-718.

[133] K.R. Kort, S. Banerjee, Shape-controlled synthesis of well-defined matlockite LnOCl (Ln: La, Ce, Gd, Dy) nanocrystals by a novel non-hydrolytic approach, Inorg Chem 50(12) (2011) 5539-44.

[134] T. Petzel, V. Marx, B. Hormann, Thermodynamics of the rhombohedral-cubic phase transition of ROF with R = Y, La, Pr, Nd, Sm-Er, Journal of Alloys and Compounds 200(1-2) (1993) 27-31.

[135] C.W.F.T. Pistorius, Effect of pressure on the rhombohedral/cubic transitions of some lanthanide oxide fluorides, Journal of the Less Common Metals 31(1) (1973) 119-124.

[136] J.H. Müller, T. Petzel, High-temperature X-ray diffraction study of the rhombohedral-cubic phase transition of ROF with R=Y, La, Pr, Nd, Sm-Er, Journal of Alloys and Compounds 224(1) (1995) 18-21.

[137] J. Heyd, G.E. Scuseria, M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential, The Journal of Chemical Physics 118(18) (2003) 8207-8215.

[138] A.B. Jorge, J. Fraxedas, A. Cantarero, A.J. Williams, J. Rodgers, J.P. Attfield, A. Fuertes, Nitrogen Doping of Ceria, Chemistry of Materials 20(5) (2008) 1682-1684.

[139] J.-S. Lee, M. Lerch, J. Maier, Nitrogen-doped zirconia: A comparison with cation stabilized zirconia, Journal of Solid State Chemistry 179(1) (2006) 270-277.

[140] G. Bergerhoff, I.D. Brown, Inorganic crystal structure database, Crystallographic Databases, International Union of Crystallography, Bonn/Cambridge/Chester, 1987, pp. 77-95.

[141] M. Kantardzic, Data Mining: Concepts, Models, Methods, and Algorithms, John Wiley & Sons, Inc., Hoboken, New Jersey, 2011.

[142] J. Graciani, A.M. Marquez, J.J. Plata, Y. Ortega, N.C. Hernandez, A. Meyer, C.M. Zicovich-Wilson, J.F. Sanz, Comparative Study on the Performance of Hybrid DFT Functionals in Highly Correlated Oxides: The Case of CeO2 and Ce2O3, J Chem Theory Comput 7(1) (2011) 56-65.

[143] A. Sharan, S. Lany, Computational discovery of stable and metastable ternary oxynitrides, J Chem Phys 154(23) (2021) 234706.

[144] A.O. Tsokol, O.I. Bodak, E.P. Marusin, M.G. Baivelman, Crystal structure of the compound $ScAl_3C_3$, Soviet Physics Crystallography 31 (1986) 467–468.

[145] P. Höhn, R. Kniep, A. Rabenau, Ba3[FeN3]: Ein neues Nitridoferrat(III) mit [CO3]2−-isosteren Anionen [FeN3]6−, Zeitschrift für Kristallographie 196(1-4) (1991) 153-158.

[146] C. Hamon, R. Marchand, Y. Laurent, J. Lang, Étude d'halogénopnictures. III. Structure de Ca2PI et Ca3PI3. Surstructures de type NaCl, Bulletin de la Société française de Minéralogie et de Cristallographie 97(1) (1974) 6-12.

[147] G. Cordier, H. Schäfer, M. Stelter, Neue Zintlphasen: Ba3GaSb3, Ca3GaAs3 und Ca3InP3 / On New Zintlphases: Ba3GaSb3, Ca3GaAs3 und Ca3InP3, Zeitschrift für Naturforschung B 40(9) (1985) 1100-1104.

[148] F. Machatschki, Präzisionsmessungen der Gitterkonstanten verschiedener Fahlerze. Formel und Struktur derselben, Zeitschrift für Kristallographie 68(1-6) (1928) 204–222.

[149] K. Klepp, H. Boller, Die Kristallstruktur von TlFe3Te3, Monatshefte für Chemie 110(3) (1979) 677-684.

[150] C.B. Pollock, H.H. Stadelmaier, The eta carbides in the Fe−W−C and Co−W−C systems, Metallurgical Transactions 1(4) (1970) 767-770.

[151] M. Ebihara, J.D. Martin, J.D. Corbett, Novel Chain and Oligomeric Condensed Cluster Phases for Gadolinium Iodides with Manganese Interstitials, Inorganic Chemistry 33(10) (1994) 2079-2084.

[152] B. Eisenmann, A. Hofmann, Crystal structure of hexapotassium di-μ-selenidobis(diselenidoaluminate), K6Al2Se6, Zeitschrift für Kristallographie - Crystalline Materials 197(1-4) (1991) 173-174.

[153] J. Kuchinke, C. Jansen, A. Lindemann, B. Krebs, Syntheses and Crystal Structures of the Novel Ternary Thioborates Na3BS3, K3BS3, and Rb3BS3, Zeitschrift für anorganische und allgemeine Chemie 627(5) (2001) 896-902.

[154] G. Dittmar, Die KristallStrukturen von K6[Ge2Te6] und K6[Sn2Te6] und ihre kristall-chemische Beziehung zum K6[Si2Te6]-Typ, Zeitschrift für anorganische und allgemeine Chemie 453(1) (1978) 68-78.

[155] I.Y. Kuznetsov, D.M. Vinnitskij, K.A. Solntsev, N.T. Kuznetsov, L.A. Butman, Crystal structure of $K_2B_6H_6$ and $Cs_2B_6H_6$, Zhurnal Neorganicheskoi Khimii 32(12) (1987) 3112-3114.

[156] M. Palazzi, Structure cristalline de l'orthotrithioarsénite trisodique Na3AsS3, Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry 32(12) (1976) 3175-3177.

[157] O.Y. Mruz, V.K. Pecharskii, A.N. Sobolev, O.I. Bodak, Crystal structure of SmNi3Ge3, Soviet Physics - Crystallography 35(1) (1990) 122-123.

[158] B.Y. Kotur, E.I. Gladyshevskii, Crystal structure of scandium-nickel silicide (Sc_3NiSi_3), Kristallografiya 28 (1983) 461-464.

[159] D. Harker, The Application of the Three-Dimensional Patterson Method and the Crystal Structures of Proustite, Ag3AsS3, and Pyrargyrite, Ag3SbS3, The Journal of Chemical Physics 4(6) (1936) 381-390.

[160] C.H. Wei, Structural analyses of tetracobalt dodecacarbonyl and tetrarhodium dodecacarbonyl. Crystallographic treatments of a disordered structure and a twinned composite, Inorganic Chemistry 8(11) (1969) 2384-2397.

[161] H.Y.P. Hong, J.C. Mikkelsen, G.W. Roland, Crystal structure of Tl3AsSe3, Materials Research Bulletin 9(4) (1974) 365-369.

[162] P. Engel, W. Nowacki, Die Kristallstruktur von Xanthokon, Ag3AsS3, Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry 24(1) (1968) 77-81.

[163] D. Fischer, B. Andriyevsky, J.C. Schön, Systematics of the allotrope formation in elemental gallium films, Materials Research Express 6(11) (2019).

[164] N. Kim, C.-H. Hsieh, J.F. Stebbins, Scandium Coordination in Solid Oxides and Stabilized Zirconia: 45Sc NMR, Chemistry of Materials 18(16) (2006) 3855-3859.

[165] G. Caimi, L. Degiorgi, N.N. Kovaleva, P. Lemmens, F.C. Chou, Infrared optical properties of the spin-1/2 quantum magnet TiOCl, Physical Review B 69(12) (2004).

[166] A.C. Komarek, T. Taetz, M.T. Fernández-Díaz, D.M. Trots, A. Möller, M. Braden, Strong magnetoelastic coupling in VOCl: Neutron and synchrotron powder x-ray diffraction study, Physical Review B 79(10) (2009).

[167] L.H. Brixner, J.F. Ackerman, C.M. Foris, Rare earth fluorescence in some LnOCl type hosts, Journal of Luminescence 26(1-2) (1981) 1-19.

[168] P. Javadian, T.K. Nielsen, D.B. Ravnsbæk, L.H. Jepsen, M. Polanski, T. Plocinski, I. Kunce, F. Besenbacher, J. Bystrzycki, T.R. Jensen, Scandium functionalized carbon aerogel: Synthesis of nanoparticles and structure of a new ScOCl and properties of NaAlH4 as a function of pore size, Journal of Solid State Chemistry 231 (2015) 190-197.

[169] W. Li, N. Miao, J. Zhou, Z. Sun, Pressure-mediated structural phase transitions and ultrawide indirect–direct bandgaps in novel rare-earth oxyhalides, Journal of Materials Chemistry C 9(2) (2021) 547-554.

[170] R. Allmann, R. Hinek, The introduction of structure types into the Inorganic Crystal Structure Database ICSD, Acta Crystallogr A 63(Pt 5) (2007) 412-7.

[171] F.A. Bannister, The crystal-structure and optical properties of matlockite (PbFCl), Mineralogical Magazine and Journal of the Mineralogical Society 23(146) (1934) 587-597.

[172] L.H. Brixner, E.P. Moore, Single-crystal refinement of the structure of LaOCl, Acta Crystallographica Section C Crystal Structure Communications 39(9) (1983) 1316-1316.

[173] H. Zhang, J.F. Banfield, Understanding Polymorphic Phase Transformation Behavior during Growth of Nanocrystalline Aggregates: Insights from TiO2, The Journal of Physical Chemistry B 104(15) (2000) 3481-3487.

[174] D. Jovanović, D. Zagorac, B. Matović, A. Zarubica, J. Zagorac, Anion substitution and influence of sulfur on the crystal structures, phase transitions, and electronic properties of mixed TiO2/TiS2 compounds, Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials 77(5) (2021) 833-847.

[175] M.A.C. Wevers, J.C. Schön, M. Jansen, Global aspects of the energy landscape of metastable crystal structures in ionic compounds, Journal of Physics: Condensed Matter 11(33) (1999) 6487-6499.

[176] P.N. Kumta, S.H. Risbud, Rare-earth chalcogenides ? an emerging class of optical materials, Journal of Materials Science 29(5) (1994) 1135-1158.

[177] H.Q. Wang, M. Batentschuk, A. Osvet, L. Pinna, C.J. Brabec, Rare-earth ion doped up-conversion materials for photovoltaic applications, Adv Mater 23(22-23) (2011) 2675-80.

[178] S. Kück, Laser-related spectroscopy of ion-doped crystals for tunable solid-state lasers, Applied Physics B 72(5) (2001) 515-562.

[179] P.K. Siwach, H.K. Singh, O.N. Srivastava, Low field magnetotransport in manganites, J Phys Condens Matter 20(27) (2008) 273201.

[180] L.D. Carlos, R.A. Ferreira, Z. Bermudez Vde, S.J. Ribeiro, Lanthanide-containing light-emitting organic-inorganic hybrids: a bet on the future, Adv Mater 21(5) (2009) 509-34.

[181] G.Y. Chen, Y. Liu, Y.G. Zhang, G. Somesfalean, Z.G. Zhang, Q. Sun, F.P. Wang, Bright white upconversion luminescence in rare-earth-ion-doped Y2O3 nanocrystals, Applied Physics Letters 91(13) (2007).

[182] M.V. Ganduglia-Pirovano, A. Hofmann, J. Sauer, Oxygen vacancies in transition metal and rare earth oxides: Current state of understanding and remaining challenges, Surface Science Reports 62(6) (2007) 219-270.

[183] H.X. Mai, Y.W. Zhang, R. Si, Z.G. Yan, L.D. Sun, L.P. You, C.H. Yan, High-quality sodium rare-earth fluoride nanocrystals: controlled synthesis and optical properties, J Am Chem Soc 128(19) (2006) 6426-36.

[184] H. Takahashi, K. Igawa, K. Arii, Y. Kamihara, M. Hirano, H. Hosono, Superconductivity at 43 K in an iron-based layered compound LaO(1-x)F(x)FeAs, Nature 453(7193) (2008) 376-8.

[185] J.S. Sanghera, I.D. Aggarwal, Active and passive chalcogenide glass optical fibers for IR applications: a review, Journal of Non-Crystalline Solids 256-257 (1999) 6-16.

[186] B. Cole, L.B. Shaw, P.C. Pureza, R. Mossadegh, J.S. Sanghera, I.D. Aggarwal, Rare-earth doped selenide glasses and fibers for active applications in the near and mid-IR, Journal of Non-Crystalline Solids 256-257 (1999) 253-259.

[187] M.F. Churbanov, B.I. Denker, B.I. Galagan, V.V. Koltashev, V.G. Plotnichenko, M.V. Sukhanov, S.E. Sverchkov, A.P. Velmuzhov, Comparison of 4.5–6 μm luminescent and lasing properties of rare earth dopants in chalcogenide glasses, Journal of Luminescence 245 (2022).

[188] G. Chen, H. Qiu, P.N. Prasad, X. Chen, Upconversion nanoparticles: design, nanochemistry, and applications in theranostics, Chem Rev 114(10) (2014) 5161-214.

[189] S.D. Jackson, Single-transverse-mode 2.5-W holmium-doped fluoride fiber laser operating at 2.86 microm, Optics Letters 29(4) (2004) 334-6.

[190] M.R. Ganjali, P. Norouzi, M. Adib, A. Ahmadalinezhad, A Novel Holmium(III) Membrane Sensor Based on N-(1-Thien-2-Ylmethylene)-1,3-Benzothiazol-2-Amine, Analytical Letters 39(6) (2006) 1075-1086.

[191] F.C. Goerigk, T. Schleid, HoF(OH)2: a fluoride-containing holmium(III) hydroxide with UCl3-type crystal structure, Acta Crystallographica Section A Foundations and Advances 77(a2) (2021) C1021-C1021.

[192] K. Das, G.A. Kumar, L. Mirandola, M. Chiriva-Internati, J. Chaudhuri, Synthesis and characterization of lanthanide-doped sodium holmium fluoride nanoparticles for potential application in photothermal therapy, Frontiers of Materials Science 13(4) (2019) 389-398.

[193] M. Sultania, J.C. Schön, D. Fischer, M. Jansen, Investigation of structural relations among the compounds in the ICSD using geometry based comparison techniques, Structural Chemistry 23(4) (2012) 1121-1129.

[194] B. Civalleri, P. D'Arco, R. Orlando, V.R. Saunders, R. Dovesi, Hartree–Fock geometry optimisation of periodic systems with the Crystal code, Chemical Physics Letters 348(1-2) (2001) 131-138.

[195] K. Doll, Analytical stress tensor and pressure calculations with the CRYSTAL code, Molecular Physics 108(3-4) (2010) 223-227.

[196] M. Dolg, H. Stoll, A. Savin, H. Preuss, Energy-adjusted pseudopotentials for the rare earth elements, Theoretica Chimica Acta 75(3) (1989) 173-194.

[197] R. Nada, C.R.A. Catlow, C. Pisani, R. Orlando, An ab-initio Hartree-Fock perturbed-cluster study of neutral defects in LiF, Modelling and Simulation in Materials Science and Engineering 1(2) (1993) 165-187.

[198] A. Tressaud, J.L. Soubeyroux, H. Touhara, G. Demazeau, F. Langlais, On a new structural type of fluorine compounds : Crystal and magnetic structures of a high pressure form of PdF2, Materials Research Bulletin 16(2) (1981) 207-214.

[199] J.C. Schon, Z. Cancarevic, M. Jansen, Structure prediction of high-pressure phases for alkali metal sulfides, J Chem Phys 121(5) (2004) 2289-304.

[200] S. Blundell, Magnetism in Condensed Matter, Oxford University Press, Oxford, 2001.

[201] G.R. Choppin, E.N. Rizkalla, Solution Chemistry of actinides and lanthanides, Handbook on the Physics and Chemistry of Rare Earths Elsevier, Amsterdam1994, pp. 559-590.

[202] C. Buyer, H. Grossholz, T. Schleid, The Structural Dimorphism of Lanthanum Oxide Fluoride Selenide La2OF2Se, Crystals 9(9) (2019).

[203] D.D. Zimmermann, H. Grossholz, S. Wolf, O. Janka, A.C. Müller, T. Schleid, Two Hexagonal Series of Lanthanoid(III) Oxide Fluoride Selenides: M6O2F8Se3 (M = La – Nd) and M2OF2Se (M = Nd, Sm, Gd – Ho), Zeitschrift für anorganische und allgemeine Chemie 641(11) (2015) 1926-1933.

[204] T. Schleid, H. Grossholz, Über Fluoridsulfide (MFS) der Lanthanide (M=La-Nd, Sm, Gd-Lu) im A-Typ mit PbFCl-Struktur, Zeitschrift für anorganische und allgemeine Chemie 627(12) (2001) 2693-2699.

[205] A. Demourgues, A. Tressaud, H. Laronze, P. Macaudière, Rare earth fluorosulfides LnSF and Ln2AF4S2 as new colour pigments, Journal of Alloys and Compounds 323-324 (2001) 223-230.

[206] R. Schmid, H. Hahn, Über Chalkogenidfluoride einiger Seltener Erden, Zeitschrift für anorganische und allgemeine Chemie 373(2) (1970) 168-175.

[207] D. Fischer, Z. Cancarevic, J.C. Schön, M. Jansen, Zur Synthese und Struktur von K3N, Zeitschrift für anorganische und allgemeine Chemie 630(1) (2004) 156-160.

[208] O. Janka, Synthesis and characterization of multinary rare-earth metal(III) fluorides and examinations to their optical behavior, University of Stuttgart, Stuttgart, Germany, 2010.

Author Biography

Milan Pejić was born on November 4, 1980, in Aleksandrovac. He completed his elementary education in Aleksandrovac, after which he continued his schooling in Kruševac, in the specialized mathematics class at "Gimnazija Kruševac". He then enrolled at the Faculty of Physics, University of Belgrade, where he graduated in 2009 with the thesis "Calculating the Electronic Structure of the DNA Molecule" under the supervision of Dr. Radomir Žikić.

In addition to studying physics, he earned an MSc in International Economics at the Faculty of Economics, University of Belgrade. He subsequently gained significant experience working in data analysis, software engineering, economic modeling, and systems administration in the financial and IT sector before enrolling in PhD studies at the Faculty of Physics, University of Belgrade, in 2019.

Since 2019, Milan has been a member of the Laboratory for Theoretical Investigation of Materials (L-TIM) at the Center for Synthesis, Processing, and Characterization for Application in Extreme Conditions (CEXTREME LAB). Since 2021, he has been employed at the Department of Materials Science at the Institute of Nuclear Sciences "Vinča", where he conducts research under the supervision of Dr. Dejan Zagorac.

His Horizon 2020 project "Energy Landscapes of Rare-Earth Compounds Containing Several Different Anions" was approved under the HPC-Europa3 Transnational Access Programme for a Pan-European Network of HPC Research Infrastructures and Laboratories for Scientific Computing. The project was funded under H2020 Call INFRAIA-2016-1, Type of Action: RIA, Grant Agreement No. 730897, which enabled the collaboration with the High-Performance Computing Center Stuttgart (HLRS), Germany.

His scientific research focuses on density-functional-theory calculations, energy-landscape exploration, machine learning methods, and the modeling of electronic, magnetic, mechanical, and vibrational properties of materials. During his PhD studies, he participated in several international conferences where he presented his PhD-related research. He is the author or coauthor of 16 SCI-indexed publications, four of which are directly related to his PhD thesis, with one additional thesis-related manuscript submitted for publication.

Publications Related to the Thesis:

1. Pejić, M., Zimmermann, D. D., Zagorac, D., Fonović, M., Zagorac, J., Schön, J. C., & Schleid, T. (2026). Structural exploration of holmium fluoride selenide (HoFSe): theory and experiment. *Journal of Physics and Chemistry of Solids*, 208, 113000.
2. Zagorac, J., Schön, J. C., Matović, B., Pejić, M., Prekajski Đorđević, M., & Zagorac, D. (2023). Computational discovery of new feasible crystal structures in Ce_3O_3N. *Crystals*, 13(5), 774.
3. Pejić, M., Zagorac, D., Zagorac, J., Matović, B., & Schön, J. C. (2022). Structure prediction via global energy landscape exploration of the ternary rare-earth compound LaOI. *Zeitschrift für anorganische und allgemeine Chemie*, 648(24), e202200308.
4. Zagorac, D., Zagorac, J., Fonović, M., Pejić, M., & Schön, J. C. (2022). Computational discovery of new modifications in scandium oxychloride (ScOCl) using a multi-methodological approach. *Zeitschrift für anorganische und allgemeine Chemie*, 648(23), e202200198.

Other Publications:

1. Škundrić, T., Zagorac, J., Pejić, M., Luković, A., & Zagorac, D. (2026). First-principles study of $CsSnX_3$ (X = Cl, F) lead-free inorganic halide perovskites: Structural and electronic properties. *Journal of Solid State Chemistry*, 362, 126144.
2. Zhukova, I., Tatarková, M., Kombamuthu, V., Zagorac, D., Pejić, M., Chlup, Z., Kovalčíková, A., Šiška, F., Hernández, F. C., Moshtaghioun, B. M., Gómez-García, D., Hosseini, N., Matović, B., Dlouhý, I., & Tatarko, P. (2026). Theoretical prediction, synthesis, and mechanical properties of non-equimolar $(Ta\text{-}Hf\text{-}Zr\text{-}Nb\text{-}Ti)B_2$ entropy-stabilised borides. *Journal of the European Ceramic Society*, 46(3), 117903.
3. Škundrić, T., Schön, J. C., Zagorac, J., Pejić, M. & Zagorac D. (2026). Unveiling crystalline modifications on the energy landscape of $Cr_3Si_3N_8$ using the multi-methodological approach. *Computational and Theoretical Chemistry*, 1256, 115632.
4. Teppala, D. T., Pejić, M., Zagorac, D., Adabifiroozjaei, E., Goyal, N., Molina-Luna, L., Mathur, S., Matović, B., & Ionescu, E. (2025). Single-source precursor synthesis of a compositionally complex early transition-metal nitride $(V, Nb, Ta, Mo, W)N_x$ and its high-temperature stability. *International Journal of Applied Ceramic Technology*, 23(1), e70120.
5. Zagorac, J., Zagorac, D., Škundrić, T., Pejić, M., Matović, B., & Schön, J. C. (2025). Theoretical insight into structural and mechanical features of $Hf_{0.5}Ta_{0.5}C$. *Processing and Application of Ceramics*, 19(2), 201–213.
6. Zagorac, D., Zagorac, J., Škundrić, T., Pejić, M., Jovanović, D., & Schön, J. C. (2024). Structure prediction and mechanical properties of tantalum carbide (TaC) at the ab initio level. *Zeitschrift für anorganische und allgemeine Chemie*, 650(22), e202400088.
7. Teppala, D. T., Bernauer, J., Rashid, A., Pejić, M., Zagorac, D., Matović, B., & Ionescu, E. (2024). Single-source precursor synthesis of a compositionally complex early transition-metal carbonitride $(Ti,Zr,Hf,Nb,Ta)N_xC_{1-x}$. *Advanced Engineering Materials*, 26(18), 2302165.
8. Matović, B., Maletaškić, J., Maksimović, V., Dimitrijević, S. P., Todorović, B., Pejić, M., Zagorac, D., Zagorac, J., Zeng, Y.-P., & Cvijović-Alagić, I. (2023). Multicomponent solid solution with pyrochlore structure. *Boletín de la Sociedad Española de Cerámica y Vidrio*, 62(6), 515–526.
9. Zagorac, J., Zagorac, D., Šrot, V., Ranđelović, M., Pejić, M., van Aken, P. A., Matović, B., & Schön, J. C. (2023). Synthesis, characterization, and electronic properties of ZnO/ZnS core/shell nanostructures investigated using a multidisciplinary approach. *Materials*, 16(1), 326.
10. Pejić, M., Pržulj, Ž., Chevizovich, D., Lazarides, N., Tsironis, G. P., & Ivić, Z. (2022). Qubit-photon bound states in superconducting metamaterials. *Physical Review B*, 105(23), 235439.
11. Zagorac, D., Zagorac, J., Pejić, M., Matović, B., & Schön, J. C. (2022). Band gap engineering of newly discovered ZnO/ZnS polytypic nanomaterials. *Nanomaterials*, 12(9), 1595.
12. Škundrić, T., Zagorac, D., Schön, J. C., Pejić, M., & Matović, B. (2021). Crystal structure prediction of the novel Cr_2SiN_4 compound via global optimization, data mining, and the PCAE method. *Crystals*, 11(8), 891.

Изјава о ауторству

Име и презиме аутора ___________Милан Пејић______________

Број индекса __________8014/2019________________

Изјављујем

да је докторска дисертација под насловом

Energy Landscapes, Crystal Structure Prediction, and Modeling of Rare Earth Ternary Compounds

(Енергетски пејзажи, предвиђање кристалних структура и моделовање тернарних једињења елемената ретких земаља)

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио ауторска права и користио интелектуалну својину других лица.

Потпис аутора

У Београду, _________________

____

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора ___Милан Пејић___

Број индекса ___8014/2019___

Студијски програм ___Физика кондензоване материје и статистичка физика___

Наслов рада ___Energy Landscapes, Crystal Structure Prediction, and Modeling of Rare Earth Ternary Compounds___
(Енергетски пејзажи, предвиђање кристалних структура и моделовање тернарних једињења елемената ретких земаља)

Ментор ___др Дејан Загорац___

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, ________________________

____

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Energy Landscapes, Crystal Structure Prediction, and Modeling of Rare Earth Ternary Compounds

(Енергетски пејзажи, предвиђање кристалних структура и моделовање тернарних једињења
елемената ретких земаља)

која је моје ауторско дело.

Дисертацију са свим прилозима предао сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прераде (CC BY-ND)

[✓ 6.] Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, ________________________

____

1. **Ауторство**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.