



Studies of superhydrides by ab-initio and machine learning accelerated random structure searching

Jean-Baptiste Charraud

► To cite this version:

Jean-Baptiste Charraud. Studies of superhydrides by ab-initio and machine learning accelerated random structure searching. Materials Science [cond-mat.mtrl-sci]. Université Paris-Saclay, 2021. English. NNT : 2021UPASF035 . tel-03405328

HAL Id: tel-03405328

<https://theses.hal.science/tel-03405328>

Submitted on 27 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Studies of Superhydrides by Ab-Initio and Machine
Learning Accelerated Random Structure Searching

*Etude de Superhydrures par Recherche Aléatoire de
Structures Couplée au Calcul Ab-Initio et Accélérée
par Machine Learning*

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 571, Sciences Chimiques: Molécules Matériaux,
Instrumentation et Biosystèmes (2MIB)

Spécialité de doctorat: Physique

Unité de recherche: Université Paris-Saclay, CEA, Laboratoire Matière sous
conditions extrêmes, 91680, Bruyères le Chatel, France.

Référent: Faculté des sciences d'Orsay

**Thèse présentée et soutenue à Paris-Saclay
le 15 Octobre 2021 par**

Jean-Baptiste Charraud

Composition du jury:

Valérie PAUL-BONCOUR Directrice de Recherche, Université Paris-Est, CNRS	Présidente
Eric BOUSQUET Chargé de Recherche, Université de Liège	Rapporteur & Examineur
Anne HEMERYCK Chargée de Recherche HDR, Université Toulouse III, LAAS-CNRS	Rapporteuse & Examinatrice
Mihai-Cosmin MARINICA Directeur de Recherche, Université Paris-Saclay, CEA	Examineur

Direction de la thèse:

Grégory GENESTE Directeur de Recherche, Université Paris-Saclay, CEA	Directeur de thèse
Marc TORRENT Chercheur CEA, Université Paris-Saclay	Co-encadrant
Jean-Bernard MAILLET Chercheur CEA, Université Paris-Saclay	Invité

Titre: Etude de Superhydrures par Recherche Aléatoire de Structures Couplée au Calcul Ab-Initio et Accélérée par Machine Learning

Mots Clés: Superhydrures, Théorie Fonctionnelle de la Densité, Machine Learning, Hydrogène

Résumé: Alliages caractérisés par la présence d'une grande quantité d'hydrogène dans leur structure cristalline, les superhydrures sont des matériaux prometteurs pour le stockage dense d'énergie et la supraconductivité à haute température. Le principal obstacle à l'utilisation des superhydrures pour des applications est la pression de synthèse considérable de l'ordre du million d'atmosphère. Pour la réduire, la connaissance de ces nouveaux matériaux doit être approfondie. Cette thèse apporte ainsi une contri-

bution à la recherche de superhydrures formés de métaux de transition. A cet effet, l'algorithme de Recherche Aléatoire de Structures couplée au calcul Ab-Initio (AIRSS) a d'abord été mis en oeuvre sur des hydrures et superhydrures de cuivre, de manganèse, d'yttrium ainsi que d'un système ternaire yttrium-fer. Le cas des superhydrures d'yttrium est un défi scientifique. Il apparaît qu'ils doivent être décrits par des structures trop complexes pour la méthode AIRSS. Pour parvenir à une meilleure description de ces composés, des outils de "machine learning" ont été développés.

Title: Studies of Superhydrides by Ab-Initio and Machine Learning Accelerated Random Structure Searching

Keywords: Superhydrides, Density Functional Theory, Machine Learning, Hydrogen

Abstract : Alloys characterized by the presence of a large amount of hydrogen in their crystal structure, superhydrides are promising materials for dense energy storage and high temperature superconductivity. The main obstacle to the use of superhydrides for applications remains their very high synthesis pressure, in the range of one million atmosphere. To reduce it, knowledge about these new materials must be deepened. This thesis thus aims at making a contribution to the search of transition metal type su-

perhydrides. For this purpose, the "Ab-Initio Random Structure Searching" (AIRSS) algorithm was firstly implemented. Studies of hydrides and superhydrides of copper, manganese, yttrium as well as a ternary yttrium-iron system were carried out. The case of yttrium superhydrides is the place of a scientific challenge, because they probably have large primitive cells, which are not accessible by the AIRSS method. To reach a better description of these compounds machine learning tools were developed and used.

Remerciements

Superhydrure, Hydrogène, Machine Learning, Calcul Ab-Initio, Simulation Quantique, Calcul HPC, Cellule à Enclume de Diamant ... le champ lexical qui aura marqué ces trois années de doctorat est bien vaste et est à l'image de cette expérience passionnante et du chemin parcouru. Cette thèse, ce n'est pas uniquement un travail purement scientifique dans un domaine bien borné, ce sont aussi des interactions, des collaborations aussi nombreuses que pluridisciplinaires. Je voudrais ainsi remercier tous ceux qui ont croisé la route de ce travail démarré il y a maintenant trois ans. Je voudrais tout d'abord saluer mon directeur de thèse, Grégory Geneste. Sa grande expertise scientifique, sa pédagogie, son optimisme et son ouverture auront joué un rôle essentiel dans cette thèse. Son intuition physique aura également été décisive pour les différentes orientations de ce projet. Mes remerciements vont également à Marc Torrent, notre collaboration pour l'implantation des fonctionnelles "Meta-GGA" dans *abinit* aura grandement enrichi cette expérience de thèse et donné des perspectives nouvelles pour mieux étudier les superhydrures par la simulation. Je voudrais également témoigner toute ma reconnaissance pour Jean-Bernard Maillet qui aura joué un rôle majeur dans les développements des nouvelles méthodes de Machine Learning servant à la prédiction de nouveaux superhydrures. Ses conseils et la collaboration riche que nous avons eue aura été une pierre angulaire de ce doctorat. D'autre part, que n'aurait été cette thèse sans ces collaborations étroites avec l'équipe des expérimentateurs de Paul Loubeyre. Ces échanges nombreux auront permis de créer un écosystème où la simulation, les algorithmes, le machine learning sont questionnés par la réalité des expériences et donnent en retours des résultats nouveaux. Je voudrais donc remercier en particulier Paul Loubeyre, Adrien Marizy, Bastien Guigue pour tout nos échanges qui auront grandement fait progresser ce travail de doctorat. Mes remerciements vont également à tous les collègues de mon laboratoire ainsi qu'à mes collègues doctorants du CEA qui auront tous, d'une certaine manière, contribué à ce cadre unique dans lequel j'ai évolué. Enfin mes remerciements vont à mes proches. Je voudrais saluer leur soutien, toute l'énergie qu'ils m'ont donnée pour persévérer et m'investir au maximum tout au long de ces trois ans. Je voudrais donc en particulier dédier cette thèse à mes parents, Clémentine, Grégoire, Lea et Céline, mais aussi à Yvette et Huguette. Merci pour tout !

Contents

1	Superhydrides: Challenges and Methods	2
1.1	Hydrogen and Superhydrides	3
1.1.1	Hydrogen and Energy Storage	4
1.1.2	The Potentialities of Superhydrides	7
1.2	Density Functional Theory	11
1.2.1	Principles	12
1.2.1.1	The Equation to Solve	12
1.2.1.2	The Theorems of DFT	13
1.2.2	Description of the Electronic Wavefunctions	15
1.2.3	Exchange and Correlation Functionals	18
1.2.3.1	Local Density Approximation (LDA)	19
1.2.3.2	Generalized Gradient Approximation (GGA)	19
1.2.3.3	Meta-GGA	19
1.2.3.4	Hybrid Functionals	20
1.2.4	Computation Choices to Search Superhydrides	21
1.2.5	Projector Augmented-Wave Approach	22
1.2.6	Meta-GGA with PAW Formalism	26
1.3	Crystal Structure Prediction Algorithms	29
1.3.1	Combining Experiments and Numerical Simulation	30
1.3.2	Formation Enthalpies and Convex Hull Representation	31
1.3.3	CSP Algorithms, General Principles	34
1.3.4	The Ab-Initio Random Structure Searching Algorithm	35
1.3.4.1	Generation Rules	36
1.3.4.2	The Structural Optimization Process	38
1.3.5	Other Algorithms	39
1.3.5.1	Particle Swarm Optimization	39
1.3.5.2	Genetic Algorithm	40
1.3.5.3	Metadynamics	43

1.3.6 CSP Algorithms: Improvement Ways	45
1.3.6.1 Representation of the Exploration Space and PES	45
1.3.6.2 Clustering	45
1.3.6.3 Machine Learning Potentials	45
1.3.7 Summary of Part 1	46
 2 Transition Metal Hydrides and Superhydrides Studied by Ab-Initio Random Structure Searching	 47
 2.1 Copper Hydrides	 48
2.1.1 State of the Art	49
2.1.2 Computational Details	50
2.1.2.1 Ab initio Random Structure Searching	50
2.1.2.2 Density-Functional Theory Calculations	51
2.1.2.3 Special Quasi Random Structure	52
2.1.3 Results: Copper Hydrides Under Pressure	52
2.1.3.1 Hydrides at Ambient Pressure	52
2.1.3.2 Hydrides under Pressure	53
2.1.3.3 Dynamical Stability	56
2.1.4 Physical Properties	56
2.1.4.1 Crystal Structures	57
2.1.4.2 Bader Charges	58
2.1.5 Conclusion	60
 2.2 Manganese Hydrides and Superhydrides	 61
2.2.1 Manganese Hydrides, State of the Art	62
2.2.2 Computational Details	63
2.2.2.1 Ab initio Random Structure Searching	63
2.2.2.2 Density-Functional Theory Calculations	63
2.2.3 Results	64
2.2.3.1 Manganese Hydriding at Ambient and low Pressure	64
2.2.4 Dynamical Stability and Nuclear Quantum Effects	68
2.2.5 Physical Properties of the Stable Structures	72
2.2.5.1 Description of Crystal Structures	72
2.2.5.2 Bader Charges	77
2.2.5.3 Electronic Structure	79
2.2.5.4 Magnetism	81
2.2.6 Conclusion	81

2.3 Yttrium Hydrides and Superhydrides	83
2.3.1 State of the Art	84
2.3.1.1 Binary Yttrium Hydrides and Superhydrides	84
2.3.1.2 Iron-Yttrium Hydrides	89
2.3.2 The AIRSS Calculations	91
2.3.2.1 Binary Yttrium Superhydrides	91
2.3.2.2 Yttrium-Iron Ternary Hydrides	98
2.3.3 Conclusion	110
2.3.4 Summary of Part 2	112
 3 Toward the Predictions of more Complex Superhydrides	 113
3.1 Machine Learning and CSP Algorithms	114
3.1.1 Numerical Potentials	115
3.1.1.1 Numerical Potentials: Main Ingredients	115
3.1.1.2 The Descriptors	116
3.1.1.3 The Main Statistical Models Employed	119
3.1.2 Spectral Neighbour Analysis Potential	120
3.1.2.1 The Bispectrum Descriptor	121
3.1.2.2 Estimating Physical Properties with SNAP	121
3.1.2.3 The Training Process	124
3.1.3 Numerical Potentials and Random Searching	126
 3.2 A First Proof of Concept on Tin	 128
3.2.1 Why Considering Tin	129
3.2.2 A First Try	130
3.2.2.1 Global Work-Flow	131
3.2.2.2 Potential Initialization	131
3.2.2.3 First Active Learning Cycle	134
3.2.2.4 Second Active Learning Cycle	137
3.2.3 A Second Try with more Converged Databases	139
3.2.3.1 Databases: Comparisons	139
3.2.3.2 First Active Learning Cycle	141
3.2.3.3 Update of the Potential	143
3.2.3.4 Second Active Learning Cycle	144
3.2.4 Predictions under Pressure	146
3.2.5 Conclusion	149
 3.3 Application to Yttrium Superhydrides	 150
3.3.1 A First Test on the YH_6 System	151
3.3.1.1 Initial Database	152

3.3.1.2 Search of the YH_6 Im-3m Phase at 200 GPa	153
3.3.1.3 Search of Possible New YH_6 Phases	155
3.3.2 Development of New Analysis Tools	160
3.3.2.1 Bispectrum Analysis	161
3.3.2.2 X-Ray Based Metric	165
3.3.3 YH_x Explorations at 147 GPa	168
3.3.3.1 Experimental X-Ray Spectrum and Metric	169
3.3.3.2 Random Searching Implementation	172
3.3.3.3 New Structures Obtained	173
3.4 Perspective for Predictions of Ternary Systems	176
3.4.1 Predict the Best Alloys for Ternary Systems	177
3.4.1.1 The Database	177
3.4.1.2 The Model Considered	179
3.4.1.3 Predict Existence and Value of the Transition Pressure	181
3.4.2 Ternary Systems and Numerical Potentials	182
3.4.3 Summary of Part 3	183
Appendices	187
A Résumé	188
B Convex Hulls and Formation Enthalpy Definition	193
C On-site Meta-GGA Contribution to Total Energy in the PAW Approach	196
D ZBL Empirical Potential	200
E Manganese Hydrides and Superhydrides	202
F Machine Learning and AIRSS	219
G Perspectives for Ternary Superhydride Predictions	226

Why this PhD Project

Improving energy storage and transport capabilities are probably among the most important challenges of the 21st century. The absence of technological solution using compact systems able to store huge amount of energy in the long term, is a strong limitation in a lot of domains. For the development of alternative and renewable energies, these improvements are vital to be able to confront a growing demand with more intermittent energy production systems (wind turbines, solar panels ...). Those systems are in fact characterized by energy production periods unsynchronized with the peaks of consumption (for instance, in winter, it is for now very difficult to use energy produced in summer by solar panels). The energy transport and the related losses are at the same time important points to consider. Indeed, For intermittent energy production systems, it is required to supply energy from where it is produced (for instance where there is wind or sun) to other regions (possibly far away) where there is not and needing energy. In other domains, for mobile applications, important questions are raised about technological choices to develop compact and light systems concentrating the most energy as possible. It applies for electric vehicles as well as for spacial applications where the conception of new energetic fuels is an active question. For now, no universal solution exists. Research of new materials able to respond to all these challenges is thus required.

In this context, my PhD aims at making a contribution in the research of new hydrogen enriched alloys called "superhydrides". These materials offer huge opportunities for energy storage under hydrogen form; and electricity transport with few losses thanks to superconductivity properties. During this PhD project at the CEA, I had to manage the numerical studies so as to guide the discovery and synthesis of new superhydrides. This work is organized in three main activities:

- Developing algorithms using quantum simulation (Ab-Initio calculation) and machine learning to predict new materials.
- Implementing these algorithms on supercomputers.
- Making the interface between these numerical studies and the teams involved in the experimental synthesis of superhydrides.

This PhD thesis focuses on the results obtained and developed tools following this strategy. The first part describes the challenges related to superhydrides and the main established numerical methods employed for the prediction of their structure. The second part presents studied systems using these methods. In particular superhydrides made of transition metal type elements are investigated. The third part is about machine learning approaches to be able to reach predictions of more complex superhydrides.

PART 1

SUPERHYDRIDES: CHALLENGES AND METHODS

Introduction

This part aims at presenting the context, challenges and tools related to the study of superhydrides. In the first chapter, the potentialities these materials may offer in the context of energy storage and supply are introduced. So as to learn more about these newly discovered compounds, the necessity to use numerical simulation is discussed in the two last chapters. In chapter 2, the principles of Density Functional Theory (DFT), which is used for highly accurate simulation of materials at the atomic scale, are presented. Then in chapter 3, the main algorithms of Crystal Structure Prediction (CSP) are described.

Chapter 1.1

Hydrogen and Superhydrides

Contents

1.1.1 Hydrogen and Energy Storage	4
1.1.2 The Potentialities of Superhydrides	7

This first chapter presents the main context related to superhydrides. The first section discusses the challenges related to energy storage and supply under hydrogen form. The second section focuses on the opportunities the superhydrides may offer in this domain.

1.1.1 Hydrogen and Energy Storage

Being the lightest and the most abundant element in the universe, hydrogen has been at the origin of a lot of scientific and technological developments since more than a century. The birth of quantum physics (e.g. Bohr model 1913), the discovery of Nuclear Magnetic Resonance (e.g. Bloch 1946) nowadays used in medicine, a great part of the progress made in astronomy, the nuclear fusion, all are striking examples of scientific and technological advances provided by this chemical element. This way of progress seems not to be stopped. Indeed hydrogen might become a cornerstone for future breakthroughs in both energy production and transport. Such a trend seems to be confirmed by the recent billion worth investment plan targeting research and industrialization of new hydrogen technologies in several European countries [1]. In fact hydrogen might be a serious candidate as an energetic vector to store and transport energy. Non-consumed energy might be stored under hydrogen form (H_2) and conversely hydrogen might be used to produce electricity or heat.

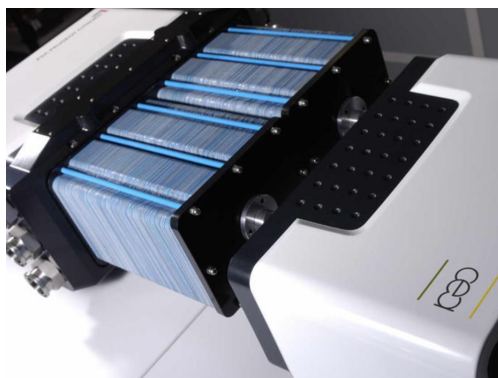


Figure 1.1.1: Hydrides and high autonomy fuel cells (CEA credits).

This reversible transformation from hydrogen to energy is made possible thanks to a fuel cell/electrolyzer device (see FIG 1.1.1). One direct application might be for instance the storage, by the electrolysis of water, of produced energy by intermittent energy production systems such as solar panels or wind turbines. This could be employed typically when energy demand is lower than the production. The hydrogen produced this way can then be employed in transport by hydrogen vehicles: it can be used in a fuel cell to generate electrical power and run an electric motor. However, one of the main remaining limitations to see a wide expansion of hydrogen technologies is the storage and supply of this element. The very low hydrogen density under normal pressure (around 11 m^3 for 1 kg of hydrogen) is the source of some difficulties. Storing a sufficient

amount of energy in the form of hydrogen with a reasonable volume requires compressing it or cooling it down at very low temperature. Typically, for mobile applications a hydrogen volumetric density of around $60 \text{ kg(H}_2\text{)}/\text{m}^3$ is commonly reached by applying a pressure of 700 bar [2]. Such density enables it to carry around 5 kg of H_2 within the car tank, giving thus in the range of 500 km autonomy. Progress in the stored energy density using this strategy will be however limited. Indeed, liquid hydrogen density at temperature below -250°C reaches a density of $71 \text{ kg(H}_2\text{)}/\text{m}^3$. It is thus the upper limit regarding the reachable density by compressing hydrogen gas.

To go well above this $71 \text{ kg(H}_2\text{)}/\text{m}^3$ limit, a new strategy needs to be considered. A promising one is to store hydrogen under solid state in alloys forming hydrides. These solid materials result from the association of a metal with hydrogen, and are thus characterized by the presence of atomic hydrogen included in the crystal structure. Their formation is made possible thanks to an electronegativity difference between the metallic atoms and hydrogen (hydrogen being more electronegative than the metal and is consequently negatively charged in the hydrides). Under ambient pressure, the stable hydrides (formation enthalpy $< -40 \text{ kJ/mol H}_2$) are mainly made of the least electronegative metals (alkaline, alkaline-earth, transition metals of the three left columns, rare-earth elements and actinides [3], FIG 1.1.2). Under ambient conditions, some of them are studied for industrial purposes [4] thanks to their large hydrogen density. Indeed, hydrides could enable important progress for hydrogen storage: in the case of magnesium dihydrides, MgH_2 , around $110 \text{ kg(H}_2\text{)}/\text{m}^3$ is reached. Even $120 \text{ kg(H}_2\text{)}/\text{m}^3$ for LiBH_4 can be obtained [5]. Another advantage in comparison with the storage under gas phase is that it avoids the risk of explosion [6].

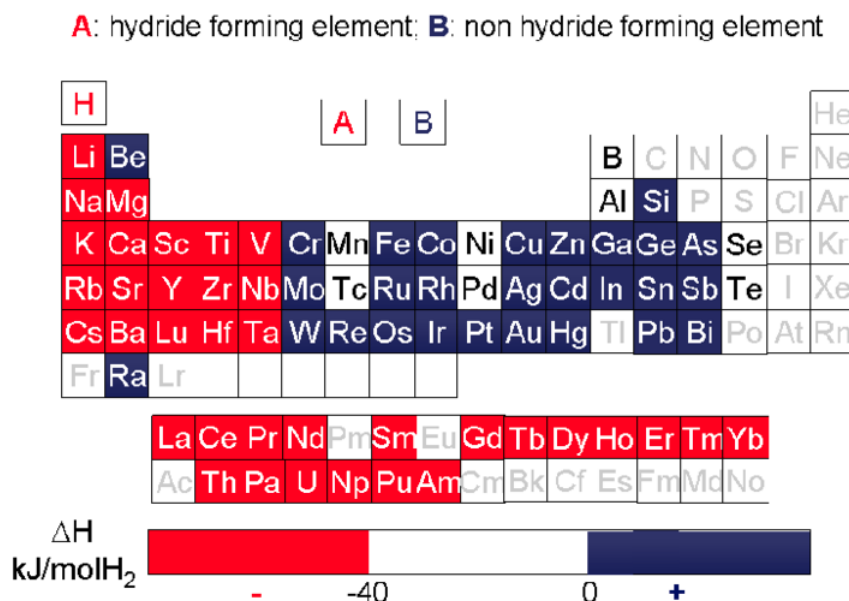


Figure 1.1.2: Hydride forming elements and non-hydride forming elements (at ambient pressure), figure taken from Ref [3].

Thanks to these properties, hydrides are considered as a high potential solution to increase autonomy of fuel cell devices and enable at the same time a size reduction. Some applications where hydrides are employed in this context are yet under development. For instance Lototskyy et al [7] propose fuel cells using hydrides for forklift. Non mobile applications, like energy supply for hotels are also studied as presented by Ref [8]. In these devices, the hydride must be heated up to release hydrogen under gas form, and then feeds the chemical reaction producing electricity and heat.

In a very different context, hydrides like ErH start to be investigated for controlled hydrogen fusion reactions. This is aimed to develop new energy resources for space explorations [9].

However, metal hydrides as a way to store hydrogen (in combination with fuel cells) do not have a wide range of applications. Indeed, hydrides have for now two important drawbacks:

- In order to be used as a hydrogen storage system, a hydride must have a negative formation enthalpy, but this enthalpy must not be too negative. Otherwise, it is necessary to heat the hydride to extract hydrogen (and thus the energy cost may be too large). If we fix at 100°C the upper limit of the temperature that could be reached for extracting hydrogen from the hydride, the formation enthalpy of the hydride must not be more negative than -58 kJ/mol H_2 [10]. Unfortunately, many hydrides have a more negative formation enthalpy: this is the case of magnesium dihydride MgH_2 , for which the formation enthalpy is -76 kJ/mol H_2 [11].

- The gravimetric density (mass of hydrogen per unit mass of the storage system) can be low in comparison with pure hydrogen gas or liquid, in particular if heavy atoms are involved in the alloy. It can thus limit the use of hydrides for applications in the mobility domain.

To enable a wider use of hydrides as hydrogen storage systems, fundamental research is still necessary. A possible way of improvement is to find new alloys which could reach higher hydrogen concentrations. This is typically a possibility offered by new hydrides recently discovered under high pressure: the superhydrides.

1.1.2 The Potentialities of Superhydrides

Superhydrides are newly discovered alloys, stable under high pressure, and characterized by a large hydrogen content in their chemical formula (e.g. FeH_5 [12], LaH_{10} [13] or CeH_9 [14]). These hydrogen-rich materials may be characterized by stoichiometries larger than what may be expected from usual valence rules, that apply at ambient pressure. Furthermore, it can be noticed that hydrogen solubility in the metal usually increases with pressure, which suggests that very hydrogen-rich compounds may be formed at sufficiently large synthesis pressures for most metals of the periodic table. These intriguing compounds, made of metal atoms combined with atomic hydrogen around, are stabilized by high pressure conditions, and are currently the subject of an intensive research, because of their very promising potentialities. Indeed, if superhydrides could survive at ambient pressure under a stable or metastable form, they could be used as efficient solid state hydrogen storage systems (see FIG 1.1.3).

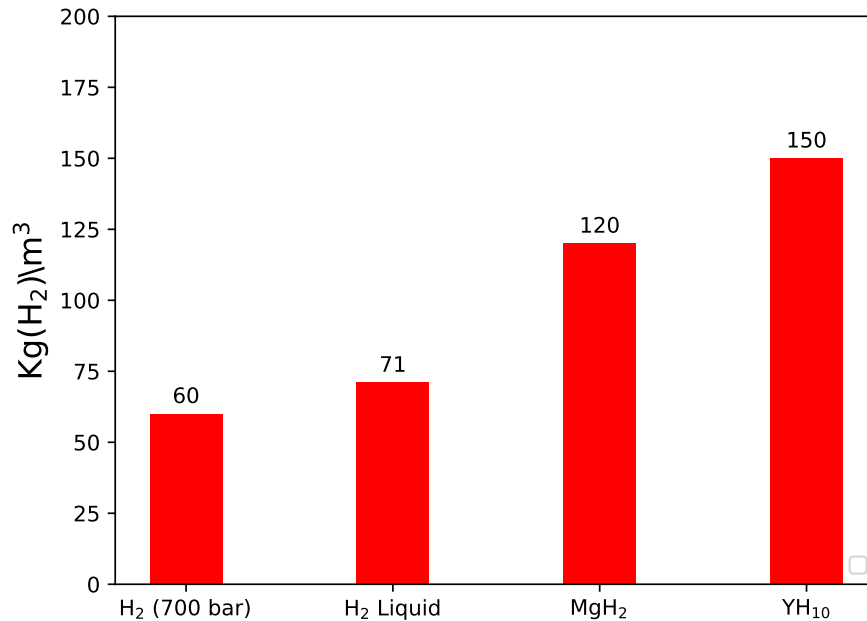


Figure 1.1.3: Volumetric density of hydrogen, in $\text{kg}(\text{H}_2)/\text{m}^3$, for different hydrogen storage materials. For the superhydride YH_{10} , the concentration is computed as if the material could remain stable at 0 GPa.

Another challenge with superhydrides is the possible obtention of conventional superconductors with high critical temperature (T_c). This is guided by Ashcroft's idea [15] according to which the high T_c superconductivity expected in metallic hydrogen (theoretically predicted and recently synthesized [16]) at large pressure (near 450 GPa [16]) could be encountered also in some metal-hydrogen alloys, but at more moderate pressure more easily accessible to diamond anvil cell experiments (typically up to 100-200 GPa), due to a chemical precompression of the hydrogen lattice by the metal atoms. It is worth noticing that not all the elements can form superhydrides and different families can be identified.

A first one is made of the alkaline and alkaline-earth metal hydrides. A good exemple is the case of lithium superhydrides where LiH_6 was predicted [17] and experimentally observed [18] above 130 GPa. Other examples are potassium KH_6 [19] or calcium CaH_6 [20] superhydrides predicted near 150 GPa. This superhydride family seems not to be an ideal one to reach high temperature superconductivity (e.g T_c around 40 K at 150 GPa for LiH_6 [21], predicted T_c of 70 K at 166 GPa for KH_6 [22]). It might be related to the fact these superhydrides are characterized by the presence of hydrogen molecules or more complex molecular units such as H_3^- [23] in the crystal structure. This leads to bad conducting properties, for instance lithium superhydrides remain insulating at least up to 215 GPa [18].

A second group is made of the transition metal element superhydrides, with the exception of the elements of the IIIB column (because their chemistry is similar to that of rare-earth elements). The series of iron hydrides, FeH_x , is emblematic: as discussed in Refs [24, 25, 12, 26, 27, 28] iron and hydrogen form FeH under a few GPa, and then successively FeH_2 at 67 GPa [25], FeH_3 at 87 GPa [25], and FeH_5 at 130 GPa [12]. In this group, no high T_c was observed (in the case of FeH_5 , a T_c was first predicted around 40-50 K [29, 26], which was then contradicted in Refs. [30, 28], that calculated a T_c close to zero).

A third group gathers the transition metal type elements of the IIIB column (Sc, Y, La) and the rare-earth ones. They have been predicted so far with the highest- T_c : yttrium superhydrides [31, 98, 99], for instance YH_6 or YH_{10} , exhibit predicted superconducting T_c s around 250-280 K [33]. They should also reach very high hydrogen stoichiometries, typically REH_6 , REH_9 or REH_{10} (RE=rare earth or IIIB transition metal). Very recently, these predictions seem to have received experimental confirmation in the case of lanthanum decahydride LaH_{10} [13, 32]. Note that these compounds REH_x adopt at large x ($\geq 5-6$) very specific crystal structures with the hydrogen atoms arranged in cages that surround the metal atoms ("hydrogen clathrate structures", see Ref [99] and see FIG 1.1.4). These high expected T_c s are probably related to the fact that this family of superhydrides is characterized by the presence of hydrogen sublattices having common points with metallic hydrogen: in these superhydrides the nearest neighbours of H atoms are H atoms, in contrast to the ionocovalent hydrides formed at ambient pressure with lower hydrogen stoichiometry. In addition, the typical H-H distances are in the range of 1.1-1.2 Å (around 1.2 Å for YH_6 and 1.15 for YH_{10} at 200 GPa [34]) and thus very close to the one observed in metallic hydrogen in comparison with the other families. Indeed, the tetragonal structure of metallic atomic hydrogen (space group $I41/amd$) artificially decompressed down to 150-200 GPa, provides with H-H distances of 1.2 Å [12].

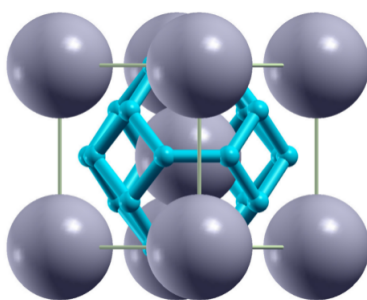


Figure 1.1.4: An illustration of a clathrate structure observed in yttrium hexahydride YH_6 . The H atoms form a sublattice with cages inserting the Y atoms. These cages are formed here by quares H_4 and hexagons H_6 .

A last group is formed by the actinide superhydrides: U [35], Ac [36] and Th [37], have been shown to form superhydrides under pressure. Uranium superhydrides UH_7 and UH_8 were experi-

mentally [38] obtained. In the case of UH_7 a synthesis pressure as small as 36 GPa is observed. It is among the lowest pressures reported so far for the synthesis of superhydrides. In the case of uranium, the predicted Tcs are not as high as observed in the former group (around 50 K predicted for UH_7 [35]). However, a higher Tc has been predicted (161 K in the 170-175 GPa range) in the case of thorium decahydride ThH_{10} [37].

Obviously, the pressures at which superhydrides are stable are not compatible with any application for hydrogen storage. However they could possibly be used to store hydrogen if they can be maintained at ambient pressure under a metastable form (i.e. the superhydride structure is dynamically stable, but has a larger enthalpy than its decomposition into any other hydride and hydrogen). Considering the applications to compact hydrogen storage, the following conditions can be listed to determine whether a new superhydride may have potentialities, if it is supposed to remain in a metastable form under ambient pressure conditions (i.e. with a negative formation enthalpy).

- The formation energy of the superhydride (with respect to metal and H_2) should be negative since no additional energy should be paid to form the hydride, even if it does not lie on the convex hull of the formation enthalpies.
- The formation energy of the superhydride must not be too negative (between 0 and -58 kJ/mol H_2 , otherwise the extraction of H_2 from the hydride may necessitate to heat above 100°C [39], see above). An ideal value for the formation energy is typically ~ -30 kJ/mol H_2 [40].
- The density of hydrogen should be above 150 kg(H_2)/m³, to have a higher density than most of the hydrides.
- The hydrogen gravimetric density has to be above 11 % [40] to aim at mobile applications for electric vehicles.

If these considerations about the potentialities of superhydrides are followed, it appears that the present investigation of several superhydride compounds under pressure could be an important step to identify the best candidates matching part of the mentioned criteria. The next chapters describe the main numerical tools which have been used in the present work for these investigations. Firstly Density Functional Theory (DFT), enabling the simulation of materials at the atomic scale with a high accuracy is presented. Secondly the Crystal Structure Prediction (CSP) algorithms are described. In particular it is shown how they can be used in combination with some experimental results.

Chapter 1.2

Density Functional Theory

Contents

1.2.1 Principles	12
1.2.1.1 The Equation to Solve	12
1.2.1.2 The Theorems of DFT	13
1.2.2 Description of the Electronic Wavefunctions	15
1.2.3 Exchange and Correlation Functionals	18
1.2.3.1 Local Density Approximation (LDA)	19
1.2.3.2 Generalized Gradient Approximation (GGA)	19
1.2.3.3 Meta-GGA	19
1.2.3.4 Hybrid Functionals	20
1.2.4 Computation Choices to Search Superhydrides	21
1.2.5 Projector Augmented-Wave Approach	22
1.2.6 Meta-GGA with PAW Formalism	26

This chapter presents the numerical methods used to simulate physical properties of materials at the atomic scale. The first section describes the main principles of Density Functional Theory (DFT). It is followed by the presentation of the commonly used exchange and correlation functionals. In particular the Meta-GGA is described. During this PhD, an implementation in the *abinit* software of the Meta-GGA functionals was made. It was performed in the framework of Projector Augmented-Wave formalism. After describing this method, the approach followed to achieve this implementation is summarized.

1.2.1 Principles

Solving directly the quantum many body problem for electrons is only possible for a few ones. Unfortunately, the system sizes compatible with such solving are much too small to deal with problems encountered in materials science. To overcome this issue, Density Functional Theory (DFT) can be considered. It is an approach which is now very popular in materials science and quantum chemistry and that is used for highly accurate simulation of the physical properties of materials. It probably leads to the best compromise between a high accuracy and the size of a system compatible with a study of relevant materials science problems. DFT can be viewed as a reformulation of the search for the ground state of an electron gas in terms of electronic density. It thus keeps the accuracy of quantum mechanics in the description of the physical properties associated with the electrons. This section gives the main theorems and equations involved.

1.2.1.1 The Equation to Solve

Let us consider a set of atoms (i. e. nuclei and electrons) described by a wavefunction Ψ . If N refers to the ions and e to the electrons, the time independent Schrödinger equation is given by:

$$[\hat{T}_e + \hat{T}_N + \hat{V}_{ee} + \hat{V}_{NN} + \hat{V}_{Ne}] \Psi = E \Psi \quad (1.2.1)$$

In Eq 1.2.1, the different terms are:

- The kinetic energy operator for the electrons \hat{T}_e and the ions \hat{T}_N .
- The interaction energy between the electrons \hat{V}_{ee} .
- The interaction energy between the ions \hat{V}_{NN} .
- The interaction energy between ions and electrons \hat{V}_{Ne} .

Based on the difference of mass between the nuclei and the electrons, the Born-Oppenheimer approximation allows to decouple the motions of these two types of particles. The lighter electrons are assumed to adjust instantaneously (adiabatically) their state to that of the nuclei. As a consequence, the total wavefunction can be written as the product of a nuclear and an electronic

wavefunction. In addition, the electronic wavefunction ψ_e becomes the solution of an "electronic" Schrödinger equation.

$$[\hat{T}_e + \hat{V}_{ee} + \hat{V}_{ext}]\psi_e = E\psi_e \quad (1.2.2)$$

In Eq 1.2.2 one has: $\hat{V}_{ext} = \hat{V}_{NN} + \hat{V}_{Ne}$, it is the ionic coulombic potential applied to the electrons. This equation has to be solved for a fixed set of ionic positions. Note that in many cases in materials science, knowledge of the electronic ground state is sufficient to describe many physical properties of the system (at low and room temperature), especially in the case of insulating systems.

1.2.1.2 The Theorems of DFT

DFT theory is based on two main principles. The first one relies on the Hohenberg and Kohn theorems which explain the crucial role played by the electronic ground state density in the determination of the ground state properties. The second one relies on the Kohn and Sham variational approach, a reformulation of the theory that leads to the so-called Kohn-Sham equations (a set of equations that allows numerical implementation of the DFT).

Hohenberg and Kohn Theorems [41]

There are two Hohenberg and Kohn (HK) theorems at the root of DFT. Let us consider a set of N_e electrons. The first theorem states that the ground state density n of the system is enough to get all its physical properties. In particular the energy of the ground state $E_0[n]$ is a functional of the electronic ground state density. More precisely, the first HK theorem demonstrates that there is a unique (i.e. that differs from any other potential by more than a constant shift) external potential V_{ext} that corresponds to a given ground state density. Let us denote by $|\psi_0[n]\rangle$ the ground state wavefunction, which is also the unique one corresponding to the density n . It is possible to define three "universal" functionals, $F_{HK}[n]$, $T_e[n]$ and $V_{ee}[n]$ such that:

$$T_e[n] = \langle \psi_0[n] | \hat{T}_e | \psi_0[n] \rangle$$

$$V_{ee}[n] = \langle \psi_0[n] | \hat{V}_{ee} | \psi_0[n] \rangle$$

$$F_{HK}[n] = T_e[n] + V_{ee}[n]$$

The second theorem fixes a variational principle. One can define for any density n and for a given external potential V_{ext} the functional:

$$E_V[n] = T_e[n] + V_{ee}[n] + \int V_{ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \quad (1.2.3)$$

The ground state density is the one that minimizes this functional.

Kohn-Sham Approach [42]

In practice the HK theorems are not applicable since the F_{HK} functional is unknown. Kohn and Sham proposed another formulation that makes DFT usable. They assume that for any real interacting electrons in an external potential there exists a fictitious system of non interacting electrons merged in a potential called V_{KS} (the Kohn-Sham potential), which has the same ground state density as the real one. This fictitious system is thus described by a set of monoelectronic wavefunctions $\psi_i(\mathbf{r})$ and monoelectronic equations called the Kohn-Sham equations. The total electronic density is related to the Kohn-Sham wavefunctions by:

$$n = \sum_i^{N_e} f_i |\psi_i|^2 \quad (1.2.4)$$

In Eq 1.2.4, the coefficients f_i are called the occupation coefficients. They are equal to 1 if the i -th state has an energy below the fermi level, 0 if its energy is above. If the energy equals the Fermi level, they can take fractional values between 0 and 1. As the monoelectronic wavefunctions are independent, they can be obtained by solving the above mentioned Kohn-Sham equations:

$$\left[-\frac{1}{2}\Delta + V_{KS}(\mathbf{r})\right]\psi_i = E_i\psi_i \quad (1.2.5)$$

In addition, following the HK variational principle applied to the fictitious system, the ground-state density minimizes the following energy functional:

$$E_V[n] = T_{KS}[n] + \int V_{KS}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \quad (1.2.6)$$

where $T_{KS}[n]$ is the kinetic energy functional associated with the fictitious system. Moreover, a more detailed expression of the functional in Eq 1.2.6 can be obtained. To do that, the Eq 1.2.3 defining the energy functional has to be considered. The functional $F_{HK}[n] = T_e[n] + V_{ee}[n]$ can be rewritten as $T_{KS}[n] + J[n] + E_{xc}[n]$ where $J[n] = \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}'$. $E_{xc}[n]$ is the exchange and correlation energy functional and is thus defined as:

$$E_{xc}[n] = T_e[n] - T_{KS}[n] + V_{ee}[n] - J[n] \quad (1.2.7)$$

This universal functional is the correction related to the fact that one replaces the wavefunction of the real system by a set of monoelectronic and fictitious ones giving the same density. Thanks to these definitions, one can derive an other expression of the energy functional:

$$E_V[n] = T_{KS}[n] + J[n] + E_{xc}[n] + \int V_{ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \quad (1.2.8)$$

Following the Hohenberg and Kohn variational principle, the ground state density $n(\mathbf{r})$ has to minimize the two energy functionals given by Eqs 1.2.6 and 1.2.8. One deduces that the two functionals have their functional derivatives in $n(\mathbf{r})$ which have to be zero. This allows to eliminate the term in T_{KS} and it leads to the expression of the Kohn-Sham potential V_{KS} :

$$V_{KS} = \frac{\delta J[n]}{\delta n} + V_{ext} + \frac{\delta E_{xc}[n]}{\delta n} \quad (1.2.9)$$

This potential can also be expressed in terms of the exchange and correlation potential V_{xc} , defined as the functional derivative of E_{xc} :

$$V_{KS}(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{ext}(\mathbf{r}) + V_{xc}[n(\mathbf{r})] \quad (1.2.10)$$

The Kohn Sham equations become:

$$\left[-\frac{1}{2}\Delta + \frac{\delta J[n]}{\delta n} + V_{ext} + \frac{\delta E_{xc}[n]}{\delta n}\right]\psi_i = E_i\psi_i \quad (1.2.11)$$

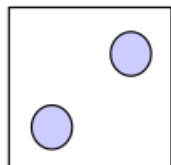
It is noticeable that Eq 1.2.11 gives theoretically the exact ground state density associated with the Schrödinger equation in Eq 1.2.2. To solve numerically these equations, three ingredients are required:

- An algorithm solving eigenvalue problems. Generally an iterative algorithm is considered because a few eigenvalues are in most cases sufficient.
- A representation of the wavefunctions (i.e. their expansion on a given basis set).
- An exchange and correlation functional.

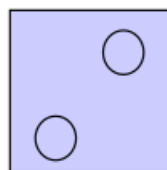
In the following sections, the two last ingredients will be more in depth discussed.

1.2.2 Description of the Electronic Wavefunctions

A first point to consider for a numerical representation of the electronic wavefunctions is a basis in which they will be expressed. Two main kinds of bases can be used: the "localized" and "delocalized" bases. Localized bases involve functions attached to a specific atom location (see FIG 1.2.1 for illustration). An example is the atomic centered orbital basis using functions made of products between a radial part and an angular part described by spherical harmonics. This kind of basis requires few elements (for each atom), in addition all the electron wavefunctions can be efficiently represented, including those highly localized (like those of the core electrons which have strong oscillations). The use of such a basis gives highly accurate results but at a price of heavy calculations. Localized type bases can however be difficult to manipulate mainly due to the fact that the basis moves with atoms. As opposed to localized bases, delocalized ones like plane-waves can be easier to manipulate as they do not depend on atom locations (see FIG 1.2.1 for illustration). The calculations using these kinds of basis functions are easy to converge and systematic. However many basis elements may be required, especially for the localized electronic wavefunctions, leading thus an important computation cost.

Localized

Spherical harmonics,
Gaussians,
Atomic orbitals, ...

Delocalized

Plane waves, ...

Figure 1.2.1: Type of basis used to describe electronic wave functions.

Moreover, solving directly Eq 1.2.5 with wavefunctions describing all the electrons (called the "all-electron wavefunctions") requires an important computation cost. It is particularly high for the core electronic states. As just mentioned, the wavefunctions associated with these states have a lot of nodes and thus strong oscillations near the atomic nuclei. This behaviour necessitates numerically to have a fine grid describing these states as well as an important number of basis functions to express the wavefunctions (and this is in particular the case if a plane-wave basis is used). At the same time, these core states do not influence the chemical properties as much as the valence ones. This remark leads to the "frozen-core electron" hypothesis. It states that they can be determined once for a given atom type and do not change when atoms are combined together to form a crystal structure. A strategy commonly used is thus to treat independently core and valence electrons. However this independent treatment can not be done directly without adaptations. Indeed, it appears the core wavefunctions influence indirectly the valence ones as they have to be orthogonal to them. As these functions have a lot of nodes, it is thus the case for the valence states near the atomic nuclei leading to an important computation cost to describe these strong oscillations. To avoid such treatment, the potential generated by an ion and its core electrons is replaced by a "pseudopotential" associated with nodless "pseudo-wavefunctions" keeping the same valence properties (see FIG 1.2.2 for illustration). The use of these pseudo-wavefunctions having a smooth (and wrong) behaviour in the core region is justified by the fact that the electrons do not take part in the chemical bonds below a certain distance from the atomic cores. This combination of approximations, frozen-core and pseudopotential eliminates the electrons which do not play any role (or negligible) in the chemical properties of a material. It thus reduces the number of orbitals which need to be computed explicitly. It also avoids dealing with the strong variations of both the potential and the valence electron wavefunctions in the core regions.

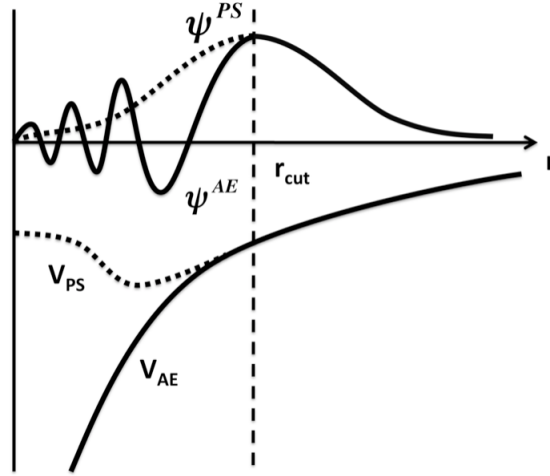


Figure 1.2.2: Pseudopotentials and pseudo-wavefunctions [43].

For this purpose, a sphere of radius r_c is determined and centered at each atom. The pseudo-wavefunctions have no node inside this sphere and coincide with the real electron wavefunctions outside (see FIG 1.2.2). More explicitly, if S is the sphere of radius r_c centered at one atom, if $|\Psi\rangle$ is a real electron wavefunction and $|\tilde{\Psi}\rangle$ a pseudized associated one, $|\tilde{\Psi}\rangle$ does not have any node in S and outside it verifies $|\tilde{\Psi}\rangle = |\Psi\rangle$. Dealing with pseudized wavefunctions instead of real ones enables decomposing them in a plane-wave basis with fewer elements, thus reducing the computation cost. It is however at a price of lower accuracy because the real potential is lost. To retrieve the long range electrostatic behaviour, a way is to impose an electron wavefunction and the pseudo associated one to have the same norms inside the sphere S of radius r_c . It means:

$$\int_S |\Psi(\mathbf{r})|^2 d\mathbf{r} = \int_S |\tilde{\Psi}(\mathbf{r})|^2 d\mathbf{r} \quad (1.2.12)$$

This constraint leads to the so-called "norm-conserving" pseudopotentials as presented by Hamann et al [44]. By this approach, the long range behaviour of the Hartree potential is recovered. However the norm-conserving condition imposes relatively strong oscillating functions requiring thus an important number of basis elements to be properly described. To overcome this difficulty, "ultrasoft pseudopotentials" have been introduced by Vanderbilt [45] which do not verify the norm-conserving condition. However the price to pay is the introduction of a fictitious charge compensation density. This additional density is introduced to compensate the electronic charge and retrieve the long range behaviour of the electrostatic potential.

Another approach introduced by Blochl [46] consists in using a linear operator between all-electron wavefunctions and pseudo ones. This approach leads to the Projector Augmented-Wave formalism (PAW) [46, 47, 48]. By combining the use of all-electron pseudo-wavefunctions and of the linear operation to get the all-electron ones, the PAW approach enables the computation

of the exact density because the information near the nuclei remains available. It is thus as accurate as an all-electron method. This requires in addition a much lower computation cost than using directly the all-electron wavefunctions. It needs however much more developments than the approach consisting in using norm-conserving or ultrasoft pseudopotentials. This approach will be described more in depth in 1.2.5.

1.2.3 Exchange and Correlation Functionals

To determine an approximation of the exchange and correlation functional, different forms can be adopted. The different kinds of functionals used are ranged in five rungs constituting the "Jacob's Ladder" [49]. The higher the rung, the more complex and accurate is the functional (and the higher the computation cost, see FIG 1.2.4). The exchange and correlation energy can be separated in two terms, one corresponding to the exchange part E_x , the other to the correlation part E_c such that $E_{xc} = E_c + E_x$. This decomposition plays an important role for the two highest rungs of the "Jacob's Ladder" including the "hybrid" functionals. The four first rungs are presented in the following subsections. The different exchange and correlation functionals are accessible for instance in the LibXC package [50, 51]. It is worth precising that the exchange and correlation potential v_{xc} is the functional derivative of $E_{xc}[n, \dots]$:

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n, \dots]}{\delta n} \quad (1.2.13)$$

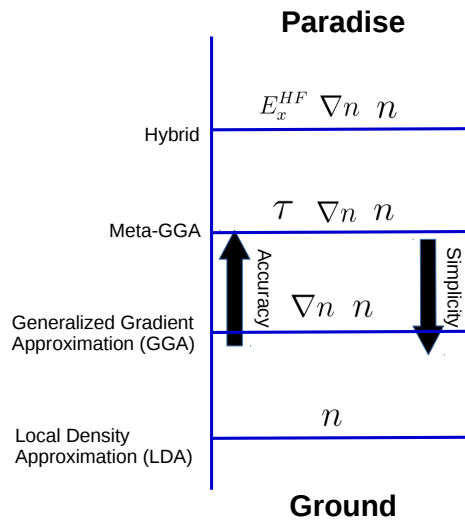


Figure 1.2.3: The four first rungs of Jacob's Ladder [49].

1.2.3.1 Local Density Approximation (LDA)

In this approximation, only the electronic density (and thus a local quantity depending only on a point \mathbf{r}) is taken into account. The exchange and correlation energy is given by:

$$E_{xc}[n] = \int \epsilon_{xc}^{homo}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r} \quad (1.2.14)$$

The associated exchange and correlation potential is expressed as:

$$v_{xc}(\mathbf{r}) = \frac{d(n\epsilon_{xc}^{homo})}{dn} \quad (1.2.15)$$

in which $\epsilon_{xc}^{homo}(n)$ is the exchange and correlation energy per electron of a homogeneous electron gas with density n . LDA is accurate for an important variety of materials and mainly for those having low spatial variations of the electronic density (see for instance computation of electronic and optic properties with boron nitride [52] or copper [53]). It appears however that LDA is not accurate for systems having strongly correlated electrons or to compute band gaps of semiconductors or insulators [54].

1.2.3.2 Generalized Gradient Approximation (GGA)

This kind of functionals includes the gradient of the density as input. It is a semi-local quantity as it depends on the density at point \mathbf{r} and its neighbourhood to compute the gradient. E_{xc} has thus the expression:

$$E_{xc}[n] = \int \epsilon_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r})]n(\mathbf{r})d\mathbf{r} \quad (1.2.16)$$

$\epsilon_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r})]$ is decomposed into:

$$\epsilon_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r})] = \epsilon_x^{homo}[n(\mathbf{r})]F_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r})] \quad (1.2.17)$$

with F_{xc} an enhancement factor. The associated exchange and correlation potential has the form:

$$v_{xc}(\mathbf{r}) = \frac{\partial(n\epsilon_{xc})}{\partial n} - \nabla \cdot \frac{\partial(n\epsilon_{xc})}{\partial(\nabla n)} \quad (1.2.18)$$

A widely used GGA functional which is also employed in this PhD is the Perdew, Burke and Ernzerhof (PBE) functional [55]. GGA is for instance accurate to estimate crystal bulk and lattice constants of transition metals [56, 57]. Transition pressure are also better described than in LDA for a lot of systems [58].

1.2.3.3 Meta-GGA

This third rung of the Jacob's Ladder consists in adding another semi-local quantity which is the kinetic energy density. Initially, Meta-GGA-like functionals were introduced to take into account the laplacian of the density in combination with the density and its gradient.

$$E_{xc}[n] = \int \epsilon_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r}), \Delta n(\mathbf{r})]n(\mathbf{r})d\mathbf{r} \quad (1.2.19)$$

For numerical reasons, the laplacian of the density can be replaced by the kinetic energy density. As introduced in [59], this quantity is defined as:

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_n f_n |\nabla \psi_n|^2 \quad (1.2.20)$$

Where ψ_i are the Kohn-Sham wavefunctions and f_n the occupation numbers. The exchange and correlation energy functional is thus expressed as:

$$E_{xc}[n] = \int \epsilon_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r}), \Delta n(\mathbf{r}), \tau(\mathbf{r})] n(\mathbf{r}) d\mathbf{r} \quad (1.2.21)$$

In this case, the exchange and correlation potential is:

$$v_{xc}^{MGA}(\mathbf{r}) = \frac{\partial(n\epsilon_{xc})}{\partial n} - \nabla \cdot \frac{\partial(n\epsilon_{xc})}{\partial \nabla n} + \Delta \left(\frac{\partial(n\epsilon_{xc})}{\partial \Delta n} \right) - \frac{1}{2} \nabla \cdot [\mu_{xc}(\mathbf{r}) \nabla] \quad (1.2.22)$$

Several Meta-GGA functionals have been introduced. The first one is called the Tao-Perdew-Staroverov-Scuseria (TPSS) functional [60] and its improvement the revTPSS [61]. TPSS and revTPSS showed however some limits regarding the description of structural phase transitions [62] or for Van der Waals interaction [63]. To address these problems, the Strongly Constrained and Appropriately Normed (SCAN) functional [64, 65] was developed. The accuracy of the Meta-GGA-type functionals was better than GGA for various studies of physical properties like adsorption phenomena [66] or for band gap predictions [67]. Examples of success include the accurate predictions of silicon or ice phases [68], the correct prediction of metal to insulator transitions and magnetic properties of cuprate [69], the successful use to study dynamics of liquid water by taking into account the Van der Waals interactions [70]. Furthermore the computation cost related to Meta-GGA is close to the one involved for GGA because the kinetic energy density is a semi-local quantity as the gradient of the density. SCAN functional suffers from numerical instabilities, to solve this problem, a regularized functional called rSCAN is introduced by Bartok et al [71]. However some constraints used to build the SCAN functional are no longer respected with the rSCAN one which leads to some transferability loss as pointed out for atomization energy calculations [72]. Furness et al proposed in [73] another functional, r2SCAN, which uses the regularization technic employed with rSCAN without reducing the number of constraints it has to fulfill in comparison with SCAN. This new functional appears to be less computation expensive and more numerically stable as SCAN while having the same accuracy as verified by S. Grimme et al [74].

1.2.3.4 Hybrid Functionals

The fourth rung of the Jacob's Ladder corresponds to the hybrid type functionals. Becke introduced the principle of these functionals [75]. It is stated that part of the functional involved in the exchange part will include Hartree Fock (HF) exchange in the form:

$$E_x^{HF} = -\frac{1}{2} \sum_{ij} \int \frac{\psi_i(\mathbf{r}_1) \psi_j^*(\mathbf{r}_1) \psi_i(\mathbf{r}_2) \psi_j^*(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (1.2.23)$$

where the ψ_i are the Kohn-Sham wavefunctions. The exchange term of Eq 1.2.23 requires to decompose the ψ_i wavefunctions in plane-waves for its computation. As a consequence an important number of basis elements is necessary (due to the products between the wavefunctions) to express E_x^{HF} requiring thus an important computation cost. It exists two main kinds of hybrid functionals:

- The "global hybrid functionals" like those of the PBE0 family [76]. These functionals have the form [77]:

$$E_{xc}^{PBE0} = \frac{1}{4}E_x^{HF} + \frac{3}{4}E_x^{PBE} + E_c^{PBE} \quad (1.2.24)$$

The HF term is mixed with GGA type ones (coming from the PBE family). This ponderation between HF and GGA terms comes from the fact the HF term is suitable to describe electronic interactions in molecules. However in crystals an electronic screening exists, which thus screens the potential associated with the HF term. The GGA terms aim thus at modeling this screening phenomenon.

- The "range separated hybrid functionals" like those of the HSE family introduced by Heyd et al [78]. Similarly as the global hybrid functionals, the screening of the HF potential observed in crystal is modeled by GGA terms. However this time is also used a range radius. Below it (short range region, noted SR), the HF term is considered. Above (long range region, noted LR), it is the GGA terms which are fully taken into account. For instance the HSE functionals have the form [77]:

$$E_{xc}^{HSE} = \frac{1}{4}E_x^{HF,SR}(\Lambda) + \frac{3}{4}E_x^{PBE,SR}(\Lambda) + E_x^{PBE,LR}(\Lambda) + E_c^{PBE} \quad (1.2.25)$$

The range radius is determined by the Λ factor and is equal to $\frac{2}{\Lambda}$. Above this radius the short range terms (denoted by SR) become negligible.

Initially, hybrid functionals use GGA-like terms combined with the HF one. However it exists nowadays more and more sophisticated functionals combining both global and range separated approaches, some of them use Meta-GGA terms (see for instance the work of Jana et al [79, 80]).

1.2.4 Computation Choices to Search Superhydrides

Following the description of the different possibilities in terms of wavefunction representations and exchange and correlation functionals, the following section discusses the choices made to reach the prediction of new hydrides and superhydrides. To search new crystal structures of superhydrides in this PhD thesis, DFT will be employed to implement methods exploring potential energy surfaces, for this purpose one needs:

- to explore the maximum number of configurations. Typically for a binary system (system made of a metallic atom type and hydrogen) with a determined hydrogen stoichiometry, around 1000 structures have to be searched. For this purpose a fast method, easy to converge and systematic will be required.

- to consider a method suitable to deal with solids and in particular crystal structures. If one wants to explore at least binary type superhydrides exhaustively, this method has to be able to simulate crystal structure unit cells having in the range of ~ 20 -30 atoms.
- to estimate the electronic gaps with a high accuracy: indeed, materials having possibly superconducting properties will be searched. To this aim, the best functionals for this kind of calculations, in terms of accuracy, remain the hybrid ones. However the computation cost associated with this kind of functionals is too high if one needs to test an important number of crystal structures. Typically, studies about binary type superhydrides where several hydrogen stoichiometries are tested require in the range of 10 million CPU hours. Which is the magnitude of computation cost accessible with the supercomputers employed during this PhD thesis. With hybrid-type functionals, the cost would be well above.
- to differentiate the stability between two crystal structures with the highest accuracy as possible. Typically, reaching a precision in the range of 10 meV/at will be important to determine at a given pressure which crystal structure candidates are really stable and which ones are not.

With all of these points, it appears that the best compromise between accuracy and minimal computation cost is to describe wavefunctions with a plane-wave basis. This choice has to be combined with the use of the Projector Augmented-Wave (PAW) formalism and of Meta-GGA-type exchange and correlation functionals. This choice did not exist in *abinit* beforehand, it was thus required to implement it. In the next sections, the PAW formalism is described as well as details about the implementation of the Meta-GGA-type functionals in *abinit* employing this formalism.

1.2.5 Projector Augmented-Wave Approach

As yet mentioned in 1.2.2 PAW establishes a link between the real all-electron wavefunctions and pseudized ones. It allows to work with pseudized quantities requiring a lower computation cost without impacting the accuracy. In the PAW framework, if $|\Psi\rangle$ is an all-electron wavefunction and $|\tilde{\Psi}\rangle$ the pseudized associated one. The link between them is determined thanks to a linear \hat{P} operator such that:

$$|\Psi\rangle = |\tilde{\Psi}\rangle + (|\Psi_S\rangle - |\tilde{\Psi}_S\rangle) = \hat{P}|\tilde{\Psi}\rangle \quad (1.2.26)$$

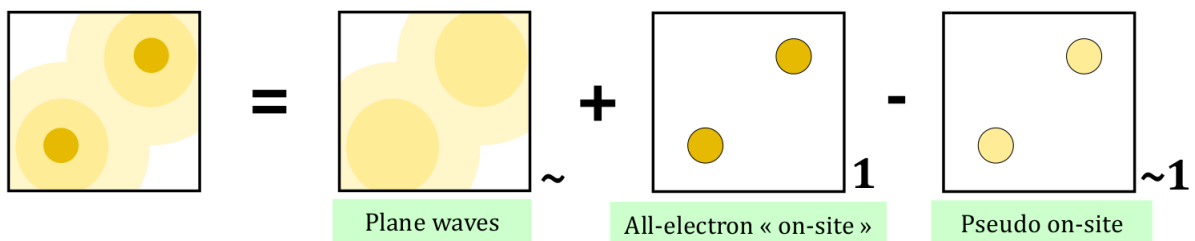


Figure 1.2.4: Illustration of the PAW linear transformation [43].

In Eq 1.2.26, $|\Psi_S\rangle$ refers to the restriction of the $|\Psi\rangle$ function to the sphere of radius r_c called the PAW sphere and similarly for $|\tilde{\Psi}_S\rangle$ with $|\tilde{\Psi}\rangle$. The PAW approach is based on four main points:

- Attached to each atom R, two local type bases are defined. A first one is called the "on-site" basis (the term "on-site" is referred to an upper 1) made of orbitals (ϕ_i^{1R}) . It is associated with an on-site pseudo-basis $(\tilde{\phi}_i^{1R})$ corresponding to smoother wavefunctions (i.e. without node) in the PAW sphere and coinciding with the on-site basis outside.
- Outside the PAW spheres, a plane-wave basis is in addition considered.
- From the on-site pseudo-basis, a set of local projectors $\langle \tilde{p}_i^{1R} |$ can be determined. It corresponds to the dual basis of $(\tilde{\phi}_i^{1R})$ such that:

$$\sum_i |\tilde{\phi}_i^{1R}\rangle \langle \tilde{p}_i^{1R}| = Id \quad (1.2.27)$$

- The core electrons give core density and core pseudo-density as well as core kinetic and core pseudo-kinetic energy densities denoted respectively n_c , \tilde{n}_c and τ_c , $\tilde{\tau}_c$. Following the frozen-core assumption, these are computed once for the isolated atom. Ionic charge density and pseudo-density n_{Zc} and \tilde{n}_{Zc} are also introduced.

These objects are built such that, a real electron wavefunction in the PAW sphere can be decomposed in the form $|\Psi_S\rangle = \sum_i |\phi_i^{1R}\rangle \langle \tilde{p}_i^{1R}| \tilde{\Psi}\rangle$ and the pseudo associated one $|\tilde{\Psi}_S\rangle = \sum_i |\tilde{\phi}_i^{1R}\rangle \langle \tilde{p}_i^{1R}| \tilde{\Psi}\rangle$. The \hat{P} operator in 1.2.26 is thus expressed as :

$$\hat{P} = 1 + \sum_{iR} (|\phi_i^{1R}\rangle - |\tilde{\phi}_i^{1R}\rangle) \langle \tilde{p}_i^{1R}| \quad (1.2.28)$$

Using these definitions and the expression of the \hat{P} operator in Eq 1.2.28, all the physical quantities represented by mean values of operators on $|\Psi\rangle$ are decomposed in the form:

$$\langle A \rangle = \langle \tilde{A} \rangle + \sum_R (\langle A^{1R} \rangle - \langle \tilde{A}^{1R} \rangle) \quad (1.2.29)$$

where:

- R indexes the atoms.
- $\langle \tilde{A} \rangle$ is the mean value of A computed on the plane-wave basis.
- $\langle A^{1R} \rangle$ is the mean value of A computed on the on-site basis (ϕ_i^{1R}) .
- $\langle \tilde{A}^{1R} \rangle$ is the mean value of A computed on the on-site pseudo-basis $(\tilde{\phi}_i^{1R})$.

Using Eq 1.2.29, the electronic density (for the valence electrons) can be expressed by applying the \hat{P} operator to the density operator:

$$n(\mathbf{r}) = \tilde{n}(\mathbf{r}) + \sum_R (n^{1R}(\mathbf{r}) - \tilde{n}^{1R}(\mathbf{r})) \quad (1.2.30)$$

In Eq 1.2.30, the different densities are:

- The all-electron pseudo-density:

$$\tilde{n}(\mathbf{r}) = \sum_n f_n \tilde{\Psi}_n^*(\mathbf{r}) \tilde{\Psi}_n(\mathbf{r}) \quad (1.2.31)$$

In this expression $\tilde{\Psi}_n$ are the all-electron pseudo-wavefunctions corresponding to different states (bands and k points) and f_n the occupation numbers of these states.

- The on-site density:

$$n^{1R}(\mathbf{r}) = \sum_{ij} \rho_{ij}^{1R} (\phi_i^{1R}(\mathbf{r}))^* \phi_j^{1R}(\mathbf{r}) \quad (1.2.32)$$

- The on-site pseudo-density:

$$\tilde{n}^{1R}(\mathbf{r}) = \sum_{ij} \rho_{ij}^{1R} (\tilde{\phi}_i^{1R}(\mathbf{r}))^* \tilde{\phi}_j^{1R}(\mathbf{r}) \quad (1.2.33)$$

In the expressions of these on-site densities in Eqs 1.2.32 and 1.2.33 is used the on-site density matrix ρ_{ij} . It is defined as:

$$\rho_{ij}^{1R} = \sum_n f_n \langle \tilde{\Psi}_n | \tilde{p}_i^{1R} \rangle \langle \tilde{p}_j^{1R} | \tilde{\Psi}_n \rangle \quad (1.2.34)$$

Similarly as with the electronic density, the \hat{P} operator can be applied to the Hamiltonien, it gives the energy:

$$E = \tilde{E} + \sum_R (E^{1R} - \tilde{E}^{1R}) \quad (1.2.35)$$

with [81] :

$$\tilde{E} = \sum_n f_n \langle \tilde{\Psi}_n | -\frac{\Delta}{2} | \tilde{\Psi}_n \rangle + E_{xc}[\tilde{n} + \tilde{n}_c] + E_H[\tilde{n} + \hat{n} + \tilde{n}_c] + \int v_H[\tilde{n}_{Zc}] \tilde{n} d\mathbf{r} \quad (1.2.36)$$

$$E^{1R} = \sum_{ij} \rho_{ij} \langle \phi_i | -\frac{\Delta}{2} | \phi_j \rangle + E_{xc}[n^{1R} + n_c] + E_H[n^{1R} + n_c] + \int v_H[n_{Zc}] n^{1R} d\mathbf{r} \quad (1.2.37)$$

$$\tilde{E}^{1R} = \sum_{ij} \rho_{ij} \langle \tilde{\phi}_i | -\frac{\Delta}{2} | \tilde{\phi}_j \rangle + E_{xc}[\tilde{n}^{1R} + \tilde{n}_c] + E_H[\tilde{n}^{1R} + \hat{n} + \tilde{n}_c] + \int v_H[\tilde{n}_{Zc}] \tilde{n}^{1R} d\mathbf{r} \quad (1.2.38)$$

In Eqs 1.2.36, 1.2.37 and 1.2.38, the potential $v_H = \frac{\delta J[n]}{\delta n} + V_{ext}$ (see Eq 1.2.11) and $E_H[n] = \int v_H[n] n d\mathbf{r}$. \hat{n} is the fictitious charge compensation density added to retrieve the electrostatic behaviour. For the demonstration these equations are written only with the electronic density but other quantities can be taken following the type of the exchange and correlation functional ($\nabla n, \Delta n, \tau$ and their associated pseudo and core quantities).

If one defines $\tilde{\rho} = |\tilde{\Psi}\rangle\langle\tilde{\Psi}|$, the PAW hamiltonian has the form [81]:

$$\frac{dE}{d\tilde{\rho}} = -\frac{1}{2}\Delta + v_{Hxc} + \sum_{ijR} |\tilde{p}_i^{1R}\rangle D_{ij}^{1R} \langle \tilde{p}_j^{1R}| \quad (1.2.39)$$

In Eq. 1.2.39, $v_{Hxc} = v_H[n] + \frac{\delta E_{xc}[n]}{\delta n} = v_H[n] + v_{xc}[n]$ (see Eq. 1.2.11), the term $\sum_{ijR} |\tilde{p}_i^{1R}\rangle D_{ij}^{1R} \langle \tilde{p}_j^{1R}|$ is specific to the PAW formulation. The PAW non-local operators D_{ij}^{1R} are expressed as:

$$D_{ij}^{1R} = \langle \phi_i^{1R} | -\frac{1}{2}\nabla + v_{Hxc}[n^{1R} + n_c] | \phi_j^R \rangle - \langle \tilde{\phi}_i^{1R} | -\frac{1}{2}\nabla + v_{Hxc}[\tilde{n}^{1R} + \tilde{n}_c] | \tilde{\phi}_j^{1R} \rangle \quad (1.2.40)$$

One can define for the exchange and correlation part:

$$D_{ijxc}^{1R} = \langle \phi_i^{1R} | v_{xc}[n^{1R} + n_c] | \phi_j^R \rangle - \langle \tilde{\phi}_i^{1R} | v_{xc}[\tilde{n}^{1R} + \tilde{n}_c] | \tilde{\phi}_j^{1R} \rangle^1 \quad (1.2.41)$$

The Eqs 1.2.36, 1.2.37, 1.2.38 and 1.2.39 show that if one changes the exchange and correlation functional, one needs to derive contributions for:

- The term of the form $E_{xc}[\tilde{n} + \tilde{n}_c]$ (see Eq 1.2.36) computed from the all-electron pseudo-wavefunctions.
- The on-site term:

$$E_{xc}[n^{1R} + n_c] - E_{xc}[\tilde{n}^{1R} + \tilde{n}_c] \quad (1.2.42)$$

(see 1.2.38, 1.2.37 and 1.2.35). To evaluate it, one needs to compute $\sum_{ij} \rho_{ij} D_{ijxc}^{1R}$. The derivation of D_{ijxc}^{1R} following the type of the exchange and correlation functional is thus important because it leads to the determination of the on-site energy term contributions as well as the form of the PAW hamiltonian.

The next section discusses the derived and implemented contributions during this PhD thesis when a Meta-GGA-like functional is considered.

¹It's worth noticing for Eq 1.2.41 the charge compensation density \hat{n} is not considered in this work in the exchange and correlation as suggested in the original article of Blochl [46]. It however might be done also as pointed out by Torrent et al [82].

1.2.6 Meta-GGA with PAW Formalism

The implementation of the Meta-GGA type functionals is based on a former work published by Sun et al [66]. In this section the implementation performed during this PhD thesis will be detailed.

As derived in [66], The expression of the exchange and correlation potential in Meta-GGA is

$$v_{xc}^{MGA}(\mathbf{r}) = \frac{\partial(n\epsilon_{xc})}{\partial n} - \nabla \cdot \frac{\partial(n\epsilon_{xc})}{\partial \nabla n} + \Delta \left(\frac{\partial(n\epsilon_{xc})}{\partial \Delta n} \right) - \frac{1}{2} \nabla \cdot [\mu_{xc}(\mathbf{r}) \nabla] \quad (1.2.43)$$

where $\mu_{xc} = \frac{\delta(n\epsilon_{xc})}{\delta \tau}$. For simplicity, the dependency of v_{xc}^{MGA} and ϵ_{xc} with $n, \nabla n, \tau$ and possibly with the laplacian of the density Δn are not explicitated when not required. In their work, Sun et al [66] derive a similar expression as Eq 1.2.43 where the laplacian of the density is not taken into account. In this equation it can be recognized:

- A "local part":

$$v_{xc}^{Loc} = \frac{\partial(n\epsilon_{xc})}{\partial n} - \nabla \cdot \frac{\partial(n\epsilon_{xc})}{\partial \nabla n} + \Delta \left(\frac{\partial(n\epsilon_{xc})}{\partial \Delta n} \right) \quad (1.2.44)$$

- A "non local" and pure Meta-GGA term which is:

$$-\frac{1}{2} \nabla \cdot [\mu_{xc}(\mathbf{r}) \nabla] \quad (1.2.45)$$

This term which arises due to the introduction of the kinetic energy density is new and necessitates new developments.

This expression of the Meta-GGA exchange and correlation potential leads to the implementation of:

- A contribution in $E_{xc}[\tilde{n}, \tilde{n}_c, \Delta \tilde{n}, \Delta \tilde{n}_c, \tilde{\tau}, \tilde{\tau}_c]$ (see Eq 1.2.36).
- A contribution in $\sum_{ij} \rho_{ij} D_{ijxc}$ (see Eq 1.2.42) to derive the on-site exchange and correlation terms.

The first contribution which has to be implemented can be written in the form:

$$\sum_n f_n \langle \tilde{\Psi}_n | -\frac{1}{2} \nabla \cdot [\mu_{xc}[\tilde{n} + \tilde{n}_c, \nabla(\tilde{n} + \tilde{n}_c), \Delta(\tilde{n} + \tilde{n}_c), \tilde{\tau} + \tilde{\tau}_c](\mathbf{r}) \nabla] | \tilde{\Psi}_n \rangle \quad (1.2.46)$$

This expression needs the core pseudo-kinetic energy density at a point \mathbf{r} , $\tilde{\tau}_c(\mathbf{r})$. This quantity is obtained by a spline operation over all the core pseudo-kinetic energy densities of each atom positioned at \mathbf{R} :

$$\tilde{\tau}_c(\mathbf{r}) = \sum_{\mathbf{R}} \tilde{\tau}_c^{at}(|\mathbf{r} - \mathbf{R}|) = \sum_{\mathbf{R}, \mathbf{G}} \tilde{\tau}_c^{at}(|\mathbf{G}|) e^{-i\mathbf{G} \cdot (\mathbf{r} - \mathbf{R})} \quad (1.2.47)$$

where $\tilde{\tau}_c^{at}(|\mathbf{r}|)$ is the atomic and spherical pseudo-kinetic energy density. To derive the contribution of Eq 1.2.46, the all-electron pseudo-wavefunctions can be decomposed in terms of plane-waves:

$$|\Psi_n\rangle = \sum_{\mathbf{G}} c_{n,\mathbf{G}} e^{-i\mathbf{G} \cdot \mathbf{r}} \quad (1.2.48)$$

One can also make the fourrier transformation of μ_{xc}

$$\mu_{xc}(\mathbf{r}) = \sum_{\mathbf{G}} \mu_{\mathbf{G}} e^{-i\mathbf{G} \cdot \mathbf{r}} \quad (1.2.49)$$

Injecting Eqs 1.2.48 and 1.2.49 in 1.2.46, it gives :

$$\langle \Psi_n | -\nabla(\mu_{xc} \nabla) | \Psi_n \rangle = \sum_{\mathbf{G}', \mathbf{G}} \mu_{\mathbf{G}} c_{n, \mathbf{G}'} c_{n, \mathbf{G}' - \mathbf{G}} \mathbf{G}' \cdot (\mathbf{G}' - \mathbf{G}) \quad (1.2.50)$$

This term is in fact computed by enforcing a double fourrier transformation over the wavefunctions, finally:

$$\sum_n f_n \langle \tilde{\Psi}_n | -\frac{1}{2} \nabla \cdot [\mu_{xc} \nabla] | \tilde{\Psi}_n \rangle = \frac{1}{2} \sum_{n, \mathbf{G}', \mathbf{G}} f_n \mu_{\mathbf{G}} c_{n, \mathbf{G}'} c_{n, \mathbf{G}' - \mathbf{G}} \mathbf{G}' \cdot (\mathbf{G}' - \mathbf{G})$$

The second contribution is specific to the PAW formalism. When considering for fixed i, j the Eq 1.2.40, it can be seen a specific Meta-GGA part is added to the PAW non-local operator D_{ij}^{1R} which is:

$$D_{ij}^{1R, MGA} = \frac{1}{2} \langle \tilde{\phi}_i^{1R} | \nabla \cdot [\mu_{xc}(\mathbf{r}) \nabla] | \tilde{\phi}_j^{1R} \rangle - \frac{1}{2} \langle \phi_i^{1R} | \nabla \cdot [\mu_{xc}(\mathbf{r}) \nabla] | \phi_j^{1R} \rangle \quad (1.2.51)$$

Integrating by part $\int_{S|r=r_c} \phi_i^R \mu_{xc}(\mathbf{r}) \nabla \phi_j^R \cdot \mathbf{u}_r ds = \int_{S|r=r_c} \tilde{\phi}_i^R \mu_{xc}(\mathbf{r}) \nabla \tilde{\phi}_j^R \cdot \mathbf{u}_r ds$ (S referring to the PAW sphere), Eq 1.2.51 becomes:

$$D_{ij}^{1R, MGA} = \frac{1}{2} \langle \nabla \phi_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \phi_j^{1R} \rangle - \frac{1}{2} \langle \nabla \tilde{\phi}_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \tilde{\phi}_j^{1R} \rangle \quad (1.2.52)$$

The contributions of each $D_{ij}^{1R, MGA}$ are added to compute the on-site PAW contribution to the energy:

$$\sum_{ijR} \rho_{ij}^{1R} D_{ij}^{1R, MGA} = \frac{1}{2} \sum_{ijR} \rho_{ij}^{1R} (\langle \nabla \phi_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \phi_j^{1R} \rangle - \langle \nabla \tilde{\phi}_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \tilde{\phi}_j^{1R} \rangle) \quad (1.2.53)$$

This contribution can be expressed as (with $d\Omega = \sin(\theta) d\theta d\phi$):

$$\begin{aligned} \frac{1}{2} \sum_{ijR} \rho_{ij}^{1R} D_{ij}^{1R, MGA} = & \frac{1}{2} \sum_{ijR} \rho_{ij}^{1R} \sum_{lm} \int \mu_{xclm}(r) \left[\left(\frac{\bar{\Phi}_i^{1R}(r) \bar{\Phi}_j^{1R}(r)}{r^2} - \frac{\tilde{\Phi}_i^{1R}(r) \tilde{\Phi}_j^{1R}(r)}{r^2} \right) G_{lm l_i m_i l_j m_j} \right. \\ & \left. + \left(\frac{\Phi_i^{1R}(r) \Phi_j^{1R}(r)}{r^2} - \frac{\tilde{\Phi}_i^{1R}(r) \tilde{\Phi}_j^{1R}(r)}{r^2} \right) H_{lm l_i m_i l_j m_j} \right] dr d\Omega \end{aligned} \quad (1.2.54)$$

In this expression are defined:

- Φ_i^{1R} and $\tilde{\Phi}_i^{1R}$ such that the elements of the on-site and on-site pseudo-bases are expressed as:

$$\begin{aligned} \phi_i^{1R}(\mathbf{r}) &= \frac{\Phi_i^{1R}(r)}{r} S_{l_i, m_i}(\theta, \phi) \\ \tilde{\phi}_i^{1R}(\mathbf{r}) &= \frac{\tilde{\Phi}_i^{1R}(r)}{r} S_{l_i, m_i}(\theta, \phi) \end{aligned} \quad (1.2.55)$$

- $\bar{\Phi}_i^{1R}$ and $\tilde{\bar{\Phi}}_i^{1R}$ such that:

$$\begin{aligned}\bar{\Phi}_i^{1R}(r) &= \frac{d\Phi^{1R}(r)}{dr} - \frac{\Phi^{1R}(r)}{r} \\ \tilde{\bar{\Phi}}_i^{1R}(r) &= \frac{d\tilde{\Phi}^{1R}(r)}{dr} - \frac{\tilde{\Phi}^{1R}(r)}{r}\end{aligned}\tag{1.2.56}$$

- μ_{xclm} is the (l, m) component of μ_{xc} in the spherical harmonic decomposition such that

$$\mu_{xclm}(r) = \int \mu_{xc}(\mathbf{r}) S_{lm}(\theta, \phi) d\Omega\tag{1.2.57}$$

- the two type of integral terms $G_{lm,l_i m_i, l_j m_j}$ and $H_{lm,l_i m_i, l_j m_j}$ such that:

$$\begin{aligned}G_{lm,l_i m_i, l_j m_j} &= \int S_{lm} S_{l_i m_i} S_{l_j m_j} d\Omega \\ H_{lm,l_i m_i, l_j m_j} &= \int S_{lm} [\nabla S_{l_i m_i} \cdot \nabla S_{l_j m_j}] d\Omega\end{aligned}\tag{1.2.58}$$

The details of the demonstration are given in appendix C.

These Meta-GGA contributions to the energy have been implemented in the *abinit* software and employed to precise the electronic gap of some new predicted superhydrides like the manganese superhydrides as it will be presented in Part 2 of this PhD thesis.

In this chapter we have presented the numerical tools to simulate the physical properties of materials at the atomic scale with a high accuracy. These methods based on DFT are the "raw material" to predict new crystal structures. The next chapter discusses thus how DFT calculations are employed to construct Crystal Structure Prediction (CSP) algorithms.

Chapter 1.3

Crystal Structure Prediction Algorithms

Contents

1.3.1	Combining Experiments and Numerical Simulation	30
1.3.2	Formation Enthalpies and Convex Hull Representation	31
1.3.3	CSP Algorithms, General Principles	34
1.3.4	The Ab-Initio Random Structure Searching Algorithm	35
1.3.4.1	Generation Rules	36
1.3.4.2	The Structural Optimization Process	38
1.3.5	Other Algorithms	39
1.3.5.1	Particle Swarm Optimization	39
1.3.5.2	Genetic Algorithm	40
1.3.5.3	Metadynamics	43
1.3.6	CSP Algorithms: Improvement Ways	45
1.3.6.1	Representation of the Exploration Space and PES	45
1.3.6.2	Clustering	45
1.3.6.3	Machine Learning Potentials	45
1.3.7	Summary of Part 1	46

This chapter presents how numerical simulation and in particular Crystal Structure Prediction (CSP) algorithms are useful tools for the study of superhydrides. As a first section, the importance of combining CSP algorithms with experimental results is explained. The second section discusses how to evaluate the stability of a hydride and superhydride from the formation enthalpy and the convex hull representation. The third section focuses specifically on the CSP algorithms used to search and predict the crystal structures of materials. Finally, different Crystal Structure Prediction algorithms are detailed in the next sections.

1.3.1 Combining Experiments and Numerical Simulation

Most superhydrides investigated and synthesized up to now are made of a single metallic atom type mixed with hydrogen (i.e. binary superhydrides). The pressures required for their synthesis (i.e. the pressures at which these compounds are thermodynamically stable) are usually in the range of 100 GPa (i.e. 1 million times the atmospheric pressure). Such extreme conditions are nowadays reachable experimentally by using diamond anvil cells. However, as already explained in Chap 1.1, it should be interesting to stabilize superhydrides at low or ambient pressure (for hydrogen storage applications for instance). A possibility might be the use of more complex alloys made of at least two different metallic atom types. Indeed, the investigation of ternary superhydrides has recently begun [83, 84]. Consequently the experimental synthesis of new superhydrides is an active research area by itself. However some difficulties need to be addressed in order to explore and identify efficiently the synthesized hydrides.

- First of all, the experimental conditions such as the high pressure at which the synthesized hydrides are stable can be challenging to reach. When they are reached, the experimental characterization of the synthesized material is not always straightforward. It necessitates to perform X-ray diffraction, and/or vibrational spectroscopy experiments while maintaining high pressure conditions. In addition some structural parameters can not be obtained from experimental means. For instance hydrogen being a light element, it weakly interacts with X-rays. As a result it is impossible to obtain any information about the hydrogen sublattice from a X-ray diffraction spectrum.
- In the case of the investigation of ternary hydrides and superhydrides, mixing two metals A and B with hydrogen may lead to a very large number of compounds, including binary hydrides (AH_x , BH_y), in addition to the desired ternary ones ($A_xB_yH_z$). The experimental identification of all these potential phases – that may appear as mixed with each other – may thus be very difficult.

To address these issues, numerical simulation can be an useful tool. It may be used to put on the track of the most promising materials and can complete the experimental informations. More precisely, the numerical simulation currently used in the present field consists in implementing Crystal Structure Prediction algorithms (CSP). Such algorithms are emerging thanks to progress of High Performance Computing (HPC), the development of softwares able to predict with a high

accuracy the physical properties of materials (ab-initio calculations), and the surge of Artificial Intelligence (AI).

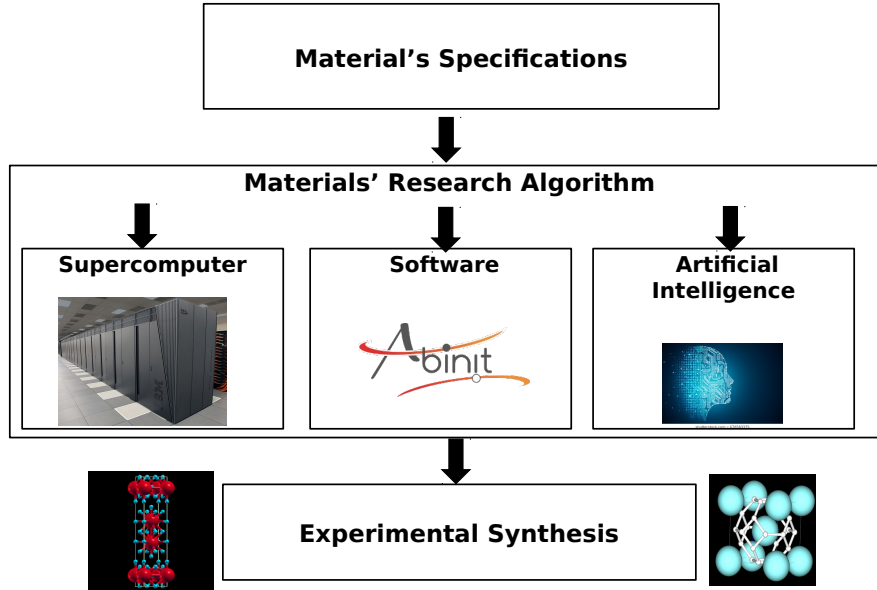


Figure 1.3.1: Predictions of materials using numerical simulation.

1.3.2 Formation Enthalpies and Convex Hull Representation

For studies of superhydrides, the hydrogen concentration of the thermodynamically stable compounds are a priori unknown, several hydrogen stoichiometries in the considered alloy have to be tested by CSP algorithms. The relative stabilities between the predicted structures are then obtained thanks to the convex hull representation. It allows to compare relative stabilities of hydrides and superhydrides made of the same alloy but having different hydrogen stoichiometries. It first consists in plotting the formation enthalpy (per atom) of the different hydrides as a function of hydrogen concentration. These formation enthalpies are calculated with respect to two reference systems, the first one being the pure metal (denoted as A) or the hydride having the lowest stoichiometry considered (denoted as AH_y), and the second one pure hydrogen in the solid phase in which it is the most stable at the considered pressure. One needs to make sure in addition that this reference hydride is thermodynamically stable. The formation enthalpies of the different hydrides are plotted for a fixed pressure. For a hydride with chemical formula AH_x (with $x > y$), the hydrogen concentration is deduced from the formation reaction $AH_y + \frac{x-y}{2} H_2 \rightarrow AH_x$:

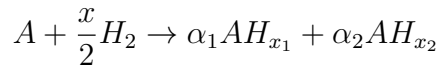
$$\frac{x - y}{1 + (x - y)} \quad (1.3.1)$$

The number of supplementary hydrogen between AH_y and AH_x is divided by the number of elementary compounds (AH_y and H). The formation enthalpy h is given by:

$$h\left(\frac{x-y}{1+(x-y)}\right) = \frac{\mathcal{H}(AH_x) - \mathcal{H}(AH_y) - \left(\frac{x-y}{2}\right)\mathcal{H}(H_2)}{C(AH_x)} \quad (1.3.2)$$

with $C(AH_x) = 1 + (x - y)$ (see Annexe B), \mathcal{H} referring to the enthalpy per formula unit. After plotting these enthalpies as a function of hydrogen concentration, a convex hull can be determined from the set of points representing each considered compound. Those on the convex hull correspond to the materials which are thermodynamically stable: they can be formed at the considered pressure and if the hydrogen concentration is appropriate.

More explicitly, suppose the pure metal A is under a hydrogen atmosphere at a given pressure with a hydrogen to metal molar ratio H/A equal to x . The convex hull plotted for the corresponding pressure points out which compounds are formed when hydrogen diffuses into the alloy. If AH_x belongs to the convex hull, it is thermodynamically stable and can be formed, provided all the kinetic barriers to the formation of the hydride can be overcome. If AH_x does not belong to the convex hull, it is not thermodynamically stable at the considered pressure. Then two compounds (denoted by AH_{x_1} and AH_{x_2}) will be formed (instead of one) which correspond to the two nearest composition to AH_x belonging to the convex hull. In one of them, the hydrogen concentration is lower than in AH_x , while it is higher in the other one. Their proportions α_1 and α_2 are then obtained by equilibrating the equation:



In the case of large hydrogen excess (which often corresponds to the experimental conditions in diamond anvil cells), the (super)hydride being on the hull having the highest hydrogen stoichiometry is formed; the hydrogen not used for the (super)hydride synthesis then remains under the form of pure H_2 . FIG 1.3.2 is an example of a convex hull representation for the manganese (super)hydrides under a pressure of 147 GPa. The two reference compounds here are MnH and H_2 in phase C2/C-24 [85]. The points representing MnH_2 , MnH_3 and MnH_7 are on the convex hull. They are thus the (super)hydrides which are thermodynamically stable at 147 GPa with the appropriate hydrogen stoichiometry. For instance for a stoichiometry $x = 2$ such that $\frac{x-y}{1+(x-y)} = \frac{2-1}{1+(2-1)} = 0.5$, the hydride MnH_2 is expected to form. On the contrary, Mn_2H_9 is not on the hull (it is represented by the point with $\frac{x-y}{1+(x-y)} = 0.77$), it is expected to decompose in MnH_3 and MnH_7 , even if the system is initially prepared with a hydrogen to metal molar ratio equal to $\frac{9}{2}$. Note that all the compounds for which the formation enthalpy is plotted have to be dynamically stable (i.e. their phonon dispersion curves at the considered pressure must not exhibit any instability). In other words, the crystal structures considered are stable with respect to any collective atomic displacement. A compound AH_x for which the formation enthalpy does not belong to the convex hull is said to be metastable (i.e. it is not the most stable state of the system $A+H$ with molar ratio $H/A=x$ but it is dynamically stable).

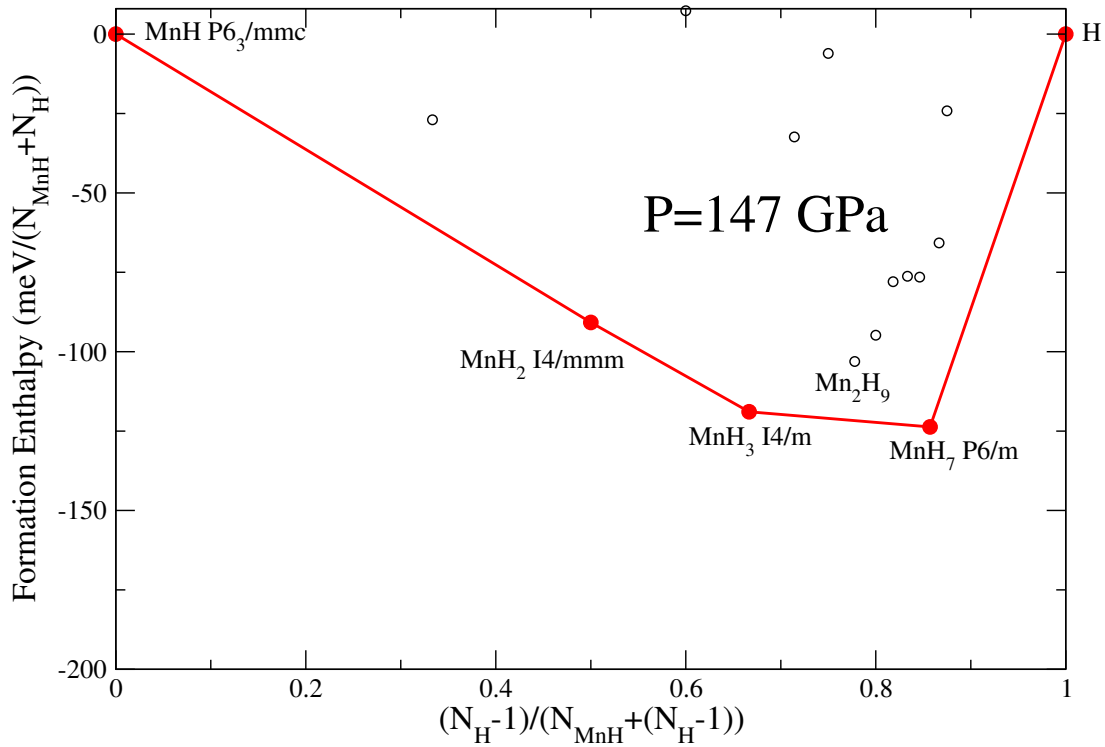


Figure 1.3.2: Formation enthalpies of the manganese hydrides at 147 GPa computed with respect to MnH and H₂ in phase C2/C-24 (the commonly accepted candidate for phase III of solid hydrogen [85]). The solid line (convex hull) joins the points corresponding to thermodynamically stable compounds. The open circles represent metastable compositions at this pressure.

As mentioned in 1.1.2, a superhydride candidate for hydrogen storage application must satisfy the following necessary conditions:

- Stable or at least metastable under low pressure conditions (reachable by industrial means).
- Release its hydrogen with a low energetic cost.

When the pressure goes to zero, the corresponding point needs to have a negative formation enthalpy (negative second coordinate see Sec 1.3.2) but close enough to zero to avoid paying too much energy for the hydrogen release.

After presenting how numerical studies can be combined with experiments to guide the synthesis of new superhydrides, and describing how formation enthalpies are calculated and how they can be used to determine the stable compositions, it is now time to focus on the tools involved for the predictions of materials. The next sections describe more in depth the Crystal Structure Prediction (CSP) algorithms employed for this purpose.

1.3.3 CSP Algorithms, General Principles

Crystal Structure Prediction (CSP) algorithms are numerical tools used for the search of new compounds. They can be employed under various conditions. These numerical tools must be combined with a method to compute the physical properties of the materials (energy, pressure), which is in practise often DFT. The output of such an algorithm is a crystal structure which can be described by a unit cell.

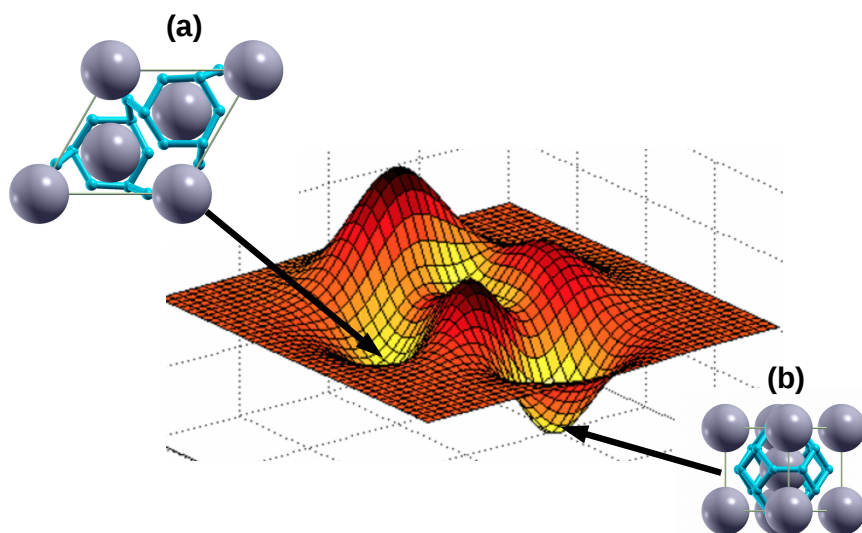


Figure 1.3.3: The PES surface (source code to plot the surface [86]). In this very simple example, (a) corresponds to a local minimum (metastable structure) while (b) corresponds to the global minimum.

A CSP algorithm aims at solving an optimization problem consisting in finding the degrees of freedom (atomic positions and cell vectors) of a given system, that minimize its energy (or its enthalpy if external pressure is applied). From a more geometrical point of view, a CSP algorithm explores a multi dimensional surface. The dimension of this surface corresponds to the number of degrees of freedom to determine. This surface is called the Potential Energy Surface (PES): it is the Born-Oppenheimer surface of the system with the electron gas considered as being in its ground state. Each point of the PES corresponds thus to a crystalline configuration of the system (atomic positions and cell vectors) with a given energy or enthalpy (which is the relevant quantity if the search is conducted under pressure). The most stable configuration is located at the global minimum (see FIG 1.3.3). The aim of the strategies chosen in a CSP algorithm is to find the global minimum and thus retrieve the most stable crystal structures.

1.3.4 The Ab-Initio Random Structure Searching Algorithm

For the flexibility it offers, the Ab-Initio Random Structure Searching (AIRSS) method is the CSP algorithm used in this PhD project. This method was developed by C. Pickard and R. J. Needs [87]. AIRSS revealed successful in the prediction of the crystal structure of several systems under pressure, for instance that of phase III [85], and Phase IV [88] of dense hydrogen, that of aluminum at Terapascal pressures [89] and that of ammonia under pressure [90]. The method was also successfully employed to predict the high pressure phases of silane SiH_4 [91] and aluminium AlH_3 hydride [92]. AIRSS was also able to predict new and more stable nitrogen phases under various pressure ranges as previously thought [93]. Last but not least, this method has been employed to predict the first superhydrides at high pressure, namely the lithium superhydrides [17]. These predictions were again confirmed by experiment later [18].

Given all these accurate predictions reached by this approach, this algorithm is the one chosen in this PhD for the search of new superhydrides. AIRSS is particularly well suited when it comes to impose constraints coming from experimental informations. It can thus be easily employed in the context of material research combining numerical simulation and experimental synthesis.

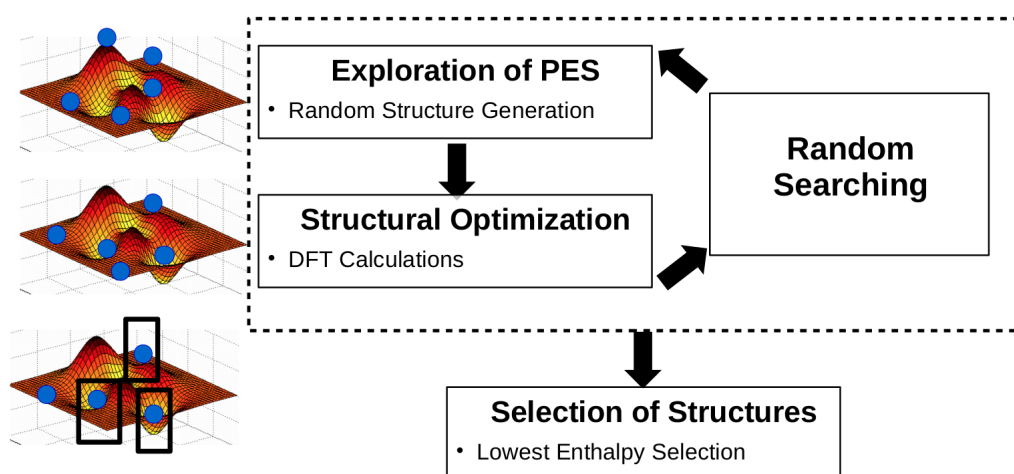


Figure 1.3.4: Main steps involved in AIRSS.

This section describes more in depth the principles used to implement it. AIRSS consists in (see FIG 1.3.4):

- Sampling the PES by randomly generating structures, which can be viewed as independent configurations of the studied system. This generation follows some rules (see Sec 1.3.4.1) aiming at focusing the search on the most relevant regions of the PES.

- Performing a structural optimization for each generated structure by using Density Functional Theory (DFT) computations. It can be viewed as moving the generated structures to the nearest local minimum of the PES.

In practise, the crystal structure of a compound with chemical formula AH_x at a given pressure is searched. The search is performed for a given number of formula units in the unit cell, starting from one and increasing this number as much as possible (typically up to four here). The number of degrees of freedom and thus the dimension of the configuration space increases with the number of formula units. Since the number of local minima exponentially increases with the dimension of the configuration space [87], the exploration becomes more and more inefficient as this dimension increases. This is why the present algorithm works efficiently for unit cells typically up to ~ 20 -30 atoms.

1.3.4.1 Generation Rules

The first step consists in a random generation of atomic configurations. However this generation can not be totally random: there are rules employed to generate the initial crystal structures. These rules are called research biases. In this subsection some of them are described more in depth.

Minimum Distance Conditions

Configurations with atoms too close to each other have to be avoided. Indeed, such configurations are not only unphysical, but also would probably lead to erroneous results when used as input of a DFT calculation. As a result a first important heuristic is to prevent two given atoms from being separated by a too small distance. To this aim minimum interatomic distances need to be fixed which do not have to be exceeded. These distance conditions depend on the chemical bonds involved. For instance a reasonable minimum distance between two hydrogen atoms can be 0.7 \AA whereas bonds between hydrogen and a metallic atom do not have to be below 1.5 \AA . Several strategies are considered to make sure that those distance conditions are fulfilled:

- A passive strategy consisting in checking after the generation of each new configuration whether all the distance conditions are fulfilled. If they are not, the current configuration is replaced by another try. This is performed until reaching a configuration fulfilling all the distance conditions. Despite the simplicity of its implementation, this method suffers however from a limitation when the number of atoms in the unit cell increases, the number of random tries can be time consuming if the distance conditions are not softened. As a consequence, a trade-off has to be found to keep both reasonable distance minimum conditions and computational cost to generate the images.
- An active strategy consisting in using repulsive empirical potential to separate the atoms whose initial positions are too close to each other (for instance ZBL potential see D). In this case very soft conditions can be initially imposed on the minimum distances. An empirical

potential is applied afterwards. Forces are computed from this potential to move apart the atoms from each other until stricter minimum distance conditions are reached. It is worth mentioning that the number of optimization steps performed by using this repulsive empirical potential needs to be limited. Otherwise, there is a risk to fall into a local or global minimum associated with this repulsive potential, and finally obtain different configurations too similar to each other.

Imposing Chemical Environments

Another heuristic can be the imposition of chemical environments (i.e. for example A surrounded by B and B surrounded by A). It can be used on the basis of experimental data or from the knowledge of similar systems as the one studied. Such bias is applied by imposing, for each atom, the chemical types of the nearest neighbour atoms, (i.e. the first coordination shell). This kind of constraint makes sense for binary systems in which ionocovalent or ionic bonds are expected (ionic crystals, ionocovalent crystals). In practise, it can be imposed to each atom of a specific chemical type (for example the metal) to be surrounded – in its first coordination shell – by atoms of another chemical type only (for instance hydrogen). In this situation, spherical shells with minimum and maximum radius centered on each atom of the the first chemical type are determined. The positions of atoms of the second chemical type are then randomly chosen inside these spherical shells.

Fixing a Sublattice

In some cases, the atomic positions of one sublattice are already known, and only those of another one need to be found: the AIRSS algorithm can then be utilized to determine the atomic positions of this sublattice. This situation arises for instance with hydrides and superhydrides. The experimentalists are indeed able to find the positions of the metallic atoms, those of the hydrogens remaining unknown (due to their very weak interactions with X rays). In this situation, the metallic atom positions can be fixed and the AIRSS algorithm is used to find positions of the hydrogen atoms only. Practically, the metallic atoms are positioned at their experimental positions, and the positions of hydrogen atoms are randomly drawn, with of course the constraint to fulfill the minimum distance conditions described above. However, both positions are afterwards structurally optimized (see Sec 1.3.4.2).

Imposing a Geometry of the Unit Cell

Some informations related to the unit cell of the crystal can be obtained thanks to experimental studies before using the AIRSS algorithm. In some cases part of the geometry of the unit cell may be available from experiment (e.g. its volume, angles between the primitive vectors or lattice parameters). These parameters can be imposed and only the other degrees of freedom are randomly chosen before structural optimization. Such situation can occur when experimentalists derive from X-ray diffraction experiments geometry candidates but are not able to find the atomic

positions. Another useful piece of information the experimentalists can retrieve is also symmetry group candidates which can be used to constrain both unit cell geometries and atomic positions.

1.3.4.2 The Structural Optimization Process

After generating a configuration by using one or several research biases, structural optimization is performed to minimize the energy of the crystal structure (or its enthalpy if the structural optimization is performed at fixed pressure). The goal of this step is to lead each random configuration to the closest local minimum on the PES. At the end of this process, when all the random configurations have been structurally optimized, we thus obtain a set of crystal structures that correspond to all the local minima of the PES that have been explored, each one being candidate to be the crystal structure adopted by the system at the considered pressure. The selected structure (i.e. the one predicted by the AIRSS as the structure of the system) is the one having the lowest energy (or enthalpy if the search is performed at non-zero pressure). In the present work, this minimization utilizes the Broyden-Fletcher-Goldfarb-Shanno (BFGS) [94] as implemented in the *abinit* code.

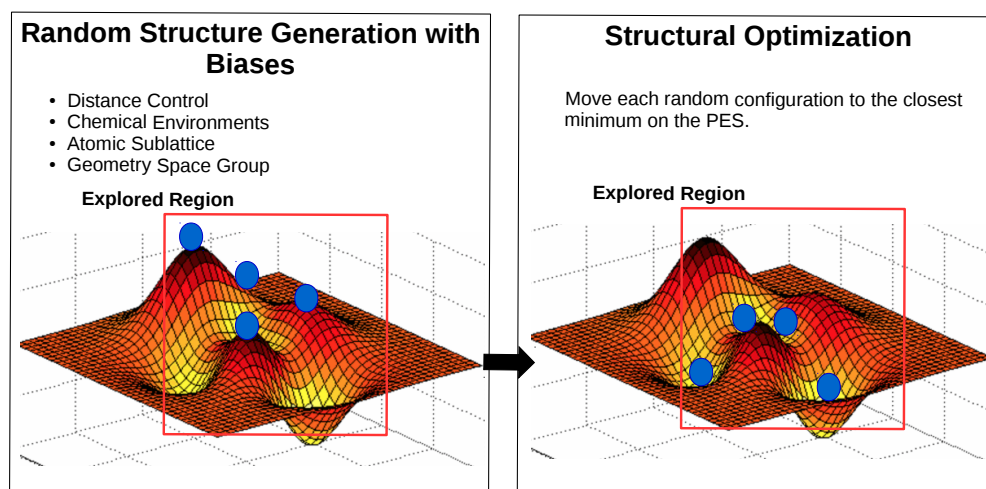


Figure 1.3.5: Random search biases and structural optimization. The biases enable to restrain the PES exploration to the most interesting regions.

We now give some details about this BFGS algorithm. It belongs to the family of quasi-Newton methods. It consists in finding iteratively the minimum of a function (here the potential energy or the enthalpy if the minimization is performed at fixed pressure) by following a direction determined by the gradient and the Hessian matrix (second derivative matrix) of the potential energy/enthalpy. However, if in DFT the gradient (opposite of the forces) can be easily calculated thanks to the Hellmann-Feynman theorem, this is not the case of the hessian, which would require either to perform finite-difference calculations or the use of Density Functional Perturbation Theory (DFPT).

The BFGS algorithm then proceeds by using an approximation of the hessian, calculated on the basis of the successive gradients. The energy (or enthalpy) being a function of all the degrees of freedom describing the unit cell, the BFGS algorithm computes iteratively atomic displacements and cell deformation leading to the local minimum of the energy/enthalpy which is the closest from the initial configuration. The k -th step of the BFGS algorithm is performed as follows :

- Compute the displacement vector D_k , by solving the equation $B_k D_k = -\nabla E(x_k)$ with E the potential energy (or enthalpy), which is a function of all the degrees of freedom of the unit cell (noted by x_k). B_k is a numerical approximation of the Hessian matrix, obtained at the end of the previous iteration.
- Find a displacement rate α_k along the D_k direction such that α_k minimizes the function in α , $E(x_k + \alpha D_k)$.
- Evaluate the quantities $s_k = \alpha_k D_k$, $x_{k+1} = x_k + s_k$ corresponding to the update of the degrees of freedom, and $y_k = \nabla E(x_{k+1}) - \nabla E(x_k)$.
- Compute an approximation of the Hessian matrix that will be used for the $k + 1$ step:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k} \quad (1.3.3)$$

1.3.5 Other Algorithms

This section briefly presents alternative algorithms to AIRSS which are commonly used in the search of new materials, like hydrides and superhydrides at high pressure. In particular, Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Metadynamics are described.

1.3.5.1 Particle Swarm Optimization

In the Particle Swarm Optimization approach (PSO), unlike the AIRSS method, the notion of evolution is used. A population of crystal configurations evolves through successive operations combining the informations of the "best" structures reached. The evolution history of each configuration is thus taken into account for the evolution of the whole population. For a given crystal structure, it integrates [95]:

- The best known position it has reached during its former evolution steps.
- The best known position reached formerly among all the configurations in the population.

Here the term "best known position" refers to a configuration having the lowest energy (or enthalpy) value. These informations are combined with random numbers to obtain displacements (for atoms and lattice vectors of the unit cell). These displacements are applied to the crystal

structures to give new configurations (see Eq. 1.3.4 FIG 1.3.6). For this purpose a metric has to be defined to compute distances between structures. If $D_i^{(p)}$ is the distance between the structure i and its best known configuration it has reached beforehand, if $D_i^{(g)}$ is the distance from the best known configuration reached among all the structures in the population, the following displacements are computed [96]:

$$\delta'_i = \delta_i + \omega_p \alpha_p D_i^{(p)} + \omega_g \alpha_g D_i^{(g)} \quad (1.3.4)$$

δ_i is a vector gathering the displacements to apply to all the degrees of freedoms. In this expression ω_p and ω_g are weight vectors imposed for the efficiency of the method. α_p and α_g are random numbers. As for AIRSS, at each evolution step, the obtained crystal structures are subjected to DFT structural optimizations. This evolution process can be moreover complemented by random generation of structures.

Such a method is implemented, see for instance [97] in the CALYPSO code. This code was for instance used to study yttrium and some rare earth hydrides and superhydrides [98, 99]. These studies highlighted the "clathrate" structures the H atoms form in some of these materials.

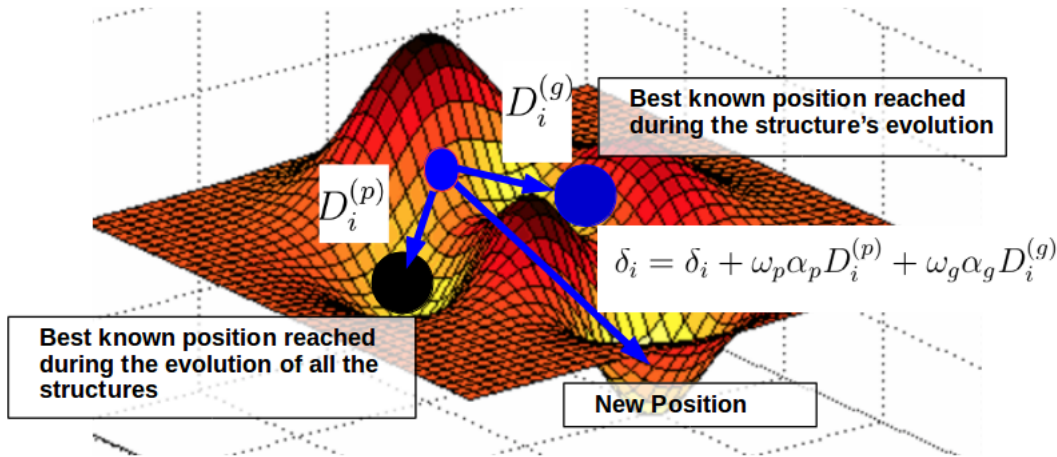


Figure 1.3.6: Evolution process within the framework of the PSO algorithm.

1.3.5.2 Genetic Algorithm

Similarly as PSO, Genetic Algorithm (GA) method includes the notion of evolution for the process of structure generation. Genetic Algorithm is made of two main steps (see FIG 1.3.7):

- A population initialization where structures are randomly generated.
- A heredity process enabling evolution of the initial population by selecting the candidates with the best properties from one generation to another. With this selection, it is expected

that part of the structures will converge to the most stable one corresponding to the energy or enthalpy minimum.

Several codes have been implemented on the basis of this method like the USPEX [96] or GATOR [100] codes.

Initialization of the population

To initialize a Genetic Algorithm, a population of configurations has to be randomly generated. The generation of these structures is a critical part of the algorithm, one needs indeed to make sure that the generated configurations are different enough from each other. Otherwise, they risk converging at the end of the algorithm to the same local minimum, missing thus others and potentially the global one. This generation step is more crucial as it is for the AIRSS algorithm because the population is generated once and all the obtained configurations will come from it. To make sure that the initial structures are separated enough between each other, a possibility is to generate atomic positions by randomly choosing symmetry groups among the 230 possible ones. Fixing different symmetry groups from one structure to another makes possible the generation of very different atomic crystal configurations.

Fitness Coefficient

After initializing the population of the structures, evolution steps are performed. Before making any heredity operation, the structures are optimized so that they move to a local energy or enthalpy minimum. Each configuration of the population is afterwards subjected to heredity operators enabling the evolution of the ensemble of the structures. These heredity operations are enforced on the basis of parent configurations which are selected using a fitness parameter. A common way to determine it is to use the energy (or enthalpy under pressure). It can be defined as [100]:

$$f_i = \frac{\epsilon_i}{\sum_j \epsilon_j} \quad (1.3.5)$$

where the sum runs over all the configurations and the ϵ coefficients are defined as

$$\epsilon_i = \frac{E_{max} - E_i}{E_{max} - E_{min}} \quad (1.3.6)$$

with E_{max} and E_{min} being the maximum and minimum energy or enthalpy values reached in the population (at the current evolution step) and E_i the energy of the considered structure.

Selection Strategies

After computing for each configuration the fitness parameter (Eq. 1.3.5 and Eq. 1.3.6), several strategies can be utilized to select parents [100]. These parent structures are combined together thanks to heredity operations (see next subsection) to give new ones. The main rules enforced for these selections are the following:

- straightforward selection, it consists in selecting only a determined number of the best configurations in terms of fitness parameter.
- roulette wheel selection, a random number r between 0 and 1 is chosen. A parent structure is selected, it corresponds to the one having the first fitness parameter for which $f_i > r$.
- tournament selection, a pool of configurations is selected randomly in the population. A determined number of structures having the highest fitness parameters in the pool is selected.
- clustering (see Sec 1.3.6) can be employed to gather the configurations in clusters following a metric. In each considered cluster, the best structures in terms of fitness parameter are selected as parents.
- PSO selection strategy, the best configuration reached for a given evolving structure C_1 and the best one obtained among the whole population during the evolution steps C_2 can be taken into account. In this process, probabilities to determine whether a structure will be combined with one of these best configurations or not is determined by the distances to them, respectively D_p and D_g . For instance in Ref [96], a probability P_p to be combined with C_1 , P_g with C_2 and P_m with none of these best configurations can be defined:

$$P_p = \frac{\psi_p r_p D_p}{\Sigma} \quad (1.3.7)$$

$$P_g = \frac{\psi_g r_g D_g}{\Sigma} \quad (1.3.8)$$

$$P_m = \frac{\omega}{\Sigma} \quad (1.3.9)$$

Here ω, r_p and r_g are random numbers, Σ is a normalization coefficient,

$$\Sigma = \omega + \psi_p r_p D_p + \psi_g r_g D_g \quad (1.3.10)$$

and ψ_p, ψ_g are weights.

Heredity Operations

From the selected parents, child crystal structures can be created by applying some heredity operations. As examples of heredity processes, it can be mentioned:

- combining slabs from different parents [96], it corresponds to cross pieces of the crystal structures of parents to generate children. In each parent structure, specific atomic environments can be selected on the basis of indicators [96].

- mutations, random atomic displacements or primitive cell deformations can be applied on a selected structure.
- soft-mutation [96]: the eigen vibration modes are computed on the structure and the atoms are then displaced following these modes.

After applying the heredity operations, the new configurations have to be checked to assess whether they have been already explored beforehand and have physical meaning. The structures succeeding in these checks are then subjected to structural optimization and will take part to the next heredity step.

Stopping Criteria

There is no systematic way to evaluate whether a genetic algorithm is converged. Some stopping criteria can be considered. An accurate one is to stop when a sufficient amount of unchanging structures from one heredity step to another is obtained.

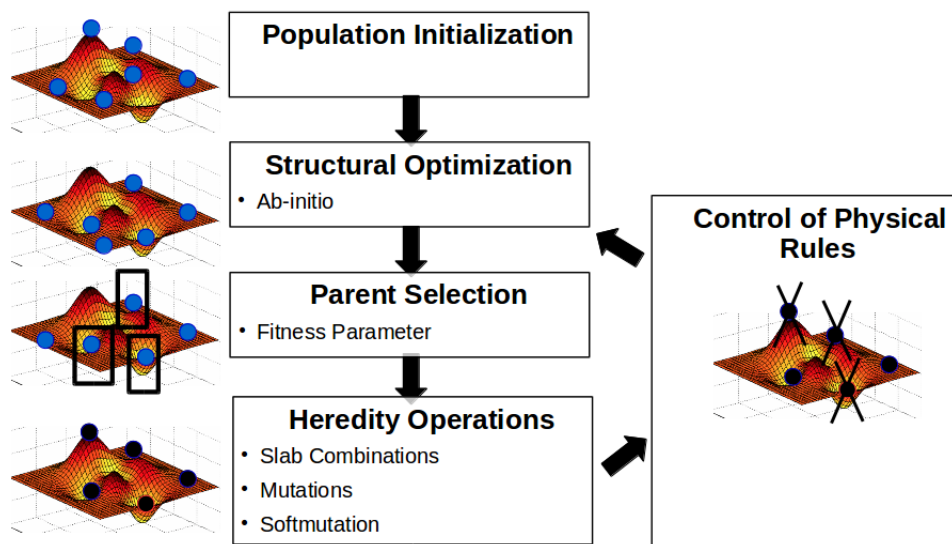


Figure 1.3.7: Work-flow of a Genetic Algorithm.

GA algorithms and in particular the USPEX software made important contributions in the search of new superhydrides. As examples it can be mentioned the works conducted by Oganov et al of iron superhydrides [24, 26] or actinide superhydrides [35, 36, 37].

1.3.5.3 Metadynamics

Metadynamics refers to a method enabling the exploration of all the minima of the PES by a given evolving crystal structure. This exploration is performed thanks to dynamics steps where the atoms move towards unexplored positions. These steps are performed so that the crystal

structure does not remain stuck in a local minimum of the PES and eventually escapes from it along the dynamics. To make this possible, collective variables have to be introduced. A given crystal structure is thus associated with specific values of them [101, 102]. On the basis of these variables, a repulsive potential is introduced and adapted through the evolution of the structure on the PES. This potential is used to derive forces which are applied to atoms for the dynamic steps. It can have the form:

$$V(P) = \sum_{P_i} \alpha(P_i) f_{P_i}(S_1(P), \dots, S_n(P)) \quad (1.3.11)$$

In Eq 1.3.11, P is the current position of the crystal structure on the PES surface where dynamic steps are enforced. The P_i s refer to the positions on the PES surface already explored by the crystal structure and the S_j s are the global variables used. $\alpha(P_i)$ are coefficients which ponderate the contribution of each f_{P_i} functions. The f_{P_i} functions are usually gaussians [103] centered at the values of the collected variables obtained at the position P_i .

$$f_{P_i}(S_1(P), \dots, S_n(P)) = \exp\left(-\sum_j \frac{(S_j(P) - S_j(P_i))^2}{2\sigma_i^2}\right) \quad (1.3.12)$$

Where in Eq 1.3.12 the sum runs over all the considered collective variables. By applying this potential, it prevents the atoms from reaching positions they have already explored. Unlike AIRSS, PSO and GA, Metadynamics makes possible the exploration of all the minima of the PES surface. However this exploration might require an unreasonable amount of time.

1.3.6 CSP Algorithms: Improvement Ways

CSP algorithms are continuously improving so as to reach the predictions of more and more complex materials. The following section sums up the main ways of improvement considered nowadays. The first one is focusing on new strategies to describe and represent the spaces where the structures are explored. The second way is investigating new machine learning based clustering tools to explore and identify the most interesting structures and atomic environments. Finally, the third way is considering the development of machine learning based interatomic potentials.

1.3.6.1 Representation of the Exploration Space and PES

So as to avoid being trapped in local minima and converge more efficiently to the low energy configurations, Pickard proposed in [104] a new approach consisting in optimizing structures in a hyperspace. This method is presented in the Geometry Optimization of Structure form Hyperspace (GOSH) framework [104]. During the structure relaxation the atoms can move in a space having more than three dimensions. It gives to the system more degrees of freedom to avoid it being stuck by potential energy barriers which might be difficult to go through if the atoms were moving in a 3d dimensional space. Other ways of improvement are aiming at representing the PES after exploring it by AIRSS. For instance, in a recent study of Lu et al [105], Stochastic Hyperspace Embedding method is employed to search new $\text{Li}_x\text{Cu}_y\text{F}_z$ compounds for batteries. This method aims at building a 2d representation of the PES surface.

1.3.6.2 Clustering

Machine learning based clustering methods such as k-means, affinity propagations [106] or agglomerative hierarchical clustering [107] are considered to generate and select the structures through the CSP processes. Like enforced in the GATOR code, these tools can be used to assess whether the generated structures at the beginning of an evolutionary cycle are not too similar. They can also be considered for the selection rules applied to select structures and make thus sure that the ensemble of selected structures is gathering sufficiently different atomic environments [107].

1.3.6.3 Machine Learning Potentials

So as to speed up the structure explorations and investigate more complex phases, machine learning is also considered to develop interatomic potentials. Such an approach is for instance employed with AIRSS. For instance Deringer et al [108] searched for phosphorus phases using a gaussian process based potential called Gaussian Approximation Potential (GAP). In a recent study to search for new Li-Si based materials for batteries, deep learning using Graph Neural Networks (GNN) is also combined with AIRSS like in the work of G. Cheon et al [109]. GNN is used in particular to evaluate forces and stress. The advances using machine learning potentials

are not limited in the context of AIRSS. Genetic algorithms are also playing important roles in these new strategies. For instance, USPEX uses machine learning based potentials called Moment Tensor Potentials (MTP) to accelerate the structure relaxations. These potentials are used in Ref [110] with an implemented active learning process to update the training databases when structures too far from what the potential has learned are explored. Such a method was used to predict complex structures of carbon and boron with more than 100 atoms. In addition a new 54-atom phase of boron is proposed. Similar approach is followed to study uranium phases [111]. Recently, an extension of USPEX called T-USPEX has been proposed in Ref [112] to predict structures at finite temperature. In this framework, structure relaxations and molecular dynamics used to compute the free energies are enforced with machine learning interatomic potentials. P-T diagrams of pure compounds like aluminium and iron are performed, as well as those of more complex systems like MgSiO_3 . Last but not least, PSO type code like CALYPSO are considering the use of machine learning potentials like Gaussian Approximation Potential (GAP). Such an approach has been for instance considered to investigate clusters of boron like B_{36} or B_{40} [113, 114]. More details about machine learning based potentials will be given in Part 3 of this PhD thesis.

In this PhD thesis, the use of machine learning interatomic potentials will be considered. These new tools will be combined with AIRSS so as to reach the prediction of more complex superhydrides. In addition, modern clustering methods employing machine learning will be investigated so as to construct active learning strategies to improve the machine learning potentials through the structure exploration process.

1.3.7 Summary of Part 1

The superhydrides are newly discovered materials, stable under high pressure and which are characterized by a hydrogen concentration which is well above that the conventional chemical rules allow (typically a H/M ratio above five). These compounds may have a high potential for the energy sector, provided these systems could be synthesized near ambient pressure. They could lead to a breakthrough for energy storage under hydrogen form as well as for the search of high temperature superconductivity. Due to the high pressure required nowadays for their synthesis, inherent difficulties related to the presence of hydrogen, and the important number of candidates to test, numerical simulation can bring a lot to the experimentalists. For this purpose, Crystal Structure Prediction (CSP) algorithms to search the crystal structures of new materials under pressure are employed. These algorithms are combined with Density Functional Theory (DFT) to calculate the PES of the system. Thanks to the success it has yet brought and the flexibility it offers to include experimental informations, the Ab-Initio Random Structure Searching (AIRSS) algorithm is considered in this PhD thesis. Using this method, the next part describes studies performed for the hydrides of some transition metal elements. These investigations aim at evaluating the possibility to form superhydrides of these elements under pressure.

PART 2

TRANSITION METAL HYDRIDES AND SUPERHYDRIDES STUDIED BY AB-INITIO RANDOM STRUCTURE SEARCHING

Introduction

This second part gathers three studies performed with the Ab-Initio Random Structure Searching (AIRSS) method to predict superhydride systems. More precisely, compounds made of transition elements (located between the IIIA and IB columns of the periodic table) are considered. These transition metal elements are worth being investigated because they could form superhydrides having high hydrogen densities (for example, FeH_5 [12] has been synthesized recently, and CrH_8 [115] has been predicted). Furthermore, a great part of the transition metal elements are not rare (at least those belonging to the first series). This could thus facilitate future potential applications. In order to explore the potentialities of transition metals as superhydride forming elements, three of them with very different electronegativities have been chosen. First, a transition element having a relatively high one, copper (Cu: 1.9 on the Pauling scale) is considered. Metals with large electronegativity are not expected to form stable hydrides but the hydriding process could possibly benefit from the high-pressure conditions. It is followed by another one with intermediate electronegativity, manganese (Mn: 1.55 on the Pauling scale). Finally a last one of relatively low electronegativity, yttrium (Y: 1.22 on the Pauling scale) is studied.

Chapter 2.1

Copper Hydrides

Contents

2.1.1 State of the Art	49
2.1.2 Computational Details	50
2.1.2.1 Ab initio Random Structure Searching	50
2.1.2.2 Density-Functional Theory Calculations	51
2.1.2.3 Special Quasi Random Structure	52
2.1.3 Results: Copper Hydrides Under Pressure	52
2.1.3.1 Hydrides at Ambient Pressure	52
2.1.3.2 Hydrides under Pressure	53
2.1.3.3 Dynamical Stability	56
2.1.4 Physical Properties	56
2.1.4.1 Crystal Structures	57
2.1.4.2 Bader Charges	58
2.1.5 Conclusion	60

This chapter presents an AIRSS study about copper hydrides so as to evaluate the possibility to form superhydrides of this element. In the first section, the state of the art regarding the different investigations achieved up to now about this system is presented. The second section discusses the AIRSS predictions as well as the dynamical stability of the predicted compounds. Finally, the last section deals with the physical properties of the predicted stable hydrides.

2.1.1 State of the Art

Copper belongs to the group 11 of the periodic table, with Silver and Gold, two noble metals. It has an electronegativity of 1.9 on the Pauling scale [116, 117], a little smaller than that of hydrogen (2.2) and very close to those of Ni (1.91), Co (1.88) and Ag (1.93). Note that the electronegativity of Au (2.54) is larger than that of H, and thus should not permit the formation of Gold hydrides. Among those elements, Ag, Au and Co do not form any stable hydride at ambient pressure, while under high pressure, only Co and Ni have been experimentally confirmed to form stable hydrides (CoH_2 above 45 GPa [118], Ni_2H_3 above 52 GPa [119] or 60 GPa [120]). Thus, copper is not expected to form stable hydrides at ambient pressure. It is not expected also to form very hydrogen-rich hydrides under high pressure.

Indeed, copper is known as being a "*non hydride forming element*" at ambient pressure [3]. The first copper hydride that has been synthesized at ambient pressure, $w\text{-CuH}$, with a wurtzite structure [121], is thermodynamically unstable, and tends to decompose into copper and gaseous dihydrogen. Its formation enthalpy, i.e. the enthalpy of the reaction $\text{Cu}(fcc) + \frac{1}{2}\text{H}_2 \rightarrow w\text{-CuH}$ is positive (+27.5 kJ/mol [122]). It is also unstable at room temperature and decomposes above -60° [123] and below 290 K [124]. $w\text{-CuH}$ has also been experimentally studied under pressure [125, 126], where it is observed to progressively decompose [127].

However, under moderate pressure, two other forms of copper hydrides have been evidenced [128]: (i) γ -hydrides, which retain the FCC lattice of copper, and in which hydrogen is inserted in the octahedral sites; (ii) ϵ -hydrides, in which the FCC lattice of copper is lost, to become a HCP lattice. The $\gamma\text{-CuH}_x$ are thus interstitial hydrides, whereas ϵ -hydrides are not.

The $\gamma\text{-CuH}_x$ have been synthesized under pressure with several hydrogen stoichiometries, ranging from $x=0.15$ to $x=1$. The first $\gamma\text{-CuH}_x$ was synthesized in a Diamond Anvil Cell (DAC) above 14.4 GPa by Burtovyy and Tkacz [129], with a hydrogen stoichiometry $x \sim 0.4$. The compound with $x=0.15$ (possessing a FCC lattice only slightly expanded with respect to Cu [130]) was observed by Donnerer et al [131], and is denoted in the literature as $\gamma_0\text{-CuH}_{0.15}$ [130], that with $x=0.4\text{-}0.5$, as $\gamma_1\text{-CuH}_{0.5}$, while that with $x=0.65$, recently synthesized by Binns et al above 30 GPa [130], has been called $\gamma_2\text{-CuH}_{0.65}$. This phase is also reported by Meier et al [132]. Finally, the phase obtained by a complete filling of all the octahedral sites of the Cu FCC lattice, $\gamma\text{-CuH}$, has been synthesized only recently by Meier et al [132], and requires a synthesis pressure higher than 110 GPa.

The ϵ copper hydride is known under the stoichiometry Cu_2H , with space group $P - 3m1$,

and was synthesized above 18.6 GPa by Donnerer et al [131] and above 35 GPa by Meier et al [132]. It has the same structure as Mn_2H [133, 134], that consists in a HCP lattice of copper, in which half of the interstitial octahedral sites of the HCP lattice are filled by hydrogen atoms. This structure is also known as the *anti-CdI₂* structure [131].

Recently, the copper hydrides have been numerically studied under high pressure using evolutionary algorithms and the Ab Initio Random Structure Searching (AIRSS) method by Xiao et al [135]: beyond the already known $\epsilon\text{-Cu}_2\text{H}$ ($P-3m1$ space group) and $\gamma\text{-CuH}$ ($Fm-3m$ space group, NaCl structure) phases, these authors predict the formation of CuH_3 ($C2/m$ space group) at 160 GPa. They found that $\epsilon\text{-Cu}_2\text{H}$ becomes thermodynamically stable at ~ 16.8 GPa, in agreement with the experiments of Donnerer et al [131], and is the only thermodynamically stable phase at 40 GPa, while $\gamma\text{-CuH}$ appears on the convex hull in the 100 GPa region, in agreement with the synthesis pressure reported by Meier et al [132],

Also, an anomalous phase of copper hydride has been observed during magnetic field-assisted electrodeposition of copper [136], which was interpreted as a combination of CuH and CuH_2 , suggesting that copper dihydride could exist under a metastable form: while the CuH phase was assigned to be the wurtzite structure (*w*- CuH), the authors suggested that the observed metastable structure of CuH_2 could be a body-centered tetragonal arrangement (space group $I4/mmm$, with lattice constants $a = 2.775 \text{ \AA}$ and $c = 3.925 \text{ \AA}$).

The AIRSS study presented in this chapter aims thus at confirming all these results and to determine whether or not copper can form hydrogen rich compounds and/or superhydrides at high pressure.

2.1.2 Computational Details

This section describes the different frameworks, methods and numerical parameters employed for the structure exploration. The way AIRSS is implemented is firstly presented. In a second time the numerical parameters involved in the DFT computations are discussed. This is followed by a discussion about the use of the Special Quasi Random Structure framework.

2.1.2.1 Ab initio Random Structure Searching

For each stoichiometry x , with x varying from 1 to 8, AIRSS was performed using one, two, three and four formula units of CuH_x in the cell. Also, the possibility of half-integer hydrogen stoichiometries is explored, Cu_2H_x , with x being an odd integer from 1 to 3, using unit cells with 1 and 2 formula units. For each considered stoichiometry and number of f.u., a set of initial atomic configurations is randomly generated and then structurally optimized by DFT. The random generation concerns atomic positions, lattice constants and cell angles. The volume distribution of these initial cells is around a reference volume (that depends on the number of f.u. and hydrogen stoichiometry), chosen close to that already known of the iron hydrides (which can be justified remarking that the ionic radius of Cu is close to that of Fe). Mainly for practical reasons, a bias

on initial interatomic distances is enforced in the random generation, to avoid initial configurations having unphysically short bond lengths. The number of structural optimizations performed for each considered stoichiometry and number of f.u. is typically a few hundred, varying from ~ 100 up to more than 1000. These structural optimizations are performed using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, and consist in an optimization of both atomic positions and cell vectors to match a chosen hydrostatic pressure. The scheme is applied at 50, 100 and 150 GPa, with in some cases several thousands of structures randomly generated and optimized (i.e. for each stoichiometry, pressure and number of formula units). The results presented here can therefore be considered as predictive for pressures up to 150 GPa. At the end of each structural optimization, the final enthalpy is computed, which allows identifying the most stable structure by comparing all the enthalpies obtained for a given hydrogen stoichiometry. In practice, the structures having a slightly higher enthalpy (typically 50 meV/atom) are also checked. In most cases, the most stable structure appears several times along the AIRSS procedure. The space group and Wyckoff positions of the selected structures are determined using the *findsym* program [137], [138]. For the random generation, the non-deterministic random number generator of R. Chandler and P. Northrup [139] is used.

2.1.2.2 Density-Functional Theory Calculations

The Density-Functional Theory (DFT) calculations are performed with the *abinit* code [140], within the Projector Augmented-Wave (PAW) formalism [81, 46]. The exchange-correlation energy functional is the Generalized Gradient Approximation as formulated by Perdew, Burke and Ernzerhof [55] throughout this chapter. The PAW atomic datasets for Cu and H are taken from the JTH table [141].

The structural optimizations at the AIRSS stage are performed using the BFGS algorithm as implemented in *abinit* as explained in Chap 1.3. They use a $6 \times 6 \times 6$ k -point mesh (and $4 \times 4 \times 4$ for the largest cells) to sample the Brillouin Zone associated with the simulation cell. A plane-wave cut-off of 15 Hartrees is employed. In order to take into account the possible formation of metallic phases, gaussian like smearing is employed with a smearing temperature of 0.0025 Ha. The optimization criterion on atomic forces is 2×10^{-3} Ha/bohr (and 2×10^{-5} Ha/bohr³ on stress tensor components). The AIRSS structural optimizations are also limited to at most 75 ionic steps. Then, the structures identified as candidates for the CuH_{*x*} compounds are re-optimized using the BFGS algorithm, at several pressures between 0.0005 Ha/bohr³ (~ 15 GPa) and 0.005 Ha/bohr³ (~ 150 GPa), using more stringent numerical parameters: a plane-wave cut-off of 30 Hartrees, a $12 \times 12 \times 12$ k -point mesh to sample the Brillouin Zone associated with the simulation cell, and an optimization criterion on atomic forces of 1.0×10^{-6} Ha/bohr (and 1×10^{-8} Ha/bohr³ on stress tensor components). Finally, in order to question the dynamical stability of the copper hydride phases, we also perform Density Functional Perturbation Theory (DFPT) calculations and plot the phonon dispersion curves of the phases in question (Cu₂H and CuH). These calculations are performed using a dense $24 \times 24 \times 24$ k -point mesh for CuH and for Cu₂H.

2.1.2.3 Special Quasi Random Structure

We have seen in the introduction that copper hydrides CuH_x with $x < 1$ may exist under moderate pressure under the form of interstitial hydrides, the H atoms occupying randomly the octahedral interstitial sites of a FCC lattice of copper ($\gamma\text{-CuH}_x$). It is thus important to include in the present work the possibility of such disordered systems. In addition the thermodynamical stability of disordered phases was taken into account thanks to the Special Quasi Random Structure (SQS) framework [142]. SQS are particular arrangements of the atomic species of different kinds in a finite-size supercell that reproduce as much as possible the spatial correlations of a perfect infinite random alloy. For this purpose, the formation enthalpy of disordered $\text{CuH}_{\frac{n}{16}}$ with n an integer between 1 and 15 was considered here and compared to that of other (ordered) crystal structures (CuH , $\text{Cu}_2\text{H}\dots$), in order to determine whether such disordered systems are stable or not. Due to the fact that pure Cu adopts a FCC structure, the atomic positions of these SQSs are taken as the ones of a FCC random binary alloy as presented in Ref [143]. In this article, the authors provide atomic positions for SQS of an A_xB_{1-x} alloy within a FCC lattice and a 32 atom supercell. In the interstitial CuH_x hydrides, the H atoms randomly occupy interstitial octahedral sites. Since the octahedral sites of a FCC lattice also form a FCC lattice, the SQS of Ref [143] can be used to model the interstitial CuH_x hydrides, with $\text{A}=\text{H}$ and $\text{B}=\emptyset$ (void).

2.1.3 Results: Copper Hydrides Under Pressure

This section describes the predictions made by AIRSS. For the computation of the formation enthalpies, the reaction between pure FCC Cu and hydrogen is considered. Under pressure and up to 117 GPa the solid hydrogen $C2/c$ -32 phase (see appendix E) is taken into account, and above, the $C2/C$ -24 one [85]. In addition to the crystal structures determined by AIRSS and the SQSs representing the disordered systems, we also considered the wurtzite structure of CuH ($w\text{-CuH}$, P63mc space group), which is the experimental known structure of CuH at ambient pressure, as explained in the introduction. This structure can be described as a HCP lattice of one of the two chemical species, in which the tetrahedral sites are occupied by the other chemical species.

2.1.3.1 Hydrides at Ambient Pressure

At ambient pressure, we did not perform AIRSS. Thus, in addition to the interstitial hydrides $\gamma\text{-CuH}_x$ and the wurtzite phase of CuH, the crystal structures obtained at 50 GPa for CuH, Cu_2H_3 and CuH_2 are considered. At 50 GPa, the lowest enthalpy structure found by AIRSS for CuH is the cubic NaCl type structure (Fm-3m). However, we found that at ambient pressure copper is not hydrided even in the form of a disordered phase (see FIG. 2.1.1). This result is coherent with the experimental investigations [132]. At 0 GPa, neither of the two candidate phases for CuH (the

wurtzite structure and the cubic rocksalt structure) is stable. It has been checked (see FIG 2.1.2) the wurtzite phase is more stable than the cubic one below 25 GPa, which is in agreement with the experiments [122].

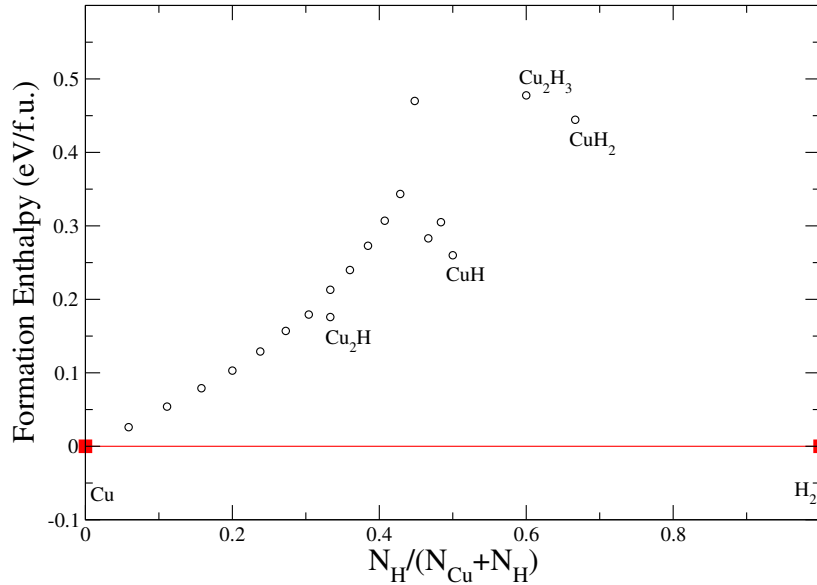


Figure 2.1.1: Formation enthalpies of interstitial disordered copper hydrides and other crystalline solids as a function of hydrogen content at 0 GPa.

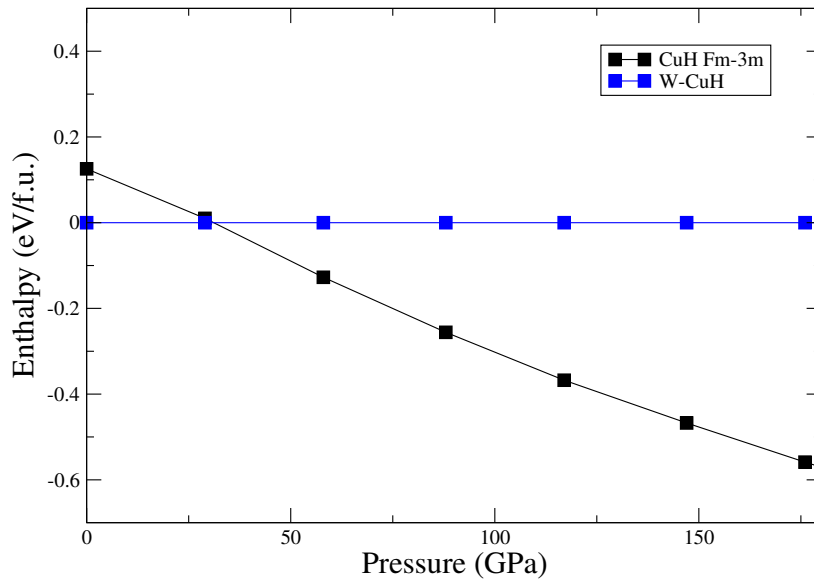


Figure 2.1.2: Stability comparison between the CuH wurtzite phase and the cubic one.

2.1.3.2 Hydrides under Pressure

The thermodynamic stability of copper hydrides is studied under pressure from 29 to 205 GPa (see FIG 2.1.3 and FIG 2.1.4). In line with experimental studies [130], Cu₂H (space group P-3m1)

is stable near 30 GPa. It remains on the convex hull at least up to 205 GPa. The Fm-3m CuH phase reaches the convex hull at higher pressure near 60 GPa. These predictions are in line with the high pressure experimental study achieved in Ref [132]. At least up to 205 GPa, no phase with hydrogen stoichiometry higher as one is expected to be stable. In contrast to the predictions of Ref [135], CuH₃ is not found thermodynamically stable. It means copper is probably not a system for which superhydrides might be obtained, at least up to ~ 200 GPa.

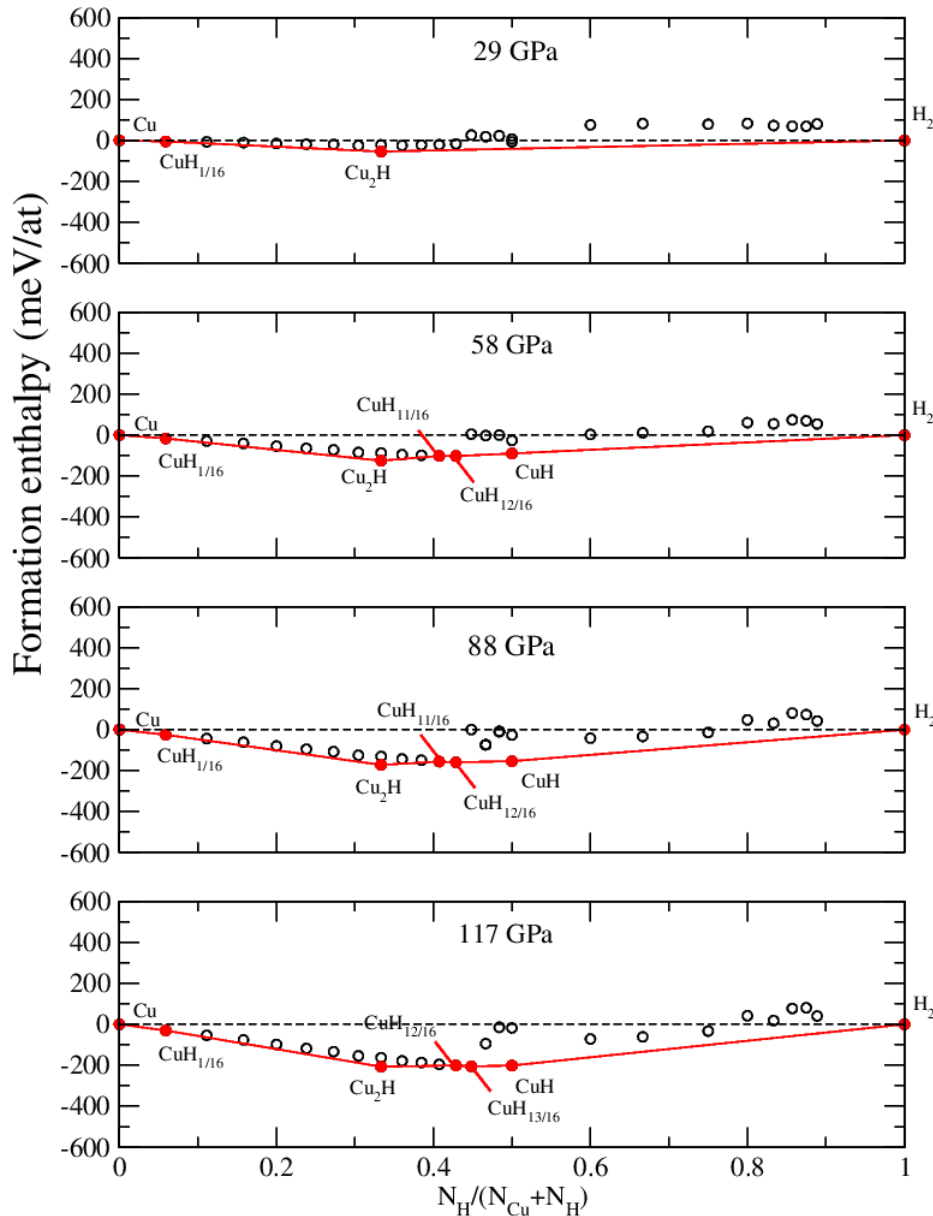


Figure 2.1.3: Formation enthalpies of CuH_x under pressure (29-117 GPa) including crystal structures and SQS phases.

The study of disordered systems (CuH_x with $x < 1$) reveals that some of them are stabilized under relatively low pressure. The formation enthalpies of the SQSs are distributed over a line between Cu and CuH. Increasingly richer hydrogen compounds are stabilized when pressure

increases. Indeed, the SQS $\text{CuH}_{\frac{1}{16}}$ starts becoming stable near 20 GPa, followed by the stabilization of $\text{CuH}_{\frac{11}{16}}$ and $\text{CuH}_{\frac{12}{16}}$ from 50 to 120 GPa with also $\text{CuH}_{\frac{13}{16}}$ near 120 GPa. From 180 GPa, it is expected these SQS phases are replaced by $\text{CuH}_{\frac{14}{16}}$ and $\text{CuH}_{\frac{15}{16}}$ on the convex hull. Note that the fact $\text{CuH}_{\frac{11}{16}}$ (i.e. $\text{CuH}_{0.69}$) is stable near 50 GPa is in line with the experimental observation of $\gamma_2\text{-CuH}_{0.65}$ in this pressure range. Indeed in Ref [130] it is pointed out the formation of $\text{CuH}_{0.64}$ in this pressure range ($\gamma_2\text{-CuH}$). Note, however, that $\gamma_1\text{-CuH}_{0.5}$ is not observed to be stable within the SQS framework, since the ϵ phase Cu_2H is always found more stable than the disordered system $\text{CuH}_{0.5}$. Also, any equivalent of $\gamma_0\text{-CuH}_{0.15}$ is found within the SQS framework. The fact that these experimental stoichiometries are not found stable when modeled by SQS phases may indicate that particular spatial correlations (not reproduced by SQSs) are present in the structures of the experimentally observed phases. We conclude that, at least up to ~ 200 GPa, the hydriding of copper is not favored by pressure beyond a H-to-metal ratio of one. CuH in its rocksalt structure appears as the most hydrided phase of copper that can be expected in this pressure range.

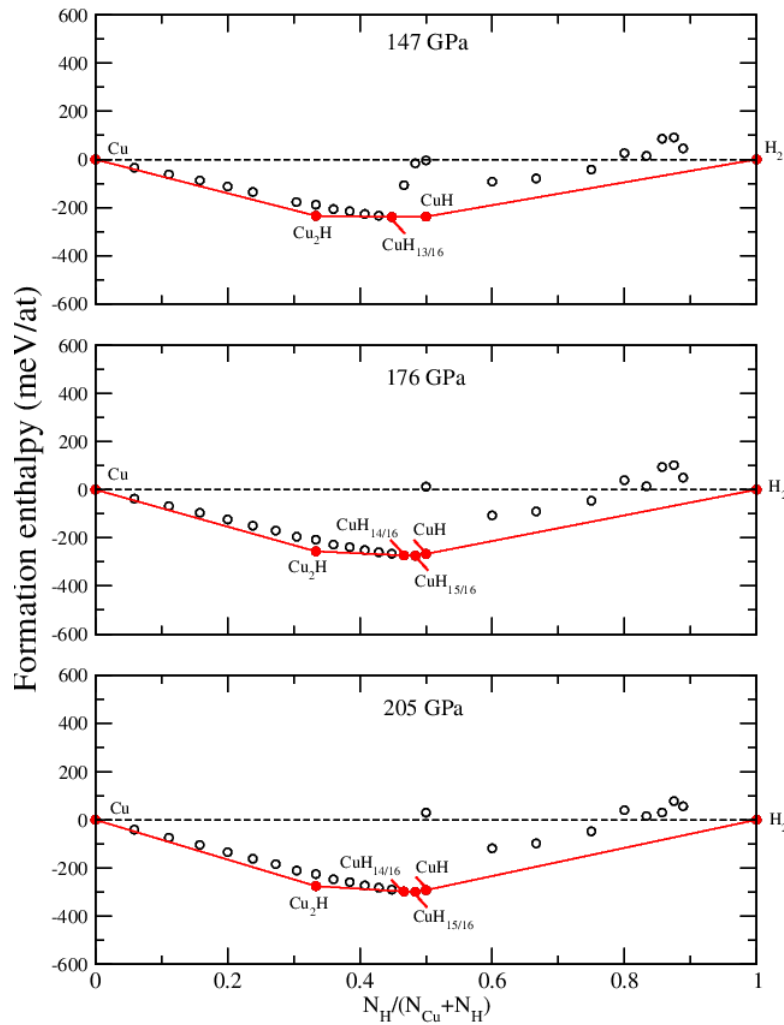


Figure 2.1.4: Formation enthalpies of CuH_x under pressure (147-205 GPa) including crystal structures and SQS phases.

2.1.3.3 Dynamical Stability

The dynamical stability of the two crystalline hydrides predicted stable under pressure (i.e. Cu_2H and CuH) is now questioned. For this purpose DFPT calculations were performed at 0, 58, 117 and 205 GPa. It is found that the Cu_2H phase is dynamically stable at the four pressures. The CuH Fm-3m phase is dynamically unstable at 0 GPa. At 58 GPa however, it is stable. FIG 2.1.5 displays the phonon dispersion curves of both Cu_2H and CuH at 0 and 117 GPa. For both hydrides, interestingly, the phonon dispersion curves are splitted in two parts. The first one, with the lowest frequencies, corresponds to vibrations of the copper sublattice. The second part, at higher frequencies to vibrations of the hydrogen one. This splitting is, of course, due to the large difference of mass between the metal and hydrogen, and is very common to hydrides with a low hydrogen-to-metal ratio [115]. FIG. 2.1.5 also illustrates the hardening of the vibration modes with increasing pressure.

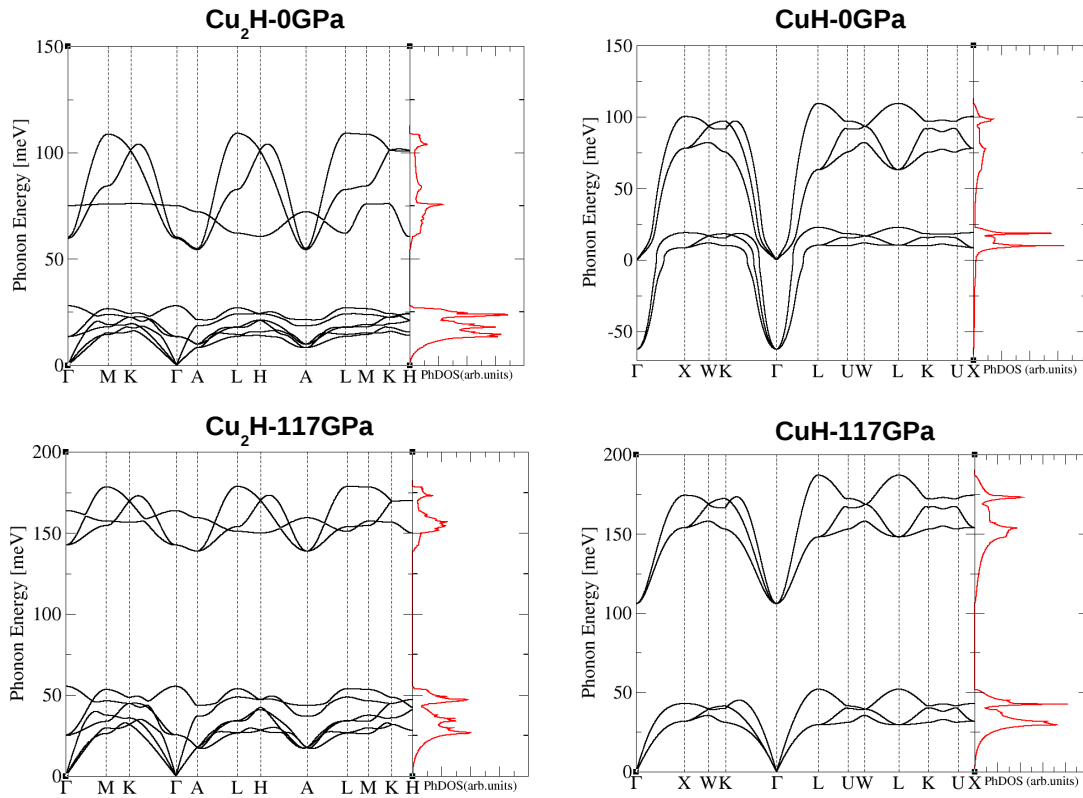


Figure 2.1.5: Phonon dispersion curves of Cu_2H P-3m1 and cubic CuH at 0 and 117 GPa.

2.1.4 Physical Properties

In this section some physical properties of Cu_2H and CuH are examined. First of all a crystal structure description of Cu_2H and CuH is made. It is followed by a discussion about the charge

transfer between Cu and H atoms using the Bader analysis.

2.1.4.1 Crystal Structures

The crystal structures of Cu_2H (a), CuH Fm-3m (b) and CuH P6₃mc (wurtzite (c)) are presented in FIG 2.1.6. Tabs 2.1.1, 2.1.2 and 2.1.3 provide the Wyckoff positions and cell parameters for the three hydrides at 117 GPa. All these compounds are interstitial type hydrides. Starting from a HCP arrangement of the Cu atoms, Cu_2H is obtained by filling with H atoms half of the octahedral sites in the (0001) planes. It is the same crystal structure as the one of Mn_2H . The CuH cubic phase is obtained by starting from the FCC structure of pure Cu and filling all the octahedral sites (rocksalt structure). Finally, in the wurtzite phase, one can recognize HCP arrangements of Cu atoms. The H atoms occupy tetrahedral sites.

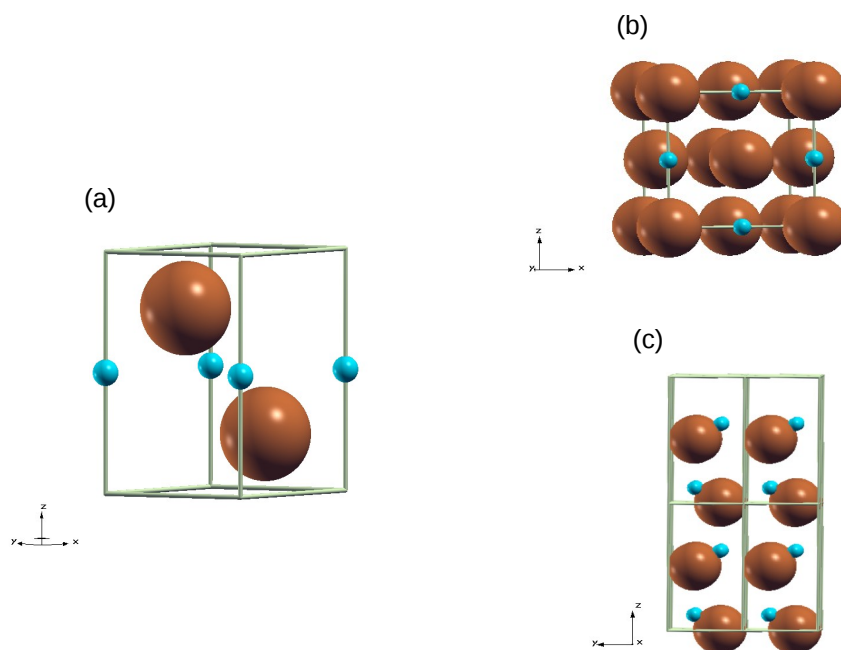


Figure 2.1.6: View of the Cu_2H P-3m1 phase (a), cubic CuH (b) and wurtzite CuH (c) at 117 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
P-3m1	117 GPa	$a=b=2.347, c=3.876$ $\alpha = \beta=90, \gamma=120^\circ$	Cu H	2d 1b	$\frac{1}{3}$ 0	$\frac{2}{3}$ 0	0.76238 $\frac{1}{2}$

Table 2.1.1: Cu_2H structure at 117 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
Fm-3m	117 GPa	$a=b=c=3.433$ $\alpha = \beta = \gamma=90^\circ$	Cu H	4a 4b	0 $\frac{1}{2}$	0 $\frac{1}{2}$	0 $\frac{1}{2}$

Table 2.1.2: CuH cubic structure at 117 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
P6 ₃ mc	117 GPa	$a=b=2.4873, c=3.9903$ $\alpha = \beta = 90, \gamma = 120$	Cu H	2b 2b	$\frac{1}{3}$ $\frac{1}{3}$	$\frac{2}{3}$ $\frac{2}{3}$	0.50300 0.12200

Table 2.1.3: CuH wurtzite structure at 117 GPa

2.1.4.2 Bader Charges

The hydriding process implies an electron transfer from the metal to the hydrogen atoms. To characterize this electron transfer in CuH_x hydrides, Bader charge analysis [145] has been performed.

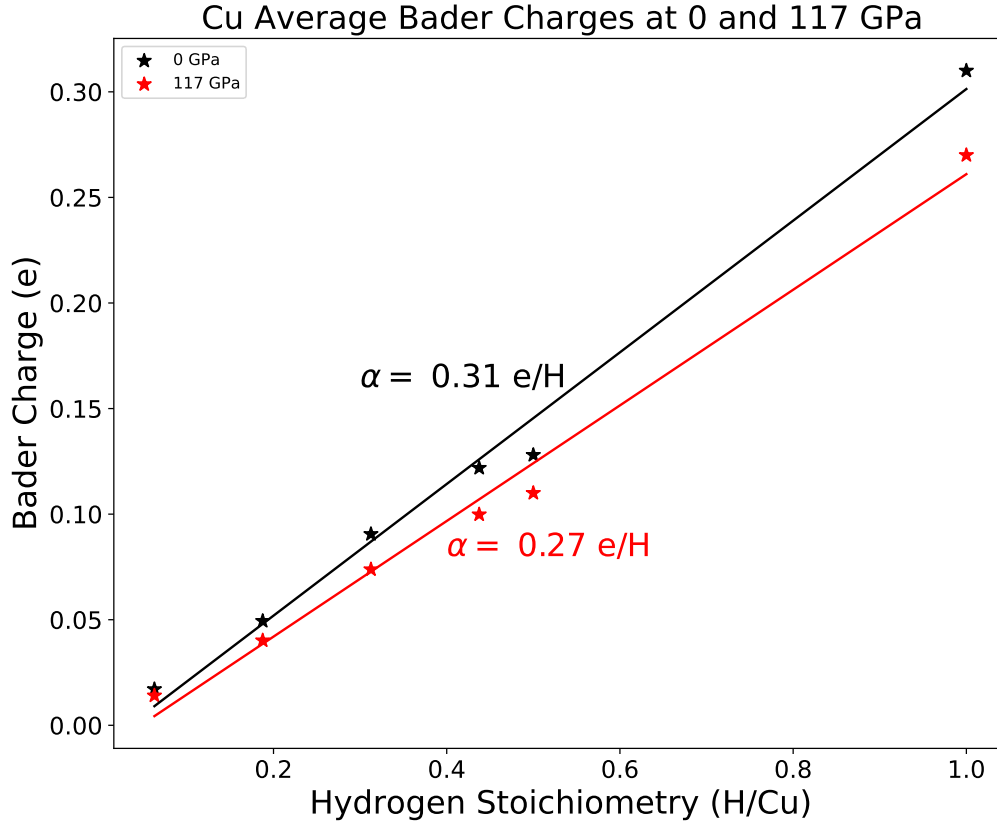


Figure 2.1.7: Average Bader charges on the Cu atoms at 0 and 117 GPa as a function of the hydrogen stoichiometry. The SQS phases $\text{CuH}_{\frac{1}{16}}$, $\text{CuH}_{\frac{3}{16}}$, $\text{CuH}_{\frac{5}{16}}$ and $\text{CuH}_{\frac{7}{16}}$ as well as the crystalline Cu_2H and CuH are considered.

FIG 2.1.7 presents the Bader charge averaged over the Cu atoms of the structure as a function of the hydrogen stoichiometry (number of H atoms by Cu). It can be observed that an electron transfer occurs from Cu to H atoms confirming that these compounds correspond effectively to hydrides (i.e. the metal Cu is positively charged, while hydrogen is negatively charged). For this calculation, the SQS phases $\text{CuH}_{\frac{1}{16}}$, $\text{CuH}_{\frac{3}{16}}$, $\text{CuH}_{\frac{5}{16}}$ and $\text{CuH}_{\frac{7}{16}}$ are considered as well as the crystal structures P-3m1 Cu_2H and CuH (wurtzite at 0 GPa and cubic at 117 GPa). This computation is performed at two pressures, 0 and 117 GPa. It can be seen that the average charge per Cu increases following a quasi linear behaviour with H stoichiometry. This quasi-linear increase can be easily interpreted assuming that each hydrogen atom catches a certain number of electrons to the copper sublattice. This number of electrons can be identified with the slope of the charge versus hydrogen concentration curve of FIG 2.1.7: it is ~ 0.31 electrons per H at 0 GPa and ~ 0.27 electrons per H at 117 GPa. It can be also remarked that the pressure tends to reduce the charge transfer.

2.1.5 Conclusion

This AIRSS study about copper hydrides is in line with most of the experimental studies concerning the stabilization of Cu_2H and cubic CuH at high pressure. In contrast to the results presented in Ref [135], the hydrogen stoichiometry reached does not exceed one at least up to ~ 200 GPa. It can be deduced that copper is not a system which favors the formation of superhydrides, as expected from the large electronegativity of this chemical element.

Chapter 2.2

Manganese Hydrides and Superhydrides

Contents

2.2.1 Manganese Hydrides, State of the Art	62
2.2.2 Computational Details	63
2.2.2.1 Ab initio Random Structure Searching	63
2.2.2.2 Density-Functional Theory Calculations	63
2.2.3 Results	64
2.2.3.1 Manganese Hydriding at Ambient and low Pressure	64
2.2.4 Dynamical Stability and Nuclear Quantum Effects	68
2.2.5 Physical Properties of the Stable Structures	72
2.2.5.1 Description of Crystal Structures	72
2.2.5.2 Bader Charges	77
2.2.5.3 Electronic Structure	79
2.2.5.4 Magnetism	81
2.2.6 Conclusion	81

This chapter presents some predictions obtained by the AIRSS algorithm on the manganese hydrides and superhydrides (MnH_x). Since this system has been well studied at low and ambient pressure, the aim is to extend its knowledge to the pressure regime currently and easily accessible to diamond anvil cells (i.e. to ~ 150 GPa). One of the main purposes is to predict whether novel forms of hydrided manganese, e.g. superhydrides, may be stable under high pressure. As a first section, the current knowledge about manganese hydrides at low pressure is presented. The second section presents the computational details of the study. The third section deals with the different predictions obtained from 0 to 150 GPa. The dynamical stability of the crystal structures found is then examined in the fourth section, as well as the contribution of the quantum Zero Point Energy (ZPE) associated with nuclear vibrations. Finally, in the fifth section, the predicted structures are described and complementary calculations (Bader charges, electronic band structures and magnetism) to characterize them are presented. Supplementary information related to the calculations is provided in appendix E.

2.2.1 Manganese Hydrides, State of the Art

Manganese is a $3d$ transition metal for which the hydriding properties have been investigated at ambient and low pressure only [148]. Manganese lies next to Fe in the periodic table, and is less electronegative (Mn: 1.55; Fe: 1.83; H: 2.2 on the Pauling scale of electronegativities [116, 117]), suggesting a larger electron transfer to hydrogen in Mn hydrides than in Fe ones. Indeed Fe has the capability to form hydrogen-rich compounds under high pressure, in which hydrogen is completely hydrided (i.e. not under a molecular form), at least up to FeH_5 [12]. Consequently, it is suggested that Mn could also form hydrogen-rich compounds, possibly superhydrides in similar conditions. At ambient pressure and up to $\sim 710^\circ\text{C}$ [152], Mn adopts a complex cubic-type structure, called α -Mn (space group $I\bar{4}3m$, 58 atoms per conventional cell) [158, 159]. Three other allotropic phases are reported with increasing temperature: the β -phase (also a complex cubic-type structure), the γ (FCC) phase and then the δ (BCC) phase [152].

Manganese hydriding was investigated in the past years at low and moderate pressures up to 30 GPa [148]. It was demonstrated that under pressure, hydrogen may be inserted in manganese, forming three types of hydrides MnH_x :

- ϵ hydrides [149, 150, 151, 152, 133, 148], having H-to-metal ratio x between 0.65 and 0.96 under pressure [153], in which the H atoms randomly occupy the interstitial octahedral sites of a HCP lattice of manganese;
- γ hydrides, with H-to-metal ratio x between 0 and 0.65 [153, 148], in which the H atoms randomly fill the interstitial octahedral sites of a FCC Mn lattice;
- for very small H-to-metal ratio $x \leq 0.1$, solid solutions of hydrogen in α - and β -Mn are also reported [152].

At moderate temperature and under pressure, formation of ϵ -MnH_{*x*} seems to be favored [152], while γ -MnH_{*x*} requires high temperature (e.g. above 900°C under 1.2 GPa according to Ref [154], above 325°C according to Ref [155]). Once formed, the HCP structure of ϵ -MnH_{*x*} (*x*=0.84) is maintained at least up to 30 GPa, while γ -MnH_{0.64} undergoes an irreversible transformation into the HCP phase with pressure [148]. This transformation from γ to ϵ was already reported by Filipek et al [156]. A double HCP phase (ϵ' -MnH_{*x*}) was also observed under pressure and at high temperature (\sim 800°C) [157]. Manganese hydrides have also been synthesized very recently under a molecular form and used to build an efficient hydrogen storage system [160]. Thus, up to now, only manganese hydrides MnH_{*x*}, with *x* < 1 have been reported in the literature, under low and moderate pressure.

2.2.2 Computational Details

This section presents the main work-flow and numerical tools employed to study the manganese hydrides and predict new structures under pressure. The main parameters involved in the implementation of the AIRSS algorithm and for the DFT computations are given.

2.2.2.1 Ab initio Random Structure Searching

In a similar way as the study of the CuH_{*x*} system 2.1.2, the AIRSS algorithm is implemented for different hydrogen stoichiometries. The compounds MnH_{*x*}, with *x* varying from 1 to 8, with one, two, three and four formula units (f.u.) per simulation cell are thus taken into account. Also, the possibility of half-integer hydrogen stoichiometries is explored, Mn₂H_{*x*}, with *x* being an odd integer from 1 to 15, using unit cells with 1 and 2 formula units. Searches have been done for pressures of:

- 50 GPa for MnH_{*x*} with *x* from 1 to 3, and Mn₂H_{*x*} with *x* odd integer from 1 to 7.
- 100 GPa for MnH_{*x*} with *x* from 1 to 8, and Mn₂H_{*x*} with *x* odd integer from 9 to 15.
- 150 GPa for MnH_{*x*} with *x* from 2 to 8.

The results presented here can therefore be considered as predictive for pressures up to 150 GPa.

2.2.2.2 Density-Functional Theory Calculations

For DFT computations, the same parameters as those used for CuH_{*x*} are applied (see Sec. 2.1.2). The dynamical stability of the predicted crystal structures for the manganese hydrides was tested by computing at several pressures the phonon dispersion curves. These calculations have been performed within the framework of the Density-Functional Perturbation Theory (DFPT), as implemented in *abinit* [146, 147]. For these calculations, the meshes used to sample the Brillouin Zone of MnH (primitive cell), MnH₂ (conventional cell), MnH₃ (conventional cell) and MnH₇

(primitive cell) are respectively $16 \times 16 \times 10$, $18 \times 18 \times 6$, $8 \times 8 \times 12$ and $8 \times 8 \times 12$ for the electronic DFPT calculation, and $8 \times 8 \times 5$, $9 \times 9 \times 3$, $4 \times 4 \times 6$ and $4 \times 4 \times 6$ for the phonon grid.

2.2.3 Results

This section describes the main predictions obtained with AIRSS. The results are presented by increasing pressure. First of all the stabilities of the manganese hydrides are studied at ambient pressure. In particular the temperature effects are taken into account. Intermediate pressure range is afterwards investigated (near 14 GPa). These investigations at moderate pressure enable to confront experimental results and predictions of the AIRSS method for the validation of the method. Finally the (super)hydrides appearing above 29 GPa are investigated.

2.2.3.1 Manganese Hydriding at Ambient and low Pressure

Hydriding at Ambient Pressure

As reported in Sec. 2.2.1, manganese hydride MnH_x has been synthesized in previous experimental studies up to $x=0.84$ [148]. The structure of this MnH_x was shown to be hexagonal closed-packed (ϵ phase, HCP, space group $P6_3/mmc$). The predictions of the AIRSS method for MnH are in agreement with this result, since the HCP structure frequently appears in the search and is found as the most stable one. MnH has been also synthesized at high temperature in the rocksalt structure (γ - MnH , space group $Fm\bar{3}m$). It has been checked, indeed, that this γ phase is less stable than the ϵ phase (this was tested up to ~ 150 GPa). Also, the double HCP phase (ϵ'), which is the low-temperature structure of FeH [161], is found less stable than the ϵ phase.

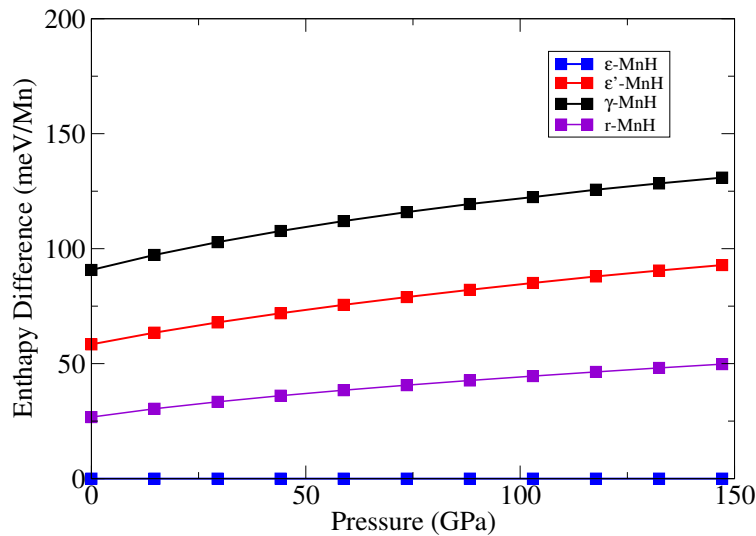
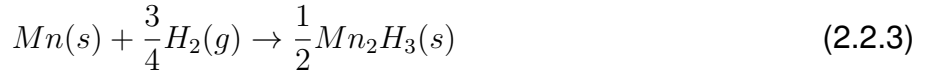
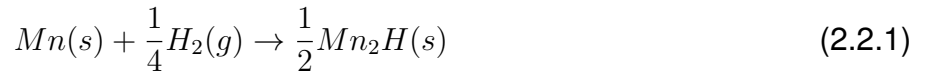


Figure 2.2.1: Enthalpy (in meV/Mn) of ϵ -MnH, ϵ' -MnH (dHCP phase), γ -MnH (FCC phase) and r-MnH (R-3m phase) as a function of pressure, taking ϵ -MnH as reference.

Note that the enthalpies of these three phases obey the order: $\epsilon < \epsilon' < \gamma$ (FIG 2.2.1), which is coherent with the experimental phase diagram proposed in Ref [157], where the dHCP phase has been observed at temperatures intermediate between those of ϵ and γ . Moreover, the AIRSS also revealed a structure with $R\bar{3}m$ space group having an enthalpy intermediate between that of ϵ and that of ϵ' (γ -MnH in FIG 2.2.1). The structure of ϵ -MnH is the same as that of CrH [115, 144].

In order to characterize Mn hydriding under ambient pressure conditions, Mn_2H and Mn_2H_3 are also considered, in the structures determined by the structural searches. Mn_2H is found to crystallize in space group $P\bar{3}m1$ (as Cu_2H [135]), and its structure can be described as that of an interstitial hydride, with the H atoms filling the interstitial octahedral sites of a Mn HCP lattice. Mn_2H_3 is found to crystallize in space group $C2/m$, as Cr_2H_3 at low pressure [115]. For pure Mn at zero pressure, no structural search is performed. The α phase is considered, which is the observed phase of Mn at ambient pressure and room temperature: it has a complex cubic structure with 58 atoms per primitive cell [158, 159]. First, the formation enthalpies ΔH_f of Mn_2H , MnH and Mn_2H_3 at zero pressure are computed, which are the energies of the following processes:



with all the solids optimized at zero pressure, and H_2 modeled as a single molecule in a large simulation box. It is obtained $\Delta H_f(Mn_2H) = -0.094$ eV/Mn, $\Delta H_f(MnH) = -0.135$ eV/Mn and $\Delta H_f(Mn_2H_3) = +0.103$ eV/Mn. The negative hydriding enthalpies reached for Mn_2H and MnH suggest these two compounds may be formed spontaneously in ambient pressure conditions and zero K. Mn_2H_3 , in contrast, can not be formed in such conditions.

At finite temperature however, it is necessary to consider the *free enthalpies* of formation of the hydrides, ΔG_f , which take into account thermal energetic and entropic effects. An approximation of ΔG_f is computed by adding to the total energy of the H_2 molecule the supplement of chemical potential, $2\Delta\mu_H$, as obtained assuming ideal gas behaviour [162].¹ By assuming the hydrogen gas pressure is $P_{H_2}=1$ bar, two numerical applications are performed, for $T=300$ K and 600 K (Tab 2.2.1). This corresponds to $\Delta\mu_H = -0.159$ and -0.380 eV, respectively. Thermal (mainly entropic) effects stabilize the gas. At 300 K, this results in a negative formation free enthalpy for Mn_2H , but positive for MnH, while at 600 K, all the free enthalpies are positive. This suggests that under $P_{H_2}=1$ bar and at room temperature, MnH is not stable, and only subhydrides with less hydrogen (such as Mn_2H) are able to form.

¹More precisely, the expression $2\Delta\mu_H(T, P_{H_2}) = H_0 - TS_0 + N_A c_p (T - T_0) - T c_p N \ln(\frac{T}{T_0}) + k_b N_A T \ln(\frac{P_{H_2}}{P^0})$ is used. In this relation $c_p = 3.5k_b$, $P^0 = 1atm$, $T_0 = 298K$, N_A is the Avogadro number, $H_0 = 8468$ J mol⁻¹, and $S_0 = 130.68$ J mol⁻¹ K⁻¹ [163].

Compound	T=0 K	T=300 K	T=600 K
Mn ₂ H	-0.094	-0.014	+0.096
MnH	-0.135	+0.024	+0.245
Mn ₂ H ₃	+0.103	+0.342	+0.673

Table 2.2.1: Formation free enthalpies ΔG_f (eV/Mn) of Mn₂H, MnH, Mn₂H₃ at ambient pressure and different temperatures. The chemical potential of the hydrogen gas is calculated assuming a hydrogen pressure $P_{H_2}=1$ bar.

Hydriding at Low Pressure

The hydriding process of Mn at ~ 14.7 GPa (0.0005 Ha/bohr³) is now investigated. At such pressure, the α phase of Mn is still used as reference for the computation of the formation enthalpies. The $C2/c$ -32 hydrogen phase (see appendix E) considered as reference for the solid phase of dense hydrogen. The formation enthalpies of the Mn hydrides found by structural searches are plotted on FIG 2.2.2. At such pressure, only Mn₂H and ϵ -MnH lie on the convex hull, suggesting that only hydrides MnH_{*x*}, with $x \leq 1$, can be formed, adopting interstitial structures with the H atoms filling the interstitial octahedral sites of a HCP lattice of Mn. It is worth highlighting that manganese hydrides have been synthesized under pressure up to 30 GPa: the hydride observed was ϵ -MnH_{*x*} (HCP structure), with $x=0.84$ [148]. Under pressure, the H-to-metal ratio in ϵ -MnH_{*x*} is reported to be lower than 1 [133]. This experimental result (i.e. only interstitial MnH_{*x*} compounds with $x \leq 1$ can be formed under low and moderate pressure conditions) is thus coherent with the present work. Note that experimental studies report the observation of disordered solutions, i.e. the H atoms are randomly distributed over the octahedral interstitial sites of the HCP lattice, at least for $x \geq 0.83$ [133]. This configurational disorder may be a cause of stabilization because of the associated configurational entropy. Such disordered states are obviously not accessible by AIRSS. At smaller H-to-metal ratio, however, e.g. for $x=0.65$, the H atoms in ϵ -MnH_{*x*} have been observed to presumably locate in one particular family of octahedral sites, namely within one (0001) plane among two. This was reported in Refs [133] and [135]: the H atoms can be viewed as forming a superstructure within the lattice of the octahedral sites, of the anti-CdI₂ type (space group $P\bar{3}m1$, the (0001) planes of octahedral sites filled with H alternate along the HCP axis). This exactly corresponds to the structure found with AIRSS for Mn₂H.

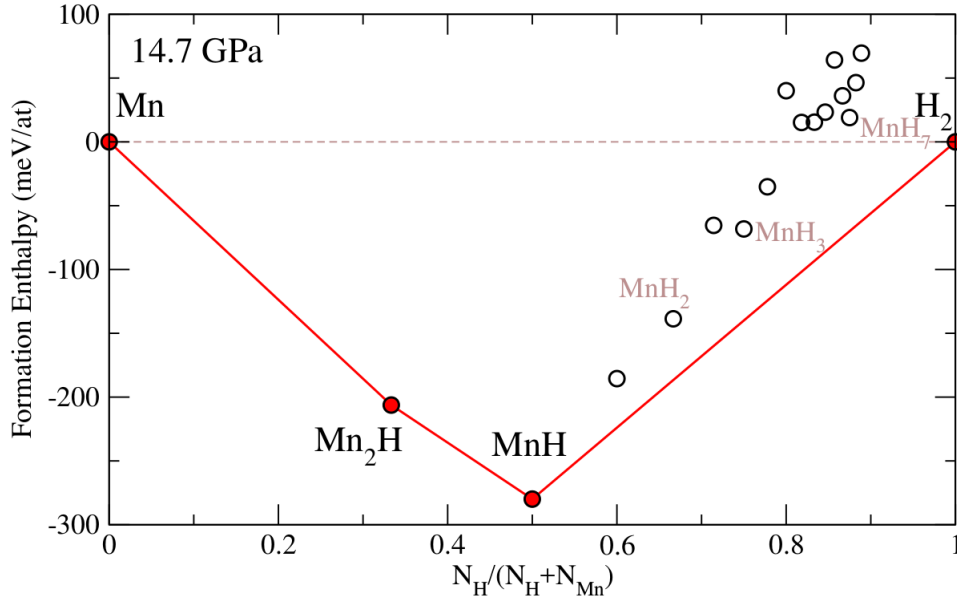
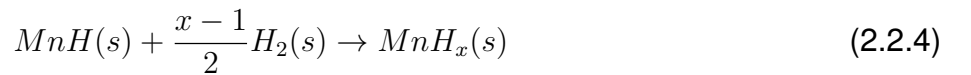


Figure 2.2.2: Formation enthalpies (eV/atom) of Mn hydrides at 14.7 GPa. The α phase of Mn is taken as reference.

Hydriding at High Pressure

The former subsection pointed out that the AIRSS algorithm gives results in agreement with experimental studies of MnH_x performed at low and intermediate pressure. It proves its accuracy regarding the studied system and allows us to be confident about the predictions that it can drive at higher pressure. The pressure range considered now lies typically between 29 and 147 GPa. The α phase of Mn is known to be stable up to 165 GPa [158]. However, the formation enthalpies of the hydrides are computed here with respect to ϵ -MnH (rather than α -Mn) and solid hydrogen ($C2/c - 32$ phase below ~ 110 GPa, $C2/c - 24$ phase [85] above), in order to make more visible the region with the large hydrogen stoichiometries. The formation enthalpy $\Delta H_f^{MnH}(\text{MnH}_x)$ with respect to MnH and H_2 is thus the enthalpy of the process:



For a hydride MnH_x , we denote by N_{MnH} the number of MnH in the formula unit MnH_x , and by N_H the number of remaining H atoms (once MnH has been removed), i.e. $N_{\text{MnH}} = 1$ and $N_H = x - 1$. In order to keep a correct meaning regarding their convex hull, these enthalpies (per formula unit), if plotted as a function of hydrogen content $N_H/(N_{\text{MnH}} + N_H)$, must be divided by $N_{\text{MnH}} + N_H (=x$ in the present case), and not by the number of atoms. This is done on FIG 2.2.3 for pressures of 29, 44, 58, 88, 117 and 147 GPa, and in Sec. 2.2.4 at 150 GPa.

FIG 2.2.3 first shows the formation enthalpies of the hydrides at 29 GPa (upper panel). At this pressure, none of the hydrides beyond MnH is stable. However, the enthalpy of three hydrides from the series has clearly started to drop: MnH_2 , MnH_3 and MnH_7 . It can be seen at 29 GPa

that, MnH_2 is very close to the convex hull, and indeed, a few GPa beyond, this manganese dihydride is the first compound containing more hydrogen than MnH that becomes stable. MnH_2 is followed by MnH_7 , that becomes stable at ~ 44 GPa (surprisingly, the enthalpy of MnH_7 drops more quickly than that of MnH_3). At 44 and 58 GPa, the two hydrides MnH_2 and MnH_7 lie on the convex hull. The structure found for MnH_2 is the same as that observed in several transition metal dihydrides (e.g. FeH_2 , CoH_2). It is a body-centered tetragonal structure with space group $I4/mmm$. The structure of MnH_7 is hexagonal (space group $P6_3/mmm$). It is characterized by hexagonal cages of atomic hydrogen bonded to Mn atoms, surrounding H_2 molecules. Hydrogen is thus only partially reduced into hydrides in MnH_7 , a part of it remaining under the molecular form (mixed hydride). Between MnH_2 and MnH_7 , no stable structure at 44 and 58 GPa are found. Note that MnH_3 , although not on the convex hull at 58 GPa, is rather close to it, and its enthalpy is decreasing with increasing pressure. At 88 GPa, MnH_3 is, indeed, now energetically stable, while MnH_7 also remains extremely stable. The structure of MnH_3 is body-centered tetragonal (space group $I4/m$) and contains four formula units per primitive cell. Then, up to 147 GPa, MnH_2 , MnH_3 and MnH_7 are the only compounds to be found on the convex hull, and MnH_7 is the most stable stoichiometry in hydrogen excess (the last hydride observed on the convex hull). At 147 GPa, MnH_8 almost lies on the convex hull. Considering how the enthalpy of this compound evolves with increasing pressure, it is highly probable that MnH_8 becomes stable for pressures beyond 150 GPa. Interestingly, the structure of this hydride contains one H_2 molecule per two f.u. (H-H distance = 0.73 Å within the molecule), just like that of MnH_7 .

2.2.4 Dynamical Stability and Nuclear Quantum Effects

To confirm the predictions made in the former section, the dynamical stability of the hydrides identified as energetically stable (ϵ - MnH , MnH_2 , MnH_3 and MnH_7) is now questioned. It is treated by computing the phonon dispersion curves within DFPT. FIG 2.2.4 displays these phonon dispersion curves, computed at 150 GPa for the four compounds, as well as the corresponding vibrational densities of states. The four hydrides are found dynamically stable at this pressure. The same calculations have also been conducted for MnH at 50 GPa, and for MnH_2 at 50 and 100 GPa (see appendix E): these two hydrides are also found dynamically stable at those pressures.

In MnH , MnH_2 and MnH_3 , the dynamics of the heavy Mn atoms and that of the light H atoms are well separated (non-overlapping peaks in the vibrational density of states), while this is not the case in MnH_7 . As said, in this compound, 6 hydrogens per f.u. are hydrided while the latter is under the molecular form ($1/2 \text{ H}_2$). The vibration modes associated with the dynamics of the six hydrogens which are under the atomic form overlap with those involving the motions of Mn. That associated with the stretching vibration of the H_2 molecule (vibron) has a very high frequency, very close to that of the vibron of free or solid molecular hydrogen ($\hbar\omega \sim 0.55$ eV).

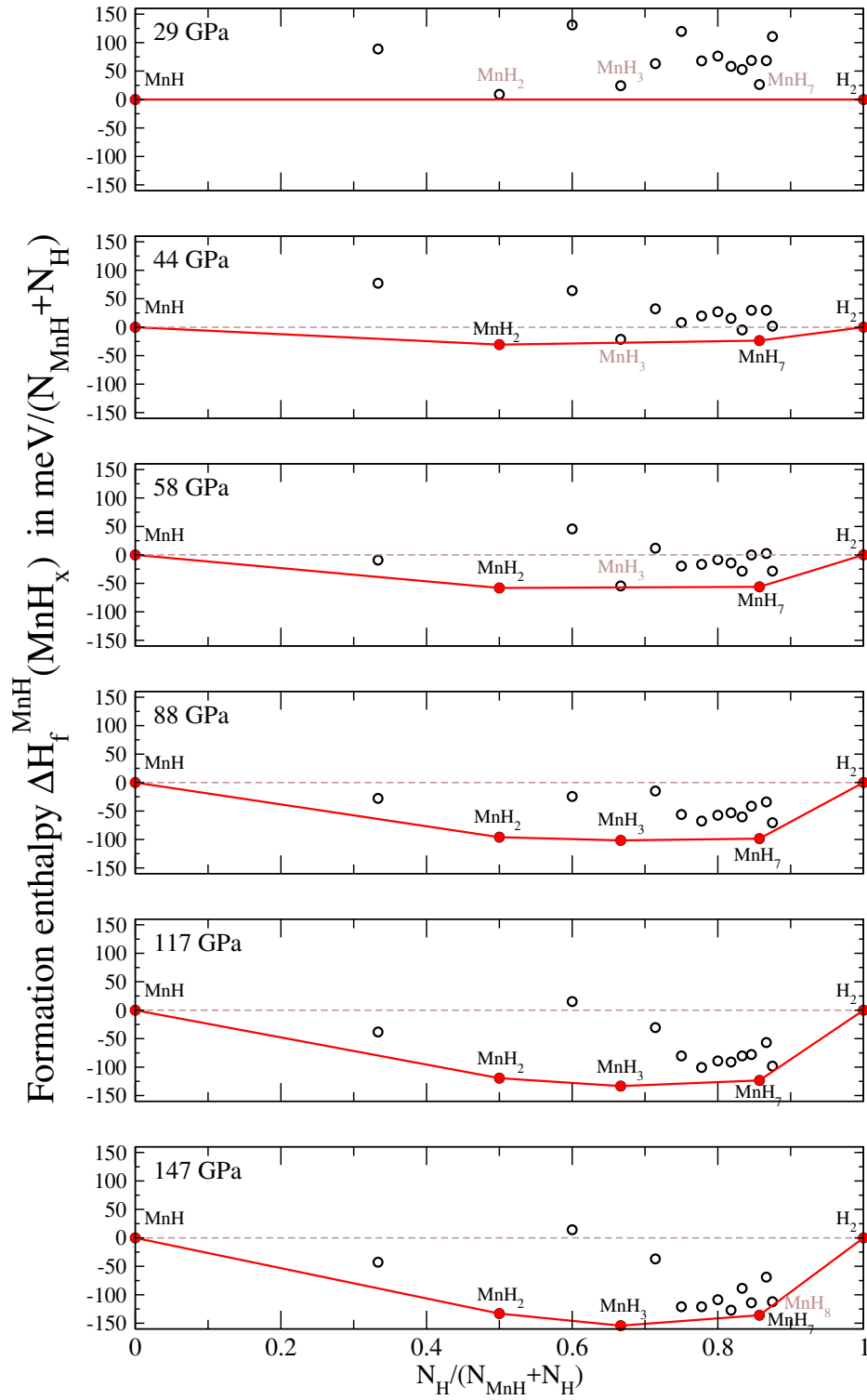


Figure 2.2.3: Formation enthalpies of the Mn hydrides with respect to $\epsilon\text{-MnH}$ and solid hydrogen, from 29 to 147 GPa.

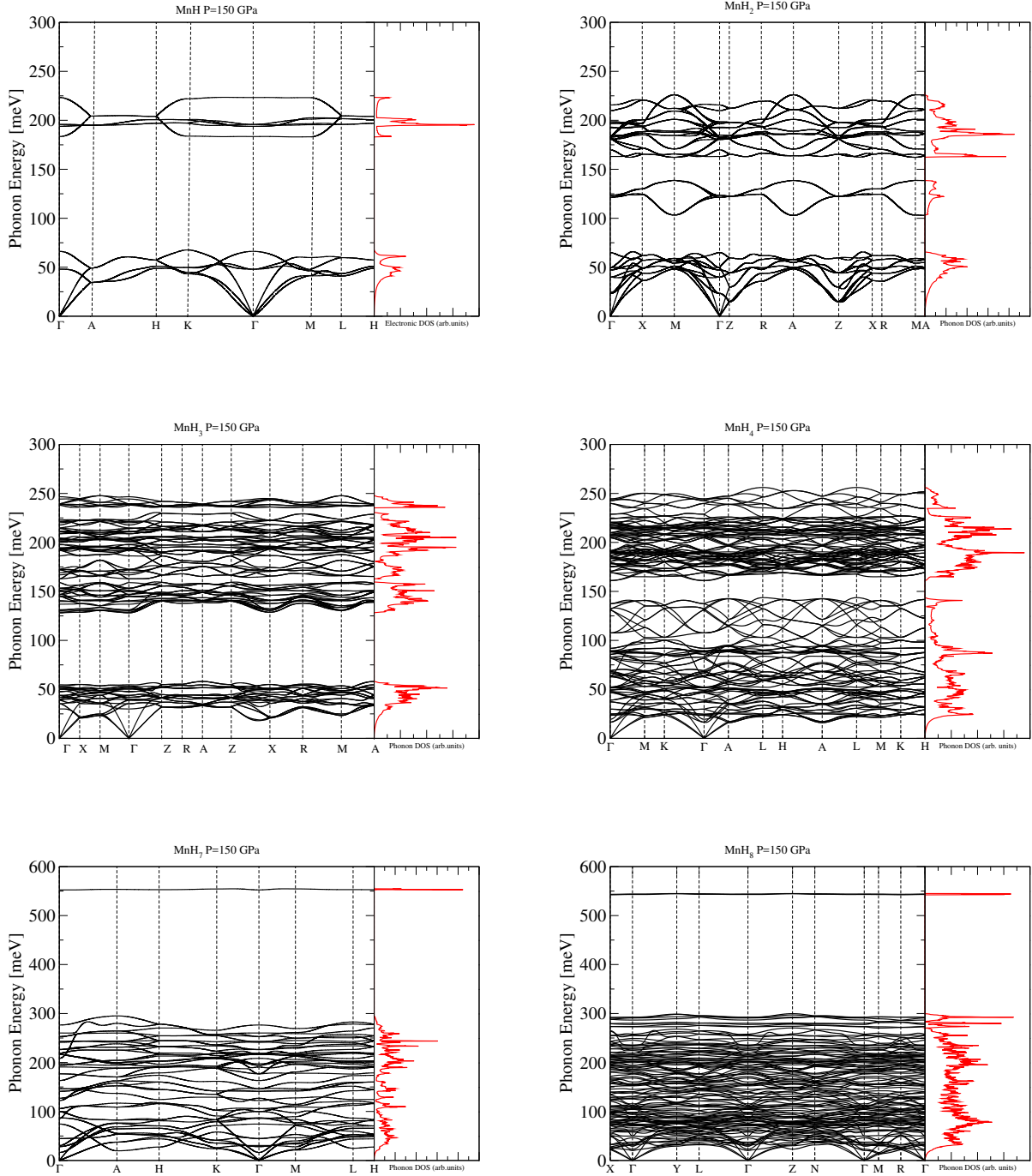


Figure 2.2.4: Phonon band structures and vibrational densities of states for MnH, MnH₂, MnH₃, MnH₄, MnH₇ and MnH₈, computed at 150 GPa. For MnH₂ (space group $I4/mmm$) and MnH₃ (space group $I4/m$), the path in the Brillouin Zone refers to the conventional (tetragonal) cell (multiplicity 2), which has been used in the DFPT calculation.

The impact of nuclear quantum effects due to atomic vibrations at 150 GPa in the harmonic approximation is also scrutinized. To this aim, the contribution of phonon Zero Point Energies (ZPE) is included to the formation enthalpies. Owing to the computational cost of the DFPT calculations needed to obtain these quantities, the calculations are limited to the structures found as being the closest to the convex hull, namely at less than ~ 30 meV. DFPT calculations were thus performed at this pressure for MnH_4 , Mn_2H_9 , MnH_5 , Mn_2H_{11} , Mn_2H_{13} , MnH_8 (and also solid molecular hydrogen in the $C2/c$ -24 phase). In three cases (Mn_2H_{11} , Mn_2H_{13} and MnH_8), the phonon dispersion curves of the structures as determined by the AIRSS exhibited unstable (imaginary frequency) modes. A slightly more stable structure was then obtained by displacing the atoms along these instabilities and structurally re-optimizing the system in a larger supercell. Phonons were recomputed on the final structures, which were found dynamically stable but in rather large primitive cells (26 atoms for Mn_2H_{11} , 60 atoms for Mn_2H_{13} , 36 atoms for MnH_8). The final enthalpies are calculated using these three new structures. Note that the subsequent stabilization is weak and that the predicted stable structures before inclusion of ZPE are unchanged. More details, as well as the phonon dispersion curves of these hydrides, can be found in appendix E. FIG 2.2.5 displays the formation enthalpies at 150 GPa, with and without the contribution of the ZPE. The main impact of ZPE is to stabilize MnH_4 and MnH_8 , two phases that now lie on the convex hull at 150 GPa. Their phonon dispersion curves are shown on FIG 2.2.4.

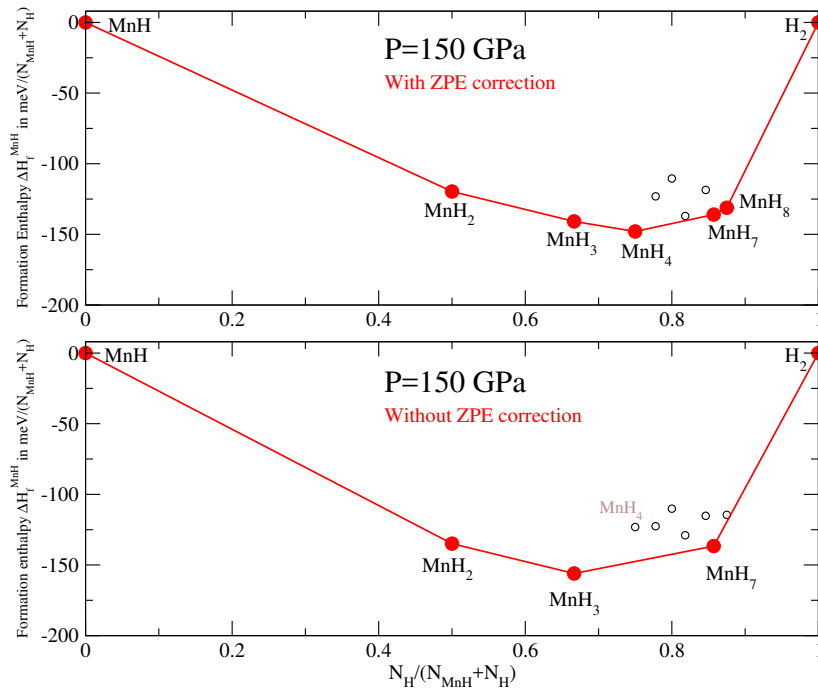


Figure 2.2.5: Formation enthalpies of the Mn hydrides with respect to ϵ -MnH and solid hydrogen at 150 GPa, including the contribution of Zero-Point motions (top panel) and without (bottom panel).

2.2.5 Physical Properties of the Stable Structures

This section describes in more depth the properties of the identified stable compounds. First of all the crystal structures of the different stable hydrides are described. It is followed by a discussion about the distribution of the atomic charges (Bader analysis), electronic and magnetic properties.

2.2.5.1 Description of Crystal Structures

Mn_2H , $\epsilon\text{-MnH}$, and MnH_2 crystallize in structures already encountered in other hydrides. Mn_2H and $\epsilon\text{-MnH}$ can be described as interstitial hydrides: starting from an hypothetical HCP Mn arrangement, Mn_2H is obtained by filling with hydrogen atoms half of the (0001) planes of octahedral sites, resulting in a $P\bar{3}m1$ space group. This structure is also called anti- CdI_2 type [133], it is also reported in Cu_2H [135]. The ϵ structure of MnH is obtained by filling all the interstitial octahedral sites with H. The two structures are shown on FIG 2.2.6 (a) and (b).

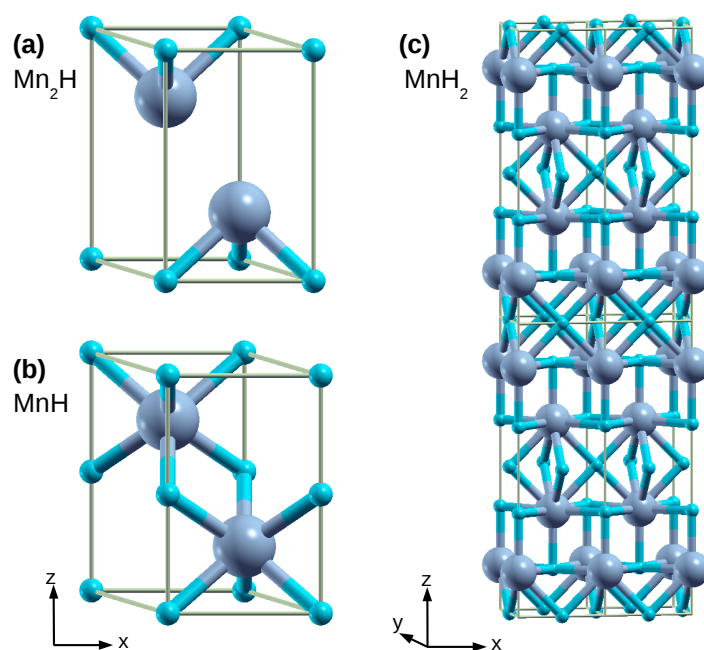


Figure 2.2.6: Crystal structure of (a) Mn_2H (space group $P\bar{3}m1$, anti- CdI_2 type), (b) $\epsilon\text{-MnH}$ (space group $P6_3/mmc$), and (c) MnH_2 (space group $I4/mmm$). The thin solid lines materialize the primitive cell for Mn_2H and MnH , the conventional unit cell for MnH_2 (multiplicity 2).

The body-centered tetragonal structure of MnH_2 is the same as that of FeH_2 [26], CoH_2 [164] (at high pressure), and NiH_2 [120]. It consists in cubic-like MnH_3 (the arrangement observed in

FeH₃ or CoH₃) slabs stacked onto each other, with two consecutive slabs being shifted in-plane by (1/2, 1/2, 0) lattice vectors of the conventional cell, which yields the body-centered symmetry (FIG 2.2.6 (c)).

MnH₃: A New Crystal Structure not Encountered in other 3d Transition Metal Hydrides

The body-centered tetragonal structure of MnH₃ (space group $I4/m$), in contrast, does not correspond to a structure encountered in other 3d transition metal hydrides: the Mn atoms form a complex sublattice that consists in the alternance of two identical planes in which the Mn atoms are arranged in squares connected by their corners. These planes are shifted in-plane with respect to each other, by (1/2,1/2,0) lattice vectors of the conventional unit cell, providing a body-centered tetragonal structure. The H atoms in MnH₃ occupy interstitial sites and belong to two subgroups that correspond to Wyckoff positions $16i$ or $8h$ (Tab 2.2.2). At 150 GPa, the H atoms in positions $16i$ have two Mn nearest neighbours at 1.57 and 1.62 Å, and two other Mn a little further at 1.72 Å.

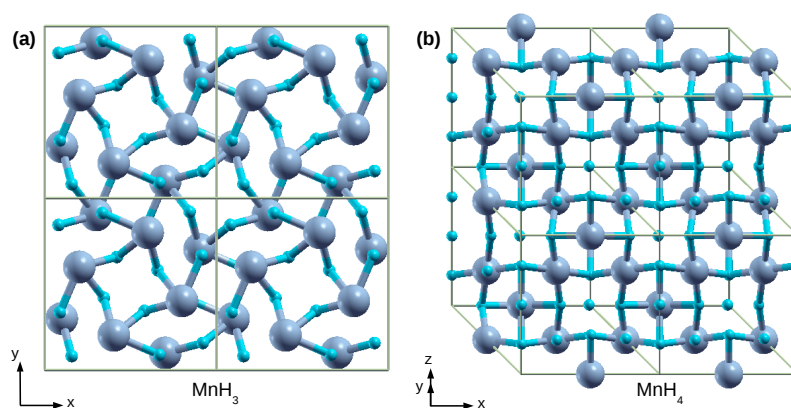


Figure 2.2.7: Crystal structure of (a) MnH₃ (space group $I4/m$) and (b) MnH₄ (space group $R32$). The thin solid lines materialize the conventional unit cell (multiplicity 2 for MnH₃ and 3 for MnH₄).

The H atoms in positions $8h$ are each surrounded by three Mn (at 1.58, 1.59 and 1.66 Å at 150 GPa), but are rather close to each other (1.39 and 1.46 Å), so that the nearest-neighbors of these H atoms are also H atoms.

They form small rectangular parallelepipeds, H_8 , that look like the ones observed in uranium polyhydrides [35], at the exception that these H_8 cages are not connected to each other in MnH_3 . Note that these short H-H distances are in line with those observed in other transition metal polyhydrides, e.g. FeH_5 , but remain significantly larger than those existing in the hydrides which are the best candidates as high- T_c superconductors at similar pressure (between ~ 1.1 and 1.2 Å in LaH_{10} and CeH_9 [14], for instance). It is thus unlikely that true chemical bonds do exist between these hydrogen atoms. The structure of MnH_3 is shown on FIG 2.2.7 (a).

MnH_7 : The Highest Stoichiometry Predicted below ~ 150 GPa

The MnH_7 superhydride is the highest stoichiometry predicted from 44 to near 150 GPa. Its structure can be described as follows: the Mn sublattice is a stacking of planes within which the Mn atoms form an hexagonal 2D lattice. The nearest-neighbor Mn atoms within the planes are bridged to each other by two H atoms; the nearest-neighbor Mn atoms between two consecutive planes are bridged to each other by three H atoms. This leaves large voids in the middle of the hexagons, which are filled by H_2 molecules (one H_2 molecule per two formula units), each H_2 molecule being oriented perpendicular to these planes. The H-H distance within the molecules is 0.73 Å at ~ 150 GPa, which is typical from the interatomic distance of the H_2 molecule. Thus, only 6 hydrogen atoms in the chemical formula MnH_7 are under the atomic form, the other one being under the molecular form, $1/2 H_2$. The primitive cell contains two formula units. The structure is shown on FIG 2.2.8 (a) and (b).

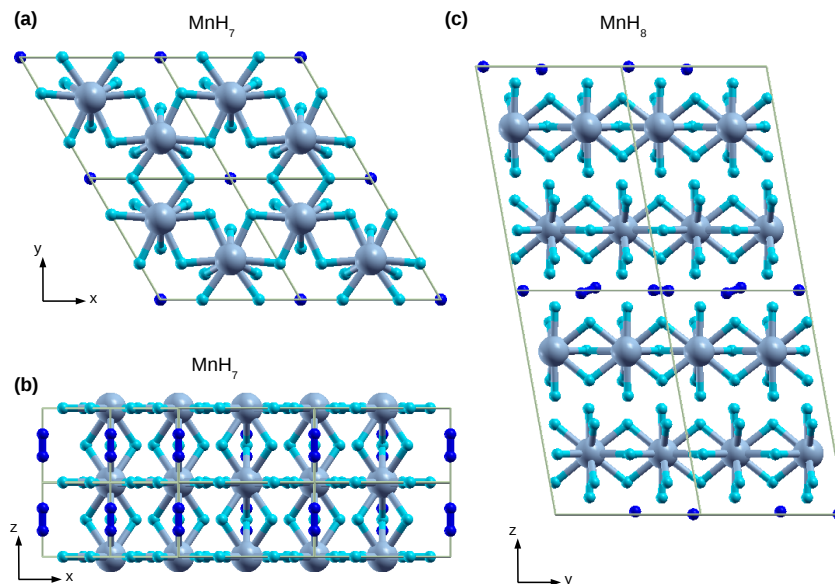


Figure 2.2.8: Crystal structure of the mixed hydrides: (a) and (b) MnH_7 (space group $P6/mmm$); (c) MnH_8 (space group $P\bar{1}$). The H_2 molecules are in dark blue, the atomic hydrogens in light blue. The thin solid lines materialize the primitive cell.

Structures stabilized with ZPE: MnH_8 and MnH_4

MnH_8 was almost stable without ZPE and stabilized by taking them into account at 150 GPa. The structure found for MnH_8 consists of Mn monolayers within which the Mn form a 2D square lattice (Tab 2.2.3 and FIG 2.2.8(c)). On both sides of each layer, the H atoms are bonded on top of Mn and in bridging positions between each pair of nearest-neighbour Mn atoms. An additional H atom also lies at the center of each square within the layer, but is placed closer to one Mn. Thus, each Mn is bonded to 11 H atoms. H_2 molecules are inserted in half of the interlayer spaces (the H-H distance in the molecule is 0.73 Å). The vibron which is the signature of the presence of molecular H_2 is clearly visible on the phonon dispersion curves at high energy (FIG 2.2.4).

The stabilization of MnH_4 by phonon ZPE, however, is more surprising. MnH_4 at this pressure adopts a rhombohedral structure (space group $R\bar{3}2$, see Tab 2.2.3), which can be seen as resulting from a slight distortion of a parent cubic structure (space group $Pm\bar{3}m$, 3 f.u. per primitive cell, see appendix E).

This structure possesses a few similarities with that of MnH_3 (see FIG 2.2.7 (b)). To provide an explanation for the stability of MnH_4 , the averaged ZPE per vibration mode as a function of hydrogen content is plotted on FIG 2.2.9 (top panel). It can be observed that this ZPE in MnH_4 is indeed significantly smaller than in MnH_3 and in the following hydrides.

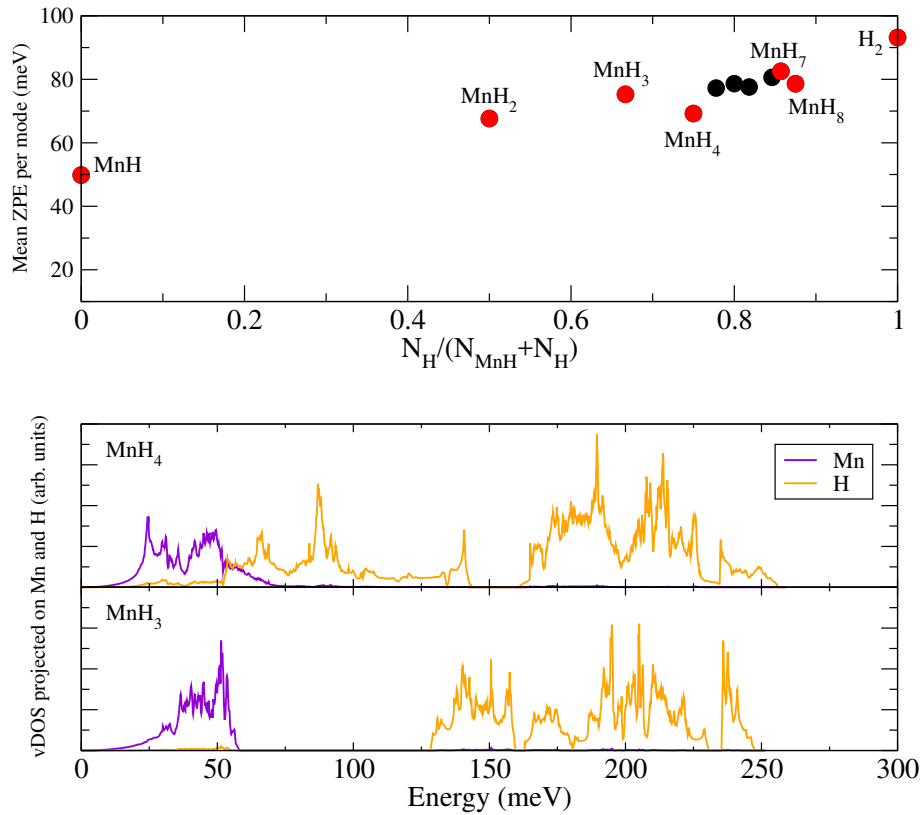


Figure 2.2.9: (Top) Zero-Point Energy of a formula unit divided by the number of modes (meV) at 150 GPa. (Bottom) Vibrational density of states of MnH_3 and MnH_4 projected on Mn and H at 150 GPa.

This behaviour may be understood from the phonon bandstructures (FIG 2.2.4) and vibrational density of states (vDOS) projected on Mn and H (FIG 2.2.9): in MnH_3 , the vDOS exhibits two well separated parts: a low-energy one ($\hbar\omega \leq 50$ meV) due to Mn motions, and a high-energy one ($\hbar\omega \geq 125$ meV) due to H motions. In MnH_4 , the vibrations due to the additional H fall in the low-energy part and overlap those of Mn, giving rise to a mixed Mn-H low-energy vDOS ($\hbar\omega \leq 150$ meV). In clear, some H atoms in MnH_4 vibrate with eigenfrequencies which are significantly smaller than in MnH_3 , which provides a smaller ZPE in MnH_4 compared to MnH_3 . The Wyckoff positions, lattice constants and angles describing these above mentioned structures are given in Tab 2.2.2 and 2.2.3 .

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
Mn ₂ H							
$P\bar{3}m1$	50 GPa	a=b=2.4564 c=3.8958 $\alpha=\beta=90$ $\gamma=120$	Mn	2d	0.33333	0.66667	0.73447
			H	1a	0	0	0
MnH							
$P6_3/mmc$	150 GPa	a=b=2.3848 c=3.7761 $\alpha=\beta=90$ $\gamma=120$	Mn	2d	0.33333	0.66667	0.75
			H	2a	0	0	0
MnH ₂							
$I4/mmm$	150 GPa	a=b=2.3782 c=7.8694 $\alpha=\beta=\gamma=90$	Mn	4e	0	0	0.84829
			H	4c	0	0.5	0
			H	4e	0	0	0.64781
MnH ₃							
$I4/m$	150 GPa	a=b=5.4981 c=3.4137 $\alpha=\beta=\gamma=90$	Mn	8h	0.59863	0.79892	0
			H	16i	0.08010	0.84326	0.21136
			H	8h	0.66512	0.08765	0
MnH ₇							
$P6/mmm$	150 GPa	a=b=4.4262 c=2.4310 $\alpha=\beta=90$ $\gamma=120$	Mn	2c	0.33333	0.66667	0
			H	2e	0	0	0.35005
			H	6m	0.21084	0.42169	0.5
			H	6j	0.70360	0	0

Table 2.2.2: Wyckoff positions, lattice constants and angles of the stable manganese hydrides (without ZPE) up to 150 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
MnH₄							
$R\bar{3}2$	150 GPa	a=b=4.9849 c=6.1041 $\alpha=\beta=90$ $\gamma=120$	Mn	9d	0.49912	0	0
			H	18f	0.51822	0.47528	0.26076
			H	9e	0.49023	0	0.5
			H	6c	0	0	0.78214
			H	3b	0	0	0.5
MnH₈							
$P\bar{1}$	150 GPa	a=2.4409 b=4.8795 c=7.7099 $\alpha=99.89$ $\beta=97.98$ $\gamma=90.71$	Mn	2i	0.85233	0.18657	0.72787
			Mn	2i	0.14462	0.31376	0.27329
			H	2i	0.03228	0.03451	0.14232
			H	2i	0.84902	0.39514	0.59780
			H	2i	0.39980	0.55217	0.01262
			H	2i	0.58944	0.08671	0.26812
			H	2i	0.74039	0.14104	0.52964
			H	2i	-0.03365	0.24839	-0.07428
			H	2i	0.42004	0.19375	0.85759
			H	2i	0.69081	0.87286	0.40232
			H	2i	-0.02140	0.46941	0.85723
			H	2i	0.14503	0.10424	0.40225
			H	2i	0.61053	0.05102	-0.00057
			H	2i	0.57369	0.30339	0.14135
			H	2i	0.30843	0.62603	0.59705
			H	2i	0.41054	0.41494	0.73201
			H	2i	0.25468	0.35725	0.47130
			H	2i	0.10737	0.25483	0.07615

Table 2.2.3: Wyckoff positions, lattice constants and angles of the manganese hydrides stabilized by ZPE at 150 GPa: MnH₄ and MnH₈.

2.2.5.2 Bader Charges

The atomic charges have been computed at 150 GPa, using the methodology proposed by Bader [145], for several stoichiometries, including energetically stable and unstable ones (MnH, MnH₂, MnH₃, MnH₄, MnH₅ and MnH₇). They are plotted on FIG 2.2.10 as a function of the number of H in the formula unit. The values are given in Tab 2.2.4 for all the symmetry-inequivalent atoms. For each structure, the charge on H has been averaged over all the hydrided atoms (i.e. those that do not belong to an H₂ molecule).

Compound	MnH	MnH ₂	MnH ₃	MnH ₄	MnH ₅	MnH ₇
Space group	$P6_3/mmc$	$I4/mmm$	$I4/m$	$R32$	$C2/m$	$P6/mmm$
Mn	0.357 (2d)	+0.466 (4e)	+0.604 (8h)	+0.528 (9d)	+0.583 (4i)	+0.563 (2c)
H	-0.357 (2a)	-0.254 (4c)	-0.205 (16i)	-0.186 (18f)	-0.152 (8j)	-0.004 (2e)
H		-0.196 (4e)	-0.211 (8h)	-0.008 (9e)	-0.145 (4h)	-0.064 (6m)
H				-0.186 (6c)	-0.058 (4i)	-0.116 (6j)
H				-0.134 (3b)	-0.069 (4i)	

Table 2.2.4: Atomic charges (in elementary charge e) as obtained from Bader analysis for the different symmetry-inequivalent atoms in MnH, MnH₂, MnH₃, MnH₄, MnH₅ and MnH₇ at 150 GPa. The numerical error on the computation of the charge is estimated to $\sim 0.01 e$.

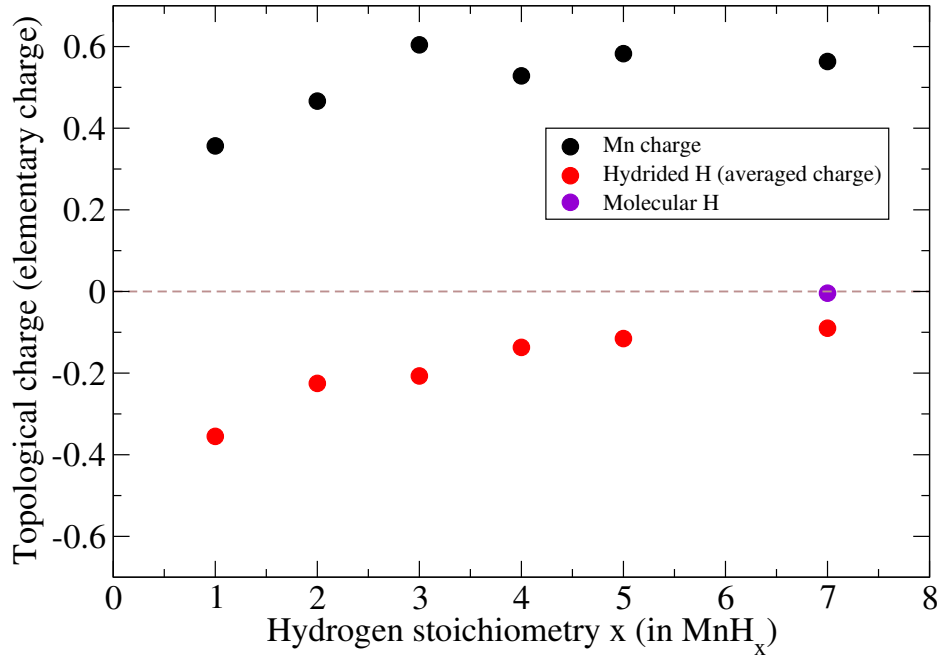


Figure 2.2.10: Topological charge of Mn and H (in units of elementary charge) as a function of hydrogen content x in MnH _{x} at 150 GPa.

First of all, the Bader analysis confirms that the studied systems are hydrides, i.e. *the hydrogens are negatively charged, and Mn is positively charged*. There is an electron transfer from Mn to H. Second, the charge of Mn is globally increasing from MnH to MnH₃, and then saturates to a value of $\sim +0.5$ - 0.6 , indicating that hydrogen tends to increase the oxidation process of Mn up to MnH₃. Beyond this hydride, Mn can not release more electrons, and the released electrons are shared between the hydrogens of the structure, excepting when some H remain molecular as in

MnH_7 . As a consequence, the charge of the hydrided hydrogens globally decreases to zero (in absolute value) with increasing hydrogen stoichiometry.

2.2.5.3 Electronic Structure

The electronic bandstructures and densities of states of MnH , MnH_2 , MnH_3 and MnH_7 , computed at 150 GPa, are shown on FIG 2.2.11. As expected, MnH , MnH_2 , and MnH_3 are metallic. MnH_7 , however, has within the GGA-PBE approximation (see Sec. 2.2.2) a very weak density of states at the Fermi level at 150 GPa, corresponding to a quasi-gap throughout the Brillouin Zone, which is closed only around the A point. At smaller pressure, the gap in MnH_7 is open, and the compound becomes a semiconductor. At 150 GPa, still within the used GGA-PBE scheme, MnH_8 is found to be a semiconductor, with a Kohn-Sham bandgap of ~ 0.5 eV.

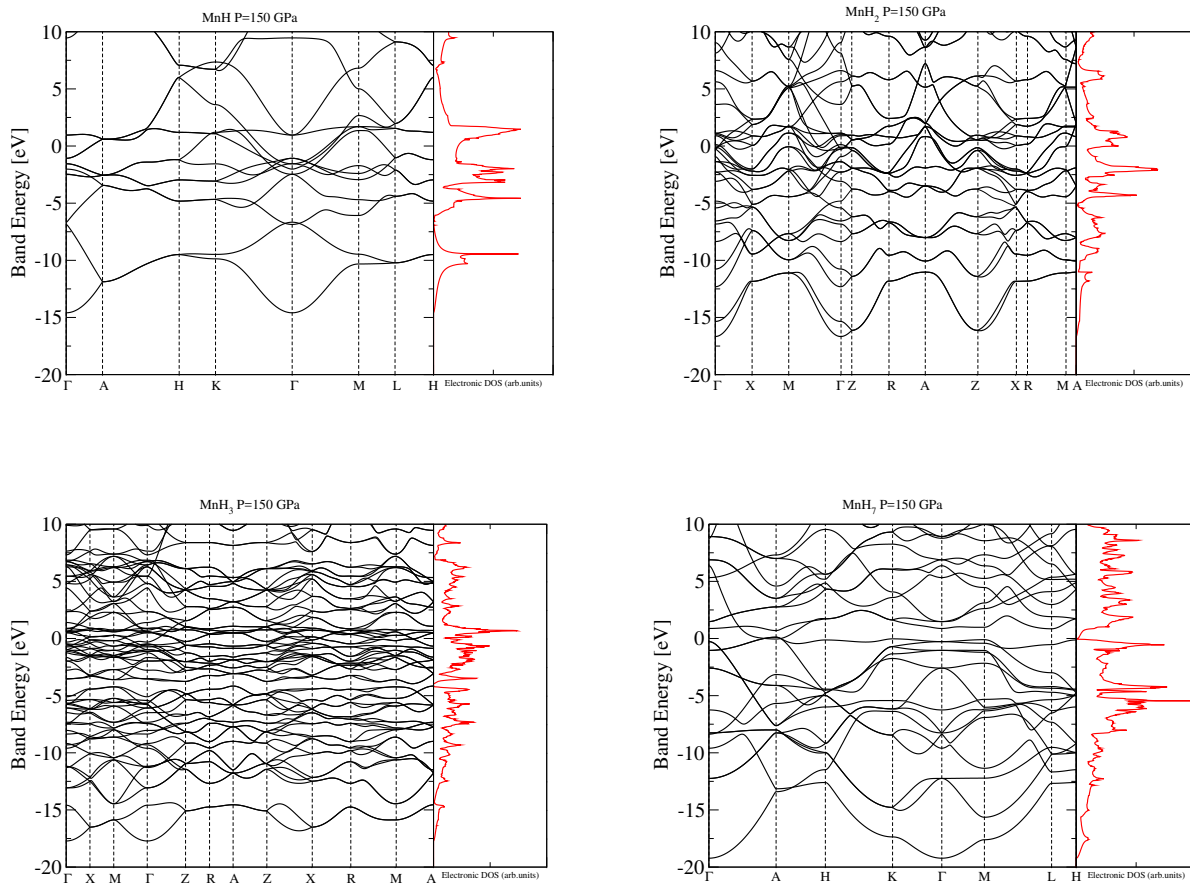


Figure 2.2.11: Electronic band structures and densities of states for MnH , MnH_2 , MnH_3 and MnH_7 computed at 150 GPa. For MnH_2 (space group $I4/mmm$) and MnH_3 (space group $I4/m$), the path in the Brillouin Zone refers to the conventional (tetragonal) cell (multiplicity 2), which has been used in the band-structure calculation. The Fermi energy is set at zero.

The GGA-PBE functional used is known to underestimate the electronic gaps. Such a problem is however partly corrected with Meta-GGA type functional. Given the fact MnH_7 has a very low density of states at the Fermi level with GGA-PBE, it might be possible that this compound has indeed an electronic gap not retrieved in GGA. In order to test this possibility, the SCAN Meta-GGA functional [64] is used. This is performed with the Meta-GGA implementation achieved in *abinit* in this PhD. FIG 2.2.12 shows the result of this calculation and the comparison with the GGA approximation. In contrast to the GGA case, the electronic bands do not cross at the Fermi level near the A point of the Brillouin zone. It opens thus the electronic gap of MnH_7 . The same calculation performed for MnH_8 with the SCAN functional gives an electronic gap of around 1 eV instead of 0.5 eV in GGA.

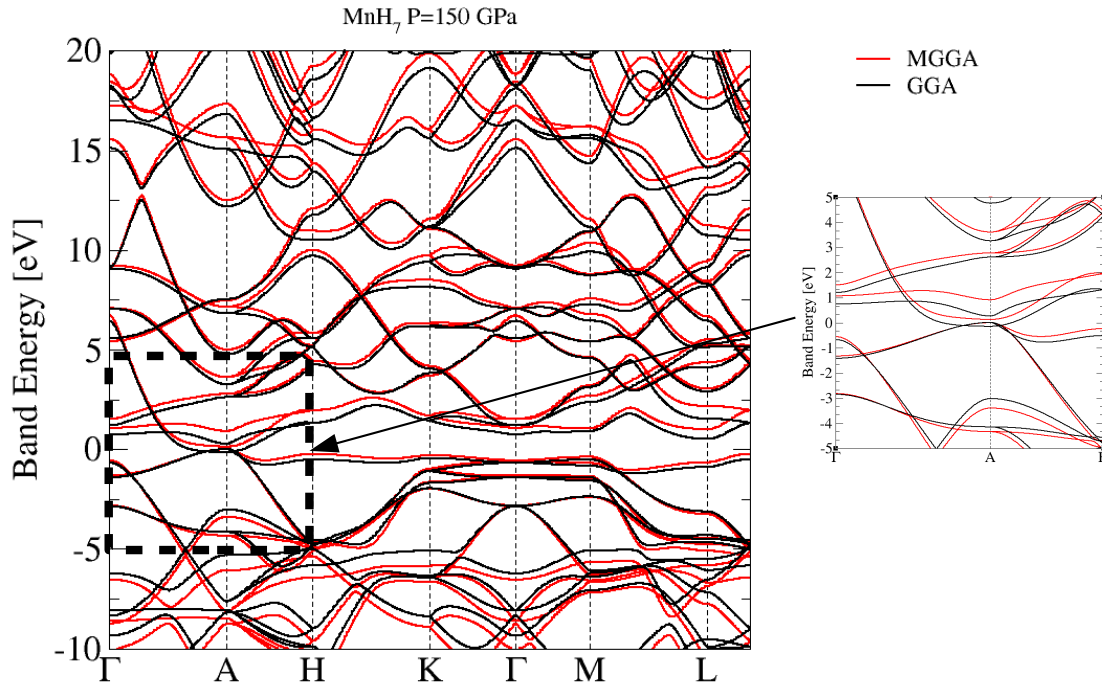


Figure 2.2.12: Electronic band structures for MnH_7 using GGA (GGA-PBE) and Meta-GGA (SCAN).

Considering these results about the electronic structures, it seems unlikely that the predicted hydrogen-rich compounds in this work are high- T_c conventional superconductors. Indeed MnH_8 is found insulating and MnH_7 has a low density of states at the Fermi level with GGA and a small electronic gap in Meta-GGA. Moreover, there is no example of predicted high- T_c superconductivity in other transition metal hydrides [165] (excluding column 3B of the periodic table: Sc, Y, La), and the best candidates for high- T_c superconductivity in high-pressure hydrides remain the rare-earth hydrides and transition metal hydrides of column 3B.

2.2.5.4 Magnetism

α -Mn is a non-collinear antiferromagnetic (AFM) solid at ambient pressure (Neel temperature 95 K [159]). An AFM order is also reported (at low pressure) in ϵ -MnH_x with a Neel temperature around 360 K [133, 166]: the AFM structure is due to the antiferromagnetic coupling between ferromagnetic (110) planes (of the hexagonal arrangement). This indicates that a magnetic order might exist in manganese hydrides, at least at low pressure. Also, γ -MnH_{0.41} is reported as an antiferromagnet at room temperature [133, 167], with a Neel temperature much above 300 K [133]. Reproducing the low-pressure AFM properties of MnH_x is, however, not the scope of the present work. It seems nonetheless that Mn hydrides exhibit a tendency to form rather stable magnetic structures, at least at low pressure. Here it has been checked whether a ferromagnetic (FM) order can exist or not in the manganese hydrides and survive with pressure.

In Mn₂H and in ϵ -MnH, no ferromagnetic order could be stabilized, regardless of the pressure, from zero to 150 GPa. In contrast, a weak FM order is found in MnH₂ at low pressure (~ 0.14 and $0.11 \mu_B/\text{Mn}$ at 0 and 15 GPa), which drops to zero between 30 and 44 GPa. At 0 and 15 GPa, this FM state of MnH₂ is however degenerate with the Non-Magnetic (NM) state (within the numerical precision of the calculation). Thus a possible FM state would not change the position of MnH₂ on the convex hull of FIG 2.2.2 (it remains unstable at these low pressures). A FM order is found stable in MnH₃ at low pressure, with magnetic moments per Mn much larger than in MnH₂. At 0 and 15 GPa respectively, the FM state in MnH₃ is more stable than the NM state by -0.17 and -0.05 eV/f.u., and the magnetic moment is 2.02 and $1.84 \mu_B/\text{Mn}$. Like in MnH₂, this is not sufficient to stabilize MnH₃ at this pressure (see FIG 2.2.2). Thus MnH₃ remains unstable despite FM at low pressure. Moreover, this FM state becomes rapidly less stable than the NM one above 25 GPa. It is possible to maintain it as a metastable state on a small range of pressure, up to ~ 30 GPa. Finally, it was not possible to stabilize a FM state in MnH₇, even at low pressure.

2.2.6 Conclusion

In this chapter, the AIRSS method has been used to identify the structures of manganese hydrides and superhydrides MnH_x up to $x=8$ in the pressure range typically accessible to diamond anvil cell experiments. Upon increasing pressure, the calculations predict the successive stabilization of several hydrides, that contain more and more hydrogen: MnH, MnH₂, MnH₃ and MnH₇. At low pressure, insertion of hydrogen stabilizes an HCP arrangement of the manganese atoms, with the H atoms filling the octahedral interstitial sites of this lattice, as in Mn₂H and MnH. This is in qualitative agreement with the experiments performed on the Mn-H system up to a few tens of GPa. The existence of manganese mono-, di- and trihydride (MnH, MnH₂ and MnH₃) is in line with some other 3d transition metal hydrides such as Fe, Co and Cr. The structure of MnH₃, with four f.u. per primitive cell, however, is rather complex and differs from other known structures of transition metal trihydrides. MnH, MnH₂ and MnH₃ are metallic compounds. The existence of a very stable MnH₇ superhydride is predicted, already present on the convex hull at ~ 44 GPa.

The crystal structure of MnH_7 contains one hydrogen molecule per two formula units. These H_2 molecules are aligned along hexagonal channels. The presence of H_2 molecules does not support the existence of high temperature superconducting properties because it demonstrates bad conducting properties. In terms of applications for hydrogen storage it might be however interesting as releasing or absorbing H_2 molecules is a much more reversible process than with atomic hydrogen because these molecules does not interact a lot with the metallic sublattice. To precise the conducting properties at high pressure (~ 150 GPa) in MnH_7 , the GGA-PBE provides an electronic structure with a quasi-gap throughout the Brillouin Zone, which is closed only around one specific point. The same calculation performed in Meta-GGA with the SCAN functional however predicts a non-zero open electronic gap. At 150 GPa, MnH_8 and MnH_4 are stabilized by phonon Zero-Point Energies. The present work should motivate experimental studies of the Mn-H system up to the high pressures accessible in diamond anvil cells.

Chapter 2.3

Yttrium Hydrides and Superhydrides

Contents

2.3.1 State of the Art	84
2.3.1.1 Binary Yttrium Hydrides and Superhydrides	84
2.3.1.2 Iron-Yttrium Hydrides	89
2.3.2 The AIRSS Calculations	91
2.3.2.1 Binary Yttrium Superhydrides	91
2.3.2.2 Yttrium-Iron Ternary Hydrides	98
2.3.3 Conclusion	110
2.3.4 Summary of Part 2	112

In this chapter, we focus on another transition metal, yttrium, and on the hydrides and superhydrides it may form under pressure. In addition to the binary hydrides YH_x , we also consider the ternary hydrides YFe_2H_x , in which iron has been added. The first section is a review of the literature about these two systems. This section is followed by the presentation of the results obtained by applying AIRSS. These results are discussed in the last section.

2.3.1 State of the Art

This section presents the state of the art regarding both systems YH_x and YFe_2H_x . For the YH_x system, the crystal structures obtained from former studies using CSP algorithms are presented and new convex hulls gathering and comparing the different predicted phases are computed. In a second time, the experimental results obtained in previous studies about the YFe_2H_x system are summarized.

2.3.1.1 Binary Yttrium Hydrides and Superhydrides

As said, the YH_x system has already been the subject of studies using CSP algorithms (PSO and GA algorithms) [98, 99, 31, 168, 169]. Experiments were also conducted [168, 170, 171]. In contrast to a wide variety of binary superhydrides where numerical predictions match experimental results, the yttrium superhydrides synthesized between 140 and 180 GPa can not be described in a satisfactory manner by the predicted crystal structures. Indeed, in this pressure range, experimental X-ray spectra indicate the existence of phases which can not be explained by predictions of CSP algorithms, or necessitate more or less complex mixtures of different predicted phases to be interpreted. Troyan et al [168] synthesized YH_x but could not explain the X-ray diffraction spectra observed by a single phase: at 166 GPa, the X-ray spectrum could be interpreted only by a mixture of Im-3m YH_6 and YH_7 (P1 and Imm2). At higher pressure (172-180 GPa) also, a mixture of Im-3m YH_6 and distorted I4/mmm YH_4 was necessary to explain the experimental spectrum. Similarly, Kong et al [170] synthesized Im-3m YH_6 and P63/mmc YH_9 in the 200-250 GPa range but should invoke the presence of an impurity phase (I4/mmm- YH_4 or other) to explain the experimental results. Finally, Snider et al [171] synthesized yttrium superhydrides in the 200 GPa range and interpreted their Raman spectra and measure of superconducting T_c , by comparison with results from ab-initio calculations, as arising from a stoichiometry close to YH_9 .

Structures of yttrium superhydrides having up to four chemical formula units were explored in the above cited studies. The space groups of the identified structures are listed below in Tab 2.3.1. To identify the most stable phases among those predicted and clarify the pressure stability ranges, the structures predicted in the literature are recomputed in this work. The convex hulls of the formation enthalpies of these various phases between 88 and 176 GPa are plotted then.

YH_x	Space Group	Reference
YH_2	P6/mmm	[98]
YH_3	I4/mmm,CmCm	[99] [169]
YH_3	P2 ₁ /m,P6 ₃ /mmc,Fm-3m	[99]
YH_3	Pnma	[98]
YH_4	I4/mmm	[99] [98] [31]
YH_6	Im-3m	[99] [98] [31]
YH_7	P1,Imm2	[168]
YH_8	Cc	[98]
YH_9	P6 ₃ m	[99]
YH_9	P6 ₃ /mmc	[99]
YH_{10}	Fm-3m	[99] [98]
YH_{12}	C2/C	[98]

Table 2.3.1: Predicted YH_x structures in the literature.

Formation Enthalpies and Stability of the Predicted Phases

In order to determine the convex hulls of the formation enthalpies of the yttrium hydrides and superhydrides, structural optimizations using the BFGS algorithm (1.3.4.2) on the crystal structures of the literature are performed. The same numerical parameters as those employed for MnH_x for the final highly converged optimizations are applied (see 2.2.2). The convex hulls are determined at four different pressures (88 GPa, 117 GPa, 147 GPa and 176 GPa).

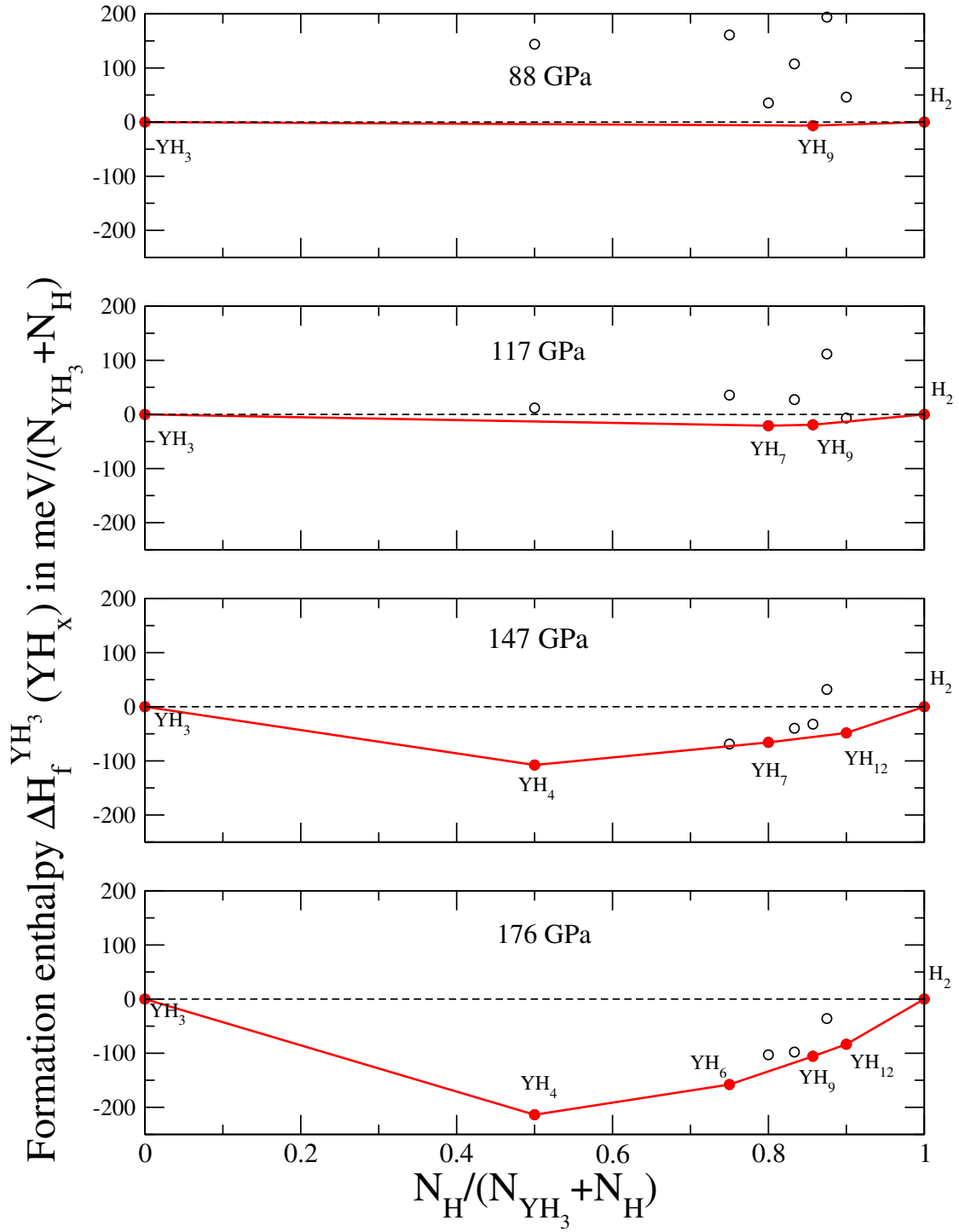


Figure 2.3.1: Formation enthalpies of the YH_x predicted in the literature. The convex hulls are plotted with red solid lines.

For the computation of the formation enthalpies, the hydride YH_3 is taken as reference in its cubic Fm-3m structure, which is the one experimentally observed. At 176 GPa however, DFT calculations point out that another structure of YH_3 , with Pnma space group, is more stable than the cubic Fm-3m phase. It is then taken as reference for the calculation of the formation enthalpies

of the hydrides at this pressure. Up to 117 GPa the C2/c-32 phase of solid hydrogen is taken as reference (see appendix E). Above this pressure, the C2/c-24 phase of solid hydrogen, which becomes more stable, is considered [85]. For compounds having the same stoichiometry, the enthalpies at a given pressure are compared to identify which structure is the most stable and thus has to be chosen for the determination of the convex hull. Tab 2.3.2 sums up the different

YH_x	176 GPa	147 GPa	117 GPa	88 GPa	Reference
YH_3	Pnma	Fm-3m	Fm-3m	Fm-3m	Pnma: [98], Fm-3m: [99] [98]
YH_4	I4/mmm	I4/mmm	I4/mmm	I4/mmm	[99] [98] [31]
YH_6	Im-3m	Im-3m	Im-3m	Im-3m	[99] [98] [31]
YH_7	Imm2	Imm2	Imm2	Imm2	[168]
YH_8	Cc	Cc	Cc	Cc	[98]
YH_9	$P6_3/mmc$	$P6_3/mmc$	$P6_3/m$	$P6_3/m$	[99]
YH_{10}	Fm-3m	Fm-3m	Fm-3m	Fm-3m	[99] [98]
YH_{12}	C2/c	C2/c	C2/c	C2/c	[98]

Table 2.3.2: Phases considered for the structures retained as the most stable ones among those predicted in the literature.

crystal structures corresponding to each stoichiometry at each pressure where the convex hulls are determined. FIG 2.3.1 presents the formation enthalpies and the convex hulls obtained for the YH_x predicted in the literature. These calculations give the following results.

- From 88 GPa to 117 GPa, YH_9 $P6_3/m$ and then YH_7 Imm2 are the only compounds with YH_3 Fm-3m appearing on the convex hull.
- From 147 GPa to 176 GPa, YH_9 $P6_3/mmc$ becomes more stable than YH_9 $P6_3/m$. However a compound containing more hydrogen than YH_9 , namely YH_{12} , becomes stable near 147 GPa and is thus the one supposed to be observed under hydrogen excess. At the same time, YH_4 I4/mmm is expected to become stable. YH_7 is still stable at this pressure.
- Near 176 GPa, YH_6 Im-3m appears on the convex hull and YH_7 Imm2 is destabilized, while YH_9 $P6_3/mmc$ becomes stable.

FIG 2.3.2 presents the different stability domains deduced from the convex hull representations.

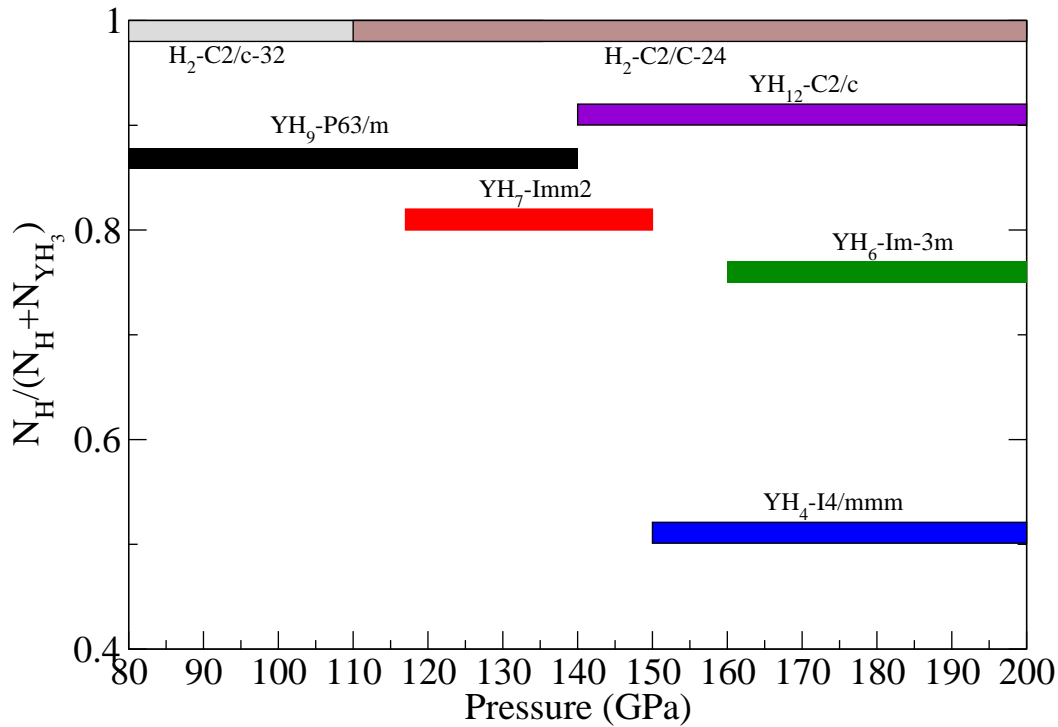


Figure 2.3.2: Stability domain determined for each YH_x phase predicted in the literature, between ~ 80 and 200 GPa. YH_3 is not shown.

All these predictions have been made thanks to GA (1.3.5.2) and PSO (1.3.5.1) methods. As explained above, recent experimental results are confirming some of the phases predicted under high pressure (i.e. near 180 GPa and above). Troyan et al reported the synthesis of the YH_6 Im-3m phase near 180 GPa [168]. Snider et al pointed out a phase near 180 GPa with might correspond to the YH_9 P6₃/mmc [171]. These two phases have been also experimentally obtained by Kong et al in the same pressure ranges [170]. Note that yttrium superhydrides are particularly challenging from the point of view of superconductivity because they have been predicted and measured with very high critical temperatures. YH_{10} has been predicted with a T_c close to 300 K near 250 GPa [98, 99], or YH_6 with T_c predicted in the range of 250 K near 150 GPa [31]. Experimental studies confirmed the high potentialities this system offers for high temperature superconductivity. For the compounds they experimentally synthesized, and identified by the authors as YH_9 P6₃/mmc, Snider et al [171] measured a T_c of 262 K at a pressure near 180 GPa. Troyan et al [168] reported a high T_c of 224 K near 170 GPa for a system identified by the authors as YH_6 Im-3m. Finally, Kong et al highlighted a T_c of 227 K near 240 GPa for a system identified as YH_6 Im-3m and 243 K near 200 GPa for a system identified as YH_9 P6₃/mmc. However, as already pointed out, none of the predicted phase could be unambiguously identified in experiments up to now, suggesting

that, either YH_x always form complex mixtures under pressure, in which several stoichiometries are present, or the YH_x crystallize in more complex structures that have not been identified by CSP methods, possibly because of their large number of f.u. in the primitive cell. In the next section 2.3.2, AIRSS is employed to make our own predictions of the YH_x crystal structures, and try to understand the discrepancies between theoretical and experimental results.

2.3.1.2 Iron-Yttrium Hydrides

Another challenge of this PhD is to start the investigations of systems which might form superhydrides under low or moderate pressures (typically a few tens of GPa). Up to now however, no binary superhydride seems to be thermodynamically stable in these conditions, with the exception of UH_7 which is formed near 40 GPa [38]. By adding another metallic element, and thus trying to form ternary hydrides, a new degree of freedom is considered, on which it should be possible to play in order to find new superhydrides, which could possibly be stabilized under lower pressure than their binary counterparts. This approach has already been used to stabilize unexpected kinds of hydrides, by combining a non-hydride forming element and a hydride-forming one. For instance it is possible to form gold hydrides (i.e. containing AuH_2^- units or Au-H-Au bonds) by adding alkaline and alkaline earths to gold, whereas gold alone does not form any stable hydride up to 300 GPa [172]. This strategy works outside the field of the superhydrides as well. The case of nitrogen is indeed emblematic. By adding a strong reductor like Li it is possible to stabilize polymerized forms of nitrogen at ambient pressure under the form of N_5 rings. This is what occurs in the LiN_5 compound [173]. Based on the knowledge of the affinity of each metal with hydrogen, ternary compounds begin to be proposed, and studied. For example, this is the case for calcium-yttrium superhydrides [83] or lithium-magnesium ones [84]. In both cases, high hydrogen stoichiometries and high critical temperatures for superconductivity are expected (i.e. $T_c=258$ K at 200 GPa for CaYH_{12} [83] and $T_c=473$ K at 250 GPa for $\text{Li}_2\text{MgH}_{16}$ [84]).

It is in this context that we have undertaken the study of iron-yttrium hydrides, with stoichiometry YFe_2H_x . Up to 60 GPa binary yttrium and iron hydrides are characterized by a hydrogen stoichiometry up to 3 (for Y) and 2 (for Fe [25]). One interesting question is whether a ternary hydride having a hydrogen stoichiometry higher than the sum of the separated binary ones can be formed in this pressure range (in other words whether a stoichiometry higher than YFe_2H_7 can be stabilized). The YFe_2H_x system was studied formerly by experiments at low pressure ranges (below 1 GPa) [176, 177, 178, 179, 180, 181, 182]. In these experiments, different YFeH_x hydrides from $x = 1.2$ to $x = 5$ having different structures were synthesized:

- For $x \leq 3.0$ the compounds are synthesized at a temperature of 408 K and a pressure lower than 1 bar.
- For $3.0 \leq x \leq 4.2$ ambient temperature and pressure are imposed.
- For $x = 5.0$, a pressure of 0.8 GPa is imposed at ambient temperature.

With increasing hydrogen concentration, it has been observed that the volume of the unit cell increases (see FIG 2.3.3) and structures which derive from an initial YFe_2 cubic one (space group Fd-3m) as reported in Tab 2.3.3 and shown in FIG 2.3.4 are observed. Indeed, with increasing hydrogen concentration, a symmetry lowering from the initial YFe_2 cubic phase is reported. It is explained by the ordering of hydrogen atoms in interstitial sites for well defined x values. More precisely, for $x = 1.2$ and $x = 1.9$ tetragonal distortions are observed (space group I-4) whereas for $x = 1.75$ the compound remains cubic with however a doubling of the cell parameter (space group I-43m). The compounds for $x = 2.5$ and $x = 2.9$ are also cubic with space group Fd-3m. Monoclinic structures (space group Pc) are found for $x = 3.5$ and $x = 4.2$ with a doubling of the b parameter for $x = 4.2$. Finally, for $x = 5$ an orthorhombic structure (space group Pmn21) is obtained. Despite not in the scope of this work, isotopic effects for hydrogen atoms were found for the highest stoichiometries. For instance YFe_2H_4 has a cell parameter 0.8% higher than that of YFe_2D_4 . This modifies the Ferromagnetic-Antiferromagnetic (FM-AFM) temperature transition: YFe_2H_4 has a FM-AFM temperature transition 50 K higher than in YFe_2D_4 [180, 183].

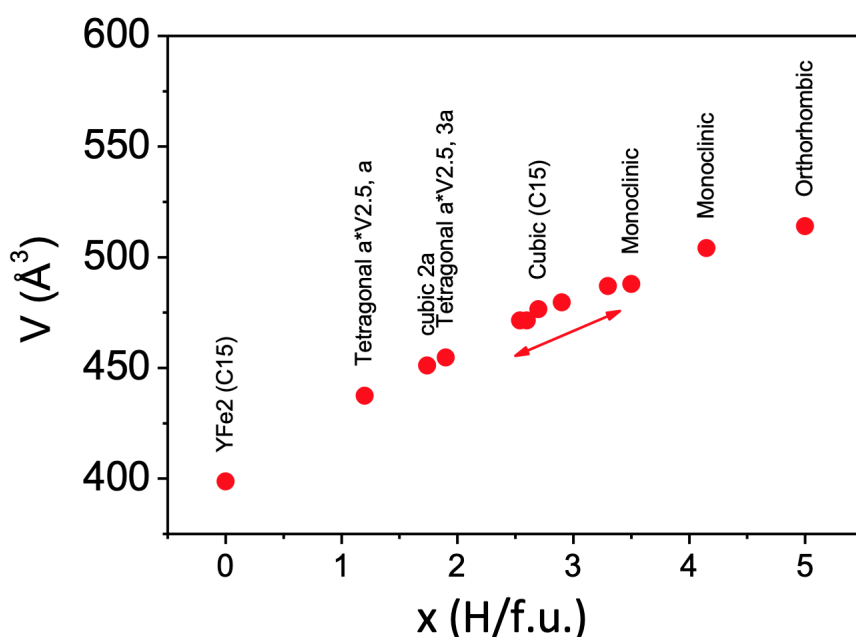


Figure 2.3.3: The different structures of YFe_2H_x experimentally found. The x-axis is the hydrogen concentration, the y-axis the volume of the unit cell (Figure from V. Paul-Boncour et al., presented in CONMAT2021, 18-22 October 2021, Valencia, Spain).

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
Fd-3m	0 GPa	$a=b=c=8.011 \text{ \AA}$ $\alpha = \beta = \gamma = 90^\circ$	Y	8a	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
			Fe	16d	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Table 2.3.3: Experimental YFe_2 structure at 0 GPa, cubic C15 structure (Fd-3m space group) [177].

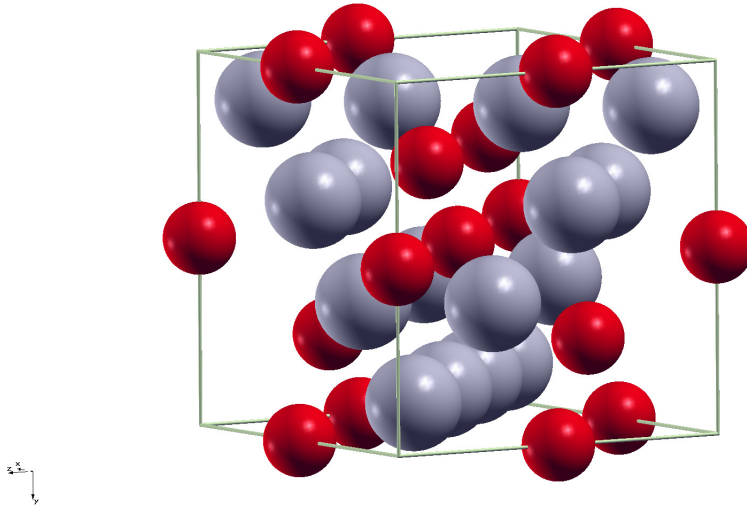


Figure 2.3.4: View of the YFe_2 crystal structure at 0 GPa, cubic C15 structure (Fd-3m space group).

The evolution of this system under high pressure has started to be investigated experimentally in the thesis of B. Guigue [184]. Hydrides were studied up to 35 GPa. Starting from the alloy YFe_2 , it has been highlighted an increase of the hydrogen stoichiometry with the formation of at least YFe_2H_8 near 30 GPa.

2.3.2 The AIRSS Calculations

Starting from the different elements presented in the former section, we now present and discuss the results obtained by AIRSS calculations on the YH_x and YFe_2H_x systems in the present work.

2.3.2.1 Binary Yttrium Superhydrides

In this subsection, the way the AIRSS algorithm is implemented to study the YH_x system is firstly precised. Phases predicted by former studies are retrieved, while new ones are discovered.

With these updated informations new convex hulls and stability ranges for the different YH_x phases are computed.

AIRSS Implementation

The AIRSS algorithm was applied to the YH_x systems with x being an integer or half integer between three and twelve. In each case, structures having one, two, three and four chemical formula units are searched. Two pressures were considered: 117 and 176 GPa. Tab 2.3.4, column 2 presents the different phases obtained. Several new structures have been identified.

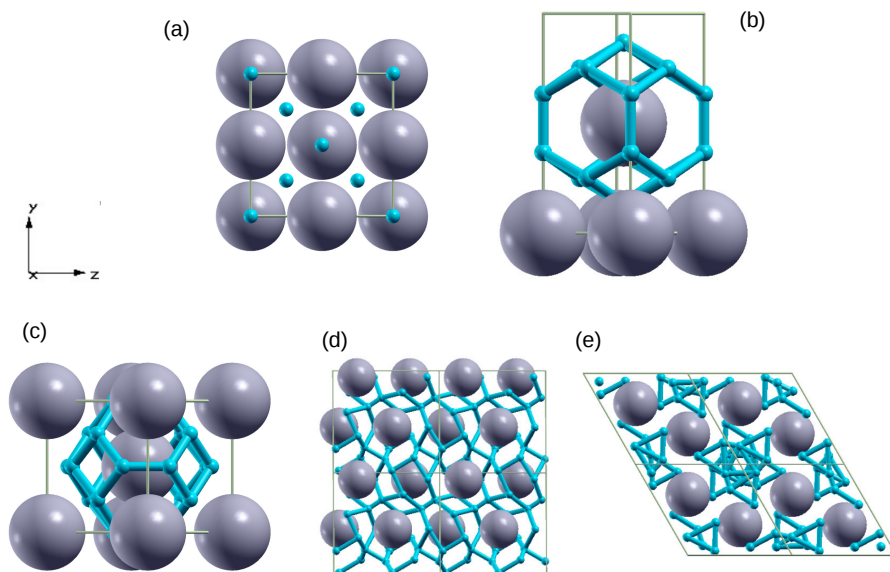
These phases have been then structurally optimized under pressure using the BFGS algorithm, with highly-converged parameters (a plane-wave cut-off of 30 Hartrees, a $12 \times 12 \times 12$ k -point mesh to sample the Brillouin Zone associated with the simulation cell, and an optimization criterion on atomic forces of 1.0×10^{-6} Ha/bohr and 1×10^{-8} Ha/bohr³ on stress tensor components), and their enthalpies computed as a function of pressure.

For each pressure, the most stable phase (among those obtained by AIRSS and the former studies) for each stoichiometry is indicated in Tab 2.3.4. The space groups written in bold are those already identified in the literature. The following remarks can be made about the AIRSS results.

- The same predictions as those of the literature are made for $\text{YH}_3, \text{YH}_4, \text{YH}_6, \text{YH}_8$. New predictions have been obtained for other stoichiometries, respectively YH_5, YH_7 and YH_{11} as well as all the half-integer ones.
- For YH_5 , a new phase containing three formula units per primitive cell having the space group P321 has been found.
- For YH_7 , an other phase of space group Cc is found more stable as the one pointed out in the literature with space group Imm2 [168].
- For YH_8 , the same structure as the one of the literature has been obtained at 176 GPa. However another new one is predicted more stable at lower pressure (with space group Cm).
- For YH_9 , a new phase at 176 GPa with space group $\text{P2}_1/\text{m}$ is predicted to be more stable than YH_9 $\text{P6}_3/\text{mmc}$ of the literature [99]. This new phase is however less stable than the YH_9 $\text{P6}_3/\text{m}$ at lower pressure (below 117 GPa).
- The YH_{10} Fm-3m phase has been retrieved but a new one is predicted to be more stable (with space group P1).
- Our AIRSS does not succeed in retrieving the YH_{12} C2/c structure described in Ref [98] or finding another one more stable. Indeed the structure found has the C2/m space group and appears as being less stable.

YH_x	Structures obtained by AIRSS	205 GPa	176 GPa	147 GPa	117 GPa	88 GPa
YH_3	Fm-3m, Pnma	Pnma	Pnma	Fm-3m	Fm-3m	Fm-3m
Y_2H_7	C2/m	C2/m	C2/m	C2/m	C2/m	C2/m
YH_4	I4/mmm	I4/mmm	I4/mmm	I4/mmm	I4/mmm	I4/mmm
Y_2H_9	P1	P1	P1	P1	P1	P1
YH_5	P321, P2 ₁ /m, P-1	P321	P321	P321	P321	P321
Y_2H_{11}	P-4m2	P-4m2	P-4m2	P-4m2	P-4m2	P-4m2
YH_6	Im-3m	Im-3m	Im-3m	Im-3m	Im-3m	Im-3m
Y_2H_{13}	P1, Cm	P1	P1	P1	Cm	Cm
YH_7	Cc	Cc	Cc	Cc	Cc	Cc
Y_2H_{15}	P1	P1	P1	P1	P1	P1
YH_8	Cc, Cm, P1	Cc	Cc	Cm	Cm	Cm
Y_2H_{17}	Amm2	Amm2	Amm2	Amm2	Amm2	Amm2
YH_9	P6₃/m, P2₁/m	P2 ₁ /m	P2 ₁ /m	P2 ₁ /m	P2 ₁ /m	P6₃/m
Y_2H_{19}	P1	P1	P1	P1	P1	P1
YH_{10}	P1, Fm-3m	P1	P1	P1	P1	P1
Y_2H_{21}	P1, Cm	P1	P1	P1	P1	P1
YH_{11}	P2 ₁ /m, P1	P2 ₁ /m	P2 ₁ /m	P2 ₁ /m	P2 ₁ /m	P2 ₁ /m
Y_2H_{23}	P1	P1	P1	P1	P1	P1
YH_{12}	C2/m, P1	C2/c	C2/c	C2/c	C2/c	C2/c

Table 2.3.4: YH_x phases retrieved by AIRSS in the present work (column 2) and those (among the structures identified by AIRSS and from the former studies) considered at each pressure for the determination of the convex hulls. The structures in bold are those already predicted in former studies [98, 169, 99, 31, 168].



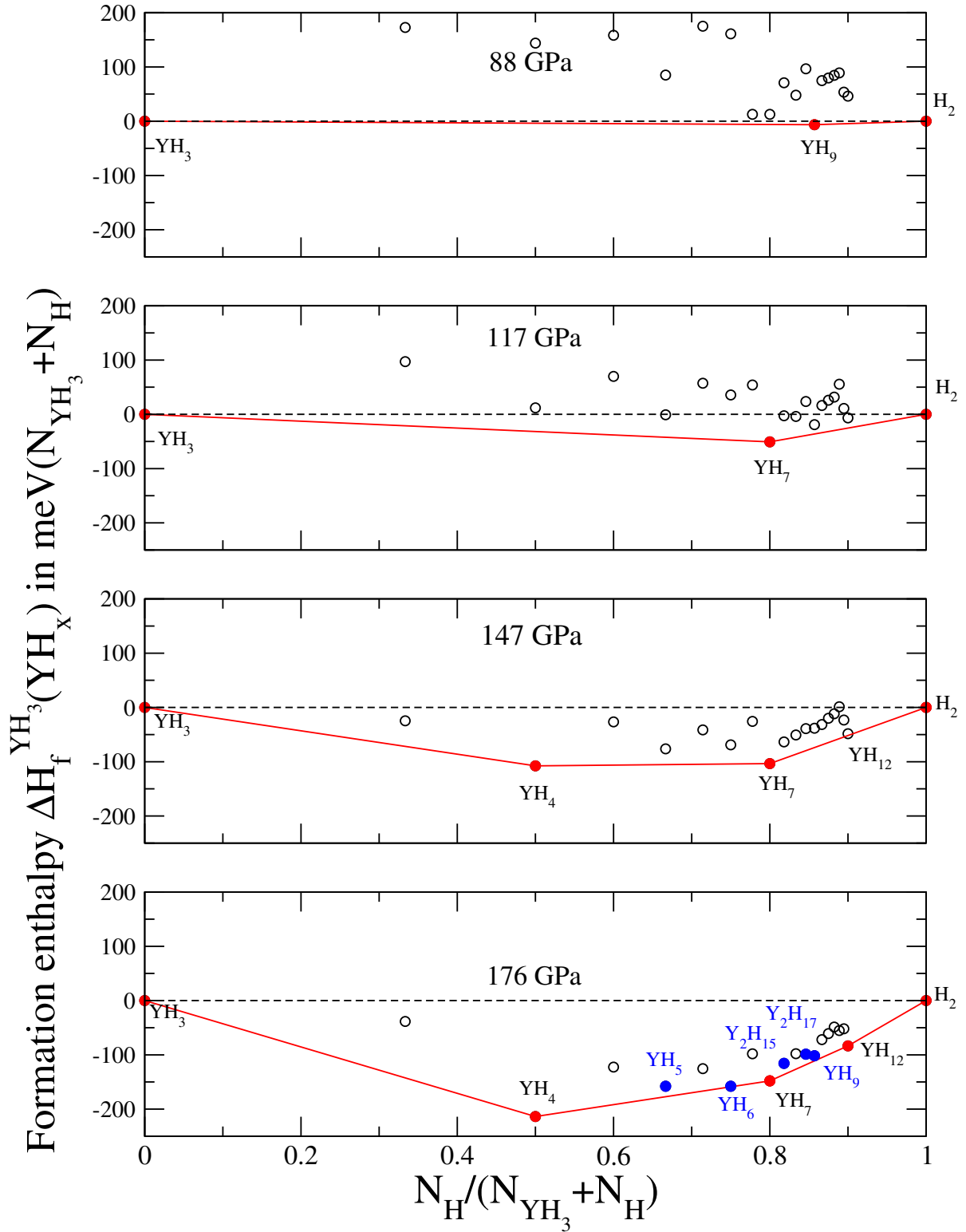


Figure 2.3.6: Formation enthalpies of the YH_x from 88 GPa to 176 GPa. The convex hulls are represented as red solid lines. Blue points correspond to stoichiometries not stable but very close to the convex hull at 176 GPa.

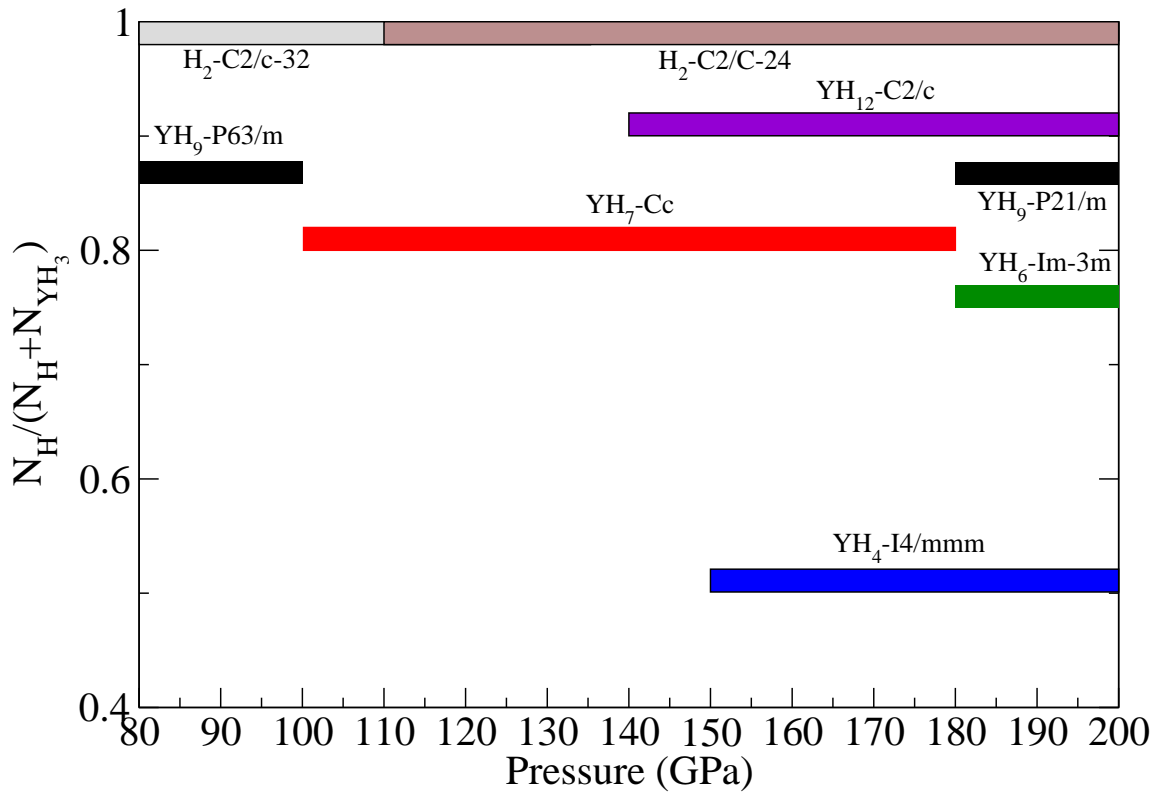


Figure 2.3.7: Stability domains of the YH_x predicted by AIRSS (the phases from the literature and those from our AIRSS study are taken into account).

Formation Enthalpies of the Hydrides and Convex Hulls

The new YH_x phases obtained are taken into account in order to determine new convex hulls of the formation enthalpies (see FIG 2.3.6). The stability domains of the compounds present on the convex hulls are given in 2.3.7. The following remarks can be made:

- The YH_7 Cc phase is predicted with a high stability. It appears on the convex hull of the formation enthalpies near 100 GPa. As a consequence, the YH_9 P6₃/m phase is no longer predicted stable at 117 GPa as well as YH_{12} .
- At 176 GPa, YH_6 is very close to the convex hull and is expected to stabilize a few GPa higher. This result is coherent with the experimental studies [168, 170] pointing out the stabilization of this phase near 180 GPa. Other structures are also found close, for instance Y_2H_{17} with space group Amm2, YH_9 with space group P2₁/m, YH_5 with space group P321, and Y_2H_{15} with space group P1.

The new structures of YH_x are shown in FIG 2.3.5. Their description in terms of Wyckoff positions and cell parameters and angles are given below in Tabs 2.3.5 and 2.3.6.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YH ₅							
P321	176 GPa	a=5.0350 b=5.0350 c=1.4175 $\alpha=90$ $\beta=90$ $\gamma=120$	Y	2d	0.333333	0.66667	0.33066
			Y	1a	0	0	0
			H	6g	0.32927	-0.07524	0.78067
			H	6g	0.06099	0.65599	0.87589
			H	3f	0.76070	0	0.5
YH ₇							
Cc	176 GPa	a=3.224 b=5.695 c=5.4279 $\alpha=90$ $\beta=90.863$ $\gamma=90.0$	Y	4a	0.39993	0.33412	0.41501
			H	4a	0.40087	0.17095	0.11041
			H	4a	0.56238	0.08630	0.19790
			H	4a	0.19375	0.40559	0.09719
			H	4a	-0.0651	0.15559	0.29919
			H	4a	0.41167	0.00350	0.40164
			H	4a	0.13296	0.08585	0.60848
			H	4a	0.84101	0.38571	0.19306
Y ₂ H ₁₇							
Amm2	176 GPa	a=5.5006 b=3.2890 c=5.9326 $\alpha=90$ $\beta=90$ $\gamma=90$	Y	2b	0.5	0	0.89074
			Y	2a	0	0	0.54717
			H	2a	0	0	0.21491
			H	4c	0.89047	0	0.88657
			H	4c	0.81596	0	0.02981
			H	8f	0.21436	0.21885	0.26688
			H	4c	0.68813	0	0.40962
			H	8f	0.30802	0.75120	0.14675
			H	2b	0.5	0	0.22499
			H	2b	0.5	0	0.56905

Table 2.3.5: Wyckoff positions, lattice constants and angles of the new predicted YH_x phases on or close to the convex hull at 176 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YH ₉							
P2 ₁ /m	176 GPa	a=3.3713 b=5.3752 c=3.4490 $\alpha=90$ $\beta=119.1530$ $\gamma=90$	Y	2e	0.33144	0.25	0.16252
			H	4f	0.70545	-0.05292	0.36898
			H	4f	0.17044	-0.06051	0.84442
			H	4f	0.84011	0.05292	0.63102
			H	4f	0.33502	0.84715	0.16847
			H	2e	-0.00105	0.25	0.49856
Y ₂ H ₁₅							
P1	176 GPa	a=3.259 b=3.305 c=5.380 $\alpha=90.044$ $\beta=90.840$ $\gamma=118.374$	Y	1a	0.30205	0.64051	0.04109
			Y	1a	0.62611	0.30708	0.55199
			H	1a	0.26803	0.67620	0.66415
			H	1a	0.70825	0.81299	0.75648
			H	1a	0.56409	0.78451	0.36843
			H	1a	0.29354	0.13644	0.84959
			H	1a	-0.03271	-0.01046	0.05667
			H	1a	0.75604	0.17459	0.87671
			H	1a	0.87618	0.13953	0.26561
			H	1a	0.33566	0.15202	0.19799
			H	1a	0.12781	0.22209	0.75854
			H	1a	0.80139	0.65552	0.85098
			H	1a	0.12384	0.31780	0.35606
			H	1a	0.87962	0.70765	0.26467
			H	1a	0.13004	0.80119	0.35867
			H	1a	-0.04867	-0.03571	0.55771
			H	1a	0.63400	0.30519	0.15905

Table 2.3.6: Wyckoff positions, lattice constants and angles of the new predicted YH_x phases on or close to the convex hull at 176 GPa.

As a conclusion, our AIRSS study of the yttrium hydrides and superhydrides allowed us to retrieve most of the results of the literature, with in addition a new phase, with YH₇ stoichiometry, that appears as being very stable between 100 and 180 GPa. However, none of the phases presented in this chapter appeared as matching with the experimental results (X-ray diffraction patterns) obtained by the experimentalists of our group [185], who synthesized yttrium hydrides in the 100-180 GPa range. It is possible that the real structures of the yttrium superhydrides possess large unit cells that would be unreachable by the AIRSS method. Note that, in the combined experimental and theoretical study of Troyan et al [168] and also in the experimental study of

kong et al [170] the X-ray diffraction experiments could not be interpreted with a single crystal structure (the X-ray diffraction pattern could only be understood on the basis of the superposition of the spectra associated with several structures). The structure of the yttrium superhydrides in the 100-180 GPa range remains therefore an open question, that probably requires to imagine and implement new methods to be elucidated. The implementation of such new approaches is the subject of part 3 of this thesis.

2.3.2.2 Yttrium-Iron Ternary Hydrides

In this section, the predictions obtained by AIRSS on the ternary system YFe_2H_x are presented and discussed. Firstly, a description of the AIRSS implementation is made. It is followed by a presentation of the different predicted stoichiometries. Finally, some physical quantities of the predicted compounds are discussed: Bader charge analysis, electronic density of states and pressure-evolution of the volume for each stoichiometry.

AIRSS Implementation

The AIRSS method is applied for pressures between 5 and 55 GPa by steps of 10 GPa. For each pressure, different (integer) hydrogen stoichiometries are considered. They are summed up in Tab 2.3.7. The generation of both the YFe_2 and hydrogen sublattices are held independently. For the YFe_2 sublattice, the crystal structure presented in Ref [175] and in Tab 2.3.3 having the Fd-3m space group, in a supercell containing 8 YFe_2 formula units, is imposed. It follows the idea that the YFe_2H_x remain interstitial hydrides. On the contrary, the hydrogen atom positions are initially determined using a random process. First of all they are randomly chosen under the constraint that distances to nearest neighbour metallic and hydrogen atoms exceed respectively 1.0 Å and 0.5 Å. A repulsive ZBL potential is afterwards employed (see appendix D) to move the atoms apart from each other until distances between hydrogen atoms and nearest neighbours of metallic and hydrogen atoms exceed respectively 1.5 Å and 0.8 Å.

Pressure	5 GPa	15 GPa	25 GPa	35 GPa	45 GPa	55 GPa
Lowest Stoichiometry Explored	YFe_2H_5	YFe_2H_5	YFe_2H_5	YFe_2H_6	YFe_2H_6	YFe_2H_6
Highest Stoichiometry Explored	YFe_2H_6	YFe_2H_7	YFe_2H_7	YFe_2H_8	YFe_2H_8	$\text{YFe}_2\text{H}_{10}$

Table 2.3.7: Pressures and stoichiometries considered for AIRSS calculations.

For a given stoichiometry at a given pressure where a structure searching is applied, the atomic configurations randomly drawn are first structurally optimized using low-precision numerical parameters: the plane-wave cut-off is set to 15 Ha, the k-point mesh to sample the Brillouin

Zone is $4 \times 4 \times 4$, and the stopping criterion for the structural optimization process is 10^{-3} Ha/bohr ($0.05 \text{ eV}/\text{\AA}$) on the atomic forces. During these optimizations, all the atom positions are relaxed, including the initial YFe_2 sublattice. Candidate structures are then selected at the end of this process as those having the lowest enthalpy. Then, we perform highly-converged calculations on the selected structures: the plane-wave cut-off is enhanced to 30 Ha, the k-point mesh to sample the Brillouin Zone is increased to $8 \times 8 \times 8$ and the structural optimization is pushed until all the cartesian components on the atomic forces are below 10^{-6} Ha/bohr. The most stable structure is then retained.

After implementing this procedure for each stoichiometry and each pressure considered, the formation enthalpies of the hydrides between 5 and 55 GPa can be plotted as a function of hydrogen content, and the convex hulls can be determined.

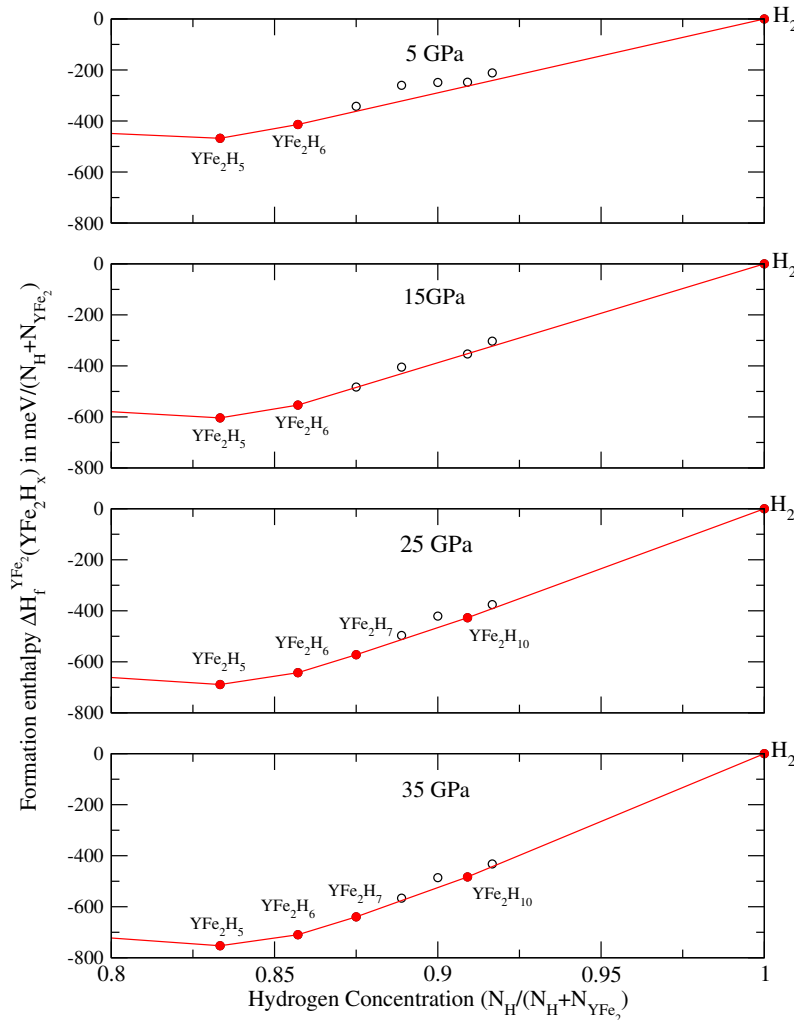


Figure 2.3.8: Formation enthalpies obtained for the YFe_2H_x system as a function of hydrogen concentration between 5 and 35 GPa. The convex hulls are plotted with red solid lines.

Stable Hydrogen Stoichiometries as a Function of Pressure

Formation enthalpies of the YFe_2H_x phases found are plotted in Figs 2.3.8 and 2.3.9 between 5 and 55 GPa, together with their convex hulls. For the computation of the formation enthalpy, the YFe_2 Fd-3m phase and the $C2/c - 32$ solid hydrogen phase are taken as references. The computed formation enthalpies correspond thus to the chemical reaction:

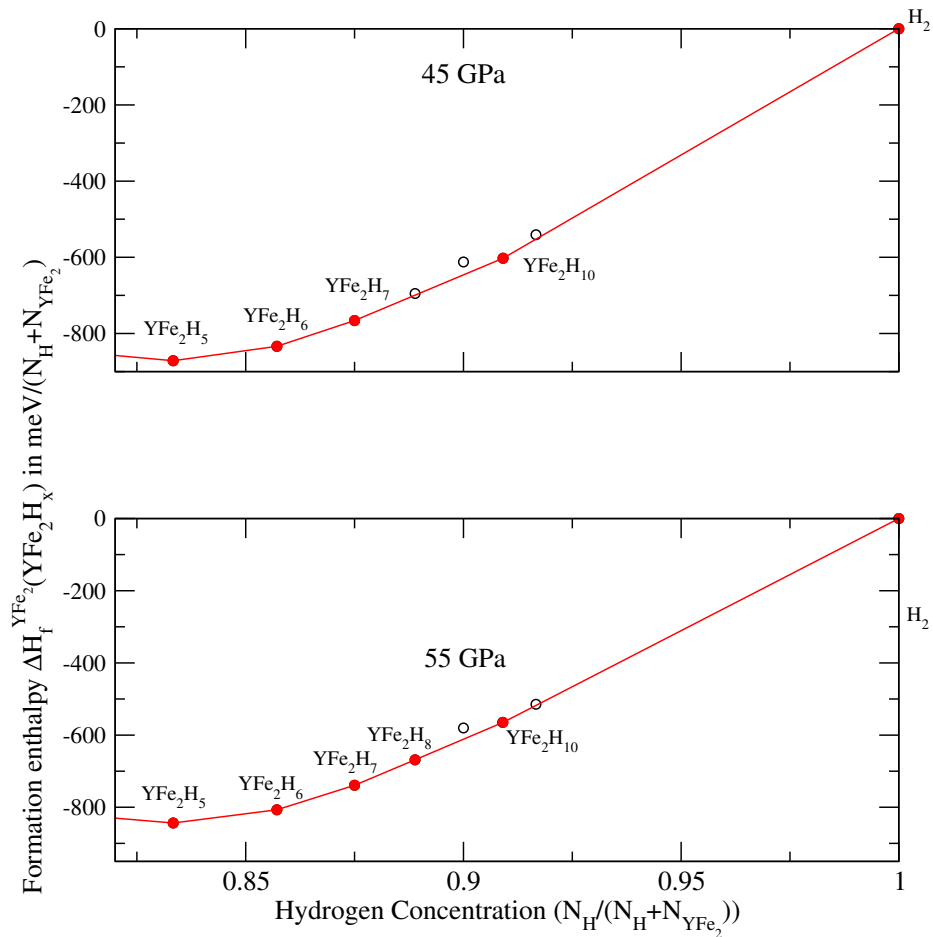
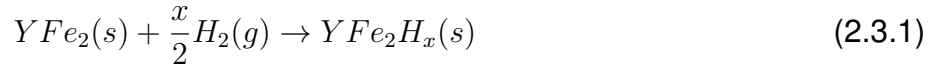


Figure 2.3.9: Formation enthalpies obtained for the YFe_2H_x system as a function of hydrogen concentration between 45 and 55 GPa. The convex hulls are plotted with red solid lines.

The stoichiometries YFe_2H_5 and YFe_2H_6 are expected to be stable at all the considered pressures. Rapidly, near 20 GPa, YFe_2H_7 and $\text{YFe}_2\text{H}_{10}$ are predicted to form (see FIG 2.3.8). $\text{YFe}_2\text{H}_{10}$ appears to be the highest stoichiometry that can be reached up to 55 GPa. It is expected to form at a pressure near ~ 25 GPa under hydrogen excess. This result suggests that this ternary hydride $\text{YFe}_2\text{H}_{10}$ enables a storage of hydrogen higher than the one that can be reached by the two separated binary compounds YH_3 and FeH_2 stable at this pressure range. It is surprisingly stabilized before YFe_2H_8 which is predicted stable near 50 GPa (see FIG 2.3.9). It is noticeable that all the compounds appearing on the convex hulls have a negative formation enthalpy, even at the lowest pressure here considered (5 GPa). It suggests that these materials might remain stable at ambient pressure after being formed at higher pressure.

Except for YFe_2H_5 which has a trigonal symmetry with space group R3, all the other compounds have triclinic symmetry (P1 space group). The metallic sublattices can be described however as a perturbed structure of the initial YFe_2 cubic one. It can be interpreted by the fact that the compounds found are interstitial hydrides where the hydrogen atoms randomly occupy some equivalent sites. As a result the disordered hydrogen sublattices found disturb the positions of the metallic atoms from the initial cubic structure. The Wyckoff positions of the predicted structures are given in Tabs 2.3.8,2.3.9,2.3.10,2.3.11,2.3.12,2.3.13.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YFe_2H_5							
R3	55 GPa	a=5.056 b=5.056 c=12.648 $\alpha=90^\circ \beta=90^\circ \gamma=120^\circ$	Y	3a	0	0	0.29072
			Y	3a	0	0	0.04418
			Fe	9b	0.50671	0.50814	0.16996
			Fe	3a	0.00000	0.00000	0.668979
			H	9b	-0.02477	0.39832	0.05274
			H	9b	0.29451	0.15509	0.16964
			H	9b	0.57722	-0.02963	0.28400
			H	3a	0.00000	0.00000	0.45485

Table 2.3.8: Wyckoff positions, lattice constants and angles of the new predicted YFe_2H_5 phase at 55 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YFe₂H₆							
P1	55 GPa	a=4.991 b=5.073 c=5.260 $\alpha=91.02$ $\beta=118.11$ $\gamma=119.11$	Y	1a	0.77276	0.65696	0.85559
			Y	1a	0.27906	0.41478	0.11868
			Fe	1a	0.02654	0.53122	0.48452
			Fe	1a	0.02189	0.03595	0.48227
			Fe	1a	0.53118	0.04561	0.48948
			Fe	1a	0.52705	0.03687	-0.00845
			H	1a	0.88462	0.03722	0.13967
			H	1a	0.60052	0.89586	0.27663
			H	1a	0.39358	0.68472	0.49491
			H	1a	0.09353	0.81159	0.33697
			H	1a	0.18469	0.06015	0.81905
			H	1a	0.32343	0.02827	0.15227
			H	1a	0.38563	0.26398	0.49194
			H	1a	0.33125	0.67589	0.83315
			H	1a	0.70883	0.39614	0.14072
			H	1a	-0.03036	0.26204	0.64416
			H	1a	0.81226	0.67490	0.49009
			H	1a	0.72515	0.21166	0.83053

Table 2.3.9: Wyckoff positions, lattice constants and angles of the new predicted YFe₂H₆ phase at 55 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YFe₂H₇							
P1	55 GPa	a=5.154 b=5.208 c=5.253 $\alpha=91.13$ $\beta=118.89$ $\gamma=119.33$	Y	1a	0.19820	0.10115	0.83272
			Y	1a	0.69407	0.84941	0.07793
			Fe	1a	-0.05091	0.47935	0.46813
			Fe	1a	-0.03521	0.48559	-0.0188
			Fe	1a	0.45532	-0.01963	0.46242
			Fe	1a	0.44100	0.47034	0.45546
			H	1a	0.23477	0.11687	0.45420
			H	1a	0.64596	0.81505	0.64044
			H	1a	0.62478	0.48227	0.79453
			H	1a	0.13696	0.66267	0.81772
			H	1a	0.28563	0.47031	0.66190
			H	1a	0.52153	0.25638	0.31914
			H	1a	0.78379	0.14214	0.81671
			H	1a	0.80190	0.12004	0.45636
			H	1a	0.80190	0.12004	0.45636
			H	1a	0.29119	0.83253	0.64158
			H	1a	0.29119	0.83253	0.64158
			H	1a	0.11889	0.83820	0.11405
			H	1a	0.79883	0.68685	0.45739
			H	1a	0.02062	0.33723	0.24723

Table 2.3.10: Wyckoff positions, lattice constants and angles of the new predicted YFe₂H₇ phase at 55 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YFe₂H₈							
P1	55 GPa	a=5.236 b=5.243 c=5.298 $\alpha=60.84$ $\beta=60.79$ $\gamma=89.93$	Y	1a	0.83756	0.57739	0.29589
			Y	1a	0.09121	0.32671	0.79548
			Fe	1a	0.48209	0.48343	0.01164
			Fe	1a	-0.01109	-0.03008	0.51803
			Fe	1a	0.46976	-0.02901	0.51867
			Fe	1a	0.44100	0.47034	0.45546
			H	1a	0.71702	0.77149	-0.08721
			H	1a	0.83668	0.61964	0.66822
			H	1a	0.47032	0.16276	0.67042
			H	1a	0.41718	0.59170	0.27477
			H	1a	0.47040	0.62895	0.66603
			H	1a	0.51009	0.30421	0.84504
			H	1a	0.28251	0.20786	0.40786
			H	1a	0.67527	0.32580	0.16051
			H	1a	0.68095	-0.03700	0.17205
			H	1a	0.83206	0.16781	0.67448
			H	1a	0.12406	-0.04781	0.17726
			H	1a	0.13805	0.33270	0.16738
			H	1a	0.79883	0.68685	0.45739
			H	1a	0.26594	0.82035	0.44345
			H	1a	0.33316	0.75987	-0.06009
			H	1a	0.10297	-0.09601	0.77387

Table 2.3.11: Wyckoff positions, lattice constants and angles of the new predicted YFe₂H₈ phase at 55 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YFe₂H₉							
P1	55 GPa	a=5.324 b=5.331 c=5.384 $\alpha=60.99^\circ$ $\beta=61.63^\circ$ $\gamma=60.89^\circ$	Y	1a	0.16453	0.15308	0.11350
			Y	1a	-0.09822	0.87437	0.83312
			Fe	1a	0.50935	0.51036	0.49212
			Fe	1a	0.00101	0.47406	0.52986
			Fe	1a	0.60413	0.01503	0.39412
			Fe	1a	0.46688	0.56519	0.01148
			H	1a	0.21793	0.86671	-0.06161
			H	1a	0.38492	0.49166	0.81887
			H	1a	0.04306	0.17949	0.79560
			H	1a	0.53192	0.84923	0.74908
			H	1a	0.82096	0.49184	0.85873
			H	1a	0.17974	0.48968	0.68373
			H	1a	0.54811	0.18958	0.08025
			H	1a	0.74606	0.69689	0.34531
			H	1a	0.23989	0.17278	0.4786
			H	1a	0.26407	0.62754	0.31964
			H	1a	0.84578	0.30272	0.50863
			H	1a	0.34625	0.85650	0.48128
			H	1a	0.66745	0.32779	0.25354
			H	1a	0.87725	0.01319	0.12389
			H	1a	0.12886	0.53178	0.16405
			H	1a	0.54263	0.72958	0.13043
			H	1a	-0.08754	0.85095	0.45966
			H	1a	0.52464	0.16539	0.65324

Table 2.3.12: Wyckoff positions, lattice constants and angles of the new predicted YFe₂H₉ phase at 55 GPa.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
YFe₂H₁₀							
P1	55 GPa	a=4.279 b=5.181 c=6.091 $\alpha=88.99^\circ$ $\beta=66.88^\circ$ $\gamma=117.20^\circ$	Y	1a	0.10478	0.40063	0.69269
			Y	1a	0.89572	0.77649	0.08622
			Fe	1a	0.49156	0.83735	0.68602
			Fe	1a	0.46541	0.61608	0.38063
			Fe	1a	-0.09866	0.07615	0.45416
			Fe	1a	0.54768	0.38555	0.05371
			H	1a	0.31202	0.41574	0.33336
			H	1a	0.66686	0.78350	0.44755
			H	1a	0.06393	0.00277	0.70550
			H	1a	0.85704	0.41630	0.39285
			H	1a	0.52471	0.13496	0.54383
			H	1a	0.39829	-0.07602	0.22876
			H	1a	0.25304	0.08689	0.44041
			H	1a	0.53894	0.50251	0.64187
			H	1a	0.62490	0.10005	0.26085
			H	1a	0.60996	0.22166	0.88299
			H	1a	0.22942	0.32721	0.08415
			H	1a	0.74780	0.55924	-0.09995
			H	1a	0.46268	-0.01284	0.85737
			H	1a	0.07688	0.77053	0.40572
			H	1a	0.18853	0.04144	-0.00721
			H	1a	0.28712	0.64622	0.85651
			H	1a	-0.00826	0.15800	0.00138
			H	1a	0.45137	0.68166	0.11475
			H	1a	-0.07178	0.26748	0.19924
			H	1a	0.84547	0.84340	0.69877

Table 2.3.13: Wyckoff positions, lattice constants and angles of the new predicted YFe₂H₁₀ phase at 55 GPa.

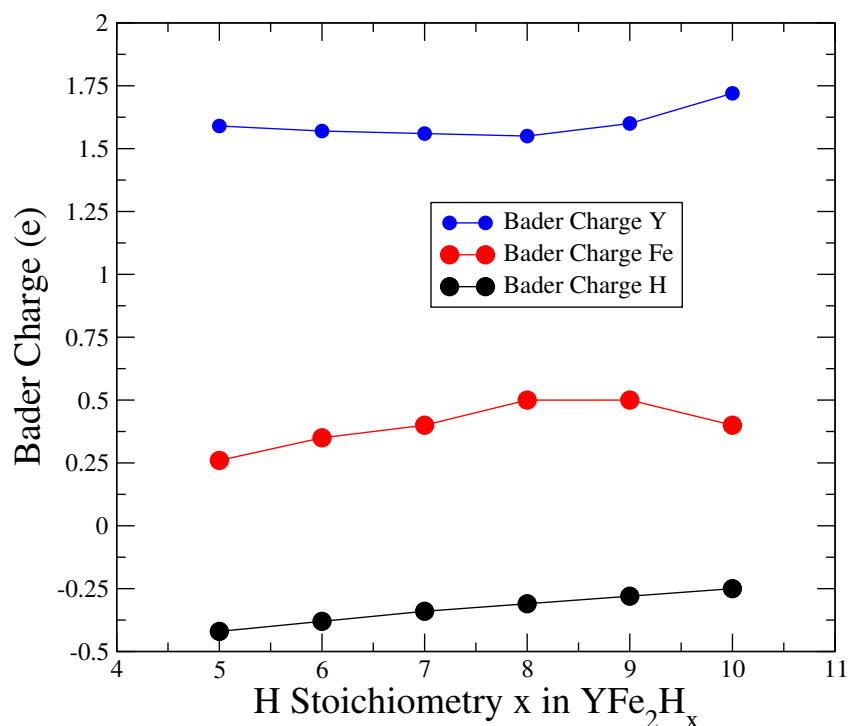


Figure 2.3.10: Bader charge evolution with hydrogen stoichiometry averaged over the unit cell at 55 GPa.

Bader Charges

The Bader charges are computed for the different hydrogen stoichiometries, at 55 GPa (see FIG 2.3.10). It can be checked that the compounds are hydrides, as expected (i.e. with an electronic transfer from metallic to hydrogen atoms). The yttriums are more oxidized than the iron atoms (atomic charge=1.5-1.75 e for yttrium and 0.25-0.5 for iron), which can be easily interpreted from the difference of electronegativity between hydrogen and the metals: on the Pauling scale, the electronegativities of Y, Fe and H are respectively 1.22, 1.83 and 2.2. In the present system, the largest difference being between Y and H, the largest electronic transfer is expected between these two elements, which is indeed found in the calculation.

Electronic Properties

The electronic density of states of all the predicted compounds are computed at 55 GPa. In all the cases, a gap in the range of 0.3-0.5 eV is found at the Fermi level. These hydrides are thus not metals. The curves in FIGs 2.3.11 and 2.3.12 display the different densities of states computed.

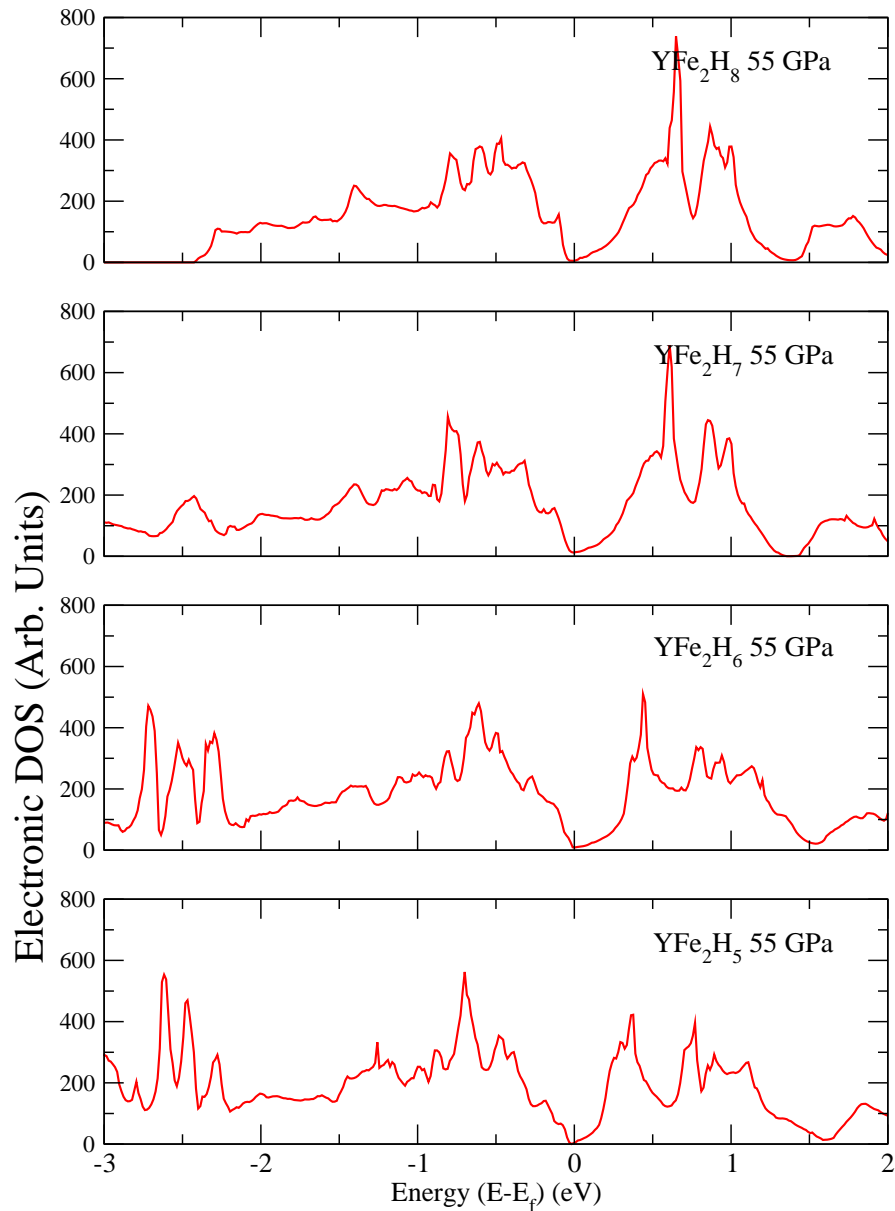


Figure 2.3.11: Electronic densities of state from YFe_2H_5 to YFe_2H_8 at 55 GPa.

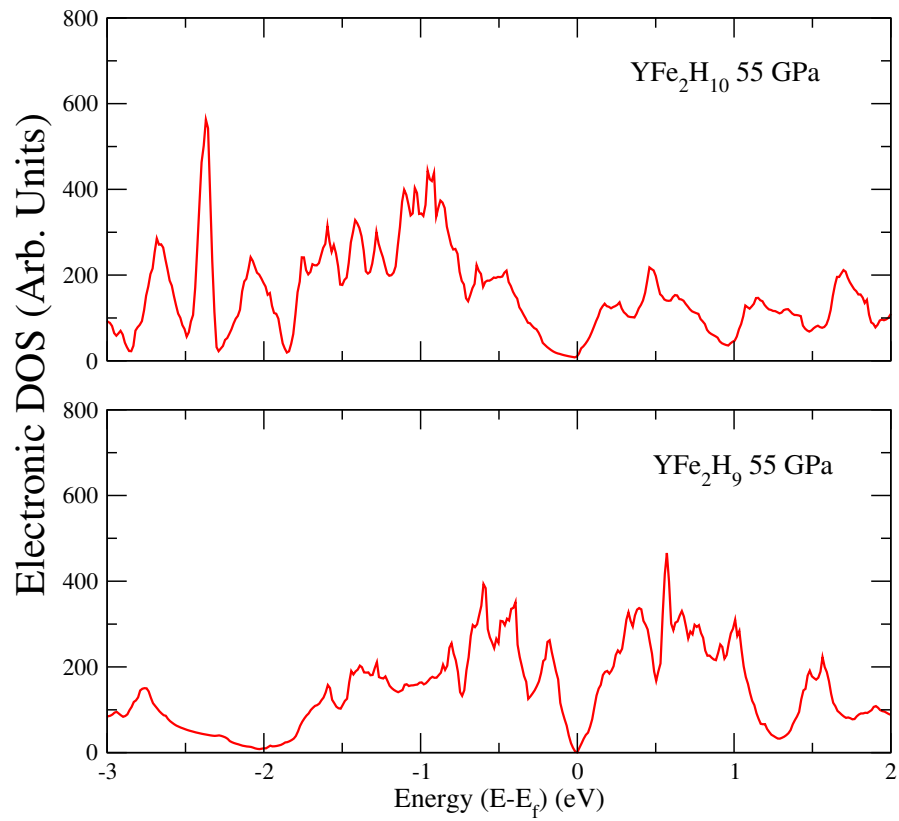


Figure 2.3.12: Electronic densities of state from YFe₂H₉ to YFe₂H₁₀ at 55 GPa.

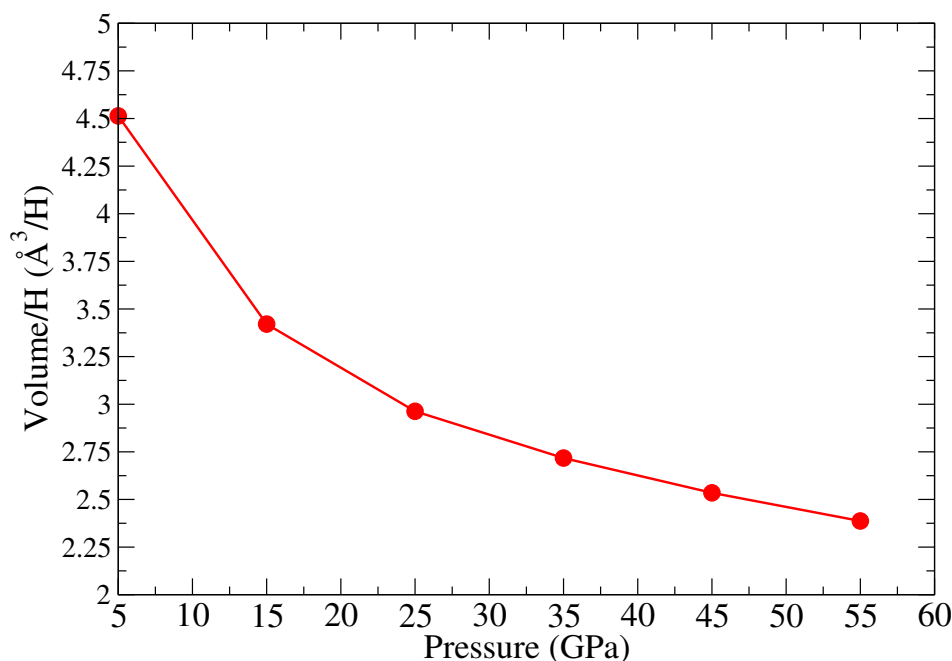


Figure 2.3.13: Additional volume per added hydrogen in the YFe_2 alloy and evolution of this additional volume value with pressure.

Volume Evolution with Hydrogen Stoichiometry

For each pressure considered, we plot the volume of the formula unit as a function of hydrogen content: the obtained curve is roughly linear, and its slope can be interpreted as the supplement of volume induced by the insertion of a supplementary hydrogen atom in the unit cell. As presented in FIG 2.3.13, the slope of this linear behaviour evolves with pressure. It goes from around $4.5 \text{ \AA}^3/\text{H}$ near 5 GPa to $2.5 \text{ \AA}^3/\text{H}$ near 50 GPa. For future experimental synthesis of this system, this information could be useful to experimentalists, in particular to determine the hydrogen content of the synthesized hydrides.

2.3.3 Conclusion

This study shows that the yttrium element can form hydrides and superhydrides with very large hydrogen content under pressure. Above 100 GPa, binary yttrium superhydrides might reach hydrogen stoichiometry higher as seven. The case of the ternary system YFe_2H_x reveals that an hydrogen-rich interstitial compound, $\text{YFe}_2\text{H}_{10}$, might be formed near ~ 25 GPa, which has to be confirmed by experiments.

The comparison between calculations and experiment points out some difficulties, in the case

of the binary yttrium system. Between 140 and 180 GPa, the phases proposed up to now by CSP methods, including the AIRSS study presented in this work, do not simply explain the experimental X-ray spectra obtained (neither those of Troyan et al [168], Kong et al [170], Snider et al [171], nor those obtained by the experimentalists of our group [185]). For instance, near 140 GPa, the X-ray diffraction pattern is characterized by peaks positioned at low diffraction angles. It suggests that the synthesized material should exhibit large interplanar distances, which could probably be obtained only with a primitive cell made of a number of chemical formula units that probably exceeds the typical size here investigated. If so, it means that the CSP algorithms coupled to DFT calculations which are limited in their exploration of configuration space to 20-30 atom per unit cell are not able to deal with the YH_x system due to the large unit cell of the latter.

A too large dimension of the configuration space is the main obstacle to the success of CSP methods when coupled to DFT. Similarly, treating by such a methods the ternary systems like YFe_2H_x in their generality may be extremely difficult due to the additional degrees of freedom associated with the introduction of a third chemical element. The study of YFe_2H_x was facilitated by the fact that the Y:Fe ratio was imposed, and the metallic sublattice known from experiments.

It is thus necessary to propose new approaches able to deal with such cases. This is the subject of part 3 of this thesis.

2.3.4 Summary of Part 2

This part confirms that transition metals such as Mn and Y are worth being considered for the synthesis of superhydrides under pressure. Some of these systems might reach important hydrogen stoichiometries like MnH_7 near 40 GPa, and probably a stoichiometry higher than seven for YH_x above 100 GPa. The inclusion of iron atoms in the YH_x system makes possible the formation of ternary interstitial type hydrides having important hydrogen concentrations at lower pressure. A compound $\text{YFe}_2\text{H}_{10}$ is thus predicted to form at just ~ 25 GPa and might remain stable at room pressure. However, if large hydrogen stoichiometries are reached for Y and Mn hydrides (which belong to the left and center of the series of transition metals), this is not the case for copper, probably because it lies on the right of the series and therefore has a too large electronegativity (Y: 1.22, Mn: 1.55, Cu: 1.9 versus 2.2 for H).

The study of the yttrium superhydrides highlights the limits of CSP algorithms when they are coupled to DFT calculations. Without experimental informations (e.g. atomic positions within the metal sublattice, lattice parameters or angles, space group, etc), such algorithms do not succeed in predicting crystal structures having a large number of atoms in the primitive cell (above 20-30). To tackle these challenges, the next part of this PhD proposes some strategies using a machine learning approach. It will be in particular applied to the YH_x system so as to propose a crystal structure explaining better the experimental results of X-ray diffraction obtained near 140 GPa.

PART 3

TOWARD THE PREDICTIONS OF MORE COMPLEX SUPERHYDRIDES

Introduction

This last part presents the strategy considered to enable research of more complex superhydrides. As illustrated with the binary yttrium hydrides and superhydrides, improvements of the CSP methods nowadays used seem required to predict compounds having huge unitary cell (typically containing more than 30 atoms). This need is also crucial for ternary systems which might form superhydrides under lower pressure ranges. For this purpose, some machine learning approaches appear being promising. The first chapter presents thus the machine learning tools considered and how they can be included in CSP algorithms. The second chapter describes a first proof of concept applied to the predictions of tin phases. The third chapter goes further and tries to employ CSP algorithm with machine learning methods for the predictions of yttrium superhydride phases. Finally, some perspectives and first studies are presented for investigations about ternary hydride and superhydride systems.

Chapter 3.1

Machine Learning and CSP Algorithms

Contents

3.1.1 Numerical Potentials	115
3.1.1.1 Numerical Potentials: Main Ingredients	115
3.1.1.2 The Descriptors	116
3.1.1.3 The Main Statistical Models Employed	119
3.1.2 Spectral Neighbour Analysis Potential	120
3.1.2.1 The Bispectrum Descriptor	121
3.1.2.2 Estimating Physical Properties with SNAP	121
3.1.2.3 The Training Process	124
3.1.3 Numerical Potentials and Random Searching	126

When complexity of the system increases, for instance when the unit cell contains many atoms, with possibly several chemical types, the PES is more complicated to explore. Indeed, as more degrees of freedom are involved, the PES becomes an extremely high-dimensional space (dimensionality crisis). From Ref [87] it is even expected that the number of local minima to visit rises exponentially with the number of atoms in the unit cell. Consequently, this exponential rise has to be accounted for by an equivalent increase in the number of generated structures during the PES sampling. At the same time, as more atoms are present in the unit cells, the computation cost for each structural optimizations increases as well (with DFT, the cost of these optimizations rises as the number of valence electrons to the power three). This last remark is particularly important for the CSP algorithms that use DFT-based calculations to perform the structural optimizations. It leads to a real limitation for the predictions of materials. To illustrate this point, if one considers a compound with a fixed stoichiometry having around 30 atoms in the primitive cell and two different chemical types, more than 1 million CPU hours would be required with AIRSS to identify the most stable configurations. As a result, it is impossible to have an exhaustive structure search in a reasonable amount of time if around one hundred atoms are simulated. This is worse if more chemical types are considered because the relative concentrations between the different chemical types become supplementary degrees of freedom. However, it is worth noticing that more than 99 % of the computation cost for structure searching is utilized for the structural optimizations using DFT [110]. One important question is then to reduce the time required to optimize the configurations and use this new available time resource to sample more in depth the PES. Some opportunities can be looked for in this direction. One of these is to use "classical" numerical potentials able to perform fast structural optimizations (here the term "classical" is used by opposition to "ab initio"). This chapter describes more in depth these potentials and how they can be combined with Random Structure Searching (RSS) to predict new materials.

3.1.1 Numerical Potentials

This section presents in detail the notion of numerical potential. Firstly the general principles underlying this notion are described. Then, the notion of descriptor is presented, including examples. Finally, some statistical models playing crucial roles in the construction of numerical potentials are discussed.

3.1.1.1 Numerical Potentials: Main Ingredients

Numerical potentials are "classical potentials" (by opposition to "ab initio") based on machine learning methods. As opposed to empirical potentials characterized by an analytical form, they offer a better flexibility because their construction relies on the training on a database. Building a numerical potential requires three main ingredients:

- A training database storing informations about configurations coming from accurate physics

computations (such as DFT computations) which constitute a set of reference data. The data typically take the form (x_i, y_i) where the x_i correspond to the configurations and y_i to physical quantities such as energy, forces and stress. The choice of the database is crucial as it determines the capabilities of the optimized numerical potential to "learn" the PES. If not enough information is stored in it, the potential may not be able to reproduce the local and global minima associated with dynamically stable crystal structures.

- A descriptor model employed to capture informations about atomic environments. The atomic configurations are represented by vectors or matrices.
- A statistical model which takes the coefficients of the descriptor representing atomic configurations as inputs. It gives as output values of the important physical quantities to be predicted (energy, forces, stress), that are associated to the atomic configuration used as input. It corresponds here to a supervised learning task.

The main descriptors and statistical models used for predictions of materials are described in the following paragraphs.

3.1.1.2 The Descriptors

In the machine learning methods, representing the atomic configurations by a multidimensional vector made of all the atomic position vectors is not appropriate, because this mathematical object does not reflect the physical invariance associated with translation, rotation and permutation of identical atoms. Thus, before using any model for physical quantity predictions, it is required to establish how the atomic configurations can be described. This is the purpose of the descriptor model chosen. A key condition to have a reliable descriptor is to enforce invariance related to permutation operations between equivalent atoms, translations or rotations. This condition comes from the fact such invariances are verified for physical laws, and also for the potential energy [187]. Several descriptors are considered in the literature. Among the most important ones, it can be mentioned:

- The Coulomb Matrix.
- The Moment Tensor Polynomials (MTP)
- The Radial and Angular Symmetric Functions.
- The Smooth Overlap of Atomic Positions (SOAP).
- The Crystal Graph Convolutional Neural Network (CGCNN).
- The Bispectrum.

These different descriptors are now briefly described one by one.

Coulomb Matrix

The Coulomb matrix was part of the first introduced descriptor models [188]. It was developed by Rupp and al [189]. It aims at describing an atomic configuration thanks to a matrix (M_{ij}) using the Coulomb interaction for the computation of the coefficients.

$$\begin{cases} M_{ij} = 0.5 \times Z_i^{2.4} \text{ for } i = j \\ \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \text{ for } i \neq j \end{cases} \quad (3.1.1)$$

The eigenvalues of this matrix are used to evaluate the similarities between different configurations. This approach was firstly applied to molecules and then generalized for periodic systems such as crystal structures [190, 191]. In particular Schutt et al [191] included the geometric properties of the primitive cell. However this descriptor suffers from different problems.

- It appears that two different systems can have the same Coulomb Matrix [192].
- In the approach proposed in Ref [191], the descriptor is not a smooth function of the coordinate system (a smooth changes of the primitive vectors do not modify smoothly the descriptor coefficients).

Moment Tensor Polynomials

Moment Tensor Polynomials (MTP) were introduced by Shapeev [193]. The main principle is to construct descriptors in the form of polynomials which are invariant when permuting their different variables (which are for instance the atomic positions of the same chemical type). For fixed μ and ν positive integers, the MTP polynomial $M_{\mu,\nu}$ has the form:

$$M_{\mu,\nu}(\mathbf{X}_1, \dots, \mathbf{X}_n) = \sum_{i=1}^n |\mathbf{X}_i|^{2\mu} \mathbf{X}_i^{\otimes \nu} \quad (3.1.2)$$

where \mathbf{X}_i are the atomic coordinates of atom i , $\mathbf{X}_i^{\otimes \nu} = \mathbf{X}_i \otimes \mathbf{X}_i \dots \otimes \mathbf{X}_i$ with \otimes the Kronecker product.

The Radial and Angular Symmetry Functions

This model was introduced by Behler and al [194, 195]. Both radial and angular informations around each atom are represented by sums of pairwise functions. For a given atom i , the radial information is given by a sum of the form [188]:

$$G^R(\mathbf{R}_i) = \sum_{j \neq i} g^R(\mathbf{R}_{ij}) f_c(r_{ij}) \quad (3.1.3)$$

where g^R is a function of the relative positions between the i -th atom and a j -th neighbour one. These relative positions are represented by the vector $\mathbf{R}_{ij} = \mathbf{R}_j - \mathbf{R}_i$. f_c is a function of the

distance between the i -th atom and a j -th neighbour one denoted by r_{ij} . In a similar way, the angular information is given by:

$$G^A(\mathbf{R}_i) = \sum_{j,k \neq i} g^A(\theta_{ijk}) f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}) \quad (3.1.4)$$

Here g^A is a function of the angle θ_{ijk} formed between $\mathbf{R}_j - \mathbf{R}_i$ and $\mathbf{R}_k - \mathbf{R}_i$ with j and k -th atoms being neighbours of i .

The Smooth Overlap of Atomic Positions (SOAP)

This descriptor was firstly introduced by Bartok and al in [196, 187]. The starting quantity used is the density of neighbouring atoms of an atom i expressed with the help of gaussians centered at each neighbour.

$$\rho_i(\mathbf{r}) = \sum_{||\mathbf{r}_i - \mathbf{r}_j|| < R_{cut}} e^{-\alpha ||\mathbf{r} - \mathbf{r}_j||^2} \quad (3.1.5)$$

where \mathbf{r}_j are the positions of the neighbouring atoms around the atom i positioned at \mathbf{r}_i . R_{cut} is a cut-off radius determining a maximum distance from the \mathbf{r}_i position used to define the neighbourhood of atom i . α is a strictly positive real number. This descriptor is based on a rotational invariant kernel defined as:

$$k(\rho, \rho') = \int d\hat{R} \int d\mathbf{r} \rho(\mathbf{r}) \rho'(\hat{R}\mathbf{r}) \quad (3.1.6)$$

This is a translational and permutational invariant quantity. In order to acquire translational and rotational invariances, \hat{R} corresponds to all the rotation operations, ρ and ρ' are two atomic densities. With this defined kernel, a similarity measurement between two local atomic environments associated respectively with the densities ρ and ρ' can be defined:

$$K(\rho, \rho') = \left[\frac{k(\rho, \rho')}{\sqrt{k(\rho, \rho) k(\rho', \rho')}} \right]^n \quad (3.1.7)$$

where n is a positive integer. This defined kernel K is then used for predictions.

Crystal Graph Convolutional Neural Networks (CGCNN)

This approach as presented by Xi and al in Ref [197] consists in representing a crystal structure by a graph containing the informations about the atoms (for example their chemical types) and their bonds. Each atom i is associated with a vector v_i . A bond between atom i and j is represented by the vector u_{ij}^k . Due to periodic boundary conditions, two atoms can be linked by several bonds. The k -index aims at indexing these bonds. For a given atom i , the vector v_i is concatenated with the v_j vectors corresponding to the neighbours of i as well as the ones representing the bonds (the u_{ij}^k). The matrix obtained associated with atom i goes through successive layers of a convolutional neural network. It gives at the output a vector v_i^C . This process is applied for all the atoms in the crystal configuration. All the obtained vectors v_i^C are concatenated and a max pooling operation (consisting in taking the highest coefficients in successive submatrices of

the input whose size are determined by a window which runs through the entire input matrix) is applied on the obtained matrix. A single vector W representing the global crystal configuration and corresponding to the expected descriptor is thus obtained. The idea of using graph combined with a neural network was also proposed in other works like Ref [198, 199]. A generalization of this approach has been introduced by Gilmer and al in Ref [200] with the Message Passing Neural Network framework.

The Bispectrum

This descriptor model as introduced by Bartok and al in Ref [201] uses the density of the neighbouring atoms decomposed in terms of 4D spherical harmonics. This descriptor consists in decomposing the atomic densities around each atom on a basis of 4D spherical harmonics. The coefficients obtained by these decompositions are used to build the descriptor. It is presented more in depth in the next section 3.1.2 because it is the one mainly used in this work..

Once the descriptor has been selected, the database is mapped onto the descriptor space, a chosen statistical method is applied in this space for learning. The main methods employed are described in the following subsection.

3.1.1.3 The Main Statistical Models Employed

In the majority of the machine learning techniques applied to materials physics, the starting point is to consider the total energy of a given configuration as the sum of local contributions of each atom and their geometric environments:

$$E = \sum_{i=1}^N E_i \quad (3.1.8)$$

In this expression, E_i is the local "atomic" energy of atom i . The machine learning models considered are then asked to derive each contribution E_i as well as those directly associated such as forces and stress. They make these predictions from the descriptors computed from each of these environments. For this purpose, a function f_α is employed. The contribution of an atom i to a target physical quantity (such as the energy) is given by $f_\alpha(d_i)$ where d_i is the descriptor computed from the environment of this atom and α being the parameters of the function (for instance in the case of a linear regression, α denotes the coefficients of the linear combination). The physical quantity (mainly the total energy) of a given crystal configuration having N atoms is then obtained by $y = \sum_j^N f_\alpha(d_j)$. This function f_α is prepared thanks to training on a crystal configuration database where the physical quantities are known (e.g. from accurate DFT calculations). This training process corresponds to the minimization of the quantity relatively to the α parameters:

$$\sum_l \left(\sum_m^N (f_\alpha(d_m^{(l)}) - y_l)^2 \right) + \lambda \|\alpha\|^2 \quad (3.1.9)$$

The first sum (with l -index) runs over all the configurations in the training database, the second sum runs (with m -index) over the atoms. y_l is the true physical quantity related to the l -th configuration. λ is a regularization coefficient introduced to avoid numerical instabilities. The f_α function can have various forms. In the case of a multidimensional linear regression, it is a linear combination of the coefficients of the descriptor. For kernel methods, f_α is a linear combination of scalar products in a high dimensional space determined by the kernel K .

$$f_\alpha(d) = \sum_j^N \beta_j K(d, d_j) \quad (3.1.10)$$

K can be for instance the one introduced above with the SOAP descriptor 3.1.1.2. Finally, in the case of a neural network, f_α is a nonlinear function. The value obtained at the output of a given layer l receiving as input the vector d (corresponding to a descriptor or to its transformation by precedent layers) is given by:

$$f_l(d) = \sigma_l(W_l d + B_l) \quad (3.1.11)$$

where W_l and B_l are the weight and bias vectors of the layer and σ_l is a non linear activation function (sigmoid, relu ...). The dimensions of W_l and B_l are determined by the number of computation units, the "neurons", forming the layer l . In the case of a multilayer perceptron, the global f_α function has the form:

$$f_\alpha(d) = f_L \circ f_{L-1} \dots \circ f_1(d) \quad (3.1.12)$$

Several types of these models combined with different kinds of descriptors can be considered, in particular:

1. Kernel based method with the SOAP descriptor like in the Gaussian Approximation Potential (GAP) [201, 108].
2. Neural network using radial and angular symmetry functions [194] or CGCNN [197].
3. Multidimensional linear regression with the use of a bispectrum descriptor like in the Spectral Neighbour Analysis Potential (SNAP) [196].

Among all the formalisms mentioned above, the Spectral Neighbour Analysis Potential (SNAP) will be mainly employed. This is the main subject of the next section.

3.1.2 Spectral Neighbour Analysis Potential

This section focuses on the Spectral Neighbour Analysis Potential (SNAP) method. First of all the bispectrum descriptor is more in depth presented. The way the predictions of the physical quantities using bispectrum are performed is described afterwards. Finally, the training process involved in the SNAP model is discussed.

3.1.2.1 The Bispectrum Descriptor

To compute the bispectrum coefficients for a given atom, the starting quantity to consider is the local atomic density as expressed in Eq 3.1.5. The main idea is to decompose the atomic density in terms of the 4D spherical harmonics (noted $U_{m,m'}^j$) [202]. The obtained coefficients from this decomposition are afterwards used to build the bispectrum matrix. As said, the density is expanded on the basis function $U_{m,m'}^j(\theta_0, \theta, \phi)$ where j is an integer or a half integer and m, m' are integers between $-j$ and j . The two angles θ and ϕ are those used for the 3D sphere. The angle θ_0 is defined by:

$$\theta_0 = \alpha_0^{max} \pi \frac{r}{R_{cut}} \quad (3.1.13)$$

In this above formula, r is the radial coordinate used in the 3D spherical coordinate system, α_0^{max} is a parameter between 0 and 1 and R_{cut} is a cut-off radius fixing the size of the atomic environments. The atomic density can thus be expressed as [202]:

$$\rho(\mathbf{r}) = \sum_{j=0, \frac{1}{2}, 1, \dots}^{j_{max}} \sum_{m=-j}^j \sum_{m'=-j}^j \alpha_{m,m'}^j U_{m,m'}^j(\theta_0, \theta, \phi) \quad (3.1.14)$$

A rotation invariant quantity can be established by the use of the $\alpha_{m,m'}^j$ coefficients. They are indeed employed to construct a third-rank matrix whose coefficients are given by:

$$B_{j_1, j_2, j_3} = \sum_{m_1, m'_1=-j_1}^{j_1} \sum_{m_2, m'_2=-j_2}^{j_2} \sum_{m_3, m'_3=-j_3}^{j_3} \alpha_{m_1, m'_1}^{j_1} C(m_1, m'_1, j_1, m_2, m'_2, j_2, m_3, m'_3, j_3) \alpha_{m_2, m'_2}^{j_2} (\alpha_{m_3, m'_3}^{j_3})^* \quad (3.1.15)$$

with $C(m_1, m'_1, j_1, m_2, m'_2, j_2, m_3, m'_3, j_3)$ are some coupling coefficients (the Wigner-3j coefficients [202]). As the introduced summations in Eq 3.1.15 can not be infinite, they are truncated at a J_{max} value. The quantities defined by Eq 3.1.15 are invariant by rotation and permutation and will be the coefficients of the bispectrum descriptor.

For a given atomic environment defined by a central atom and its neighbours (using the cut-off radius R_{cut}), a third-rank matrix is computed whose size is arbitrarily defined by the J_{max} integer. Some relations between the B_{j_1, j_2, j_3} can be deduced [202] to come up with a bispectrum matrix made of $J = \frac{(J_{max}+1)(J_{max}+2)(J_{max}+\frac{3}{2})}{3}$ distinct components.

3.1.2.2 Estimating Physical Properties with SNAP

In the case of the SNAP method, for a given atom i , the bispectrum of its environment is computed and all the coefficients are gathered inside a vector of size $K \times 1$ (K being the number of bispectrum coefficients used to describe the atomic environments). Then the coefficients of the

above introduced third-rank matrix are stored (see Eq. 3.1.15) in a single vector:

$$\mathbf{B}^{(i)} = \begin{bmatrix} B_1^{(i)} \\ \cdot \\ \cdot \\ B_j^{(i)} \\ \cdot \\ \cdot \\ B_K^{(i)} \end{bmatrix} \quad (3.1.16)$$

By summing all the k -th coefficients of the bispectrum vectors associated with all the N atoms of a crystal structure, the coefficients $B_k = \sum_i^N B_k^{(i)}$ can be defined. In what follows, the B_k components are stored in a vector :

$$\mathbf{B} = \begin{bmatrix} B_1 \\ \cdot \\ B_k \\ \cdot \\ B_K \end{bmatrix} \quad (3.1.17)$$

For a given atomic configuration denoted by \mathbf{c} , the aim is to obtain some predictions about the $3N + 7$ following quantities:

- The total energy of the configuration, E_{tot} (1 component).
- The forces on all the atoms $F_l^{(i)}$ ($3N$ components).
- The virial $W_{l,p}$ (6 components with the Voigt notation).

These are stored in the $3N + 7$ prediction vector:

$$\mathbf{F}_c = \begin{bmatrix} E_{tot} \\ \cdot \\ \cdot \\ F_l^{(i)} \\ \cdot \\ \cdot \\ W_{1,1} \\ W_{2,2} \\ \cdot \\ W_{2,1} \end{bmatrix} \quad (3.1.18)$$

This vector is computed thanks to the relation

$$\mathbf{A}_c \cdot \mathbf{T} = \mathbf{F}_c \quad (3.1.19)$$

where \mathbf{T} is a $K + 1$ vector:

$$\mathbf{T} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_K \end{bmatrix} \quad (3.1.20)$$

and \mathbf{A}_c is the feature matrix of the configuration c , i.e. the part that contains the bispectrum coefficients, whereas \mathbf{T} contains the coefficients to be found by the training process. This matrix is built up by expressing all the aimed quantities for prediction in terms of the \mathbf{T} vector. The starting point is the energy. The local energy of the i -th atom is expressed as:

$$E^{(i)} = \beta_0 + \sum_{k=1}^J \beta_k B_k^{(i)} \quad (3.1.21)$$

The total energy of a given configuration made of N atoms is thus written as:

$$\begin{aligned} E_{tot} &= \sum_i^N E^{(i)} = N\beta_0 + \sum_i^N \sum_k^J \beta_k B_k^{(i)} \\ &= N\beta_0 + \sum_k^J \beta_k \sum_i^N B_k^{(i)} \\ &= N\beta_0 + \sum_k^J \beta_k B_k \end{aligned}$$

It can be deduced:

$$E = [N, \mathbf{B}^T] \cdot \mathbf{T} \quad (3.1.22)$$

where $[N, \mathbf{B}^T]$ is a $K + 1$ sized line-vector. Forces and virial can be derived from the expression of the energy. The l -th component of the force acting on the i -th atom is given by the energy derivative :

$$\begin{aligned} F_l^{(i)} &= -\frac{\partial E_{tot}}{\partial r_l^{(i)}} \\ &= -\sum_k^J \beta_k \frac{\partial B_k}{\partial r_l^{(i)}} \end{aligned}$$

Using the \mathbf{B} matrix, it gives:

$$F_l^{(i)} = [0, -\frac{\partial \mathbf{B}^T}{\partial r_l^{(i)}}] \cdot \mathbf{T} \quad (3.1.23)$$

The components of the virial defined as:

$$W_{l,p} = \sum_i^N r_l^{(i)} \cdot F_p^{(i)}$$

$$\begin{aligned}
&= - \sum_i^N r_l^{(i)} \sum_k^J \beta_k \frac{\partial B_k}{\partial r_p^{(i)}} \\
&= - \sum_k^J \beta_k \sum_i^N r_l^{(i)} \frac{\partial B_k}{\partial r_p^{(i)}}
\end{aligned}$$

can be also expressed in a more compact way:

$$W_{l,p} = [0, - \sum_i^N r_l^{(i)} \frac{\partial \mathbf{B}^T}{\partial r_p^{(i)}}] \cdot \mathbf{T} \quad (3.1.24)$$

Given these relations, the $(3N + 7) \times (K + 1)$ sized feature matrix of the c configuration \mathbf{A}_c can be deduced:

$$\mathbf{A}_c = \begin{bmatrix} N & \mathbf{B}^T \\ 0 & \cdot \\ 0 & -\frac{\partial \mathbf{B}^T}{\partial r_l^{(i)}} \\ 0 & \cdot \\ 0 & -\sum_i^N r_l^{(i)} \frac{\partial \mathbf{B}^T}{\partial r_p^{(i)}} \\ 0 & \cdot \end{bmatrix} \quad (3.1.25)$$

3.1.2.3 The Training Process

Training a SNAP type numerical potential first consists in fixing two types of hyperparameters:

- Global hyperparameters for the computation of the bispectrum: it regroups the cut-off radius determining the size of the local atomic environments (R_{cut}) around each atom as well as the size of the decomposition of the bispectrum (K). If several atom types are considered, a radius associated to each of them needs to be defined as well as a weight.
- An ensemble of weights gathered in a diagonal matrix, \mathbf{O} . It ponderates the different features of each atomic configuration involved in the training database, as well as the type of the physical quantity.

These hyperparameters have to be determined so as to fix the regression vector \mathbf{T} as defined in Eq. 3.1.19. Suppose the global hyperparameters are fixed. The determination of this \mathbf{T} vector is achieved as follows. Each configuration gives its own \mathbf{A}_c feature matrix. These matrices are built up as described above in Eq. 3.1.25 by considering all the atomic environments of a given structure. Suppose the database contains C configurations, these configuration matrices can be stored into a single one of size $C(3N + 7) \times (K + 1)$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \cdot \\ \mathbf{A}_i \\ \cdot \\ \mathbf{A}_C \end{bmatrix} \quad (3.1.26)$$

It turns out that some configurations or some features are not as relevant as others. This relevancy discrimination can be enforced by the use of the $C(3N + 7) \times C(3N + 7)$ diagonal weight matrix \mathbf{O} . It is made of different weights on its diagonal associated with each feature of each configuration in the database.

$$\mathbf{O} = \begin{bmatrix} \omega_1^{(1)} & 0 & . & 0 \\ 0 & \omega_2^{(1)} & . & . \\ 0 & \dots & . & . \\ 0 & 0 & \dots & \omega_{3N+7}^{(C)} \end{bmatrix} \quad (3.1.27)$$

For a given matrix \mathbf{O} , the \mathbf{T} vector can be determined by solving the equation:

$$[\mathbf{O}.\mathbf{A}].\mathbf{T} = \mathbf{O}.\mathbf{F} \quad (3.1.28)$$

where \mathbf{F} is the $C(3N + 7) \times 1$ vector made of the $(3N + 7) \times 1$ sized \mathbf{F}_C vectors of each configuration as defined in the former subsection (see Eq. 3.1.18).

To determine the global hyperparameters as well as the weights of \mathbf{O} and \mathbf{T} , a differential evolution algorithm [203] is used. This optimization process is organized as follows:

- Fix a coefficient set of the \mathbf{O} matrix as well as global hyperparameters if required.
- Solve in \mathbf{T} the equation (using the ordinary least square estimator):

$$[\mathbf{O}.\mathbf{A}].\mathbf{T} = \mathbf{O}.\mathbf{F}$$

on a training set of configurations.

- Compute the Root Mean Squared Error (RMSE) on a validation set. It is evaluated by comparing the values predicted by the model (using the \mathbf{T} vector derived at the former step and Eq. 3.1.19) and the real ones coming from DFT calculations.
- Update the set of \mathbf{O} matrix coefficients and, if required, the global hyperparameters. It is performed from the RMSE computed value at the former step and the rules enforced by the differential evolution algorithm employed [203].
- Reiterate this cycle with the updated \mathbf{O} matrix and if required the updated global hyperparameters.

This optimization algorithm is used twice. First of all on a reduced database, the global hyperparameters have to be determined. When they are fixed, another optimization procedure on the database is enforced to find a \mathbf{T} vector and an associated \mathbf{O} matrix. After the last optimization iteration, the last resolution of

$$[\mathbf{O}.\mathbf{A}].\mathbf{T} = \mathbf{O}.\mathbf{F}$$

finally gives the \mathbf{T} vector for predictions of physical quantities performed thanks to Eq. 3.1.19.

3.1.3 Numerical Potentials and Random Searching

To combine numerical potentials with CSP algorithms, an active learning process is considered where the training databases of the numerical potentials are updated throughout the exploration process. With such a process, the aim is to improve by successive cycles of structure searches and training the numerical potentials. The different steps involved are presented in FIG 3.1.1. Firstly, a numerical potential is trained on an initial database gathering random atomic configurations of the studied system. The properties of these different atomic configurations (energy, forces, stress) are computed accurately thanks to DFT computations. These calculations do not optimize the crystal structures but take the atomic configurations and evaluate the targeted physical quantities. These types of DFT calculations will be named "single point" DFT computations in what follows. After training the numerical potential on this initial database, an active learning process starts where the numerical potential learns new informations about the system. An active learning cycle is decomposed between:

- A random searching step where crystal structures are randomly generated and structurally optimized by the numerical potential. This step is similar to the one performed with AIRSS, the main difference being the use of numerical potential instead of DFT. This enables to explore many more structures (and also larger supercells) than AIRSS thanks to a very small computational time associated with structural relaxation using numerical potentials.
- A selection step aiming at selecting the best structures among those explored. The selection criteria can be based on the energy (or enthalpy) predicted by the numerical potential. Other possibilities are the use of descriptions of atomic environments (like bispectrum) or the simulation of X-ray spectra.
- Single point DFT calculations are enforced on the selected structures. It aims at computing their physical properties and including these new informations in the training database.
- With this updated database, a new numerical potential is trained. It is then employed to explore again crystal structures in a random searching process.
- At the end of the active learning cycle (when the potential does not find unexplored structure among those having the lowest predicted enthalpies), a few best structure candidates explored by the numerical potential are selected. Full DFT structural optimizations are then applied. The structures having the lowest enthalpies correspond to the final predictions.

The implementation of the above presented active learning process is performed with *abinit* for the different DFT computations. For the structural optimizations using the numerical potentials, *lammps* is employed [204].

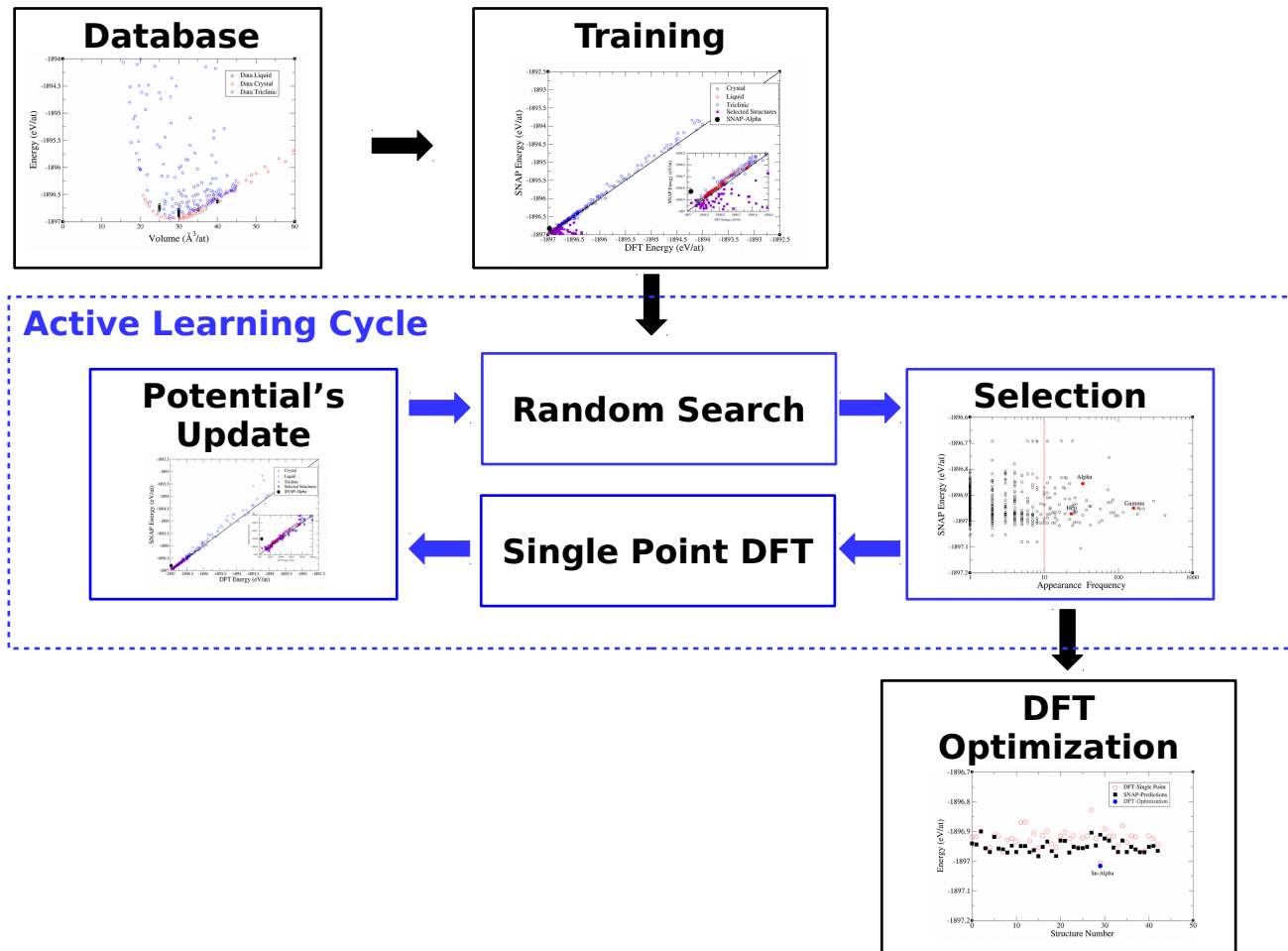


Figure 3.1.1: The active learning process of RSS combined with numerical potentials. Usually, the initial database contains structurally optimized configurations of the system corresponding to the relevant pressure range of referenced structures.

This chapter presented the main principles associated with the numerical potentials and how they can be included in crystal structure prediction. In the active learning process presented, the numerical potentials will be used to speed up the structure searches. They will enable to push back the exponential wall related to the dimensionality surge of the PES when structures having more atoms and of different types in the unit cells have to be explored. In addition, the sampling problems due to the fact this exponential surge has to be followed by the number of generated structures might be solved for a wide variety of systems. Indeed, as the structural optimizations with numerical potentials cost far less than those performed with DFT, several orders of magnitude will be gained in terms of the number of structures generated and explored. The next chapter is about a proof of concept of this approach.

Chapter 3.2

A First Proof of Concept on Tin

Contents

3.2.1 Why Considering Tin	129
3.2.2 A First Try	130
3.2.2.1 Global Work-Flow	131
3.2.2.2 Potential Initialization	131
3.2.2.3 First Active Learning Cycle	134
3.2.2.4 Second Active Learning Cycle	137
3.2.3 A Second Try with more Converged Databases	139
3.2.3.1 Databases: Comparisons	139
3.2.3.2 First Active Learning Cycle	141
3.2.3.3 Update of the Potential	143
3.2.3.4 Second Active Learning Cycle	144
3.2.4 Predictions under Pressure	146
3.2.5 Conclusion	149

In this chapter, a proof of concept regarding the application of numerical potentials to CSP algorithms is presented. The Spectral Neighbour Analysis Potential (SNAP) as introduced in 3.1.2 is considered. This study consists in retrieving the most stable phases of tin at 0 K and under different pressures. Beside the challenge of predicting the good crystal structure using numerical potentials, this part also aims at deriving a work-flow to obtain the most stable structures of a compound in a systematic way. As mentioned in 3.1.1.1, numerical potentials are machine learning methods and thus rely on a training database. This database is built thanks to DFT calculations, whose precision relies on a correct choice of several numerical parameters. Their influence on the efficiency of the potential is also studied.

3.2.1 Why Considering Tin

As a first step to develop CSP algorithms using machine learning, a pure compound is considered. As said, tin is the system chosen for this work. It has a relatively complex phase diagram in which five different phases are competing with each other. These phases are [206, 207, 208]:

- The α phase with space group Fd-3m (Wyckoff position 8a). It is the most stable phase at 0 GPa and 0 K.
- The β phase with space group I41/amd. It is the one appearing at 0 GPa and at non zero temperature.
- The γ phase with space group I4/mmm (Wyckoff position 2a). It is the most stable phase under moderate pressure at 0 K.
- The HCP phase with space group P6₃/mmc. This phase is stable at 0 K and high pressure (above 160 GPa).
- The BCC phase with the space group Im-3m. BCC and γ phases have close stability under pressure (typically above 25 GPa).

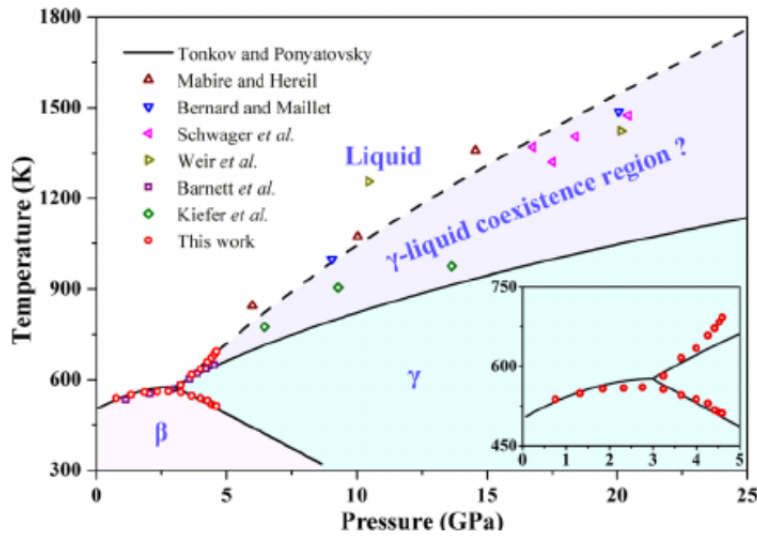


Figure 3.2.1: Pressure-temperature phase diagram of tin up to 25 GPa and between 300 and 1800 K [205].

Another difficulty related to this diagram is the particularity of the α phase. It has indeed covalent chemical bonds whereas metallic ones are observed for the others. Thus, the α phase is semi-conducting while all the other phases are metallic. Determining whether a numerical potential is able to predict the α phase without learning it is a real challenge. Indeed, if it works, it means that the numerical potential can reproduce a change in the electronic structure of a material (whereas the initial structures forming the database all have the electronic structure of a given type). In order to assess the accuracy of the predictions obtained with this approach, the results are compared with reference DFT calculations presented in the appendix (see F). In this work, three pressure points are considered which are 0, 20 and 100 GPa. The aim is thus to obtain at 0 GPa the α phase and under pressure the γ one (see FIG 3.2.1 for the different phases and stability domains), without inserting these phases in the initial training database. If the developed approach is able to identify the correct phases of such complex phase diagram, it will encourage to pursue its improvement towards predictions of superhydrides.

3.2.2 A First Try

In this section, a first attempt to predict the phases of tin at 0 GPa and under pressure is presented. The global work-flow is firstly reminded. This is followed by a description of the initial database and of the way the potential is initialized. Finally, the results obtained through the active learning process are discussed.

3.2.2.1 Global Work-Flow

The work-flow followed to train a numerical potential for crystal structure predictions is mainly based on the active learning process presented in 3.1.3. As formerly explained, two kinds of DFT calculations are involved:

- Single-point DFT computations (DFTSP) which consist in evaluating the physical properties of a structure without modifying it. They are used to build up and update the training databases.
- DFT structural optimizations (DFTO-1) as employed after the active learning process. They are used to identify the most stable phases among those selected.

The calculations are also enforced by using the Generalized Gradient Approximation (GGA-PBE). For these two kinds of calculations, different sets of numerical parameters are considered. For both, a 20 Ha plane-wave cut-off determining the size of the plane-wave basis is fixed. The PAW atomic database from JTH table is taken for tin. A k-point mesh of $4 \times 4 \times 4$ is used for single point DFT computations (DFTSP) whereas $12 \times 12 \times 12$ is employed for DFT structural optimizations (DFTO-1). The DFT self-consistent cycles in both cases use a criterium on the forces (between two consecutive DFT iterations) to determine their convergence. The cycle is stopped when the difference in the maximal force between two consecutive iterations is below 0.02 times this maximal force (twice consecutively). For DFT structural optimizations, the BFGS algorithm (see 1.3.4.2) is enforced with a stopping criteria of 10^{-6} Ha/bohr³. Outside the active learning process, a third kind of DFT computations is performed once (denoted DFTO-2). It corresponds to the structural optimizations under different pressure values of the different phases of tin. The purpose of these calculations is to have precise DFT values regarding these phases under pressure. It gives a basis of comparisons between SNAP predictions and the expected phases at a given pressure. For these calculations, the k-point mesh used is $24 \times 24 \times 24$, the other parameters being those applied in (DFTO-1). The DFTO-2 parameters are those considered in the results presented in F.

3.2.2.2 Potential Initialization

This section describes how the potential is initialized before starting the active learning cycles. First of all the initial database is described. It is followed by a description of the training process of the potential.

The Initial Architecture of the Database

An initial database has to be built up to train the numerical potential for the first random searching. For this purpose, several types of atomic configurations are considered.

- A first kind is made of "liquid" configurations. These structures are inserted in a supercell having an orthorhombic shape, with disordered atomic configurations made of 64 atoms. As

the atoms do not occupy high symmetry positions, they thus feel forces whose values are distributed on a wide range. This part of the training database is designed so as to have a lot of informations related to forces. To generate this part of the database, different values of the atomic density are considered, corresponding to atomic volumes of: 25,30,35,40 $\text{\AA}^3/\text{at}$.

- A second part is made of simple crystal structures (Faced-Centered Cubic, Simple Cubic, note that the crystal structures corresponding to α , β and γ phases are intentionally excluded from the database). These structures are generated under different atomic densities corresponding to atomic volumes between 20 and 60 $\text{\AA}^3/\text{at}$ (cold curves). This enables to learn how the energy evolves with atomic density (or equivalently with pressure).
- A third part is made of triclinic structures with one atom inside the primitive cell. These configurations have random primitive cell geometries (random angles and cell parameters) corresponding to a distribution of atomic volumes between 20 $\text{\AA}^3/\text{at}$ and 40 $\text{\AA}^3/\text{at}$. This last part of the database can be utilized to compensate for the fact that liquid phases have only orthorhombic supercells and thus improve the transferability of the potential for structures which do not have this property. In particular, phases having non-hydrostatic stress tensors will be learned on this kind of configurations as well as the elasticity properties.

DFT calculations (using DFTSP parameters) are performed on these structures to compute their physical properties. The energy, stress and forces are stored in the database. The distribution in the plane (atomic energy, atomic volume) of the training database is shown in FIG 3.2.2. The black dots represent liquid-type configurations. They are concentrated on vertical lines related to the specific atomic density used for their generation. The red dots correspond to high-symmetry crystal structures (cold equations of states). The blue dots show the distribution for the triclinic configurations. They are sampling a wide range in terms of stress tensor components and pressure. As a result they are distributed on a large range of energy per atom (up to 3 eV difference).

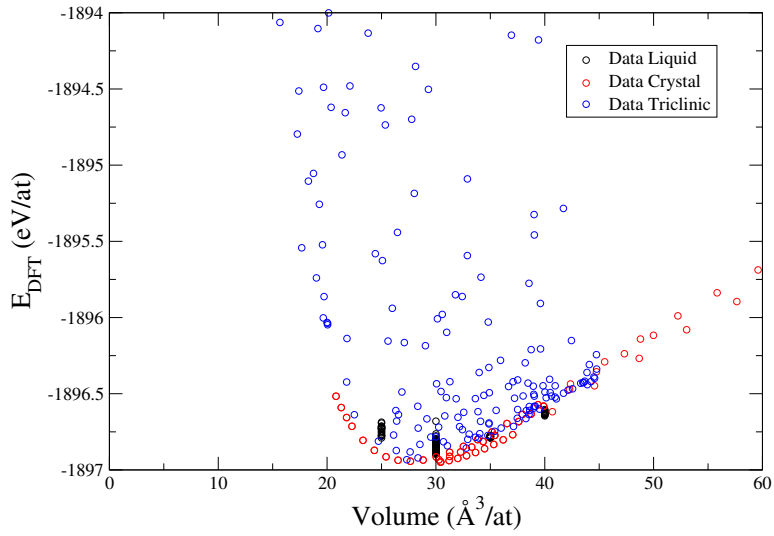


Figure 3.2.2: Energy volume per atom distribution of the training database.

Training on the Initial Database

After generating this initial database, the numerical potential has to be trained on it. This training step is enforced following the different steps described in 3.1.2.3. An optimum cut-off radius (i.e. with the lowest Root Mean Squared Error (RMSE) on the initial training database) (R_{cut}) of 5.5 \AA is found for the computations of the descriptors. The weights derived from the training process enabled a Root Mean Squared Error (RMSE) between the energies predicted by SNAP and those obtained by DFT of 54.43 meV/at for the structures stored in the initial training database.

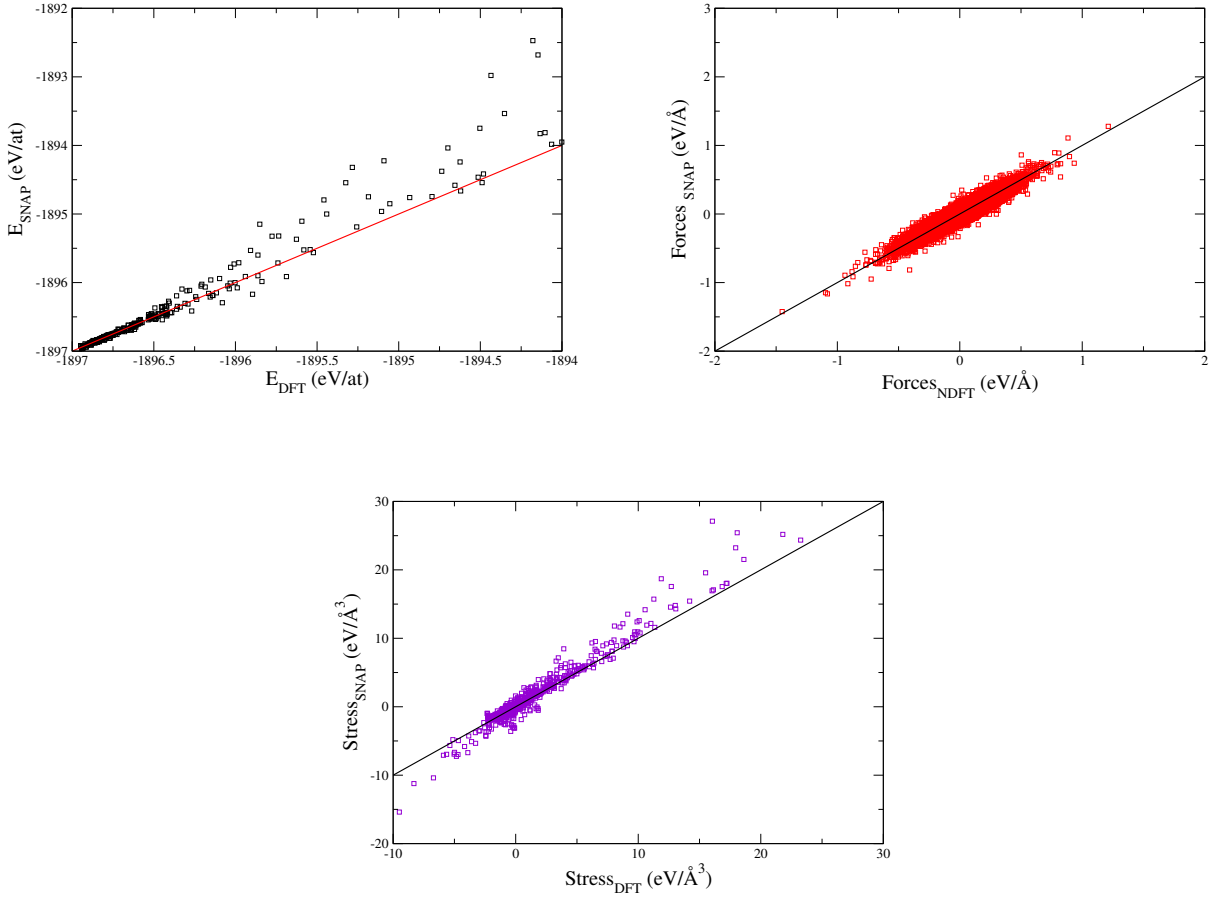


Figure 3.2.3: Energy, force and stress correlations on the initial training database.

	Energy (meV/at)	Forces (meV/Å)	Stress (meV/Å ³)
RMSE (meV/at)	54.43	22.41	49.51

Table 3.2.1: RMSE values for energy, forces and stress on the initial training database.

Correlations of the numerical potential for energy, forces and stress of the structures in the training database are given in FIG 3.2.3 and Tab 3.2.1. The ability of the potential to reproduce cold curves is presented in appendix F. After this first training process, the numerical potential is now ready to perform a first exploration of crystal structures. The way these crystal structures are generated is presented in depth in the appendix (F). The next section describes in detail the results of this first exploration achieved by the numerical potential.

3.2.2.3 First Active Learning Cycle

10 000 crystal structures having up to 12 atoms in their unit cell are generated following the process described in F. These crystal configurations are then structurally optimized by the

numerical potential. The different phases obtained are then analyzed in terms of appearance frequency and predicted energy.

Random Searching with SNAP

The 10 000 generated structures are optimized by the SNAP potential at a pressure of 0 GPa. Using a symmetry analysis of the structures with the *findsym* [137, 138] program, three configurations (α, β, γ) making part of the phase diagram are identified. Note that these structures were not included in the initial database, and correspond to predictions of the numerical potential. Tab 3.2.2 provides the appearance frequencies obtained for each of them. In addition comparisons between the energy values given by SNAP and that obtained by DFT are given (these DFT calculations use the DFTO-2 parameters). At this step, it can be said that the numerical potential is good enough to retrieve some phases that were not included in its training database, but not accurate enough to predict the correct DFT energetic order ($\alpha < \beta < \gamma$) between these phases (see Tab 3.2.2). Moreover, it predicts as more stable other structures which do not make part of the phase diagram of tin (see FIG 3.2.4).

	α	β	γ
Appearance Frequency (%)	0.42	0.8	2.58
SNAP Energy (eV/at)	-1896.9898	-1897.0808	-1896.9882
Ref. DFT Energy (eV/at)	-1897.0160	-1896.9695	-1896.9579
SNAP-DFT (meV/at)	26.15	-111.30	-30.300

Table 3.2.2: Appearance frequencies and comparisons between the α, β and γ structures given by SNAP and those obtained at 0 GPa by DFT using the DFTO-2 parameters .

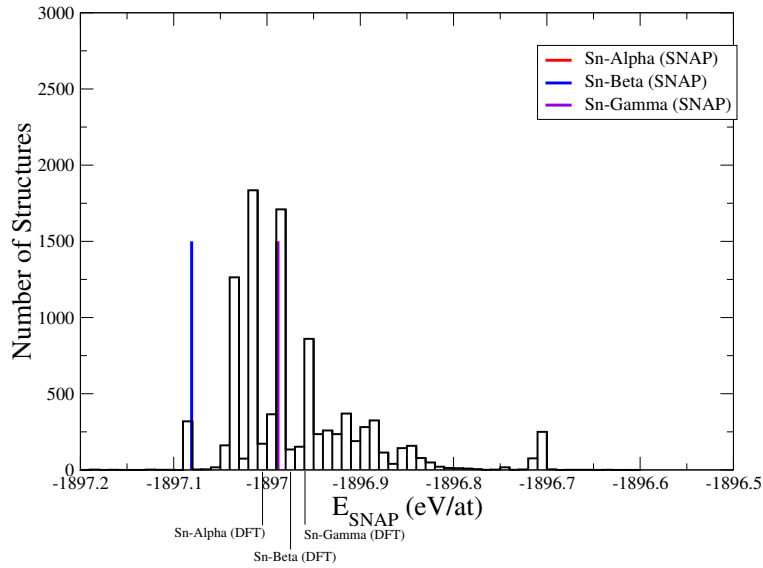


Figure 3.2.4: Energy per atom distribution obtained after random searching with SNAP. The vertical lines are the energy values predicted by SNAP and computed for the α , β and γ phases.

Energy and Appearance Frequency Distributions

To go further, one needs to quantify exactly how many distinct configurations are obtained after the first random search with SNAP. For this purpose, the structures are gathered in clusters. Each cluster is determined by a symmetry group and a set of Wyckoff positions and consequently represents similar or close crystal configurations. After identifying all the obtained distinct clusters, they are filtered to retrieve the most relevant ones. The criterion used is based on the appearance frequency. Indeed for this first selection, the energies predicted by SNAP are not considered as accurate enough to be used as a filter to select relevant configurations. The potential is in fact not accurate enough for energy predictions at the beginning of the active learning process. FIG 3.2.5 represents the obtained distribution of configurations in the plane (SNAP energy, appearance frequency). The red points highlight the identified phases (α, β, γ) belonging to the phase diagram of tin. A threshold frequency of 0.1 % (10 structures over 10 000 explored) can be set so as to select the 100 most frequent configurations. The three identified phases belong to these selected structures. Single-point DFT computations using the parameters of DFTSP are performed on these 100 crystal structures. This information is added to the database to train another numerical potential.

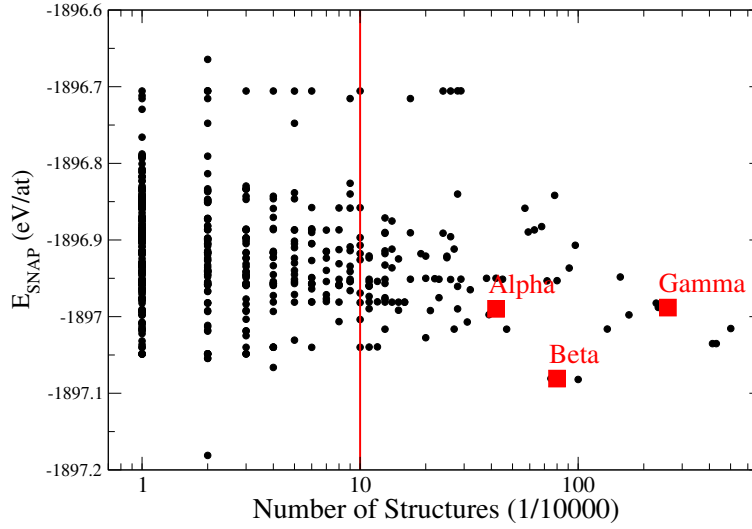


Figure 3.2.5: Distribution of the clusters in the plane (energy per atom, appearance frequency) each point is associated with a given space group and a set of Wyckoff positions.

3.2.2.4 Second Active Learning Cycle

The new updated potential is used to perform a second random search with still structural optimizations imposing a final pressure of 0 GPa. This new potential is considered accurate enough so the energy values of the optimized structures can be used as a filter together with the appearance frequency to select a few structures. FIG 3.2.6 shows in the plane (SNAP energy, appearance frequency) the distribution of the structures after the SNAP random searching.

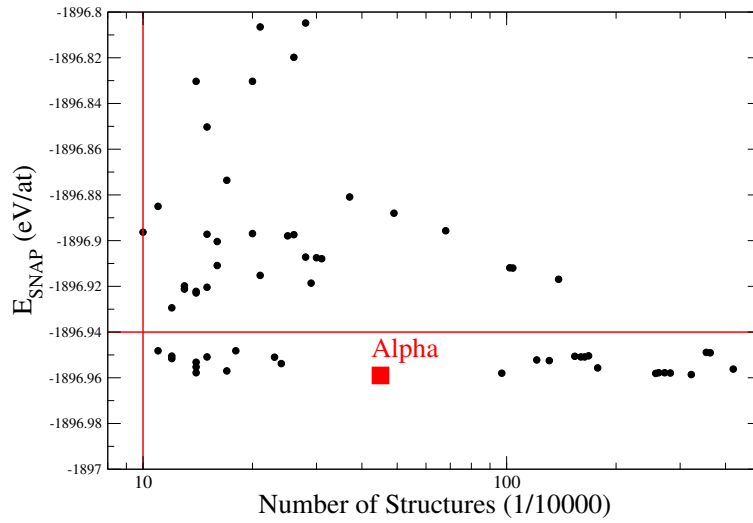


Figure 3.2.6: Energy per atom and appearance frequency distribution after the second random searching with SNAP. The structures appearing with a frequency higher than 0.1 % are represented. Note that β and γ phases are not obtained.

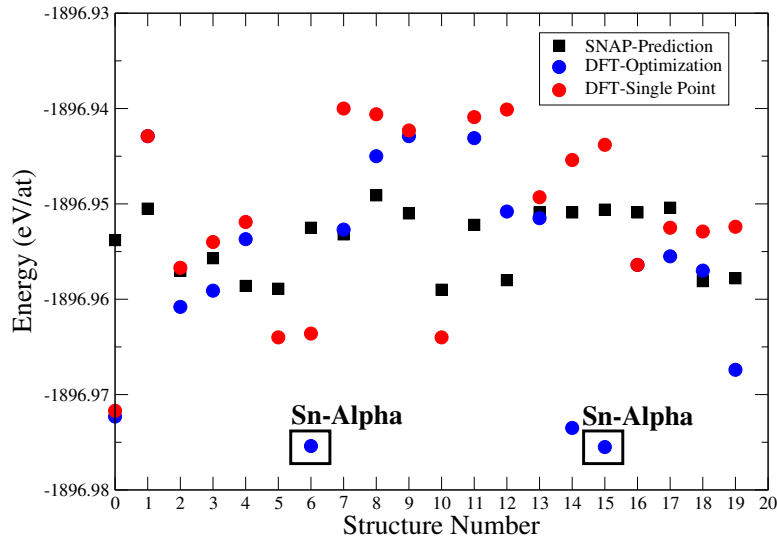


Figure 3.2.7: Comparison of the SNAP predicted energies, those obtained by single point DFT and those by DFT structural optimizations of the selected structures.

To select a dozen of structures, a threshold frequency of 0.1 % is used and combined with an energy threshold of -1896.96 eV/at (the horizontal line corresponds to the energy filter applied, the vertical one to the appearance frequency filter). A region is thus delimited with these filters. It is the one including among the most stable structures according to SNAP and also the most frequent. In this case, among all the known phases, only the α phase is included in it. These configurations

are then optimized thanks to DFT using the DFTO-1 parameters. FIG 3.2.7 presents for each of the selected structures the predicted SNAP energy (black square), the DFT value of the energy before optimization (red circle), and the energy reached after structural optimization at 0 GPa (blue circle). Following the DFT structural optimization, two structures appear with the lowest energy. It has been checked they can both be identified as the α phase. As shown in FIG 3.2.7, the α phase identified by the numerical potential is not the structure having the lowest energy among those selected when single-point DFT computations are applied. Full DFT optimizations are needed to retrieve the most stable structure. This is explained by the fact that the structures given directly by the numerical potential are not under a pressure of 0 GPa (for DFT) due to the error of SNAP on the stress.

It can be concluded that this random searching with SNAP gives the correct phase at 0 GPa. This result is encouraging because a phase very different (from the point of view of crystal and electronic structures) from the structures stored initially in the training database has been found (the tin α phase is the only one to be insulating). It gives thus the hope to build a method which can be transferable (in the sense able to make correct predictions out of the envelope of the training database). However, attempts to implement the predictions under pressure were not successful, in particular due to an important error made by the numerical potential to evaluate pressure and enthalpy of the structures. In order to investigate the source of this error, the following section aims at evaluating the influence of the DFT convergence parameters used to compute the training databases. The potential obtained after this study will be afterwards employed to perform predictions under non-zero pressure.

3.2.3 A Second Try with more Converged Databases

In the former section, the ability to find a correct crystal structure at 0 GPa using the numerical potential approach has been proven. This section aims now at testing whether the numerical potential itself – and its predictions – is converged with the numerical parameters used to compute by DFT the physical quantities associated with the database.

3.2.3.1 Databases: Comparisons

This paragraph aims at describing the differences between the former and the new more converged database (both are composed of the same crystal structures). The new database is characterized by calculations made with a denser k-point mesh. A mesh of $12 \times 12 \times 12$ instead of $4 \times 4 \times 4$ in the former database is applied to the "liquid" and crystal type structures. The triclinic type configurations are treated apart with a k-point mesh of $24 \times 24 \times 24$ (whereas a k-point mesh of $4 \times 4 \times 4$ was used formerly). The convergence criterium to stop the self-consistent DFT cycles is still employed on the forces. However, the force differences between two consecutive self-consistent iterations (in Ha/bohr) have to be lower than the maximum force value (of the current

DFT cycle) times a tolerance factor of 10^{-5} .

$$\delta F < 10^{-5} \times F_{max}$$

This is much more drastic than the criterion imposed formerly. Indeed, for the former database, this ratio was 10^{-2} instead of 10^{-5} .

Better evaluation of the stress is enabled furthermore thanks to a smoothing operation (called the smearing) of the plane-wave contributions having an energy close to the cut-off energy value (set to 20 Ha). As a result the plane-waves having a kinetic energy closer as 0.5 Ha are subjected to this smoothing operation. In the former database, this was not enforced. The Tabs 3.2.3 and 3.2.4 sum up the main differences regarding the convergence parameters employed. These new DFT parameters are denoted by DFTSP-CV.

The differences between the two databases in terms of energy per atom and pressure are given in Tab 3.2.5 and FIG 3.2.8. For each type of structure in the database (crystal, liquid and triclinic), the average differences (in absolute value) are given. The most important differences are reached for the triclinic configurations, and a high convergence criterium may appear relevant. These structures are indeed distributed over a wide range of pressures (see FIG 3.2.2).

	k-point mesh	Forces, convergence criteria (Ha/bohr)	Plane-wave cut-off smearing (Ha)
Crystal	$4 \times 4 \times 4$	$\delta F < 2 \times 10^{-2} \times F_{max}$	0
Liquid	$4 \times 4 \times 4$	$\delta F < 2 \times 10^{-2} \times F_{max}$	0
Triclinic	$4 \times 4 \times 4$	$\delta F < 2 \times 10^{-2} \times F_{max}$	0

Table 3.2.3: DFT numerical parameters of the former database (corresponding to the DFTSP set).

	k-point mesh	Forces, convergence criteria	Plane-wave cut-off smearing (Ha)
Crystal	$12 \times 12 \times 12$	$\delta F < 10^{-5} \times F_{max}$	0.5
Liquid	$12 \times 12 \times 12$	$\delta F < 10^{-5} \times F_{max}$	0.5
Triclinic	$24 \times 24 \times 24$	$\delta F < 10^{-5} \times F_{max}$	0.5

Table 3.2.4: DFT numerical parameters for the new database characterized by the DFTSP-CV set.

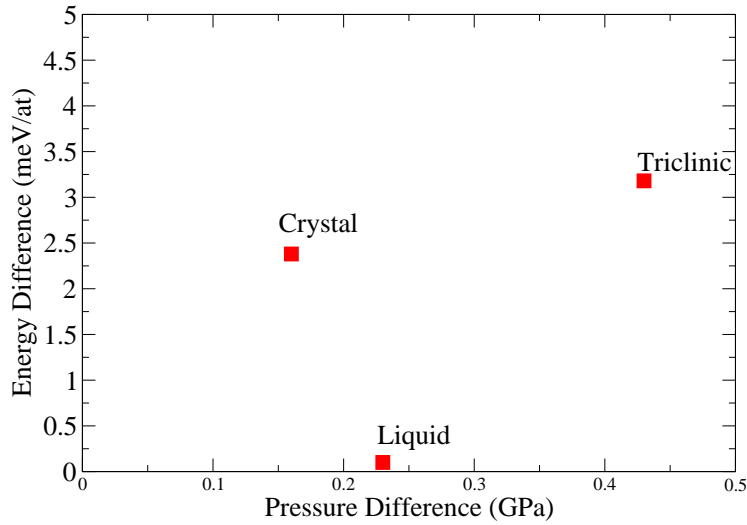


Figure 3.2.8: Absolute energy and pressure differences between the first and the more converged databases.

	Energy Differences (meV/at)	Pressure Differences (Ha/bohr ³)	Pressure Differences (GPa)
Crystal	2.38	9.22×10^{-4}	0.16
Liquid	0.1	1.38×10^{-3}	0.23
Triclinic	3.18	2.35×10^{-3}	0.43

Table 3.2.5: Absolute energy and pressure differences between the first (using DFTSP set) and the more converged (using DFTSP-CV set) databases.

3.2.3.2 First Active Learning Cycle

The 10 000 structures generated formerly are used to perform a first random search with a numerical potential trained on the initial more converged database. A pressure of 0 GPa is imposed. Appearance frequency of 0.1 % is taken as a filter to select the structures (see FIG 3.2.9). In the selected region, three known phases are retrieved: the α , HCP and γ ones. According to the reference DFT computations (see FIG F.5), it corresponds to the three first phases in terms of stability at this pressure (see FIG F).

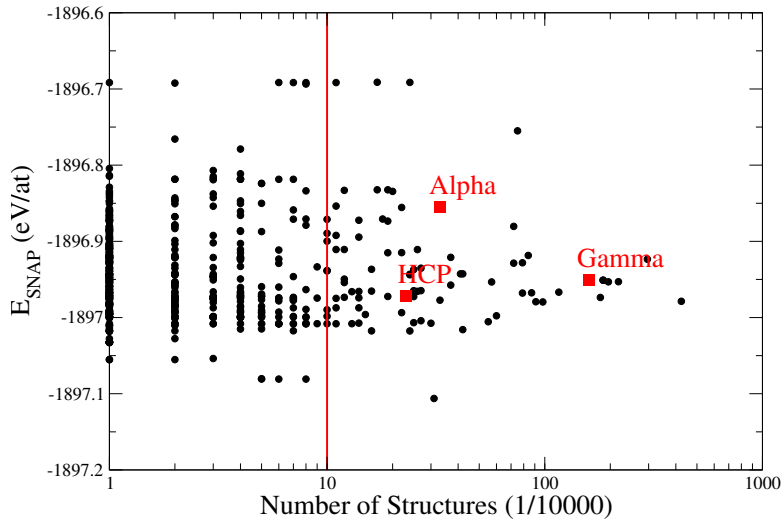


Figure 3.2.9: Appearance and SNAP energy distribution of the structures after the first random search with SNAP.

It is noticeable that the HCP phase is retrieved instead of the β one formerly obtained. Tab 3.2.6 presents the appearance frequencies of these retrieved phases. The differences between the energies predicted by SNAP values and the DFT ones (using DFTO-2) expected for these phases at 0 GPa are given. It can be observed that the difference between the SNAP and DFT

	α	Hcp	γ
Appearance Frequency (%)	0.33	0.23	1.6
SNAP Energy (eV/at)	-1896.8554	-1896.9721	-1896.9504
Ref DFT Energy (eV/at)	-1897.0160	-1896.9564	-1896.9579
SNAP-DFT (meV/at)	160.52	-15.72	7.500

Table 3.2.6: Appearance frequencies and differences between the SNAP predicted energy values and the DFT expected ones at 0 GPa for α , HCP and γ phases.

energies for the α phase is more important than formerly (see 3.2.2). The prediction is however better for the γ phase. The appearance frequencies reached are equivalent with those obtained beforehand (for the α and γ phases, see FIG 3.2.2). Using the above mentioned appearance frequency filter, 87 structures are selected for single-point DFT computations. The results are added to the initial database. For these DFT calculations the DFTSP-CV set is used. The updated database is now employed to train another numerical potential. The effect of this update on the efficiency of the potential is illustrated in the following paragraph.

3.2.3.3 Update of the Potential

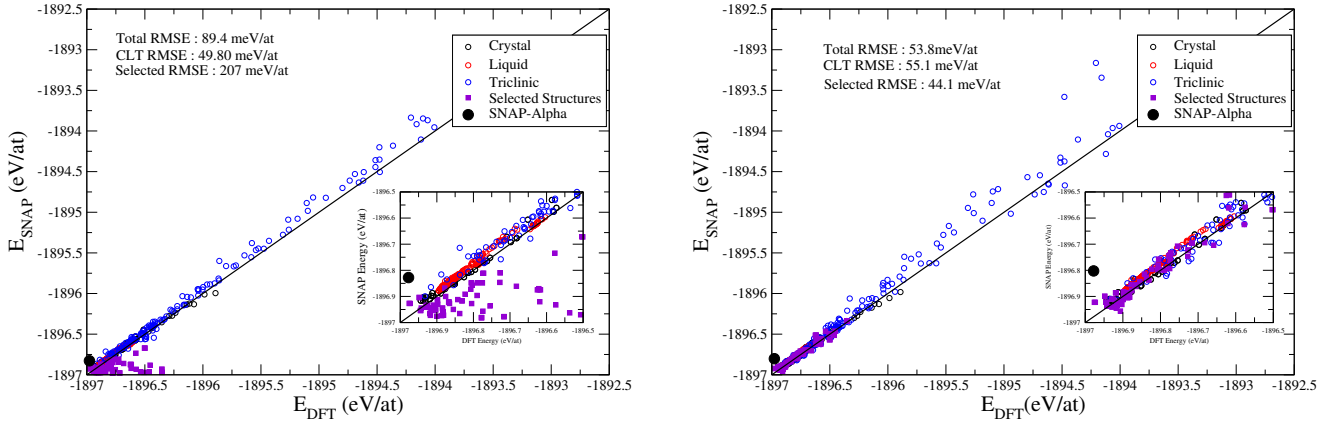


Figure 3.2.10: Left: SNAP and DFT energy correlations of the initial potential. Right: SNAP and DFT energy correlations of the updated potential.

FIGs 3.2.10 and 3.2.10 show the correlations DFT energy/ SNAP energy (in eV/at) for the initial numerical potential and for the one obtained after adding to the starting database the selected structures. For the first potential, correlations are good for the structures initially included in the training database (RMSE of 49.80 meV/at), the RMSE on unlearned structures is high (RMSE of 207 meV/at). It proves that the selected structures correspond to unknown configurations for the initial potential (meaning far from reference structures included in the database). It is thus worth including them into the database. After training the potential on the updated database, good correlations for the selected structures are reached (RMSE of 44.1 meV/at). This update then enabled it to build a numerical potential which improves its knowledge of the system. The new numerical potential obtained is afterwards employed to enforce a second active learning cycle. The RMSE on the initial database, the updated one and the one obtained on the selected structures added to the initial database are given in Tab 3.2.7.

	Initial Potential	Updated Potential
Updated database, RMSE (meV/at)	89.4	53.8
Initial database, RMSE (meV/at)	49.80	55.1
Selected configuration database, RMSE (meV/at)	207	44.1

Table 3.2.7: RMSE values on the updated database, the initial one and on the selected structure database added to the initial database (Selected RMSE).

3.2.3.4 Second Active Learning Cycle

The numerical potential obtained from the updated database is used for another random search at 0 GPa. To select a few dozen structures, a SNAP energy filter is applied in addition to the one based on appearance frequency. The structures appearing more than 10 times over 10 000 and having a SNAP energy below -1896.85 eV/at are selected. It leads to a selection region made of 40 structures, where the three phases α , β and γ are found (see graph 3.2.11).

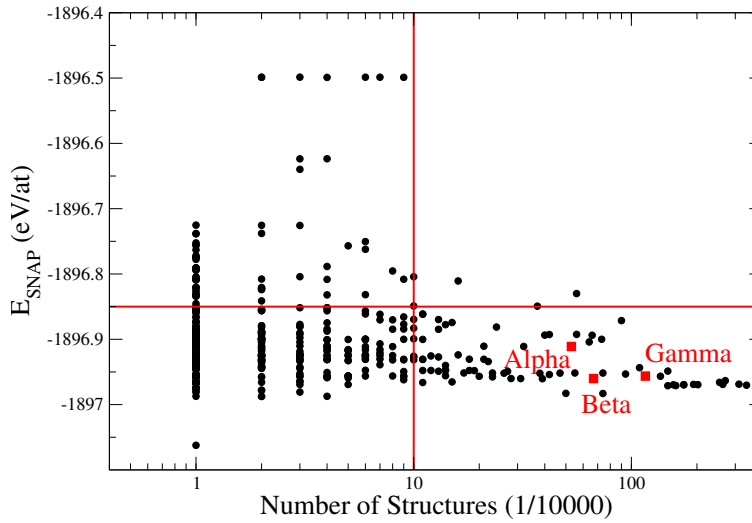


Figure 3.2.11: Appearance frequency and SNAP energy distribution of the structures after the second random search with SNAP.

Single-point DFT computations (using DFTSP-CV) are enforced on the selected structures. The SNAP energy predictions and the values obtained by DFT are compared in FIG 3.2.12. In addition, the selected phases are optimized by DFT using the DFTSP-CV parameter set. The single-point DFT computations performed on the selected structures give directly the α phase identified by the potential as the most stable one (see FIG 3.2.12). This was not obtained formerly with the less converged database. This result can be interpreted by the fact that the numerical potential reaches a better accuracy in terms of pressure and in particular for the one applied to optimize the structures (here 0 GPa). Indeed, it can be highlighted that the energy values between single-point DFT and DFT optimizations are in most cases close between each other in FIG 3.2.12. This higher precision is due to the better convergence reached for stress fields and pressures, in particular for the triclinic structures stored in the database. As a result, the α phase obtained after SNAP optimization is close to the one expected at 0 GPa. The final DFT optimizations on the selected structure finally confirm that the α phase as the most stable one.

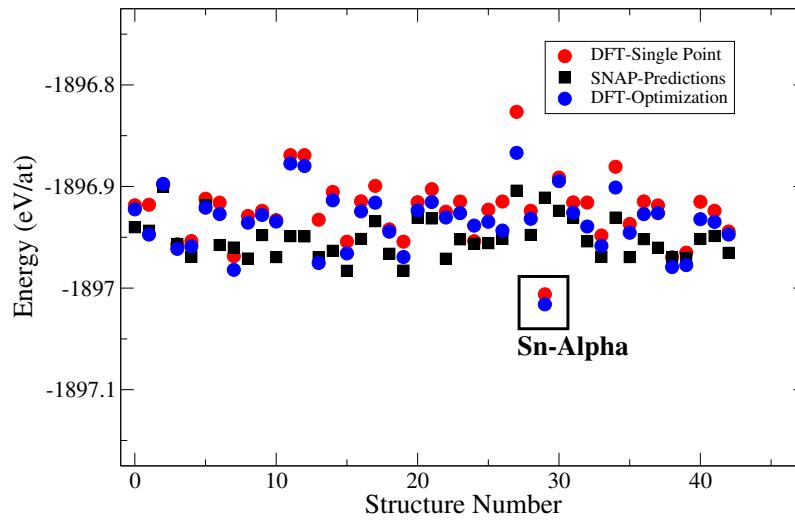


Figure 3.2.12: Comparisons of the energies predicted, obtained by single-point DFT and after DFT structural optimization of the selected structures. The RMSE between SNAP predictions and DFT single points is 21 meV/at.

To illustrate the fact that the obtained potential predicts more accurately the pressure to impose for structural optimization. FIG 3.2.13 shows the absolute differences in terms of DFT pressure and volume between the reference α phase and those predicted using the less converged training database (SNAP) and the more converged one (SNAP-CV). It can be highlighted that the α structure predicted by SNAP-CV is closer to the DFT results. Tab 3.2.8 presents the single point DFT values for the volume and pressure computed on the α phase predicted by SNAP-CV.

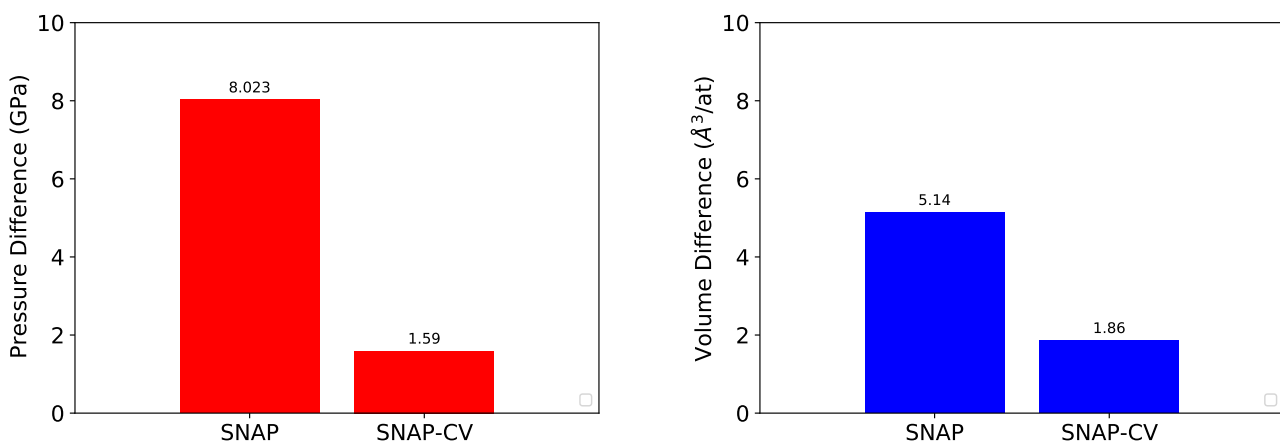


Figure 3.2.13: Comparison of DFT single-point physical quantities (pressure and volume) between the α phases predicted by random searching and the reference one at 0 GPa. SNAP refers to the search with the less converged database, SNAP-CV to the one trained with the more converged database.

	Volume ($\text{\AA}^3/\text{at}$)	Pressure (GPa)
SNAP-CV Phase	38.7176	-1.59
Ref. DFT Phase	36.86	0

Table 3.2.8: Comparisons of the predicted α phase by SNAP-CV and the reference DFT one at 0 GPa.

This study confirms that, the k-point mesh, the convergence criteria for the forces and the plane-wave cut-off smearing applied to compute the training databases all influence strongly the quality of the SNAP potential. It also gave indications of how much the DFT calculations of the training database have to be converged to obtain an accurate numerical potential.

3.2.4 Predictions under Pressure

The obtained potential trained on the more converged database and updated at the end of the active learning cycles is now used to predict structures under pressure. In particular, structure predictions are performed at 20 GPa and in the high-pressure domain (100 GPa).

	Volume ($\text{\AA}^3/\text{at}$)	Pressure (GPa)
SNAP-CV Phase	22.06157	19.49
Ref. DFT Phase	22.01	20

Table 3.2.9: Comparison of the predicted γ phase by the SNAP-CV numerical potential and the reference DFT one at 20 GPa.

Structure Predictions at 20 GPa

Random searching using the SNAP potential is applied on the 10 000 structures formerly generated by applying an external pressure of 20 GPa. The phases having an appearance frequency higher as 10 over 10 000 and a predicted enthalpy below -1893.9 eV/at are selected. The BCC and γ phases are retrieved. It corresponds to the two most stable phases at this pressure ($\gamma < \text{BCC}$).

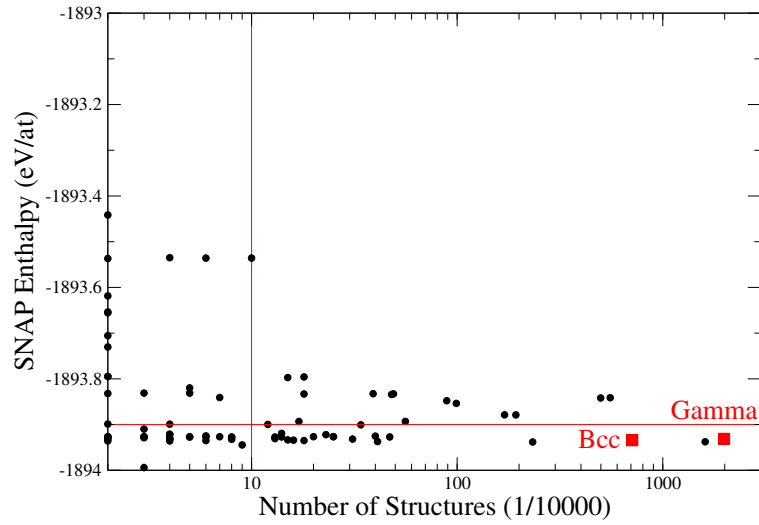


Figure 3.2.14: Appearance frequency and SNAP enthalpy distribution following the random searching (at 20 GPa) using the numerical potential obtained after the two active learning cycles.

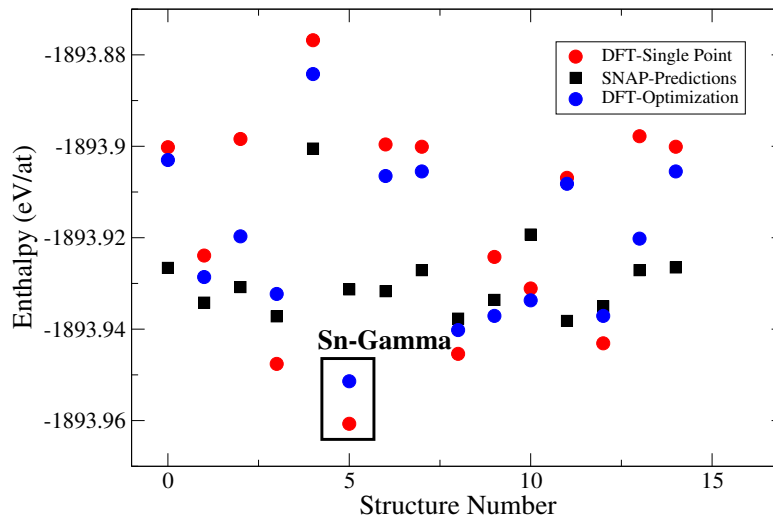


Figure 3.2.15: Comparisons of the enthalpies predicted by SNAP, obtained by single point DFT and after DFT structural optimization of the selected structures. The RMSE between SNAP predicted energies and single-points DFT is 23 meV/at.

The selected region defined by the filters on both appearance frequency and SNAP enthalpy is made of 14 distinct crystal configurations. The differences for volume and pressure between the γ predicted structure and the expected one at 20 GPa (and computed by DFT) are presented in Tab 3.2.9. The structure predicted is again close to the reference one. Final DFT optimizations on the selected structures give the γ phase as the most stable one (see FIG 3.2.15).

Structure Predictions at High Pressure

As mentioned above, the numerical potential using the highly-converged database appears as better to evaluate the properties of tin under pressure. It is then tested to retrieve phases under high pressure at 100 GPa. The BCC, γ and HCP phases are obtained with γ and BCC phases predicted to be more stable as the HCP one (see 3.2.16). It is exactly the expected results.

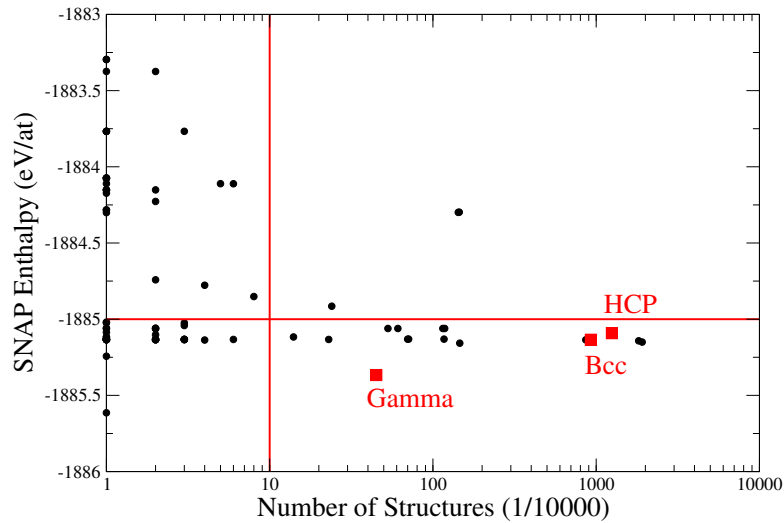


Figure 3.2.16: Appearance and SNAP enthalpy distribution following the random searching (at 100 GPa) using the numerical potential obtained after the two active learning cycles.

3.2.5 Conclusion

The machine learning approach for crystal structure prediction as developed in this chapter provides with rather satisfactory results in the case of tin. The reference phases have been retrieved at 0, 20 and 100 GPa although not incorporated in the initial training database. It is in addition worth mentioning that 10 000 tin structures are explored by just using around 2 CPU hours (notice this remark does not take into account the cost of the training database which is however paid once and requires only single-point DFT computations having a much lower cost than DFT structural optimizations). AIRSS would require in a range of 10^6 CPU hours if the DFTO-1 parameter set would be employed. This study was also the opportunity to assess the importance of controlling the convergence of the training databases in terms of DFT numerical parameters. In particular, after improving the convergence it can be noticed that:

- By comparing the predictions made at 0 GPa, the α phase identified after the second round of SNAP random searching is closer to the one expected (computed by reference DFT calculations). In particular, if one considers the structures searched by SNAP and selected by the filters. The enthalpies obtained by single point DFT computations are closer with those of the structures optimized by full and precise DFT calculations. The same remark works for the atomic volume. It is due to the fact that the numerical potential trained on the more converged database is more accurate in terms of external pressure to apply.
- Accurate phase identifications are obtained under pressure. Moreover, the transferability of the potential to higher pressure domains at 100 GPa works nicely. Such predictions did not work with the numerical potential trained on the first and less converged training database. The phase expected at 100 GPa is successfully retrieved.

These improvements come from the fact that a better precision is reached on the training database for the stress tensor components, in particular for the triclinic type crystal structures.

From a more general point of view, the fact that the α phase (having by far a different electronic structure as the phases initially learnt) has been successfully retrieved supports that numerical potentials and the active learning process developed here are able to perform structure predictions. It highlights thus the capability of transferability. It is complementary to the reproducibility, i.e. the capability of the numerical potential to reproduce with a high accuracy (RMSE with a few meV/at) the physical quantities (e.g. energy) of structures being not too far from those of the training database. This ability will be confirmed in what follows in the case of a superhydride system of YH_x .

Following this proof of concept of a RSS algorithm which combines DFT calculations and numerical potential on a pure compound. The extension of the method to the research of superhydrides is presented in the next part. The binary system YH_x is considered for these developments.

Chapter 3.3

Application to Yttrium Superhydrides

Contents

3.3.1 A First Test on the YH_6 System	151
3.3.1.1 Initial Database	152
3.3.1.2 Search of the YH_6 Im-3m Phase at 200 GPa	153
3.3.1.3 Search of Possible New YH_6 Phases	155
3.3.2 Development of New Analysis Tools	160
3.3.2.1 Bispectrum Analysis	161
3.3.2.2 X-Ray Based Metric	165
3.3.3 YH_x Explorations at 147 GPa	168
3.3.3.1 Experimental X-Ray Spectrum and Metric	169
3.3.3.2 Random Searching Implementation	172
3.3.3.3 New Structures Obtained	173

This chapter presents the application of the numerical potential approach to the binary yttrium superhydride YH_x system. The difficulties encountered to explain the X-ray diffraction patterns of some experimental YH_x phases with the CSP algorithms using DFT are probably related with the fact that these approaches can not work with structures having more than ~ 20 -30 atoms in their primitive cell. Indeed, above this size, the number of local minima in the configuration space (each one corresponding to a crystal structure), increases exponentially with the number of atoms in the unit cell. This number becomes thus too large to efficiently explore the PES with Ab-Initio methods. One faces an exponential wall of the number of structures to explore or equivalently the number of dimensions of the PES. A main motivation of this chapter is thus to propose solutions to push back this exponential wall and reach successful predictions of complex materials. The first section is about a first test on a particular stoichiometry, namely YH_6 . Random search with the SNAP numerical potential is performed at 200 GPa where the most stable phase is known from previous studies. The main purpose is to check the transferability of the approach, in particular whether it is able to identify the correct structure without including it in the training database. The second section discusses analysis tools to identify the most interesting structures among the thousands explored by a numerical potential. Finally the last section details the implementation of the tools developed to try to identify of a new YH_x phase that could explain the experimental X-ray spectrum near 150 GPa.

3.3.1 A First Test on the YH_6 System

This section presents the implementation of the numerical potential approach to search the structure of YH_6 phases at 200 GPa. The main purpose is to validate the transferability of the method by retrieving the correct expected phase without putting it in the training database, as successfully done for tin. We remind that this expected structure is a cubic phase with space group Im-3m [31]. In a second time, random searching in association with the numerical potential is performed in order to identify possible new structures of this superhydride.

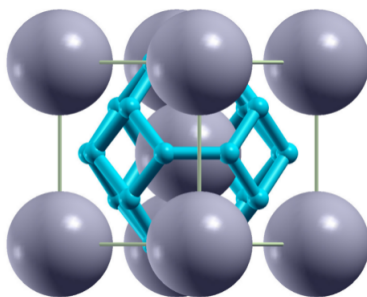


Figure 3.3.1: The YH_6 Im-3m phase at 200 GPa.

3.3.1.1 Initial Database

First of all, an initial database has to be built. Three main types of atomic configurations are considered.

- Disordered configurations: 281 structures are generated. They are initially made of up to 42 atoms (6 formula units). Each of these structures contains 1 to 2 yttriums with 3 to 20 hydrogen atoms. They are used to "learn" atomic forces.
- Crystal type configurations corresponding to different phases identified by the former study performed by DFT (see 2.3): eleven crystal structures (computed at different pressure values) are taken into account. In particular, two YH_6 structures of space group Im-3m (note that in a first step, these structures will not be considered to train the numerical potential because one checks its ability to retrieve the Im-3m structure without learning it), four YH_3 structures of space group Fm-3m , three YH_4 of space group I4/mmm and two YH_9 structures of space group $\text{P2}_1/\text{m}$.
- Triclinic configurations (130 structures). They are generated to learn stress.

To build this database, a criterion for the interatomic distances is imposed. Here, the shortest distance is imposed to be larger than 1 Å. As a result, no configuration having hydrogen molecules is stored in the database (we remind that the interatomic distance in the H_2 molecule is ~ 0.75 Å). FIG 3.3.2 presents the pressure distribution of the configurations stored in the initial database. It can be remarked that no structure at pressure around 200 GPa is present in it. Identifying correct phases at this pressure will thus be a supplementary validation of the transferability of the method.

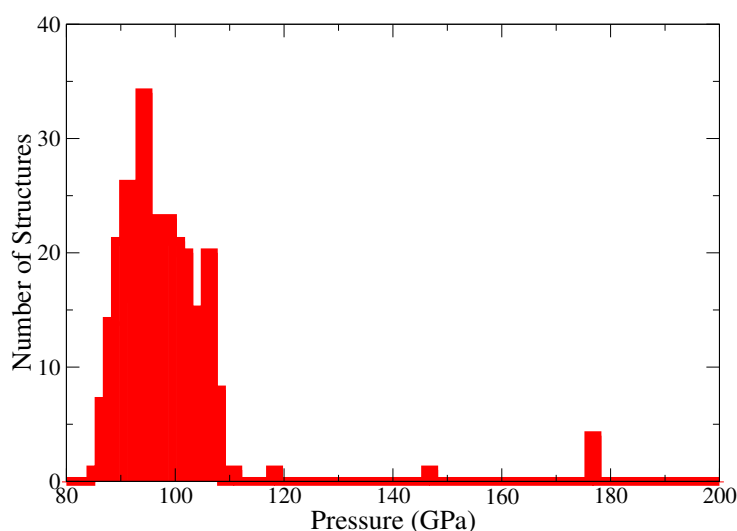


Figure 3.3.2: Pressure distribution of the structures stored in the initial database.

To perform the different DFT computations (single-point and structural optimizations), the following parameters are used. The Generalized Gradient Approximation (GGA) is employed for the exchange and correlation energy (GGA-PBE). The calculations are performed in the framework of the PAW formalism, with PAW datasets for Y and H taken from the JTH table [141]. The plane-wave cut-off is 30 Ha. For the single-point DFT computations a k-point mesh of $8 \times 8 \times 8$ is used. A k-point mesh $12 \times 12 \times 12$ is employed for the DFT structural optimizations. Between two consecutive self-consistent cycles, the differences between maximal forces has to be below the maximal force of the current cycle times 10^{-5} ($\delta F < 10^{-5} F_{max}$). The BFGS algorithm is applied by imposing 10^{-6} Ha/Bohr³ as stopping criterion with up to 100 iterations.

3.3.1.2 Search of the YH_6 Im-3m Phase at 200 GPa

A first validation consists in finding the expected Im-3m phase (see 2.3.7) without including the related structures in the learning database. A numerical potential is thus trained on a filtered initial database where all the YH_6 structures linked to the Im-3m phase are removed.

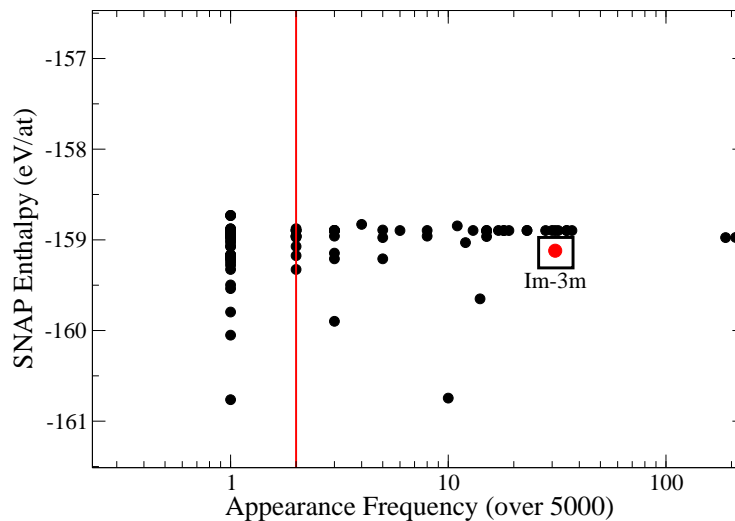


Figure 3.3.3: SNAP enthalpy and appearance frequency distribution of the structure clusters after the SNAP random searching.

A random search is performed on 5000 configurations having one chemical formula unit. The configurations randomly generated are structurally optimized with SNAP, with a final pressure fixed to 200 GPa. As made with tin (see 3.2.6), the different configurations obtained are gathered in clusters of similar structures sharing the same space group and set of Wyckoff positions. The number of structures in each cluster is called the appearance frequency. FIG 3.3.3 shows the distribution in the plan (SNAP enthalpy, appearance frequency). Those having an appearance

frequency larger than 2 (over 5000) are selected for single-point DFT calculations. It can be observed that a phase having the Im-3m space group appears around 30 times, exhibiting one of the most important frequencies. Tab 3.3.1 shows the differences between the Im-3m phase identified by the random searching process with SNAP, before and after the DFT structural optimization at 200 GPa (the structure so obtained is indeed structurally optimized by DFT). The expected Im-3m phase is confirmed after DFT structural optimization. The differences observed in Tab 3.3.1 are due to the fact that the numerical potential made an error on the pressure. The DFT pressure of the phase identified by the random searching before DFT optimization is indeed 208 GPa instead of 200 GPa.

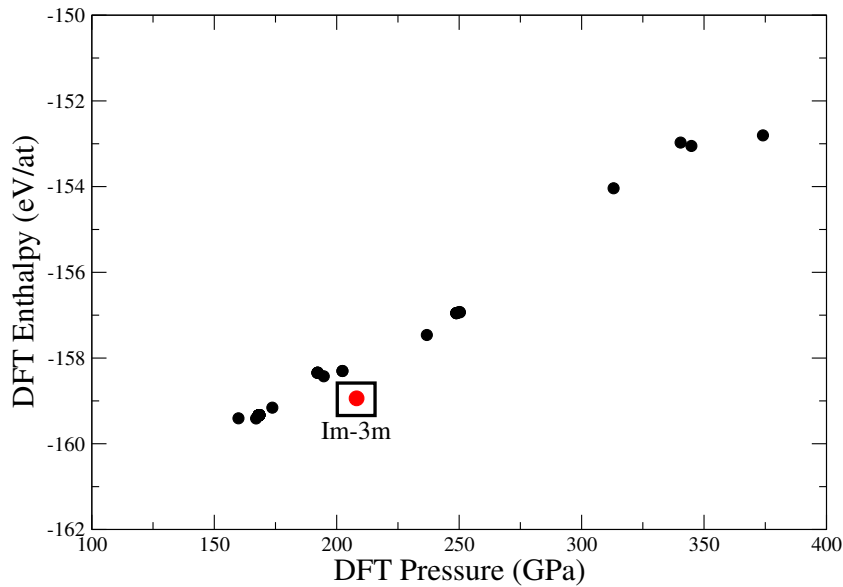


Figure 3.3.4: DFT enthalpy and DFT pressure distributions of the selected structures after single-point DFT computation.

From each cluster selected, one representative structure is chosen to perform single-point DFT calculations. FIG 3.3.4 presents the distribution of these selected structures in terms of DFT enthalpy and DFT pressure. For the configurations having a pressure near 200 GPa, the Im-3m phase is the most stable one. However for some structures, the numerical potential made important errors on the pressure. In particular, some of them have in fact a lower DFT pressure and at the same time a lower DFT enthalpy, hence lower enthalpy values are probably related to lower pressure applied. It may indicate the possibility to find other phases with a lower (or at least very close enthalpy) than the Im-3m phase at 200 GPa. To answer this question, another random search with a numerical potential having more information in the training database (and thus the structures related to the Im-3m phase) is performed. This is the purpose of the following

	Im-3m Phase Before DFT Optimization	Im-3m phase After Full DFT Optimization (200 GPa)
DFT Enthalpy (eV/at)	-158.811933	-158.96665
Lattice Parameters (\AA) of the Conventional Cell	a,b,c=3.497	a,b,c=3.5105
Volume ($\text{\AA}^3/\text{f.u}$)	21.38	21.63
DFT pressure (GPa)	208	200

Table 3.3.1: Comparison between the Im-3m phases predicted by SNAP random searching before and after full DFT structural optimization at 200 GPa. The conventional unit cell has a multiplicity equal to 2 and contains thus 14 atoms.

paragraph. We can conclude at this point that the Im-3m phase has been correctly predicted by a SNAP numerical potential.

3.3.1.3 Search of Possible New YH_6 Phases

In this paragraph a random search at 200 GPa is performed on 5000 atomic configurations of YH_6 made of one single chemical formula (7 atoms). The aim is to confirm, among these structures, whether the Im-3m phase is the most stable one. The purpose of this study being to identify possible unknown phases, it is needed to convey the more information to the numerical potential during the training. As a consequence, the structures related to the YH_6 Im-3m phases are included in the training database, and "learnt". For this structural exploration, an active learning process is employed (see 3.1.3). At the end of the second active learning cycle, the different phases selected are subjected to full DFT optimization at 200 GPa.

First Random Searching

A numerical potential is trained on the full initial database (including the structures related to the Im-3m phase), G1. 5000 structures of YH_6 having one formula unit are generated. After the random search using G1, 121 different clusters of structures sharing the same space group are obtained (see 3.3.5). The clusters having more than two structures are selected for DFT computations. Among the selected clusters, one associated to the Im-3m space group can be identified. It appears more than 300 times, which corresponds to a high frequency. It can thus be noticed that including the Im-3m phase in the training database increases by a factor 10 its appearance frequency (see 3.3.3 for comparison).

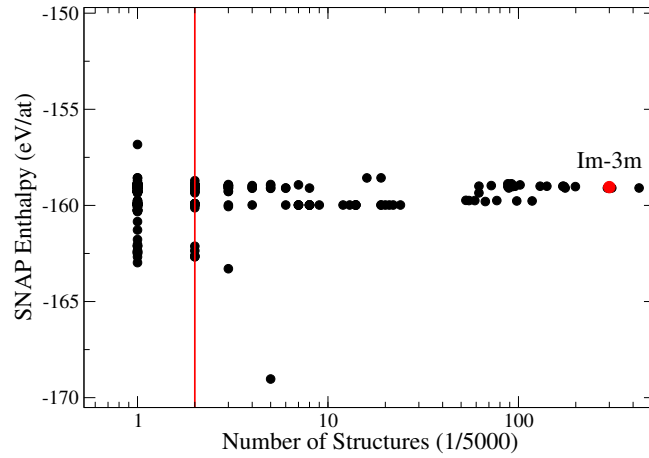


Figure 3.3.5: SNAP enthalpy and appearance frequency distribution of the structure clusters after the random search with G1. The selected group of structures are those located on the right of the red vertical line.

By taking one representative structure for each selected cluster, FIG 3.3.6 presents the distribution in the plan (enthalpy, pressure) computed by DFT of these phases (single-point DFT calculations). It can be noticed that the potential makes sometimes important errors on the pressure. The phase of space group Im-3m has however the lowest enthalpy near 200 GPa.

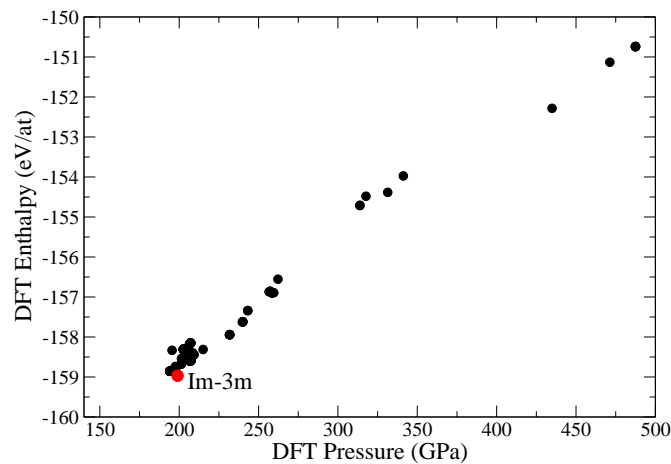


Figure 3.3.6: DFT enthalpy and pressure distribution of the selected structures after single-point DFT computation.

Update of the Potential

The selected structures are added to the training database and another potential is trained on it (called G2). As an illustration of the active learning process, FIG 3.3.7 presents the correlations of the stress for the potential G1 and the new one G2. The SNAP G1 potential exhibits almost no correlation for the selected structures. It means that the selected structures are far from the training database and are thus required for learning. These correlations are deeply improved with G2 with a RMSE of a few eV/Ha³ for the selected phases. The potential improves its knowledge about the studied system.

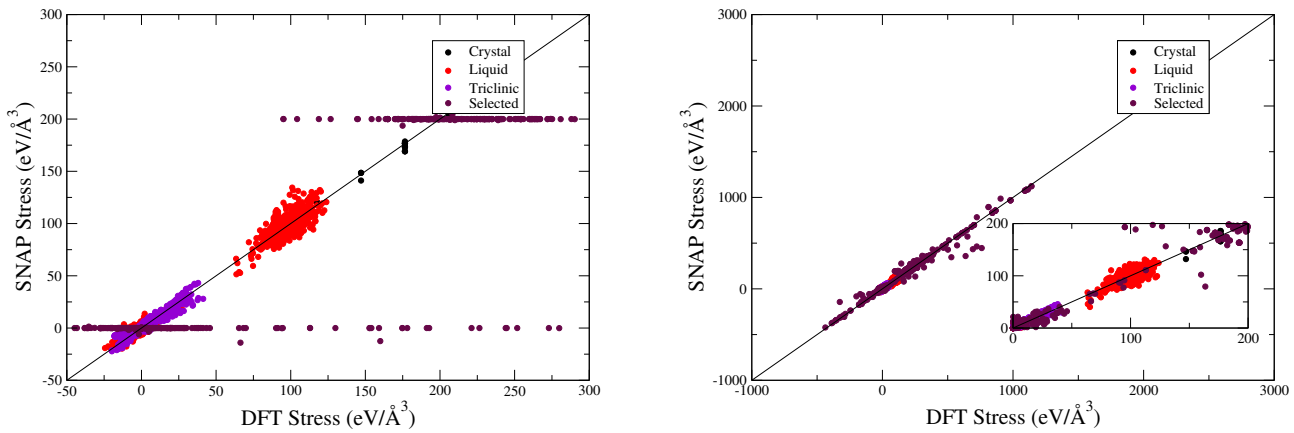


Figure 3.3.7: Stress correlations between SNAP and DFT values with G1 (left) and G2 (right) numerical potentials.

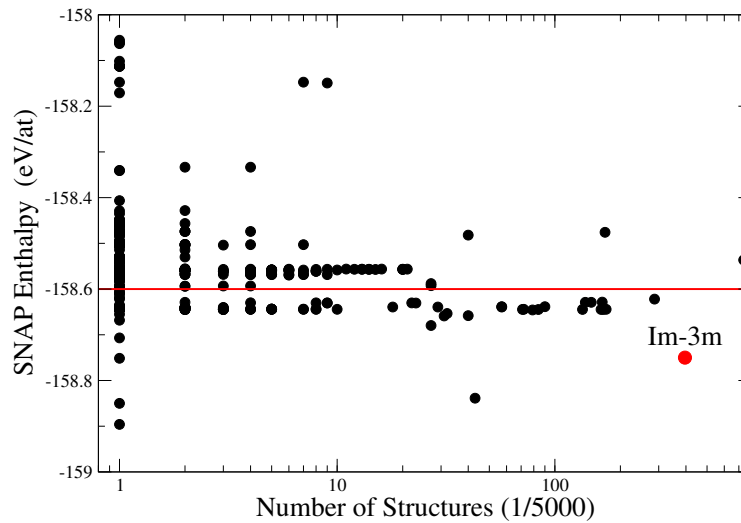


Figure 3.3.8: SNAP enthalpy and appearance frequency distribution of the structure clusters after random search using G2.

Second Random Searching

The random searching is performed again using the G2 potential on the 5000 formerly generated configurations. A filter on the predicted SNAP enthalpies is applied: the groups of structures below -158.6 eV/at are selected for single point DFT computations. This energy threshold is fixed to select half of the obtained clusters. The graph 3.3.8 shows the distribution in terms of SNAP enthalpy and appearance frequency of the structure clusters after the random search using G2. The structures having the Im-3m space group are included in the selection region and are among the most stable (according to the predicted enthalpy). The appearance frequency of the Im-3m phase goes up from around 300 with G1 to 400 structures with G2.

Structural Optimization

Choosing again one structure in each selected cluster, the graph 3.3.9 presents the distribution in terms of DFT pressure and DFT enthalpy. It appears that some of them have both a lower DFT pressure and DFT enthalpy as the Im-3m phase and thus might be more stable. To answer this question, full DFT structural optimizations at 200 GPa have to be made on a few and representative structures. For this purpose, the DFT enthalpies obtained are used to select the configurations. The structures having a DFT enthalpy lower as -158.5 eV/at are taken into account for structural optimizations with DFT (see FIG 3.3.9). This energy threshold is chosen to select the structures having a DFT pressure near 200 GPa (see FIG 3.3.9). By considering these selected structures, the different space groups are analyzed with *findsym*.

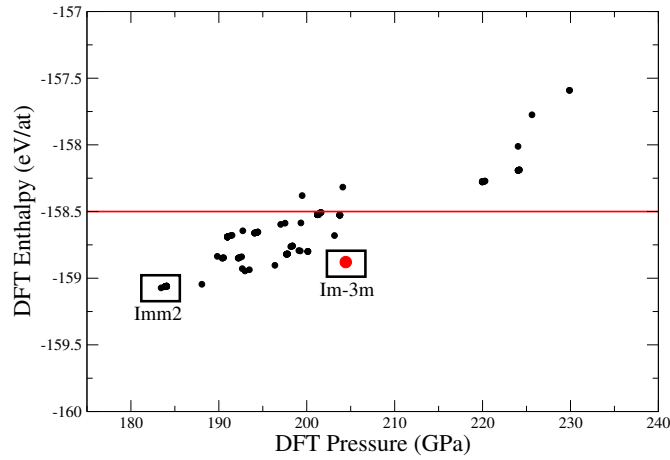


Figure 3.3.9: Distribution of the structures computed by single-point DFT computations in the plan (pressure, energy).

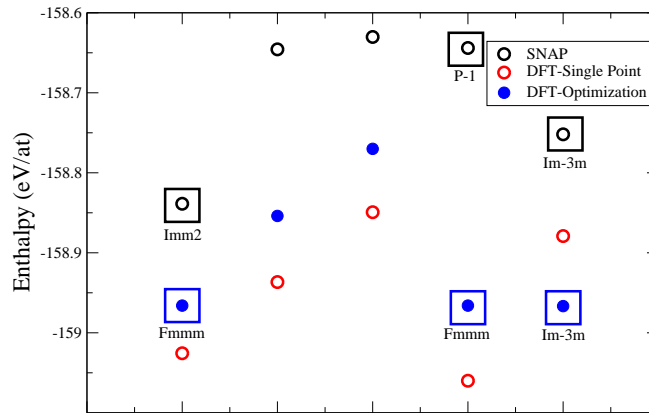


Figure 3.3.10: Enthalpies of the selected structures, as obtained by (i) SNAP structural optimizations at 200 GPa, (ii) single-point DFT calculations using the geometries optimized with SNAP, and (iii) full DFT structural optimizations at 200 GPa. Note that single-point DFT enthalpy values can be lower than those obtained after full DFT optimizations. It is explained by the fact that the numerical potential makes error for the pressure which has consequences on the DFT enthalpy value.

Five different space groups are identified, a structure of each of these is taken for full DFT structural optimization. At the end of it, two structures converge to the same space group Fmmm.

These structures had respectively the Imm2 and P-1 symmetry before the DFT optimization. It appears in FIG 3.3.10 that the Fmmm phase has an enthalpy very close to that of Im-3m one in the range of 8 meV/at higher (which remains however a significant difference as the precision of the DFT computations with the convergence parameters chosen here is in the range of 1 meV/at). Tab 3.3.2 provides the Wyckoff positions of this Fmmm phase at 200 GPa. On the contrary to the Im-3m phase having a cubic geometry, this Fmmm phase is orthorhombic.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
<i>Fmmm</i>	200 GPa	a=4.95 b=4.98 c=3.50 $\alpha = \beta = \gamma = 90$	Y	4b	0	0	0.5
			H	16o	0.37545	0.37517	0
			H	8f	0.25	0.25	0.25

Table 3.3.2: Wyckoff positions and lattice parameters and angles of the Fmmm phase at 200 GPa.

This first section was the opportunity to validate the transferability of the method and a first try to search for new YH_6 structures. However these searches remained simple as only structures having one chemical formula in their unit cell were investigated. The coming challenge with the numerical potentials is to enable the exploration of more complex and potentially more stable phases. As a preliminary step, new analysis tools to identify the best structure candidates have to be developed. This is the purpose of the next section.

3.3.2 Development of New Analysis Tools

The structural optimizations using a numerical potential are not as precise as those made with DFT (during the structural optimizations by a numerical potential, error is made for the pressure and forces applied to the atoms in comparison with reference values obtained by DFT). As a result, when such a potential is employed to explore more complex phases (i.e. having more than one chemical formula unit in the unit cell), the direct space group analysis to build clusters (using *findsym* [137, 138] program) does not always work. In some cases, the structural optimizations using SNAP lead to artificially low-symmetry configurations which can be only identified as having P1 as space group. This is a real problem because it means when several thousands or millions of structures will be explored by a numerical potential, it will be impossible to identify the most interesting ones. If an efficient selection of structures (to update the training databases or to enforce DFT structural optimizations) can not be performed, the explorations made will be unusable as it will not be possible to structurally optimize by DFT all the structures explored, due to their very large number. In parallel to the development of the active learning process using numerical potential to search more complex structures (having typically up to 100 atoms in the unit cell), new tools have thus to be developed so as to distinguish the structures between each other for candidate identification. These tools have to be more flexible with respect to the less precise structural

optimizations a numerical potential enforces on a given structure. To this aim, two analysis methods are presented in this section. The first one is based on a bispectrum descriptor (see 3.1.2.1) coupled with a clustering algorithm. The second one uses X-ray spectrum simulation.

3.3.2.1 Bispectrum Analysis

To test the possibility to identify groups of similar structures based on bispectrum analysis, a SNAP random search enforced on the single chemical formula unit (Z1) YH_6 structures is firstly considered. In addition, random searching with structures having two (Z2) and three (Z3) chemical formula units will be implemented in a second time to extend the use of bispectrum analysis to more complex phases. To build another numerical potential able to search phases with several chemical formula units, new structures are added to the training database presented in 3.3.1.1. These configurations correspond to different steps of DFT optimizations applied to phases having three chemical formula units of YH_6 . This additional part of the database aims at learning the first steps of a structural optimization for structures made typically of more than 10 atoms. The new potential obtained is more general than the former one. It makes possible the optimization of YH_6 structures having more than one chemical formula unit. However it is worth noticing that it is a little bit less precise for the optimizations of the Z1 structures. Indeed it finds Z1 structures having the space group Im-3 instead of the Im-3m one. However when full DFT optimizations are applied on them, the correct Im-3m phase is obtained.

In the case of the exploration of the Z1 YH_6 phases, the direct symmetry analysis with *findsym* still works. This provides a basis of comparisons to evaluate the accuracy of a clustering method. The bispectrum vectors of all the yttrium atoms are computed with $J=55$ 3.1.2.1. It consists thus of vectors having a dimension of 55. The clustering process is enforced by an unsupervised machine learning method called the Gaussian Mixture Model (GMM) using the sklearn python library (see [209]). It aims at distributing the 55 dimensioned bispectrum vectors in clusters. An important hyperparameter to fix is the number of clusters. A silhouette score analysis (see [210]) points out that a number of 30 clusters could be considered.

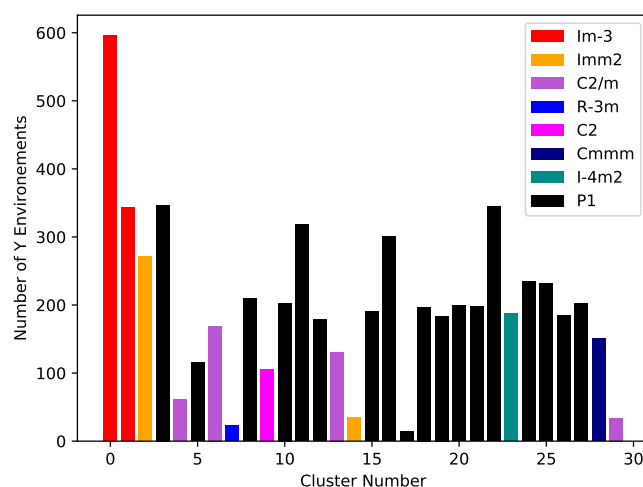


Figure 3.3.11: Distribution of the Z1 structures between the 30 clusters.

The GMM clustering on bispectrum vectors identified clusters of structures. It appears by verifying with *findsym* that the structures belonging to one given cluster have the same symmetry space group. It is noticeable that two distinct clusters can have the same space group (like Im-3, Imm2 or C2/m for instance). It can be explained by two different reasons:

- Two clusters associated to the same space group are made of structures having different sets of Wyckoff positions. This is the case for the clusters made of structures having the C2/m space group.
- Two clusters can be distinct due to small variations of the atomic environments after the SNAP structural optimization (probably related to the fact that the numerical potential identified two close but different local minima). However space groups and Wyckoff positions remain the same. This is the case encountered for clusters associated to the Im-3 and Imm2 space groups.

Seven different space groups (without including P1) are identified. In particular, two clusters are made of structures having the Im-3 space group (as mentioned above, it is the structures which give the Im-3m phase after DFT optimization). FIG 3.3.12 presents how the 30 clusters identified are distributed. For visualization purpose, a projection in the plan of the first and second component of the bispectrum vectors (B_0, B_1), is presented in FIG 3.3.12. The colored ellipses highlight the clusters associated to a specific space group (different from P1). Their centers are determined by the mean of the bispectrum coefficients, width and height are fixed by the variance following each direction.

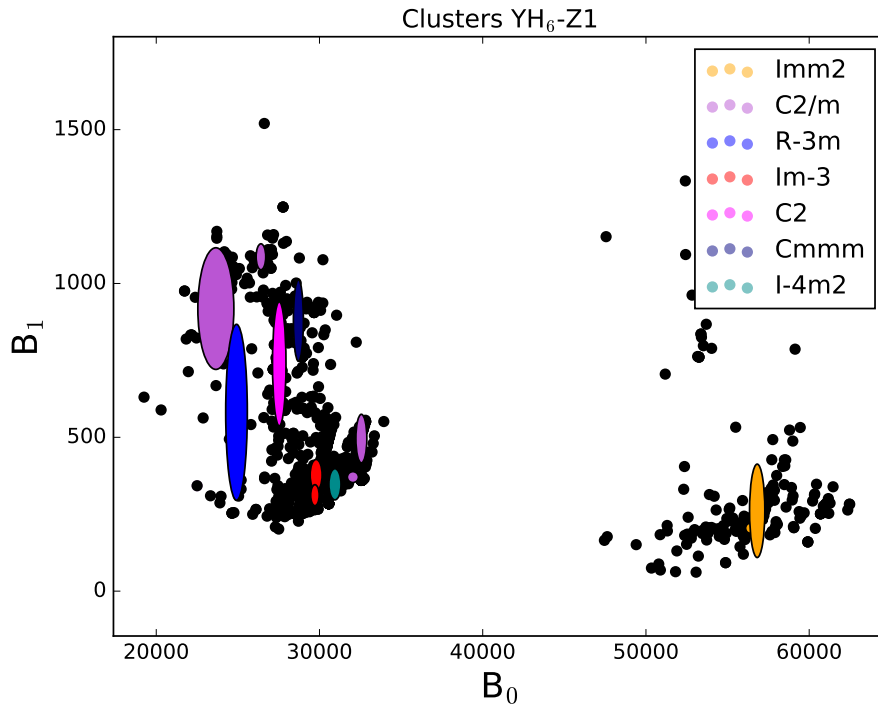


Figure 3.3.12: Distribution of the Z1 clusters projected in the plan (B_0, B_1) for visualization purpose.

Following this first work on the Z1 structures which shows the ability of the method to gather similar phases in clusters, the challenge is now to extend it so as to analyze searches performed on more complex structures. The main task here is to identify structures which give the Im-3m phase after DFT optimization. The bispectrum vectors of each environment of the yttrium atoms are computed after two SNAP random searchings involving 3000 structures of two (Z2) and three (Z3) chemical formula units respectively. In this case, each structure gives two (for Z2) and three (for Z3) bispectrum vectors. These random searchings share the fact that the direct symmetry analysis (with *findsym*) fails, all the explored structures have only the P1 space group. The distributions of the bispectrum vectors projected in the plan (B_0, B_1) are shown in FIG 3.3.13. In addition, clusters identified by Z1 random searching and associated to specific space groups are represented in the form of ellipses. It can be noticed these regions are occupied by Z2 and Z3 environments.

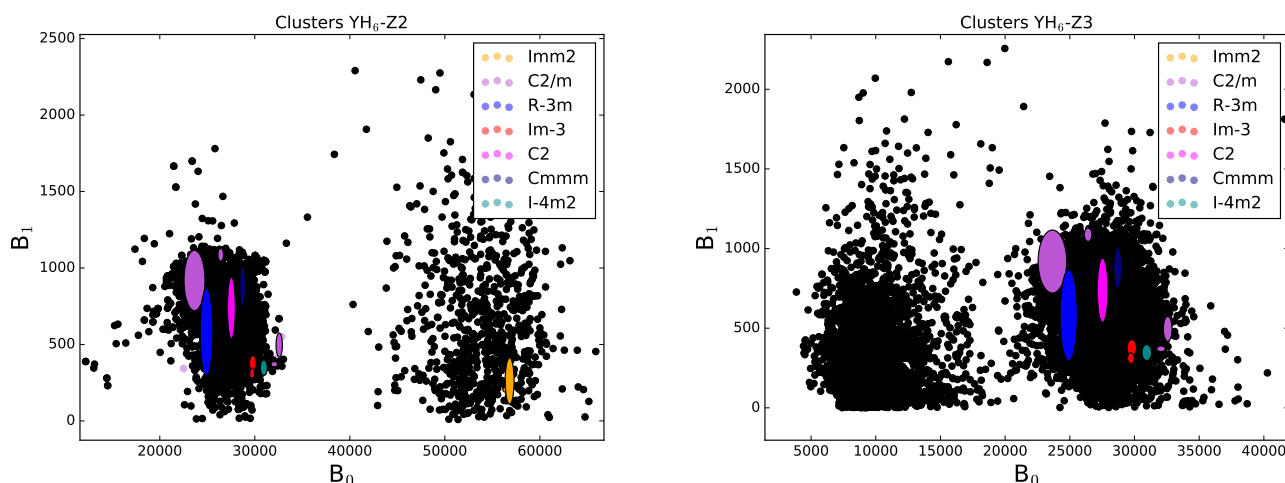


Figure 3.3.13: Distribution of the clusters in the plan (B_0, B_1) for the Z2 (left) and Z3 (right) yttrium environments, the colored ellipses represent the clusters identified by Z1 random searching and associated with a specific space group. It is worth mentioning that the Imm2 region is not retrieved with Z3 because this Imm2 phase is not dynamically stable in Z3 with the numerical potential used.

To identify structures optimized by SNAP which give the phase Im-3m after DFT optimization, an euclidean distance on the 55 dimensioned vectors is used. For this purpose, the two clusters associated to the Im-3 space group (and thus made of Z1 structures giving the Im-3m space group after DFT optimization) are considered. The barycenter of these clusters is computed. From it, a group of the 100 closest bispectrum vectors is selected. Among these environments, 10 structures having all their yttrium atoms in this group are chosen. They initially have the P1 space group. After full DFT optimization the Im-3m phase is obtained. The same work is conducted on the Z3 atomic environments. The first hundred bispectrum vectors closest to the Im-3 region are selected. Five structures having all their yttrium atoms in this group are taken and give after full DFT optimization the Im-3m phase.

Thanks to the use of a metric combined with bispectrum analysis, it was thus possible to identify the P1 structures the closest from the Im-3m phase to finally retrieve it with DFT optimization. In addition, this result proves that the expected phase is effectively obtained by random search using SNAP numerical potential for structures having two and three chemical formula units. These tests show that bispectrum can be considered as an analysis tool to select structures of interest after random searching explorations. Combined with a clustering method, it can be employed to identify groups of similar atomic environments without the use of a direct symmetry analysis. When it is used with an euclidean metric, it gives the possibility to identify atomic environments close or far from reference ones and thus select structures following the distance criterium considered.

3.3.2.2 X-Ray Based Metric

X-ray spectrum is the main tool employed to experimentally determine a crystal structure of a material. The main idea here is to exploit the X-ray spectrum to select phases explored by random searching such that experimental information can be used to filter the numerical results. The selection strategy might be based on information coming from experiments. This approach is presented here in two steps. Firstly, it is applied on a random search involving single chemical formula YH_6 structures. Using a metric taking into account the positions of the peaks of the experimental X-ray spectrum, the aim is to identify which structures give the Im-3m phase after DFT optimization. The second test consists in identifying by the same way a structure explored in a random search of YH_6 phases made of six chemical formula units, and to see whether it gives the Im-3m space group after DFT optimization.

The Metric Considered

For these tests, X-ray spectrum presented in FIG 3.3.14 of the Im-3m phase at 200 GPa is considered (note that the intensities are rescaled in the way that a value of 100 is associated to the maximum). The X-ray wavelength taken is $\lambda = 0.3738 \text{ \AA}$. The metric which is employed to evaluate distances between X-ray spectra is based on the positions of the peaks. For the reference X-ray spectrum, the main positions of the peaks (noted p_i) are placed for 2θ equal 8.65, 12.23, 15.0, 17.33 and 19.39 degrees. For a given spectrum having its positions of the peaks in q_j , the distance with the reference is given by:

$$\frac{1}{N} \sum_i \sum_j (1 - e^{-\frac{(p_i - q_j)^2}{\sigma^2}})$$

Where N is the number of (i,j) couples. This metric needs two hyper-parameters.

- A minimum intensity threshold to determine which maximum in the X-ray spectrum is a peak. It is fixed to 15 arb. units.
- The scale parameter σ , it is fixed here to 0.005 degree.

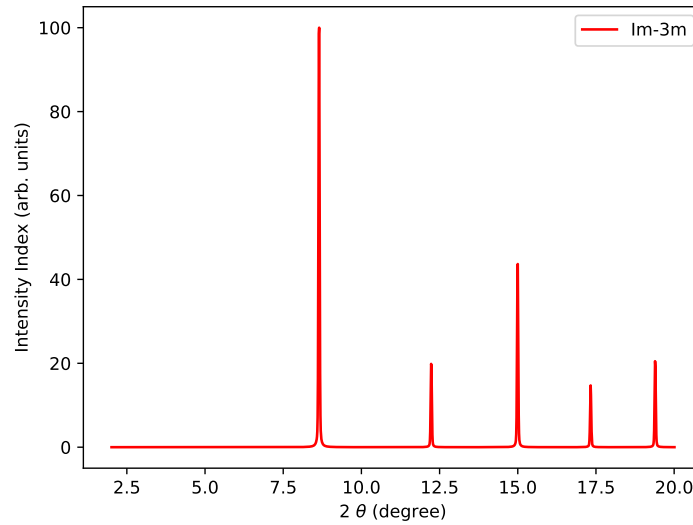


Figure 3.3.14: Reference X-ray spectrum of the Im-3m phase at 200 GPa computed with *cryscalc* [211] software.

As a first validation, 500 YH_6 structures having one chemical formula unit are explored by random searching using SNAP numerical potential. For each of these structures, the metric defined above is applied. FIG 3.3.15 presents the distribution of the distances. The red bars correspond to the explored structures giving the Im-3m phase when DFT optimizations are applied. It can be noticed that this metric enables a bimodal distribution where the structures close to the correct phase are well separated from the other configurations labelled P1 (no symmetry) with *findsym*. It is worth noticing that with the metric defined here, a small variation of the positions of the peaks from the target structure (characterized by the q_j s positions) can change the metric value from zero to a value close to one. It is the reason why the values of the metric in itself is not really meaningful, the main point is to have one able to distinguish the closest structures to the target one by highlighting bimodal distributions.

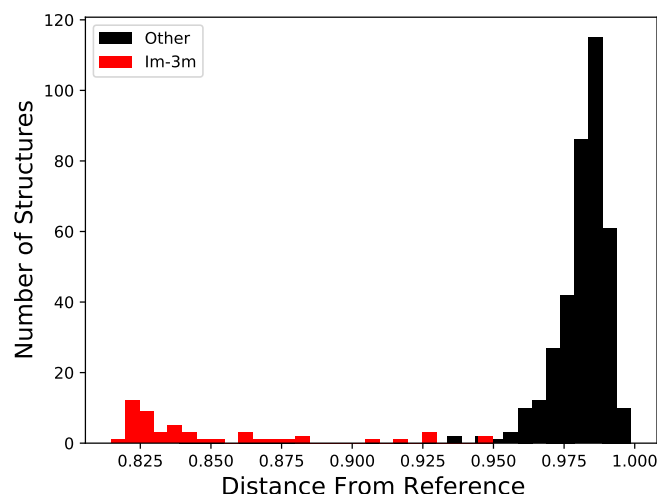


Figure 3.3.15: X-ray spectrum distance distribution of the simulated YH_6 Z1 structures. The structures giving the Im-3m space group after DFT optimization (red bars) are well separated from the others.

The second test consists in identifying a P1 phase which gives the Im-3m phase among YH_6 structures having six chemical formula units. For this purpose, 3000 structures are explored by SNAP random searching. FIG 3.3.16 presents the distribution of the distances. It can be seen that a few dozen of structures have a distance below 0.95 arb. units and thus are completely separated from the others.

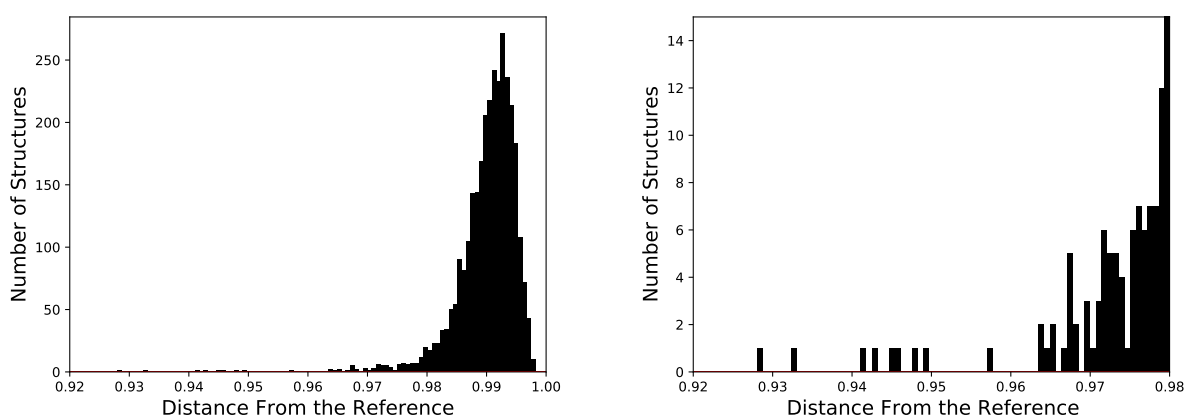


Figure 3.3.16: X-ray spectrum distance distribution of the YH_6 Z6 structures explored. The right figure highlights the configurations closest to the reference and separated from the others.

One of the structures having its distance below 0.94 arb. units is taken. The simulated X-ray spectrum with the reference one is shown in FIG 3.3.17. One can see that the locations of the peaks are close to those of the reference.

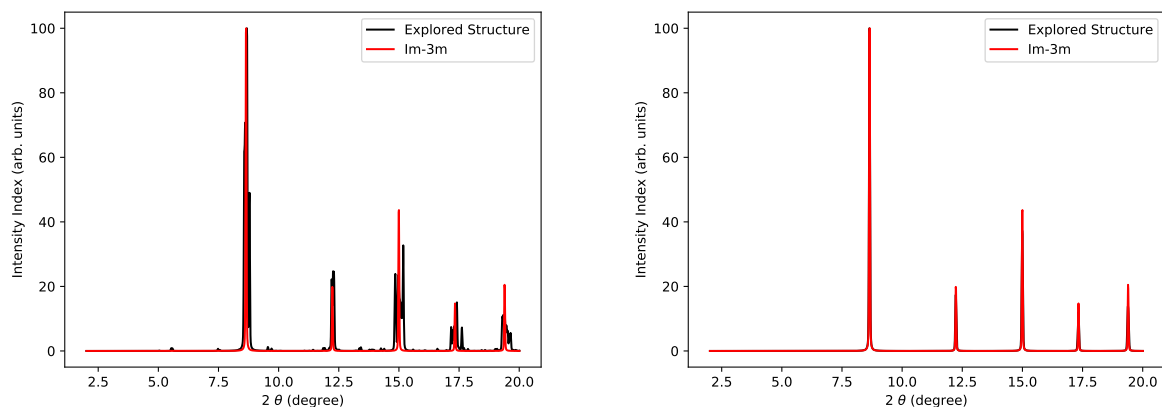


Figure 3.3.17: Comparison between the X-ray spectrum of the selected YH_6 Z6 and the reference one before (left) and after (right) the DFT optimization.

This phase has initially the P1 space group. A full DFT optimization is enforced on it and the space group Im-3m phase is obtained at the end. The figure 3.3.17 presents the X-ray spectrum of the structure after its DFT optimization. It is completely superposed with the reference one.

These two tests confirmed that X-ray spectrum analysis is an additional tool to select structure candidates without using a direct symmetry analysis. From a reference, it was possible to identify the phases closest to the targeted one. This approach could then be enforced for random searching where structures are explored to find a phase matching a given experimental X-ray spectrum.

This section presented two analysis tools to select structures after a random searching involving possibly complex phases. The first one is based on description of atomic environments and consists of gathering similar structures (thanks to a clustering algorithm), or selecting those having a particular known atomic environment (by the use of a high dimensional euclidean metric). The second one is constructed on more global information thanks to X-ray spectrum simulations (or experiment). This method is particularly suitable to include experimental information (the positions of the peaks for instance) in the random searching process. Random searching using numerical potential as well as the analysis tools presented here are now applied to explore unknown phases. The main purpose of these explorations is to reach a better explanation of the YH_x X-ray spectrum obtained near 150 GPa. All these developments are presented in the next section.

3.3.3 YH_x Explorations at 147 GPa

As discussed in Chap 2.3, the identification of the structure of binary yttrium superhydrides is a real scientific challenge. Indeed, despite all the numerical exhaustive studies using CSP

algorithms which have been performed, including the one of this PhD thesis, some experimental phases are not still explained by the predicted structures. It appears that in some pressure domain like near 150 GPa, the experimental X-ray spectrum [168, 170, 171, 185] is not retrieved when X-ray spectra are simulated on the structures predicted. The point explaining this failure is probably due to the fact the synthesized phases are more complex as those yet explored (i.e. they probably possess a unit cell larger than the ones accessible by CSP methods coupled to Ab-Initio calculations). At this point, all the necessary tools (numerical potentials, active learning process, analysis method for structure selection) to explore more complex binary yttrium superhydride structures are now developed and can be employed for new predictions. This section presents the results obtained following the application of these tools. In this section, YH_x structures which might explain the experimental spectrum are searched at a pressure of 147 GPa. Firstly, some of the experimental particularities of the experimental X-ray spectrum and the metrics considered to select the most interesting explored phases are discussed. Secondly details about how the random searching process is performed are described. Finally some new structure candidates are presented.

3.3.3.1 Experimental X-Ray Spectrum and Metric

Discrepancies between DFT Predictions and Experiment

According to the AIRSS study presented in Chap 2.3, the expected stoichiometry near 150 GPa might be YH_7 ($\text{YH}_7\text{-Cc}$). However, all the phases predicted nowadays can not explain the experimental X-ray spectrum obtained at this pressure (see FIG 3.3.18). In particular, it appears that the phase obtained experimentally is characterized by diffraction peaks positioned at low angles. If a X-ray spectrum wavelength of 0.3737 \AA is used, some peaks typically are positioned below 4 degrees, which corresponds to an interplanar distance above 2.7 \AA , (see FIG 3.3.18). None of the YH_x phases predicted by AIRSS owns diffraction peaks in this domain. As a result, a phase with a large number of chemical formula might explain this X-ray spectrum. Such complexity of the system is however hardly accessible by full DFT random searching. The use of a numerical potential coupled with bispectrum and X-ray spectrum analysis is thus suitable to reach more complex structure explorations.

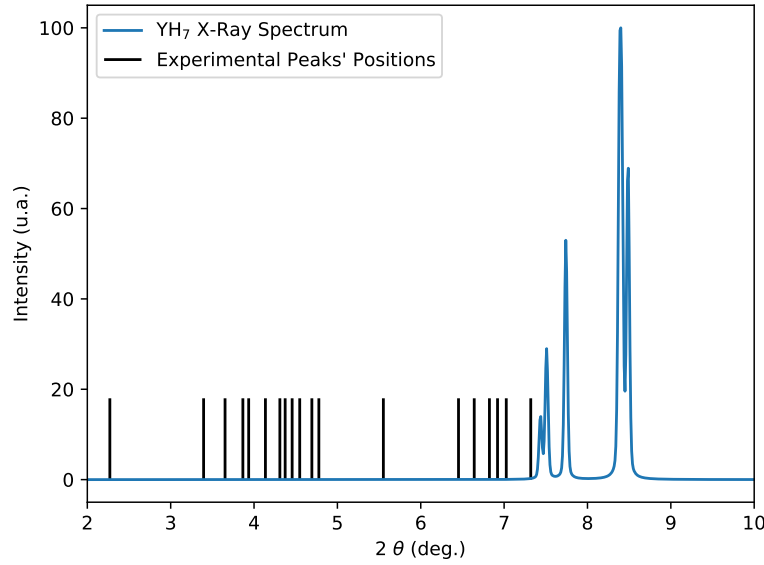


Figure 3.3.18: In black, some characteristic positions of the peaks of the phase synthesized near 150 GPa [185]. These positions are obtained by taking as input the experimental interplanar D-spacing data and a X-ray wavelength of 0.3738 \AA . The curve is the simulated X-ray spectrum of the $\text{YH}_7\text{-Cc}$ phase. As seen this X-ray spectrum does not match the experimental positions of the peaks.

Definitions of Metrics

In order to select the phase candidates, metrics based on the diffraction positions of the peaks are defined in a similar way as for the YH_6 explorations (see 3.3.2.2). However the positions of the peaks are weighted. Two different metrics are thus defined. A first Long Range (LR) one which focuses on the low diffraction angles (ie the long atomic distances). For this metric, the positions of the peaks below 4 degrees are overweighted in comparison to the others. This metric plays a crucial role. Indeed, as none of the phases predicted up to now have diffraction peaks in the domain of 2θ below 4 degrees, an important aim will be to explore and identify structures having some diffraction peaks located in this range. A complementary Short Range (SR) metric prioritizing shorter distances (i.e. the peaks associated with higher angles) is also introduced. These two metrics are written in the form:

$$D_X = \frac{1}{N} \sum_i \sum_j (1 - \omega(p_i) e^{-\frac{(p_i - q_j)^2}{\sigma}}) \quad (3.3.1)$$

in Eq. 3.3.1, p_i are the experimental positions of the peaks, q_j are those of a given studied structure. The function weight ω associated to each p_i position and the σ coefficient both differ

from one metric to the other. For the LR metric these quantities are :

$$\begin{cases} \sigma = 18 \\ \omega(\theta) = 1 \text{ if } \theta \leq 4 \\ \omega(\theta) = e^{-\frac{(4-\theta)^2}{\sigma}} \text{ if } \theta > 4 \end{cases} \quad (3.3.2)$$

for the SR metric, they write:

$$\begin{cases} \sigma = 10 \\ \omega(\theta) = \frac{e^{\frac{\theta}{\sigma}} - 1}{e^{\frac{4}{\sigma}} - 1} \text{ if } \theta \leq 4 \\ \omega(\theta) = 1 \text{ if } \theta > 4 \end{cases} \quad (3.3.3)$$

FIG 3.3.19 presents the LR and SR ponderations as well as the experimental positions of the peaks. The σ coefficients have been chosen with the following ideas:

- For the LR metric, the peaks above 4 degrees, and in particular those shortly above, need to have a small but non zero contribution in comparison with the peaks below 4 degree. Typically a weight in the range of 10 % is imposed for the first peak above 4 degrees. This is why such a compromise is made possible by fixing $\sigma = 18$.
- For the SR metric, as we are in general more interested in identifying structures with peaks positioned below 4 degrees, the contribution of the peaks located in this domain do not have to decrease too fast. Typically the peaks associated with the lowest 2θ need to contribute in the range of 10 %. It is why a $\sigma = 10$ is imposed.

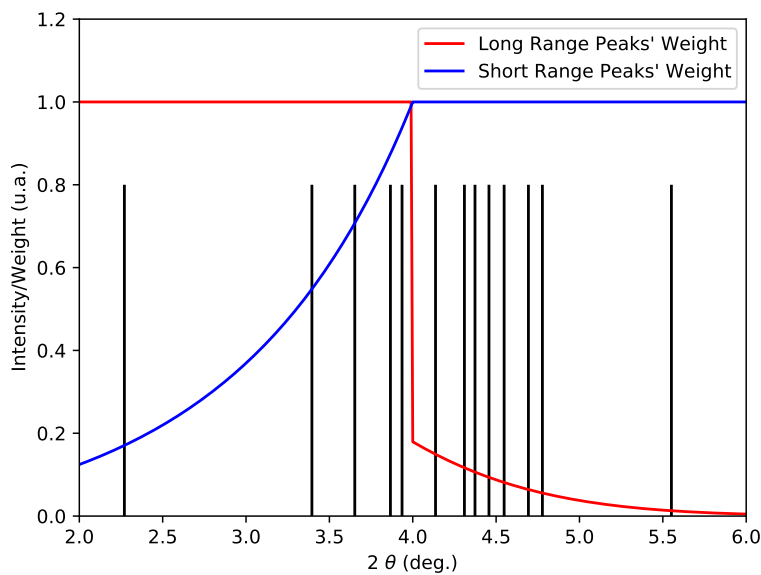


Figure 3.3.19: Characteristic positions of the peaks of the phase synthesized near 150 GPa and weighting associated respectively to the LR metric (blue) and the SR one (red).

3.3.3.2 Random Searching Implementation

Using these metrics, structures are selected from a full exploration with numerical potential following this scheme:

Organization of the Explorations

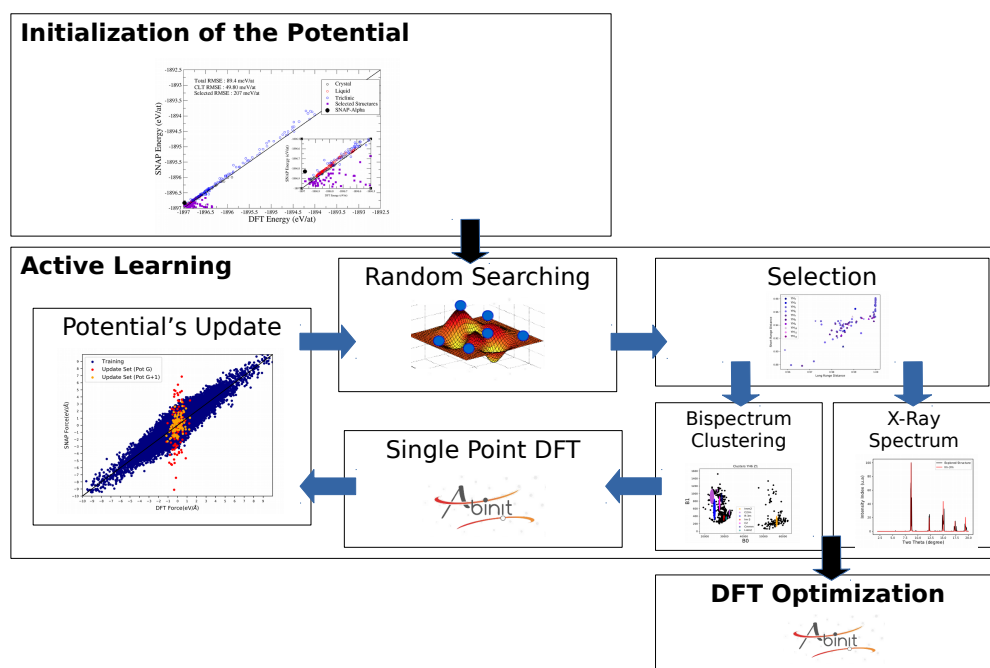


Figure 3.3.20: Active learning process applied to YH_x structural search.

The phases from YH_4 to YH_7 are explored using the active learning process employing mainly a SNAP type numerical potential. For complementary structural searches, Neural Network (NN) [194] type numerical potentials are also used. As presented in 3.1.1.3, this last numerical potential uses as descriptor symmetry functions and as a statistical model a neural network to enforce predictions. The neural network is made of two fully connected hidden layers of 40 neurons each. NN potentials can reach better accuracy for prediction of the physical quantities than with SNAP. However they require higher computation cost for training (around 5 times that of SNAP) as well as for the structural optimizations (around 3 times that of SNAP). For this reason NN potentials are used as final explorations after several cycles of active learning made with SNAP. Indeed, for each stoichiometry, structures from four to ten chemical formula units are searched. For a given stoichiometry and a given number of chemical formula units, 40 000 structures are explored in four steps. For the three first ones, the structure explorations are enforced using the

SNAP numerical potentials. From one of these steps to the other, a new potential is trained on an updated database including the informations of the new structures. At the fourth step, an exploration of 10000 structures is performed using NN potentials. All these explorations give rise to visit around 1.12 million of configurations.

How the Structures are Selected

In the active learning process, structures are selected for two purposes (see FIG 3.3.20):

- Some phases are considered as they have chemical environments worth taking into account to train a more accurate numerical potential on an updated database. To select them, the bispectrum analysis coupled with a clustering algorithm (GMM) is employed (see 3.3.2.1). To this aim, the 500 structures having the lowest predicted enthalpies are considered. The bispectrum of each yttrium atom is computed. These bispectrums are afterwards gathered in different clusters associated with similar chemical environments. Among the 500 structures, a set of 100 phases covering the maximum number of the clusters are selected. Single-point DFT computations are performed on them and the computed physical properties are added to the training database. This selection process is motivated to increase the diversity of the training database so as to improve the transferability of the numerical potential.
- Some phases are selected for subsequent DFT structural optimizations because they exhibit a simulated X-ray spectrum closer to the experimental one following the metrics defined above (see 3.3.3.1). After each structure exploration with a numerical potential, the 40 structures having the lowest LR metric values are taken. DFT optimizations are run and the one having the lowest enthalpy is finally considered as a structure candidate to explain the experimental X-ray spectrum.

In each of the explorations, these two selection modes are applied. In particular, all the final candidate phases and those obtained formerly by AIRSS are compared with each other using the LR and SR metrics (see FIG 3.3.21).

3.3.3.3 New Structures Obtained

The different phases identified by the successive searches with SNAP and NN numerical potentials as well as that found with AIRSS are compared in terms of the LR and SR metric values. They are computed after precise DFT optimizations at 147 GPa. FIG 3.3.21 presents how these structures are distributed.

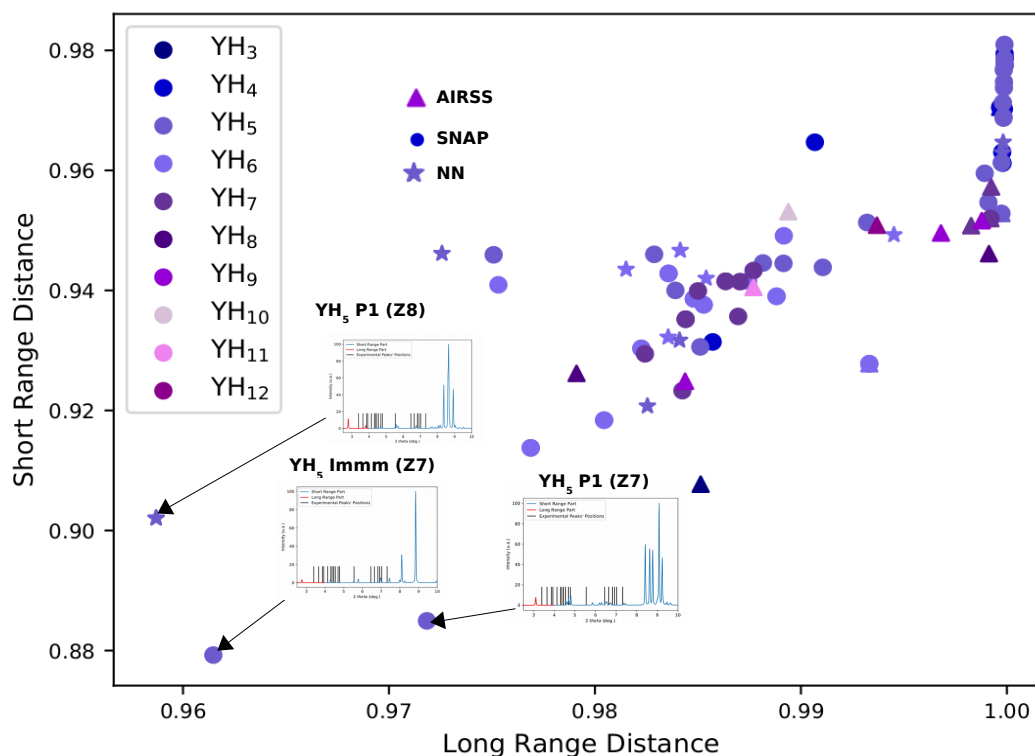
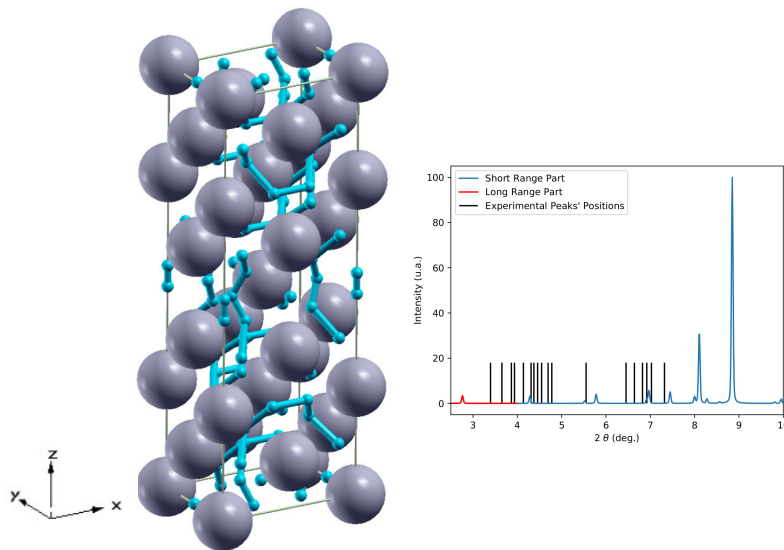


Figure 3.3.21: LR and SR metric value distribution of the phase candidates after DFT optimization. The triangles represent the phases obtained by AIRSS, the circles those get by SNAP numerical potentials and finally the stars correspond to the structures identified with NN numerical potentials.

A phase of YH_5 made of seven chemical formula units has the lowest LR and SR metric values. It was found by a SNAP numerical potential. This phase is tetragonal and has the Immm space group. Very close to it, another P1 phase of YH_5 made of eight chemical formula units was identified by a NN numerical potential.

Space group	Pressure	Lattice constants (\AA) and angles (degree)	Atom	Wyckoff position	x	y	z
<i>Immm</i>	147 GPa	a=5.295 b=3.828 c=15.560 $\alpha = \beta = \gamma = 90$	Y	4j	0.50000	0.00000	0.62940
			Y	4j	0.50000	0.00000	0.89065
			Y	2a	0.00000	0.00000	0.00000
			Y	4i	0.00000	0.00000	0.74572
			H	8m	0.86787	0.00000	0.37878
			H	16o	0.75115	0.78639	0.44251
			H	8m	0.36842	0.00000	0.24705
			H	16o	0.74232	0.23278	0.18422
			H	4i	0.00000	0.00000	0.52468
			H	8m	0.13067	0.00000	0.87213
			H	8n	0.32070	0.21189	0.00000
			H	2d	0.50000	0.00000	0.50000

Table 3.3.3: Wyckoff positions of the YH_5 *Immm* phaseFigure 3.3.22: View of the YH_5 *Immm* structure and the associated X-ray spectrum.

The Wyckoff positions of the YH_5 *Immm* phase are given at 147 GPa in Tab 3.3.3. This phase is unfortunately not on the convex hull, its formation enthalpy being 60 meV/at higher. In FIG 3.3.22 are shown a view of this structure as well as its simulated X-ray spectrum. This phase is characterized by yttrium slabs spaced by an interplanar distance in the range of 2.5 \AA . However its simulated X-ray spectrum still appears very different from the experimental one. Despite the fact it exhibits diffraction peaks at much lower angles than the phases identified by AIRSS.

Chapter 3.4

Perspective for Predictions of Ternary Systems

Contents

3.4.1 Predict the Best Alloys for Ternary Systems	177
3.4.1.1 The Database	177
3.4.1.2 The Model Considered	179
3.4.1.3 Predict Existence and Value of the Transition Pressure	181
3.4.2 Ternary Systems and Numerical Potentials	182
3.4.3 Summary of Part 3	183

Apart from the YFe_2H_x study, this PhD was essentially focused on binary superhydrides. This chapter presents some ways and perspectives to go further, in particular for investigations of ternary systems. As said, such compounds are hardly accessible by full ab-initio CSP algorithms due to the very large number of possibilities to test (i.e. the larger dimension of their configuration space). In fact, after choosing a set of two metallic elements, it is required to find the most stable alloy (i.e. the relative concentration of each element) and the crystal structures formed when it is combined with hydrogen. In the first section, an approach to isolate the best alloys to form superhydride systems is presented. The second section discusses briefly how the use of the numerical potentials might help in these investigations of ternary superhydrides.

3.4.1 Predict the Best Alloys for Ternary Systems

This first approach consists in taking into account all the published results about binary superhydride systems in a pressure range between 0 and 200 GPa. It aims at identifying the best descriptors or features enabling to represent an alloy candidate for the formation of superhydrides. In addition, correlations which might be extended to the ternary systems afterwards are searched. The main purpose here is to derive a machine learning model able to predict whether a given alloy can form a superhydride or not. If it does, this model is expected to give pressure estimations for which the system goes from a hydride to a superhydride. More generally this work has to be seen as an attempt to easily identify system candidates among several thousand ones to form superhydrides. It is worth noticing that predictions performed here have to be taken with caution due to the few information about ternary superhydrides. The approach is expected to be improved when more knowledge about these systems will be available. In this section, the database used is firstly described. Secondly, the model proposed is presented. Finally, the predictions of the model are discussed.

3.4.1.1 The Database

The initial database considered (DtB0) contains all the published results (mainly the highest hydrogen stoichiometries reached and pressure synthesis) about the binary (super)hydride systems (see appendix to see the complete database G). FIG 3.4.1 shows in the plan (atomic number, pressure) how the data are distributed.

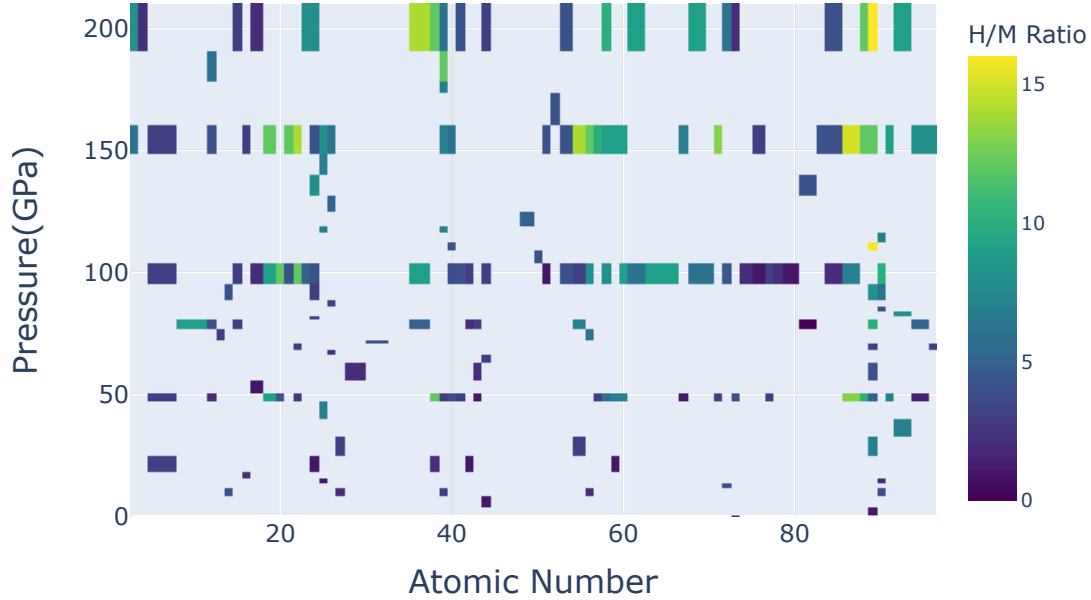


Figure 3.4.1: Distribution of the data points (corresponding to the binary systems) in the plan (atomic number, pressure).

The problem is to determine whether a given metallic alloy is able to form a superhydride at a given pressure. It is worth noticing that for a lot of systems forming superhydrides, the stoichiometry evolution with pressure is not smooth (considering the highest possible stable hydrogen stoichiometry). A jump is frequently encountered at a certain pressure (e.g. UH_3 to UH_7 at 36 GPa [38] or FeH_3 to FeH_5 at 130 GPa [12]). This pressure is called the transition pressure and is here arbitrarily defined as the one where the number of hydrogen atoms compared to metallic ones (H/M ratio) reaches five. Predicting the existence of this pressure (i.e. the system forms a superhydride or not) as well as its value is thus an important step towards superhydride predictions. For this purpose, a supervised learning approach is considered.

In order to train a supervised machine learning model, the database should contain labeled data. For this purpose label to each data in DtB0 is assigned as followed:

- If the system reaches a H/M ratio above five at a given pressure, the alloy is labeled as forming a superhydride and a value of the transition pressure is given.
- On the contrary if the H/M ratio remains strictly below five in the pressure range considered (ie 0-200 GPa), it is labeled as not favorable to form a superhydride.

3.4.1.2 The Model Considered

The machine learning model takes as input physical properties characterizing a given alloy (the features). The output is a label (0 or 1) indicating whether it forms a superhydride at a given pressure (1) or not (0). If it does, a transition pressure value is evaluated a second time. We have selected two features (among others) to describe the alloys and its capacity to form a superhydride:

- The electronegativity (Pauling scale) of the element forming the alloy.
- The crystal radius corresponding to an effective size of the chemical environment formed by the element in a crystal structure.

In the case of a ternary alloy, effective electronegativity and crystal radius are defined using the formula:

$$\chi_{eff} = x_1\chi_1 + x_2\chi_2 \quad (3.4.1)$$

$$R_{eff} = (x_1R_1^3 + x_2R_2^3)^{\frac{1}{3}} \quad (3.4.2)$$

In Eq. 3.4.1 and 3.4.2, x_i are the relative concentrations of each element involved in the alloy. The E_i and R_i are their electronegativities and crystal radii.

A first task is thus to decide whether a superhydride might be formed for a given alloy. In order to do so, a Support Vector Machine Model (SVM, see [212]) is employed. It is used to perform a binary classification in the plan (electronegativity, crystal radius). This task is enforced by determining domains characterized by the presence (or absence) of a transition pressure.

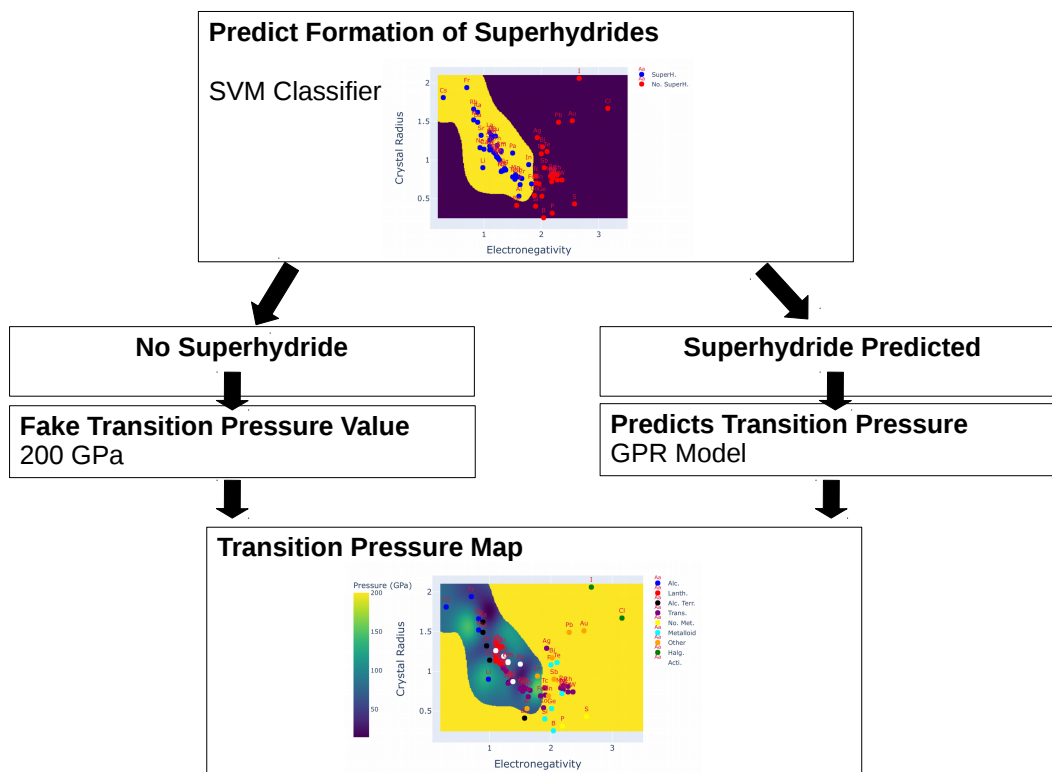


Figure 3.4.2: Work-flow for the determination of transition pressure.

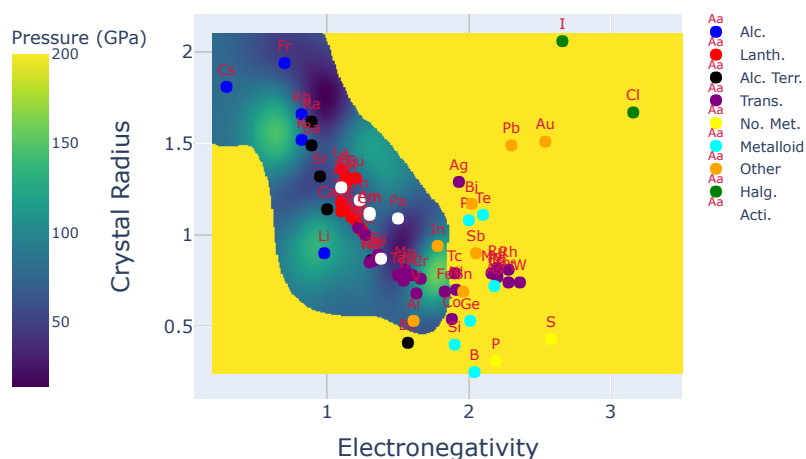


Figure 3.4.3: Distribution in the plan (electronegativity, crystal radius) of the pure alloys forming binary (super)hydrides. An interest region gathering all the systems forming superhydrides is identified.

The electronegativity and crystal radius of an alloy defines its location in this plan and its adherence to one domain indicates the existence of a transition. If such property is present, a Gaussian Process Regression (GPR, see [213]) model is considered to estimate a value of the pressure associated with this transition. As for the binary classification using SVM, both electronegativity and crystal radius are taken as features. If no transition pressure is expected, a fake value of 200 GPa is associated with the alloy. The figure 3.4.2 shows the work-flow considered for the estimation of the transition pressure.

3.4.1.3 Predict Existence and Value of the Transition Pressure

Application to Binary (Super)hydrides

The model combining SVM and GPR models is trained on all the informations coming from DtB0 (See appendix for details about the training process G). A region characterized by an electronegativity below 1.8 and crystal radius above 0.5 \AA in the plan electronegativity and crystal radius appears to be favorable for the formation of superhydrides. Indeed this region gathers all the known binary systems forming superhydrides. It can be interpreted by the fact that an alloy has to be made of elements having a sufficiently small electronegativity in comparison with that of hydrogen. This is required to enable the electronic transfer from the metallic to the hydrogen sublattices. At the same time the effective crystal radius needs to be large enough to free space for hydrogen atoms so that they can be included in the metallic sublattice. It is however required to be careful with this role of the crystal radius because apart from Be, no element has a smaller radius as 0.5 and an electronegativity below 1.9. Moreover, there exists a correlation between electronegativity and atomic radius [214] that could prevent from considering these two parameters as fully independent.

Application to Ternary (Super)hydrides

The model is now applied to some ternary alloys so as to evaluate the possibility to form ternary superhydride systems. The task equivalently consists of positioning the alloys in FIG 3.4.3 using the effective relations presented in Eq. 3.4.1 and 3.4.2. The FIG 3.4.4 shows the distribution of some ternary systems (the list of the ternary alloys considered is given in the appendix at the end of the superhydride database see G). In red are highlighted some recently investigated systems or probably worth studying further (yttrium-calcium (YCa) in [83], lithium-silicon (LiSi) in [215], carbon-sulfur (CS) in [216], yttrium-iron (YF_2) investigated in [177, 184] and in this PhD, lanthanum-yttrium (LaY)).

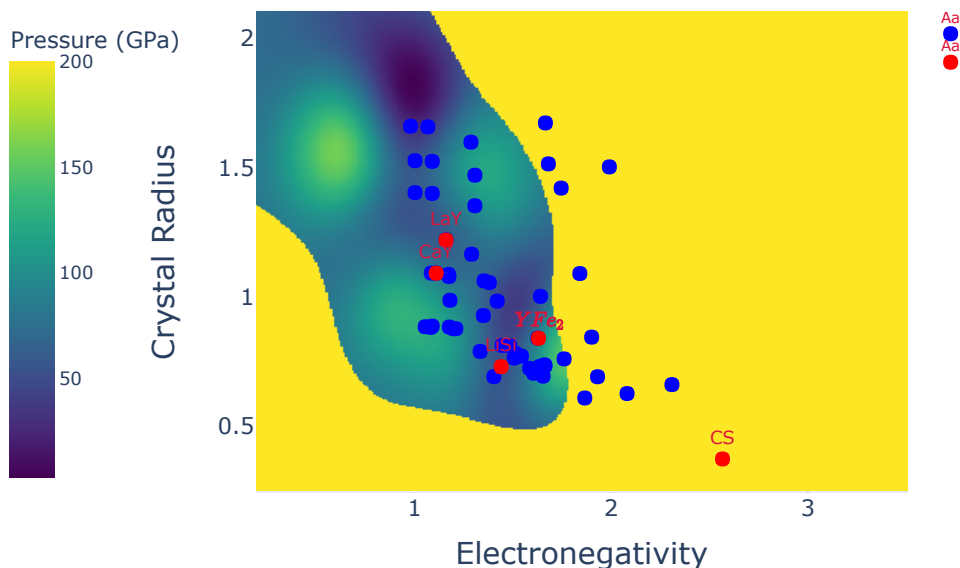


Figure 3.4.4: Distribution of some ternary alloys in the plan (electronegativity, crystal radius). The red points are systems which should be interesting to investigate experimentally so as to confirm the presented model. The blue points to other systems listed in the database of the appendix.

According to the model, the systems LaY, CaY, LiSi and YFe_2 are candidates to form superhydrides on the contrary to the CS one.

3.4.2 Ternary Systems and Numerical Potentials

The former section presented a first approach so as to identify ternary alloy candidates forming superhydrides. It might be used to select the best metallic element combinations. A second task is to determine the crystal structures formed. The use of full AIRSS will be limited due to the fact that three atom types are involved. In addition several relative concentrations of the metallic elements have to be assessed. As a result, the use of numerical potentials might be a key tool to perform an extensive search. For this purpose the work conducted on binary systems could thus be extended to ternary ones. It would require building a database with DFT calculations on ternary structures. Such an approach could help to improve the knowledge about more complex (super)hydrides which might be stabilized at lower pressure.

3.4.3 Summary of Part 3

This part presented developments and applications of a CSP algorithm combining Ab-Initio calculations and machine learning methods. The machine learning part is implemented thanks to the use of numerical potentials and an active learning process. This method has been first validated for a pure compound, tin. In a second time a validation has been achieved on a superhydride system (YH_6 at 200 GPa). This second verification step was also the opportunity to introduce analysis methods so as to select the most interesting phases among several thousands explored, enabling the structural analysis of large systems. A first method is based on the description of local atomic environments thanks to the bispectrum representation. In this strategy an unsupervised machine learning model (Gaussian Mixture) is applied to gather similar atomic environments. A second method is based on the simulated X-ray spectrum.

The new CSP algorithm and analysis methods have been applied to search new YH_x phases that might explain the experimental X-ray spectrum. A complex and symmetric phase of YH_5 having 42 atoms in its primitive cell has been identified and exhibits an X-ray spectrum as close as possible from experiments. Although this phase does not appear on the convex hull and does not fully explain the experiment. It confirms that the method is able to explore complex phases which are not accessible by CSP algorithms based only on DFT. This approach combined with additional experimental information will probably enable the identification of the complex yttrium phase appearing in the sample near 140 GPa.

Another challenge related to the search of complex phases of superhydrides concerns ternary systems. Using all the results known about the binary superhydride systems, a first machine learning based model is proposed. It aims at pointing out the best metallic atom combinations which might form superhydrides. We show that the electronegativity and crystal radius can be used as ingredients for describing the formation of superhydrides. A domain appears in the low electronegativities and high crystal radii that seems to be favorable for the formation of superhydrides. This approach is then considered as a possible preliminary step before applying a CSP algorithm to determine the most stable phases. Such a CSP algorithm might use the machine learning methods as developed in this thesis.

GENERAL CONCLUSION

This PhD is a contribution to the investigation of the newly discovered superhydrides. It was also the opportunity to develop methods combining Ab-Initio calculations and machine learning methods for more complex predictions of materials.

First of all, studies of some transition metal hydrides and superhydrides have been studied. Using the AIRSS algorithm copper, manganese, yttrium and iron-yttrium hydrides and superhydrides were investigated. From these searches, two main results can be reminded:

- A rich hydrogen superhydride MnH_7 is expected to form at a rather low pressure in comparison with other systems having similar H/M ratio. Indeed, the predicted pressure synthesis is 40 GPa instead of more than 100 GPa for the majority of the superhydrides with the same hydrogen stoichiometry.
- A ternary compound $\text{YFe}_2\text{H}_{10}$ is expected near 25 GPa and might have a negative formation enthalpy at ambient pressure, indicating that it might remain stable at room pressure. If this is verified, the reachable hydrogen storage density would be in the range of $200 \text{ KgH}_2/\text{m}^3$ (around 2.8 times the liquid density of hydrogen).

These results should motivate experimental investigations. They could first check these predictions. In a second time strategies to stabilize these systems under low pressure could be investigated. FIG 3.4.5 presents in the plan (volumetric hydrogen density, gravimetric density) how the different studied phases having a negative formation enthalpy at 0 GPa are distributed. They are worth being considered as they are those which might remain stable at ambient pressure after being formed at higher pressure. As shown, none of the copper hydrides have a negative formation enthalpy as well as the manganese ones having a H/M ratio strictly above one. It is worth highlighting the YFe_2H_x system is the one for which the highest hydrogen concentrations might be obtained at ambient pressure. Referring to the conditions formulated in 1.1.2, it can be seen that the phases YH_{10} , YFe_2H_8 , YFe_2H_9 , $\text{YFe}_2\text{H}_{10}$ are those for which the conditions are the best fulfilled. In particular YH_{10} is close to the gravimetric density of 11 % which is targetted for mobile applications. Despite being far from this criterion, the YFe_2H_x phases have very high hydrogen density in the range of $200 \text{ KgH}_2/\text{m}^3$. They have an enthalpy formation which is negative without being too large in absolute value (and thus might easily release their hydrogen for fuel cell supply). This system might be interesting to consider for non mobile hydrogen storage applications.

It is also worth highlighting the work conducted to improve the CSP algorithms. The main purpose was to reach predictions of complex materials (having more atoms in the primitive cell or more distinct types than what is accessible by DFT methods). At the beginning of this PhD, the

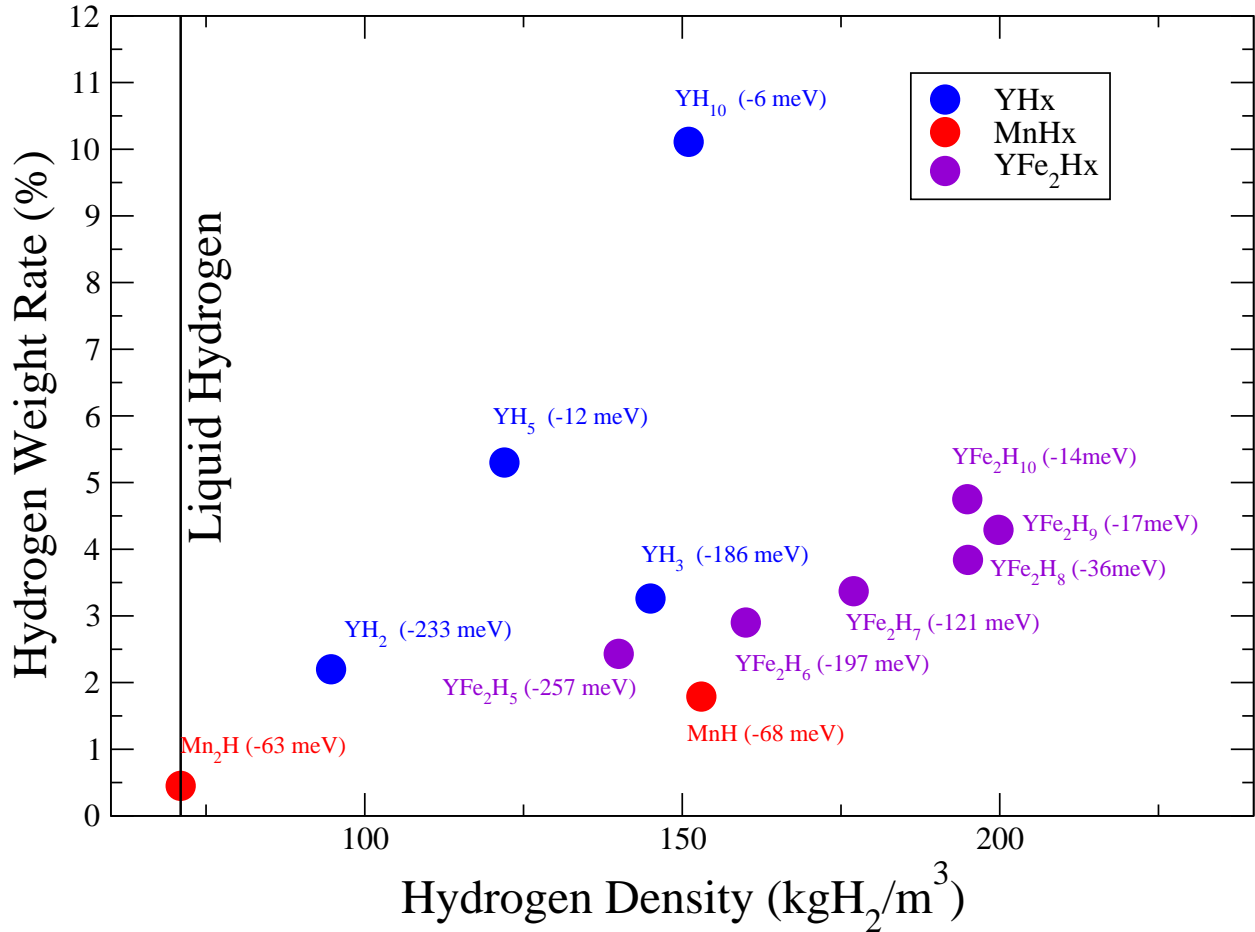


Figure 3.4.5: Distribution in the plan (volumetric hydrogen density, gravimetric density) of the different identified phases having a negative formation enthalpy at 0 GPa.

AIRSS algorithm enabled the exploration of configurations having a fixed number of atoms below ~ 20 -30 and of two different types (a metal and hydrogen) in the primitive cell, the generation and structural optimization of around 1000 configurations for a given pressure and number of formula units. It required a computation time in the range of 1 million CPU hours. The development of machine learning methods combined with the CSP algorithms enabled the explorations of more complex phases having in the range of 100 atoms in the primitive cell. The application of this method to the YH_x system reveals that more than 1 millions configurations can be explored for a computation time in the range of 10 000 CPU hours (mainly due to the DFT calculations for both training database computations and structural optimizations of the few phase candidates). Thanks to this approach, complex YH_x phases which can not be reached by AIRSS have been identified and are in better agreement (despite not completely) with the experimental results.

The use of machine learning made possible the development of CSP methods reaching a much more important complexity than the one accessible with CSP methods coupled to DFT. This approach will be very important to study ternary or even more complex hydrides and superhydrides so as to bring their promising properties back to ambient conditions. Beyond the scope of superhydrides, this progress might give access to the prediction and the study of more complex materials offering new applications and opportunities in a lot of domains.

Appendices

Résumé

Alliages caractérisés par la présence d'hydrogène dans leurs structures cristallines, les hydrures sont des matériaux prometteurs pour le stockage de l'hydrogène sous forme solide, et donc pour le stockage de l'énergie sous la forme d'hydrogène. Il y a quelques années le fort potentiel de cette classe de matériaux a été renforcé suite à la découverte de la stabilité, sous très haute pression, d'hydrures particuliers appelés les "superhydrures". Ces nouveaux composés se caractérisent par d'importantes concentrations en hydrogène supérieures à ce qu'autorisent les règles usuelles de la chimie à pression ambiante (par exemple FeH_5 [12] ou LiH_6 [17, 18]). Des concentrations dépassant d'un facteur trois la densité de l'hydrogène liquide à pression ambiante et d'un facteur quatre celles aujourd'hui atteintes dans l'industrie pourraient être franchies. Ce qui marquerait une rupture importante dans les technologies de stockage d'hydrogène, si de tels matériaux pouvaient être stabilisés à pression ambiante. La majorité des superhydrures étudiés et synthétisés sont des composés dits "binaires" où un élément métallique est combiné avec de l'hydrogène. Ces premières investigations ont permis de faire émerger différentes tendances. En particulier, il a été mis en évidence que les propriétés de ces superhydrures varient selon la nature du métal considéré dans l'alliage. Les propriétés les plus remarquables restent celles obtenues avec les terres rares ou les éléments de transition de la colonne IIIB de la classification périodique (par exemple, avec les superhydrures de lanthane [13] LaH_{10} , de cerium [14] CeH_9 , d'yttrium [99] YH_{10} ...). Les plus hautes stoechiométries en hydrogène ont en effet été atteintes pour cette catégorie de systèmes, ainsi qu'une supraconductivité à haute température (par exemple avec LaH_{10} , qui possède une température critique de supraconductivité mesurée à 215 K vers 150 GPa [13, 32]). Cette dernière propriété provient du fait que le sous-réseau d'hydrogène possède des propriétés analogues à celles de l'hydrogène métallique prédit comme étant supraconducteur à température ambiante [15]. Généralement moins favorables pour la supraconductivité à haute température (à l'exception de ceux de la colonne IIIB), certains métaux de transition (fer, manganèse, chrome ...) sont tout aussi prometteurs pour le stockage d'hydrogène et sont également plus facilement exploitables (car plus abondants ou plus facilement accessibles dans la croûte terrestre) que les terres rares. Ils pourraient ainsi être plus adaptés pour des applications industrielles. Le principal frein à l'emploi des superhydrures pour des applications reste cependant les

pressions considérables à imposer pour les synthétiser. Ces pressions sont en effet aujourd'hui de l'ordre du million d'atmosphère. Pour les abaisser, il apparaît nécessaire d'approfondir les connaissances sur ces nouveaux matériaux. D'autre part, en plus des pressions extrêmes à atteindre s'ajoutent des difficultés expérimentales pour évaluer les quantités d'hydrogène présentes dans les alliages ainsi que les positions des atomes d'hydrogène (du fait d'une faible interaction entre rayons X et hydrogène). Pour un alliage donné, binaire ou ternaire, le nombre de stœchiométries et structures possibles étant grand (et d'autant plus si plusieurs éléments métalliques sont combinés entre eux), il est essentiel d'employer la simulation numérique et des algorithmes de prédiction de structures cristallines pour guider les expérimentateurs vers les meilleurs candidats. Cette thèse vise ainsi à apporter une contribution à l'étude d'hydrures et de superhydrures de métaux de transition, par la mise en oeuvre d'études numériques. Elle a également pour objectif de proposer de nouvelles méthodes numériques pour rechercher et prédire des superhydrures plus complexes qui pourraient peut-être être ramenés à des pressions plus faibles sous une forme métastable.

Pour la prédiction de nouveaux superhydrures de métaux de transition, l'algorithme de Recherche Aléatoire de Structures couplée au Calcul Ab-Initio (AIRSS) a été mis en oeuvre [87]. Cette méthode s'appuie sur la théorie fonctionnelle de la densité (DFT) et permet de trouver, pour un système donné, dans des conditions de pression fixées, la structure cristalline la plus stable. L'algorithme est basé sur une exploration qui consiste à générer initialement des configurations selon différentes règles pour balayer au mieux l'ensemble des structures ayant une signification physique. Ces structures subissent ensuite des "optimisations structurales" qui consistent à modifier les positions atomiques et la géométrie de la maille, de façon à minimiser leur énergie (ou leur enthalpie lorsque l'optimisation est faite à pression fixée). Ces optimisations s'appuient sur le calcul DFT pour évaluer les différentes grandeurs physiques (énergie, forces et contraintes). L'algorithme AIRSS est un type particulier d'algorithme de prédiction de structures cristallines, d'autres types existent comme les algorithmes évolutionnaire [97] ou génétique par exemple [96].

La méthode AIRSS a été appliquée dans ce travail sur les hydrures et superhydrures de cuivre (CuH_x), manganèse (MnH_x), yttrium (YH_x) ainsi que sur un système "ternaire" (où l'alliage est composé d'atomes métalliques de deux types différents) yttrium-fer (YFe_2H_x). Tous les calculs Ab-Initio ont été réalisés avec le code *abinit* [140]. Contrairement au cuivre pour lequel aucun nouveau superhydrure n'a été prédit jusqu'à environ 200 GPa, le manganèse et l'yttrium s'avèrent être des systèmes particulièrement riches. Concernant le manganèse, de nouveaux composés très riches en hydrogène comme MnH_3 , MnH_4 , MnH_7 et MnH_8 ont été prédits. Ces prédictions incluent la confirmation d'une stabilité énergétique par l'évaluation des enthalpies de formation ainsi que dynamique par le calcul des modes de phonons. En particulier la stabilisation vers 150 GPa de MnH_4 et MnH_8 a été obtenue en prenant en compte la contribution des énergies de point zéro associées à la quantification des modes de vibration. Egalement, le caractère isolant ou faiblement métallique de certains de ces superhydrures a été mis en évidence. Cette propriété a été confirmée sur MnH_7 grâce à l'implantation, au cours de ce travail de thèse, d'une

fonctionnelle d'échange-corrélation de type "Meta-GGA" dans le code *abinit*. Pour l'yttrium, une nouvelle phase non prédite par les études précédentes [98, 169, 99, 31, 168], de stœchiométrie YH_7 avec un large domaine de stabilité (entre 100 et 180 GPa) a pu être identifiée. D'autre part, l'inclusion d'atomes de fer pour former les composés ternaires YFe_2H_x donne lieu à la formation d'hydrures à forte concentration en hydrogène qui se formeraient à des pressions modérées (vers 25 GPa) et pourraient rester stables à pression ambiante. Ce qui pourrait ouvrir la voie vers la synthèse de systèmes prometteurs pour le stockage d'hydrogène à haute densité. Les composés d'hydrures et superhydrures binaires d'yttrium ainsi que l'étude des hydrures ternaires yttrium-fer montrent les limites des algorithmes de prédiction de structures cristallines de matériaux tels que la méthode AIRSS, qui utilise essentiellement du calcul DFT. Ces méthodes nécessitent des coûts importants en temps de calcul notamment du fait des optimisations structurales employant la DFT. Mener une étude complète d'un superhydrure binaire nécessite de l'ordre de dix millions d'heures CPU et 99 % de ce temps est alloué aux optimisations structurales par la DFT. Ainsi, les investigations pour YFe_2H_x ont été facilitées par le fait que le rapport Y:Fe était connu ainsi que la géométrie du sous réseau métallique initial. Sans ces informations, le coût en temps de calcul nécessaire aurait été trop important (au-delà de la centaine de millions d'heures CPU probablement). L'étude exhaustive d'un système ternaire (avec prédictions des structures cristallines pour les principales stœchiométries possibles) sans information expérimentale est donc difficilement faisable en un temps raisonnable. Concernant les hydrures et superhydrures binaires d'yttrium, les phases synthétisées à certaines pressions ne sont expliquées par aucune structure prédite par la simulation [168, 170, 171, 185]. Il est probable qu'elles doivent être décrites par une maille primitive contenant un grand nombre d'atomes et ont ainsi une complexité qui n'est pas encore accessible par la méthode AIRSS, qui fonctionne bien pour des mailles contenant jusqu'à typiquement 20-30 atomes. Pour progresser dans la prédiction de matériaux plus complexes, des approches combinant le calcul DFT et des techniques de "machine learning" sont développées dans cette thèse. L'objectif principal de ces développements est d'être capable de proposer des structures cristallines pour les phases de superhydrures binaires d'yttrium plus complexes, et qui soient susceptibles d'expliquer mieux l'expérience aux alentours de 140 GPa.

Pour permettre l'exploration à la fois d'un plus grand nombre de structures (typiquement pour passer de quelques milliers à quelques millions) qui contiennent également davantage d'atomes dans leur maille primitive (pour passer de mailles ayant au maximum 30 atomes à des mailles en ayant environ 100), la voie des potentiels numériques a été investiguée. Les potentiels numériques sont des modèles de machine learning qui permettent de prédire avec un faible coût de calcul les propriétés physiques des matériaux. Pour ce faire ils ont besoin (i) d'une base de données d'entraînement regroupant différentes configurations atomiques et leurs propriétés calculées par la DFT (énergie, forces, contraintes), (ii) d'un descripteur qui est un modèle mathématique servant à traduire les environnements atomiques en vecteurs, et (iii) d'un modèle statistique qui prend en entrée les vecteurs issus du descripteur et donne en sortie des prédictions sur les propriétés physiques (essentiellement énergie, forces et contraintes). Un type de poten-

tiel principalement utilisé dans cette thèse est le potentiel "Spectral Neighbour Analysis Potential" (SNAP) qui combine comme descripteur le "bispectre", qui consiste à faire une décomposition des densités atomiques autour de chaque atome sur une base d'harmoniques sphériques. Le modèle statistique est une régression linéaire généralisée [202, 201]. En complément un potentiel combinant des fonctions de symétrie et des réseaux de neurones a été également utilisé [194, 195]. La combinaison des calculs DFT (pour le calcul précis de la base de données) avec l'utilisation de potentiels numériques pour la recherche de structures est mise en oeuvre par une approche de type "apprentissage actif". Suivant cette démarche, des recherches de structures successives sont menées avec un faible coût de calcul (quelques heures CPU) par des potentiels numériques qui apprennent sur le système étudié en enrichissant leurs bases de données d'entraînement avec les nouvelles structures trouvées lors des explorations précédentes. Une première validation de cette méthode a été réalisée sur l'étain. Un premier objectif, qui a été validé, était de retrouver à 0 GPa la phase α . La difficulté de cette prédiction résidait dans le fait que cette phase est semi-conductrice et possède des liaisons covalentes entre ses atomes contrairement aux autres phases connues et de celles constituant initialement la base de données d'entraînement du potentiel (qui sont métalliques et donc très différentes de la phase α). Sa prédiction prouve la bonne transférabilité du potentiel numérique obtenu par l'approche de type apprentissage actif car une phase ayant des propriétés électroniques très différentes de celles initialement apprises a été retrouvée. Dans une seconde étape la structure cristalline stable sous haute pression (100 GPa), la phase γ , a été correctement prédite. Suite à ce premier test sur un matériau constitué d'atomes d'un seul élément chimique, une autre validation consistait à mener des prédictions sur une phase d'un superhydruure connu. Ainsi les structures de superhydruures binaires d'yttrium de stoechiométrie YH_6 à 200 GPa ont été explorées. Le but était de retrouver la phase de groupe d'espace $Im-3m$ déjà prédite par les méthodes de recherche de structures couplées à la DFT, et synthétisée [99, 168, 170]. Cette structure a été correctement retrouvée en utilisant les potentiels numériques. De plus une nouvelle phase légèrement plus instable (de quelques meV/at) a pu être identifiée. Cette étude a d'autre part démontré la nécessité d'améliorer les méthodes de sélection des structures explorées. En effet, bien que les potentiels numériques permettent de mener des recherches sur plusieurs millions de configurations contenant plus d'atomes dans leur maille cristalline (que dans le cas de la méthode AIRSS) avec un coût de calcul très faible (quelques dizaines d'heures CPU sans tenir compte du coût des calculs DFT des bases d'entraînement). Les optimisations structurales restent cependant moins précises qu'avec la DFT (erreur sur les contraintes et les forces par rapport à la DFT). Ceci oblige à développer de nouvelles approches plus robustes vis-à-vis des imprécisions numériques des optimisations structurales. Dans ce but, deux méthodes ont été développées et utilisées dans cette thèse. Une première s'appuie sur la description des environnements atomiques par le bispectre (le descripteur employé par le potentiel SNAP) et un algorithme de "clustering", les mélanges gaussiens de densité. La seconde méthode consiste à simuler les spectres de diffraction X des structures explorées et de les comparer avec des spectres de référence (comme ceux provenant de l'expérience). La comparaison est effectuée sur la base d'une métrique évaluant une distance entre les spectres. Ces deux

méthodes ont pu être testées avec succès pour retrouver la phase de groupe d'espace Im-3m du superhydrures YH_6 simulée dans une supercellule contenant 42 atomes (6 formules unités) à l'aide du potentiel numérique.

Suite à ces différentes validations, la méthode d'apprentissage actif combinant potentiels numériques, calculs DFT et les deux nouvelles techniques de sélection des structures ont été mises en oeuvre pour rechercher de nouvelles phases de superhydrures d'yttrium. 1.12 millions de structures allant de la stœchiométrie YH_4 à YH_7 et ayant de 20 à 80 atomes dans la maille élémentaire ont été explorées pour un coût de calcul de l'ordre de 10 000 heures CPU. La même étude avec l'algorithme AIRSS aurait nécessité de l'ordre de 10^9 heures CPU. De plus de nouvelles phases inaccessibles en terme de nombre d'atomes par la méthode AIRSS et expliquant mieux (bien qu'encore imparfaitement) le spectre expérimental ont été trouvées. En particulier, une structure de symétrie tétragonale avec 42 atomes a pu être identifiée.

Ce travail de thèse se termine par quelques perspectives afin d'aller vers l'identification de nouveaux superhydrures ternaires. Notamment une étude visant à identifier les combinaisons d'éléments les plus intéressantes pour former des alliages de superhydrures ternaires est proposée. Pour ce faire un modèle prospectif de machine learning s'appuyant sur une base de données d'hydrures et de superhydrures binaires et ternaires déjà connus est présenté afin de donner des premières tendances pour cibler les éléments et combinaisons d'éléments les plus prometteurs.

Appendix **B**

Convex Hulls and Formation Enthalpy Definition

The function $C(AHx)$ to determine the second components of the points has to be defined properly. Indeed, if it appears AHx is not on the convex hull and between two stable compounds AHx_1 and AHx_2 , one have the inequality (H being the formation enthalpy of a compound):

$$H\left(\frac{x-y}{1+(x-y)}\right) > \alpha H\left(\frac{x_1-y}{1+(x_1-y)}\right) + \beta H\left(\frac{x_2-y}{1+(x_2-y)}\right) \quad (B.1)$$

with:

$$\alpha \frac{x_1-y}{1+(x_1-y)} + \beta \frac{x_2-y}{1+(x_2-y)} = \frac{x-y}{1+(x-y)} \quad (B.2)$$

The inequality B.1 has to imply moreover the instability condition:

$$\mathcal{H}(AHx) > \frac{x-x_2}{x_1-x_2} \mathcal{H}(AHx_1) + \frac{x_1-x}{x_1-x_2} \mathcal{H}(AHx_2) \quad (B.3)$$

Expressing the fact the compound AH_x is less stable than the mixture made of AH_{x_1} and AH_{x_2} .

It comes from the fact the instability of AH_x between AH_{x_1} and AH_{x_2} fosters the chemical reaction:

$$AHx = \frac{x-x_2}{x_1-x_2} AHx_1 + \frac{x_1-x}{x_1-x_2} AHx_2$$

We can moreover from this equation deduce:

$$x = x_1 \frac{x-x_2}{x_1-x_2} + x_2 \frac{x_1-x}{x_1-x_2} \quad (B.4)$$

From Eqs B.2 and B.4 it can be obtained:

$$\frac{x-y}{1+(x-y)} = \left[\frac{1+(x_1-y)}{1+(x-y)} \frac{x-x_2}{x_1-x_2} \right] \frac{x_1-y}{1+(x_1-y)} + \left[\frac{(1+(x_2-y))}{1+(x-y)} \frac{x_1-x}{x_1-x_2} \right] \frac{x_2-y}{1+(x_2-y)}$$

one can thus identify:

$$\alpha = \frac{1+(x_1-y)}{1+(x-y)} \frac{x-x_2}{x_1-x_2}$$

and

$$\beta = \frac{(1+(x_2-y))}{1+(x-y)} \frac{x_1-x}{x_1-x_2}$$

Given the expression of α and β and the instability condition of Eq B.3, one can deduce the choice of $C(AHx) = 1+(x-y)$ enables the conservation of the condition of Eq B.3. Indeed, if one uses the definition of the formations enthalpy (H) and the obtained expressions of α and β it implies:

$$H\left(\frac{x}{1+(x-y)}\right) > \alpha H\left(\frac{x_1-y}{1+(x_1-y)}\right) + \beta H\left(\frac{x_2-y}{1+(x_2-y)}\right)$$

$$\begin{aligned}
&=> \frac{\mathcal{H}(AHx) - \mathcal{H}(AHy) - (x - y)\mathcal{H}(H)}{1 + (x - y)} > \frac{1 + (x_1 - y)}{1 + (x - y)} \frac{x - x_2}{x_1 - x_2} \frac{\mathcal{H}(AHx_1) - \mathcal{H}(AHy) - (x_1 - y)\mathcal{H}(H)}{1 + (x_1 - y)} \\
&\quad + \frac{(1 + (x_2 - y))}{1 + (x - y)} \frac{x_1 - x}{x_1 - x_2} \frac{\mathcal{H}(AHx_2) - \mathcal{H}(AHy) - (x_2 - y)\mathcal{H}(H)}{1 + (x_2 - y)} \\
&=> \mathcal{H}(AHx) - \mathcal{H}(AHy) - (x - y)\mathcal{H}(H) > (1 + (x_1 - y)) \frac{x - x_2}{x_1 - x_2} \frac{\mathcal{H}(AHx_1) - \mathcal{H}(AHy) - (x_1 - y)\mathcal{H}(H)}{1 + (x_1 - y)} \\
&\quad + (1 + (x_2 - y)) \frac{x_1 - x}{x_1 - x_2} \frac{\mathcal{H}(AHx_2) - \mathcal{H}(AHy) - (x_2 - y)\mathcal{H}(H)}{1 + (x_2 - y)}
\end{aligned}$$

Thus:

$$\begin{aligned}
&\mathcal{H}(AHx) - \mathcal{H}(AHy) - (x - y)\mathcal{H}(H) > \\
&\frac{x - x_2}{x_1 - x_2} (\mathcal{H}(AHx_1) - \mathcal{H}(AHy) - (x_1 - y)\mathcal{H}(H)) + \\
&\frac{x_1 - x}{x_1 - x_2} (\mathcal{H}(AHx_2) - \mathcal{H}(AHy) - (x_2 - y)\mathcal{H}(H))
\end{aligned}$$

gives clearly:

$$\mathcal{H}(AHx) > \frac{x - x_2}{x_1 - x_2} \mathcal{H}(AHx_1) + \frac{x_1 - x}{x_1 - x_2} \mathcal{H}(AHx_2)$$

In the following studies, the definition of:

$$H\left(\frac{x - y}{1 + (x - y)}\right) = \frac{\mathcal{H}(AHx) - \mathcal{H}(AHy) - (x - y)\mathcal{H}(H)}{1 + (x - y)}$$

will be employed.

Appendix C

On-site Meta-GGA Contribution to Total Energy in the PAW Approach

We remind first the following quantities:

- Φ_i^{1R} and $\tilde{\Phi}_i^{1R}$ such that the elements of the on-site and on-site pseudo-bases are expressed as:

$$\begin{aligned}\phi_i^{1R}(\mathbf{r}) &= \frac{\Phi_i^{1R}(r)}{r} S_{l_i, m_i}(\theta, \phi) \\ \tilde{\phi}_i^{1R}(\mathbf{r}) &= \frac{\tilde{\Phi}_i^{1R}(r)}{r} S_{l_i, m_i}(\theta, \phi)\end{aligned}\tag{C.1}$$

- $\bar{\Phi}_i^{1R}$ and $\tilde{\bar{\Phi}}_i^{1R}$ such that:

$$\begin{aligned}\bar{\Phi}_i^{1R}(r) &= \frac{d\Phi_i^{1R}(r)}{dr} - \frac{\Phi_i^{1R}(r)}{r} \\ \tilde{\bar{\Phi}}_i^{1R}(r) &= \frac{d\tilde{\Phi}_i^{1R}(r)}{dr} - \frac{\tilde{\Phi}_i^{1R}(r)}{r}\end{aligned}\tag{C.2}$$

- μ_{xclm} is the (l, m) component of μ_{xc} in the spherical harmonic decomposition such that

$$\mu_{xclm}(r) = \int \mu_{xc}(\mathbf{r}) S_{lm}(\theta, \phi) d\Omega\tag{C.3}$$

- the two type of integral terms $G_{lm, l_i m_i, l_j m_j}$ and $H_{lm, l_i m_i, l_j m_j}$ such that:

$$\begin{aligned}G_{lm, l_i m_i, l_j m_j} &= \int S_{lm} S_{l_i m_i} S_{l_j m_j} d\Omega \\ H_{lm, l_i m_i, l_j m_j} &= \int S_{lm} [\nabla S_{l_i m_i} \cdot \nabla S_{l_j m_j}] d\Omega\end{aligned}\tag{C.4}$$

The first quantity to derive is the on-site kinetic energy density τ^{1R} , for a given atom R. If one remembers the general expression of the kinetic energy density is:

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_n f_n |\nabla \Psi_n|^2\tag{C.5}$$

τ^{1R} can be expressed in a similar way as n^{1R} (see Eq. 1.2.32):

$$\tau^{1R}(\mathbf{r}) = \frac{1}{2} \sum_{ij} \rho_{ij}^{1R} \nabla \phi_i^{1R} \nabla \phi_j^{1R}\tag{C.6}$$

Usually, this quantity is expressed by its moments $\tau_{lm}^{1R}(r)$ such that:

$$\tau^{1R}(\mathbf{r}) = \frac{1}{2} \sum_{lm} \tau_{lm}^{1R}(r) S_{lm}(\theta, \phi)\tag{C.7}$$

and

$$\tau_{lm}^{1R}(r) = \int \tau^{1R}(\mathbf{r}) S_{lm}(\theta, \phi) d\Omega\tag{C.8}$$

The gradient of the on-site basis has to be expressed in terms of spherical harmonics. It gives:

$$\nabla \phi_i^{1R}(\mathbf{r}) = \frac{\bar{\Phi}_i^{1R}(r)}{r} S_{l_i m_i}(\theta, \phi) \mathbf{u}_r + \frac{\Phi_i^{1R}(r)}{r} \nabla S_{l_i m_i}(\theta, \phi) \quad (\text{C.9})$$

Injecting Eqs C.6 and C.9 in C.8 gives (the θ, ϕ dependency of the S_{lm} are not written for more simplicity):

$$\tau_{lm}^{1R}(r) = \frac{1}{2} \int \sum_{ij} \rho_{ij}^{1R} \left[\frac{\bar{\Phi}_i^{1R}(r)}{r} S_{l_i m_i} \mathbf{u}_r + \frac{\Phi_i^{1R}(r)}{r} \nabla S_{l_i m_i} \right] \cdot \left[\frac{\bar{\Phi}_j^{1R}(r)}{r} S_{l_j m_j} \mathbf{u}_r + \frac{\Phi_j^{1R}(r)}{r} \nabla S_{l_j m_j} \right] S_{lm} d\Omega \quad (\text{C.10})$$

It's worth remarking the gradient of a spherical harmonic is orthogonal to the radial direction: $\nabla S_{lm}(\theta, \phi) \cdot \mathbf{u}_r = 0$. It can be deduced:

$$\tau_{lm}^{1R}(r) = \frac{1}{2} \sum_{ij} \rho_{ij} \left[\frac{\bar{\Phi}_i^{1R}(r) \bar{\Phi}_j^{1R}(r)}{r^2} \int S_{l_i m_i} S_{l_j m_j} S_{lm} d\Omega + \frac{\Phi_i^{1R}(r) \Phi_j^{1R}(r)}{r^2} \int S_{lm} [\nabla S_{l_i m_i} \cdot \nabla S_{l_j m_j}] d\Omega \right] \quad (\text{C.11})$$

The final τ_{lm}^{1R} expression is:

$$\tau_{lm}^{1R}(r) = \frac{1}{2} \sum_{ij} \frac{\rho_{ij}^{1R}}{r^2} [\bar{\Phi}_i^{1R}(r) \Phi_j^{1R}(r) G_{lm, l_i m_i, l_j m_j} + \Phi_i^{1R}(r) \bar{\Phi}_j^{1R}(r) H_{lm, l_i m_i, l_j m_j}] \quad (\text{C.12})$$

As shown in Eq 1.2.52, Meta-GGA on-site contribution coming from the D_{ij}^{1R} operators introduces terms in the form:

$$\langle \nabla \phi_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \phi_j^{1R} \rangle \quad (\text{C.13})$$

As for τ^{1R} , the potential μ_{xc} is expressed by its moments:

$$\mu_{xclm}(r) = \int \mu_{xc}(\mathbf{r}) S_{lm}(\theta, \phi) d\Omega \quad (\text{C.14})$$

Using Eqs C.14 and C.9, the Eq.C.13 becomes:

$$\begin{aligned} \langle \nabla \phi_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \phi_j^{1R} \rangle &= \int \mu_{xc} \left[\frac{\bar{\Phi}_i^{1R}(r)}{r} S_{l_i m_i} \mathbf{u}_r + \frac{\Phi_i^{1R}(r)}{r} \nabla S_{l_i m_i} \right] \cdot \left[\frac{\bar{\Phi}_j^{1R}(r)}{r} S_{l_j m_j} \mathbf{u}_r + \frac{\Phi_j^{1R}(r)}{r} \nabla S_{l_j m_j} \right] dr d\Omega \\ &= \int \sum_{lm} \mu_{xclm}(r) S_{lm} \left[\frac{\bar{\Phi}_i^{1R}(r)}{r} S_{l_i m_i} \mathbf{u}_r + \frac{\Phi_i^{1R}(r)}{r} \nabla S_{l_i m_i} \right] \cdot \left[\frac{\bar{\Phi}_j^{1R}(r)}{r} S_{l_j m_j} \mathbf{u}_r + \frac{\Phi_j^{1R}(r)}{r} \nabla S_{l_j m_j} \right] dr d\Omega \end{aligned}$$

It gives finally:

$$\langle \nabla \phi_i^{1R} | \mu_{xc}(\mathbf{r}) | \nabla \phi_j^{1R} \rangle = \sum_{lm} \int \mu_{xclm}(r) \left[\frac{\bar{\Phi}_i^{1R}(r) \bar{\Phi}_j^{1R}(r)}{r^2} G_{lm, l_i m_i, l_j m_j} + \frac{\Phi_i^{1R}(r) \Phi_j^{1R}(r)}{r^2} H_{lm, l_i m_i, l_j m_j} \right] dr d\Omega \quad (\text{C.15})$$

By adding the pseudo terms, the total on-site Meta-GGA contribution to the energy is thus:

$$\begin{aligned} \frac{1}{2} \sum_{ijR} \rho_{ij}^{1R} D_{ij}^{1R, MGA} = & \frac{1}{2} \sum_{ijR} \rho_{ij}^{1R} \sum_{lm} \int \mu_{xclm}(r) \left[\left(\frac{\bar{\Phi}_i^{1R}(r) \bar{\Phi}_j^{1R}(r)}{r^2} - \frac{\tilde{\Phi}_i^{1R}(r) \tilde{\Phi}_j^{1R}(r)}{r^2} \right) G_{lm l_i m_i l_j m_j} \right. \\ & \left. + \left(\frac{\Phi_i^{1R}(r) \Phi_j^{1R}(r)}{r^2} - \frac{\tilde{\Phi}_i^{1R}(r) \tilde{\Phi}_j^{1R}(r)}{r^2} \right) H_{lm l_i m_i l_j m_j} \right] dr d\Omega \end{aligned} \quad (C.16)$$

Appendix D

ZBL Empirical Potential

Given a crystal configuration where the atoms have the positions \mathbf{R}_i , ZBL potential is computed by the form [204]:

$$E_{ij}^{ZBL} = \frac{1}{4\pi\epsilon_0} \frac{Z_i Z_j}{\|\mathbf{R}_i - \mathbf{R}_j\|} \phi\left(\frac{\|\mathbf{R}_i - \mathbf{R}_j\|}{a}\right) + S(\|\mathbf{R}_i - \mathbf{R}_j\|)$$

with

$$a = \frac{0.46850}{Z_i^{0.23} + Z_j^{0.23}}$$

and

$$\phi(x) = 0.1817e^{-3.19980x} + 0.50986e^{-0.94229x} + 0.28022e^{-0.40286x} + 0.02817e^{-0.20162x}$$

The S function ramps the energy and forces derived from ZBL potential to zero between an inner and outer cut off radii. This process being held smoothly. The ZBL energy of the atomic configuration in its whole is given by summing all the pairs' interactions:

$$E^{ZBL} = \sum_{ij} E_{ij}^{ZBL}$$

Appendix	E
----------	----------

Manganese Hydrides and Superhydrides

Crystal Structures of 3d Transition Metal Hydrides

Tab. E.1 gathers the crystal structures of 3d transition metal hydrides, as observed or predicted, including the ones found in the present work. Note that scandium hydrides are excluded from the table, because column IIIB transition metals hydrides are closer to the family of rare-earth hydrides [230].

Compound	Space group	Structure	Example
M_2H	<i>P-3m1</i> <i>I4₁/amd</i> <i>C2/m</i>		Cu_2H [2], Mn_2H^* V_2H (low P) [3] V_2H (high P) [3]
MH	<i>P6₃/mmc</i> <i>Fm-3m</i> <i>P6₃/mmc</i> <i>R-3m</i>	hcp (ϵ) rocksalt (γ) double hcp (ϵ')	CrH [4,5], MnH^* CoH [6], NiH [7,8], CuH [2] FeH [9] VH [3]
M_2H_3	<i>C2/m</i> <i>I2/m</i> <i>P2₁/m</i>		Ni_2H_3 [8], Cr_2H_3 (low P) [4,5] Ni_2H_3 [7] Cr_2H_3 (high P) [4]
MH_2	<i>I4/mmm</i> <i>Fm-3m</i> <i>Pnma</i>	fluorite	FeH_2 [10], CoH_2 (high P) [6], NiH_2 [8], MnH_2^* CoH_2 (low P) [6], VH_2 (low P) [3], TiH_2 (distorted) [11] CrH_2 [4,5], VH_2 (high P)
MH_3	<i>Pm-3m</i> <i>P6₃/mmc</i> <i>C2/m</i> <i>Fm-3m</i> <i>I4/m</i>	Simple cubic fcc	FeH_3 [12,9,10], CoH_3 [6] CrH_3 [4] CuH_3 [2] VH_3 [3] MnH_3^*
MH_4	<i>I4/mmm</i> <i>R32</i>		CrH_4 [4] MnH_4^*
MH_5	<i>I4/mmm</i> <i>P6₃/mmc</i>		FeH_5 [13,10] VH_5 [3]
MH_6	<i>Cmmm</i>		FeH_6 [10]
MH_7	<i>P6/mmm</i>		MnH_7^*
MH_8	<i>I41md</i> <i>C2/m</i>		CrH_8 [4] VH_8 [3]

* present work.

Figure E.1: Crystal structures (observed or predicted) for 3d transition metal (excluding Sc) hydrides and superhydrides.

The Different Structures of MnH (ϵ , ϵ' , γ , r)

We have explicitly checked that the ϵ (hcp) phase of MnH is more stable than the ϵ' (dhcp) and the γ (fcc rocksalt) phases over the pressure range considered in this work (from 0 to 150 GPa, FIG E.2). Moreover, the ϵ' (dhcp) phase is more stable than the γ phase. In addition, our AIRSS on MnH have provided a structure with the R-3m space group, r-MnH (and 3 f.u./primitive cells). This phase is less stable than ϵ but is the second one in the stability order. Its structure also consists of a stacking of hexagonal Mn planes.

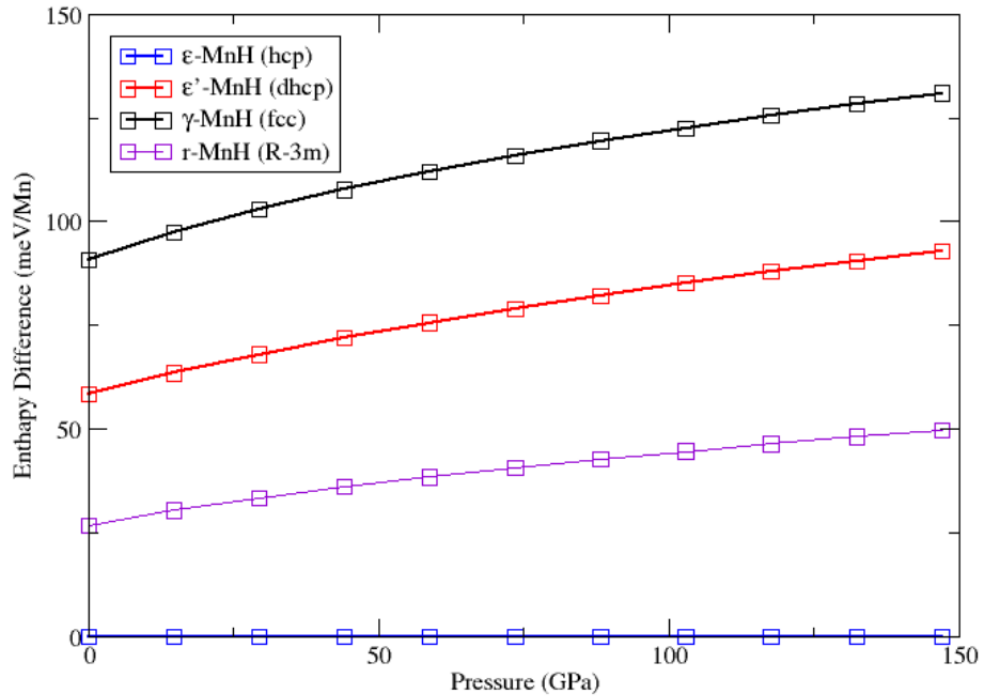


Figure E.2: Enthalpy of ϵ -MnH (hcp phase), ϵ' -MnH (dhcp phase), γ -MnH (fcc phase) and r-MnH (space group R-3m) as a function of pressure. The enthalpy of ϵ -MnH is taken as reference.

The atomic positions, lattice constants and angles of this R-3m phase at 50 GPa are:

$$a=b=2.5122 \text{ \AA}, c=17.973 \text{ \AA}$$

$$\alpha=\beta=90^\circ, \gamma=120^\circ$$

$$\text{Mn (6c): } 0 \ 0 \ 0.2791$$

$$\text{Mn (3b): } 0 \ 0 \ 0.5$$

$$\text{H (6c): } 0 \ 0 \ 0.11292$$

$$\text{H (3a): } 0 \ 0 \ 0$$

The stacking of the hexagonal Mn planes in this R-3m phase is as follows:

-A-B-C-B-C-A-C-A-B-

FIG E.3 shows the crystal structure of r-MnH.

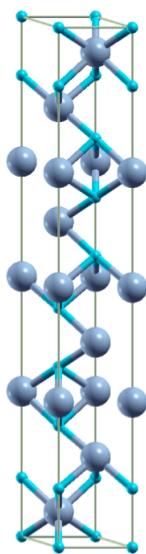


Figure E.3: Enthalpy of ϵ -MnH (hcp phase), ϵ' -MnH (dhcp phase), γ -MnH (fcc phase) and r-MnH (space group R-3m) as a function of pressure. The enthalpy of ϵ -MnH is taken as reference.

Crystal Structures of Unstable Phases

This section describes the crystal structure of the compounds that are found energetically unstable in the present work (Tab. E.13). Some of them are displayed on the following figures.

Mn_2H_3

Mn_2H_3 is found to crystallize in space group C2/m. The Wyckoff positions are the same as in the structure of C2/m- Cr_2H_3 [?]. A view of the structure is shown on FIG E.4.

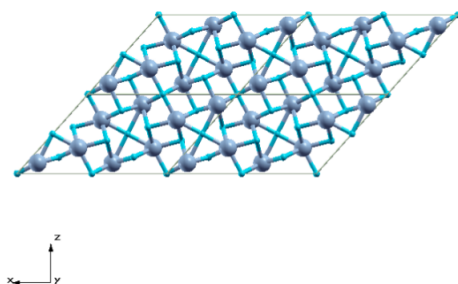


Figure E.4: Crystal structure of Mn_2H_3 (energetically unstable compound). Space group C2/m.

MnH_4

MnH_4 is found to adopt a P4/nmm space group at low pressure, and a rhombohedral structure (space group R32) at high pressure, which can be seen as a small distortion from a parent cubic

structure (Pm-3m space group). The low-pressure structure consists of a stacking of bilayers, with each bilayer being a rocksalt-type MnH arrangement truncated along [001]. Hydrogen is inserted in the space between the bilayers, with some H atoms on top of Mn, and other H atoms in bridge positions above a Mn-Mn line (like in the structure of Pepin2017Pepin2017Pepin2017FeH5). At high pressure, the nearly cubic structure of MnH₄ consists in a complex arrangement that contains 3 f.u. per primitive cell (FIG E.5 , Tab. E.13): the Mn atoms (Wyckoff position 3c) occupy the centers of the faces of the (primitive) cubic cell, one H is at the center of the cell (position 1b), other H atoms occupy the middle of the edge (position 3d), and 8 H atoms occupy positions near the 8 corners of the cubic cell (position 8g). The same structural unit “H8 cube” as in MnH₃ can be found in this high-pressure structure of MnH₄, however with an additional H atom at the center of the cube. These two structures of MnH₄ do not contain molecular hydrogen.

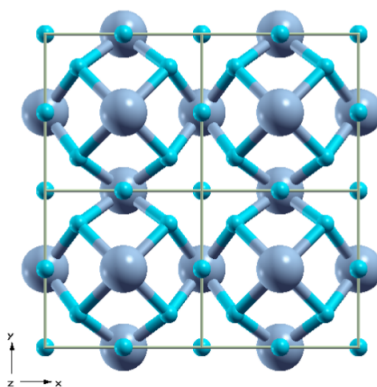


Figure E.5: High-pressure crystal structure of MnH₄, shown here as a cubic crystal (space group Pm-3m). The true structure possesses a slight rhombohedral distortion from the cubic structure (see main text)

Mn₂H₉

Mn₂H₉ adopts a body-centered orthorhombic structure (space group Immm) that can be described as a stacking of Mn monolayers within which the H atoms are bonded to three (or two) Mn. Pairs of H atoms, or single H atoms, play the role of bridges between the layers. The structure does not contain molecular hydrogen (FIG E.6).

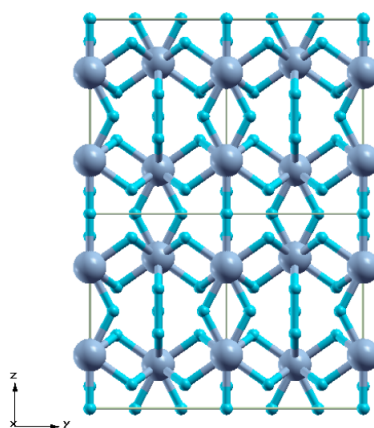


Figure E.6: Crystal structure of Mn_2H_9 (energetically unstable compound). Space group Immm.

MnH_5

MnH_5 is found to adopt a monoclinic structure with space group C2/m, that can be described as a stacking of Mn monolayers within which the H atoms are either bonded to three (or two) Mn just above the layer, or placed on top of Mn atoms. Pairs of H atoms play the role of bridges between the layers. The structure does not contain molecular hydrogen (FIG E.7). This structure differs from that of Mn_2H_9 only by the H atoms placed on top of the Mn.

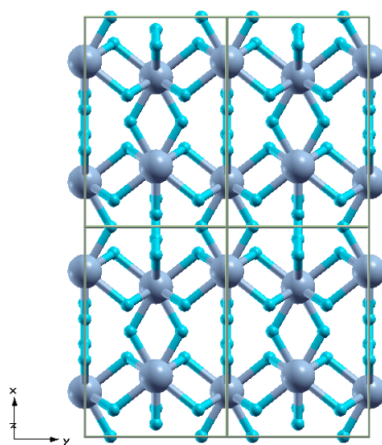


Figure E.7: Crystal structure of MnH_5 (energetically unstable compound). Space group C2/m.

Mn_2H_{11}

Mn_2H_{11} is found to adopt a monoclinic structure with space group Pc. We note the presence of H_2 units with an interatomic distance of 1.17 Å, visible on FIG E.8.

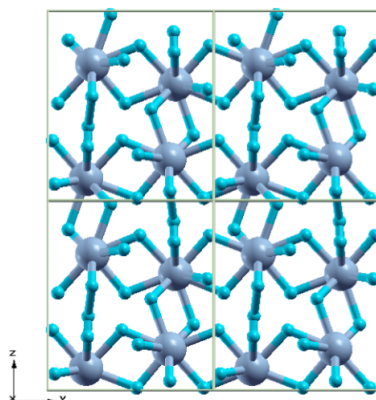


Figure E.8: Crystal structure of Mn_2H_{11} (energetically unstable compound). Space group Pc.

MnH_6

The structure of MnH_6 is monoclinic and evolves from space group P21/m at low pressure to space group Cm at high pressure. At low pressure, the orthorhombic structure consists of a stacking of Mn monolayers (the Mn forming a quasi-hexagonal 2D lattice), with H atoms slightly above the layers and bonded to three Mn, and H atoms placed on top of Mn, either alone or by pairs (bondlength 0.88 Å at 44 GPa), see FIG E.9. The structure does not contain true molecular hydrogen. At high pressure, the monoclinic structure is complex (3 f.u. per primitive cell) but does not contain H_2 molecules.

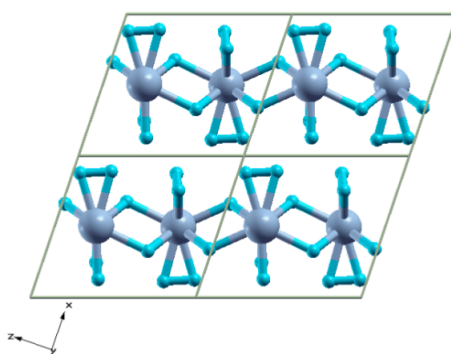


Figure E.9: Crystal structure of MnH_6 (energetically unstable compound). Space group P2₁/m.

Mn_2H_{13}

Mn_2H_{13} is a mixed hydride (i.e. it contains hydrided hydrogen and molecular hydrogen). Its structure (shown on FIG E.10) is close to that of MnH_7 .

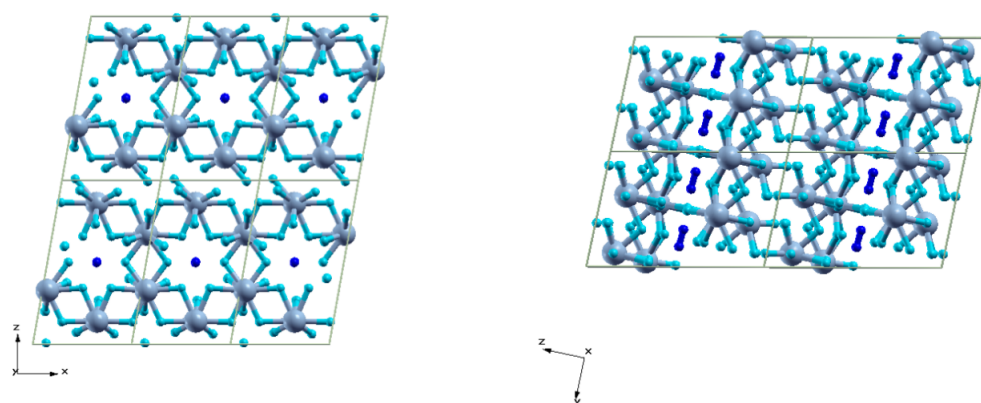


Figure E.10: Crystal structure of Mn_2H_{13} (energetically unstable compound). Space group P-1. It contains one H_2 molecule per two formula unit of Mn_2H_{13} . The hydrogens involved in the molecule are in dark blue.

Compound	Space Group	Pressure (GPa)	Lattice (Å) constants and angles (degree)	Atom	Wyckoff position	x	y	z
Mn₂H₃	<i>C2/m</i>	50	a=8.4956 b=2.4937 c=6.0291 $\alpha=90$ $\beta=131.80$ $\gamma=90$	Mn	4i	0.25477	0	0.67488
				Mn	4i	0.07328	0	0.86516
				H	4i	0.59630	0	0.41603
				H	2c	0	0	0.5
				H	4i	0.32083	0	0.20911
				H	2b	0	0.5	0
Mn₂H₅	<i>P-1</i>	50	a=3.5908 b=4.5350 c=4.5781 $\alpha=100.16$ $\beta=112.85$ $\gamma=113.12$	Mn	2i	0.75050	0.21940	0.28890
				Mn	2i	-0.07193	0.75411	0.10371
				H	2i	0.19982	0.55964	0.32208
				H	1f	0.5	0	0.5
				H	2i	0.22332	0.18166	0.27060
				H	2i	0.56223	0.21898	0.88921
				H	1b	0	0	0.5
				H	2i	0.63415	0.54289	0.28433
Mn₂H₇	<i>Cm</i>	50	a=3.5777 b=3.5865 c=5.5981 $\alpha=90$ $\beta=101.61$ $\gamma=90$	Mn	2a	0.50793	0	0.48432
				Mn	2a	-0.09206	0	0.16074
				H	4b	0.30725	0.23776	0.66798
				H	4b	0.60348	0.26782	-0.03014
				H	2a	0.42171	0	0.17807
				H	2a	-0.00817	0	0.45635
				H	2a	0.63894	0	0.77032

Figure E.11

MnH₄	<i>Pm-3m</i>	150	a=3.5244 b=3.5244 c=3.5244 $\alpha=90$ $\beta=90$ $\gamma=90$	Mn H H H	3d 8g 3d 1b	0.5 0.78211 0.5 0.5	0 0.78211 0 0.5	0 0.78211 0 0.5
Mn₂H₉	<i>Immm</i>	150	a=2.4156 b=4.2722 c=6.0835 $\alpha=90$ $\beta=90$ $\gamma=90$	Mn H H H H	4i 8l 4h 4j 2a	0 0 0 0.5 0	0 0.69606 0.83065 0 0	0.27422 0.87637 0.5 0.11506 0
MnH₅	<i>C2/m</i>	150	a=6.6802 b=4.2844 c=2.4458 $\alpha=90$ $\beta=108.05$ $\gamma=90$	Mn H H H H	4i 8j 4h 4i 4i	0.78715 0.87354 0 0.36441 0.44052	0 0.20168 0.17995 0 0	0.30599 0.88815 0.5 0.31769 0.79982
Mn₂H₁₁	<i>Pc</i>	150	a=2.4810 b=4.9532 c=6.0765 $\alpha=90$ $\beta=111.07$ $\gamma=90$	Mn Mn H H H H H H H H H H	2a 2a 2a 2a 2a 2a 2a 2a 2a 2a 2a 2a	0.04183 0.19022 0.41667 0.12391 0.28348 0.66818 -0.08412 0.50719 0.72511 0.56048 0.81113 0.33071 -0.03803	0.76949 0.73090 0.62185 0.24231 -0.02727 0.23870 -0.06006 -0.03985 0.54358 0.26539 0.87739 0.44689 0.46432	0.61432 0.21634 0.48063 0.39009 0.81585 0.28191 0.01240 0.58666 0.24152 0.54909 0.34594 0.31651 0.51248
MnH₆	<i>P2₁/m</i>	50	a=4.2465 b=2.5644 c=4.7194 $\alpha=90$ $\beta=108.88$ $\gamma=90$	Mn H H H H H H	2e 2e 2e 2e 2e 2e 2e	0.50422 0.33746 0.12434 0.88967 0.32237 0.75569 0.87754	0.25 0.25 0.25 0.25 0.25 0.25 0.25	0.74944 0.01726 0.62689 0.77142 0.38519 0.32489 -0.04392
Mn₂H₁₃	<i>P-1</i>	150	a=4.3839 b=4.9108 c=7.5528 $\alpha=78.20$ $\beta=79.84$ $\gamma=88.68$	Mn Mn Mn Mn H H H	2i 2i 2i 2i 2i 2i 2i	0.60608 -0.05000 0.39473 0.05006 0.86801 0.74836 0.24089	0.11980 0.45125 0.38051 0.04894 0.62157 0.39070 0.78663	0.14169 0.67274 0.85780 0.32708 -0.00768 0.86288 0.36849

Figure E.12

				H	2i	0.31217	0.43740	0.65778
				H	2i	0.50124	0.82407	0.50228
				H	2i	0.41885	0.87914	0.07194
				H	2i	0.40471	0.67712	0.73316
				H	2i	0.19615	0.62232	-0.06195
				H	2i	-0.06646	0.65655	0.80017
				H	2i	0.40064	0.05833	0.34298
				H	2i	0.85812	0.78110	0.40500
				H	2i	0.03135	0.38745	0.86452
				H	2i	0.21046	-0.01421	0.49999
				H	2i	0.13144	0.87912	0.00816
				H	2i	0.25227	0.10786	0.13703
				H	2i	0.76669	0.71532	0.62396
				H	2i	0.68842	0.03584	0.34221
				H	2i	0.49949	0.67673	0.49720
				H	2i	0.58029	0.62035	-0.07276
				H	2i	0.58280	0.82344	0.26738
				H	2i	0.80367	0.87716	0.06303
				H	2i	0.06194	0.84338	0.20012
				H	2i	0.59880	0.46923	0.65734
				H	2i	0.14801	0.71678	0.60104
				H	2i	-0.03076	0.11497	0.13500
				H	2i	0.78924	0.48863	0.50014
Mn₂H₁₅	<i>P-1</i>	100	a=2.4922 b=2.5162 c=7.6762 α =85.71 β =85.53 γ =89.95	Mn	2i	0.03707	0.54834	0.22014
				H	2i	-0.03727	-0.03049	0.65211
				H	2i	0.45831	0.31976	-0.09160
				H	2i	0.01458	0.49053	0.41835
				H	2i	0.55946	0.13671	0.22403
				H	2i	0.39882	0.89410	0.50875
				H	2i	0.52261	0.58766	0.34749
				H	1c	0	0.5	0
				H	2i	0.89939	-0.0750	-0.9278

Figure E.13: Atomic positions, lattice constants and angles of some (energetically) unstable structures of manganese hydrides. The nearly cubic structure of MnH₄ is also given.

High Pressure Phases of Dense Hydrogen Taken as References

In the pressure range considered in this work (from 0 to 150 GPa), dense hydrogen is in phase I at room temperature, and in phases I and II at low temperatures. Around 150 GPa, dense hydrogen transits to the so-called phase III. In these three phases, hydrogen is under the molecular form H_2 , and is an insulating solid. The best candidate for phase III is the C2/c-24 phase [85]. However, phases I and II are extremely difficult to describe by DFT because the dynamics of the molecules in these phases is affected by very large thermal and quantum fluctuations ([?],[?]). In particular, the molecules in phase I are free rotors, i.e. they do not keep a well-defined orientation and rotate under the effect of thermal and quantum effects. To model dense hydrogen in the low-pressure region, we use here the C2/c-32 phase (Tab. E.16). This phase is monoclinic, it is very close to the hexagonal P63/m phase described in Ref [85] but is slightly more stable. Beyond 117 GPa, the C2/c-24 phase [85] becomes more stable than C2/c-32. Thus we use C2/c-24 as reference for the calculation of formation enthalpies for pressures above 117 GPa.

Compound	Space Group	Pressure (GPa)	Lattice (Å) constants and angles (degree)	Atom	Wyckoff position	x	y	z
H_2	C2/c	47	a=7.1426 b=4.1369 c=6.6473 $\alpha=90$ $\beta=90.18$ $\gamma=90$	H	8f	0.40576	0.26817	0.32177
				H	8f	0.42196	0.23874	0.43002
				H	8f	0.71149	0.00421	0.38919
				H	8f	0.11140	0.49172	0.36220
				H	8f	0.84763	0.07416	0.60181
				H	8f	0.31044	0.42400	0.64677
				H	8f	0.55383	0.32706	0.63164
				H	8f	0.61516	0.18419	0.62075

Figure E.14

Equations of States

FIG E.15 shows the pressure evolution of the volume in the stable hydrides. From the volume of Mn_2H , MnH , MnH_2 and MnH_3 , we deduce at each pressure the volume increment per hydrogen atom, which is plotted on FIG E.16 as a function of pressure. Note that MnH_7 has been excluded from the list because hydrogen in MnH_7 is partly molecular. This volume increment per hydrogen atom is in line with the one observed in iron hydrides.

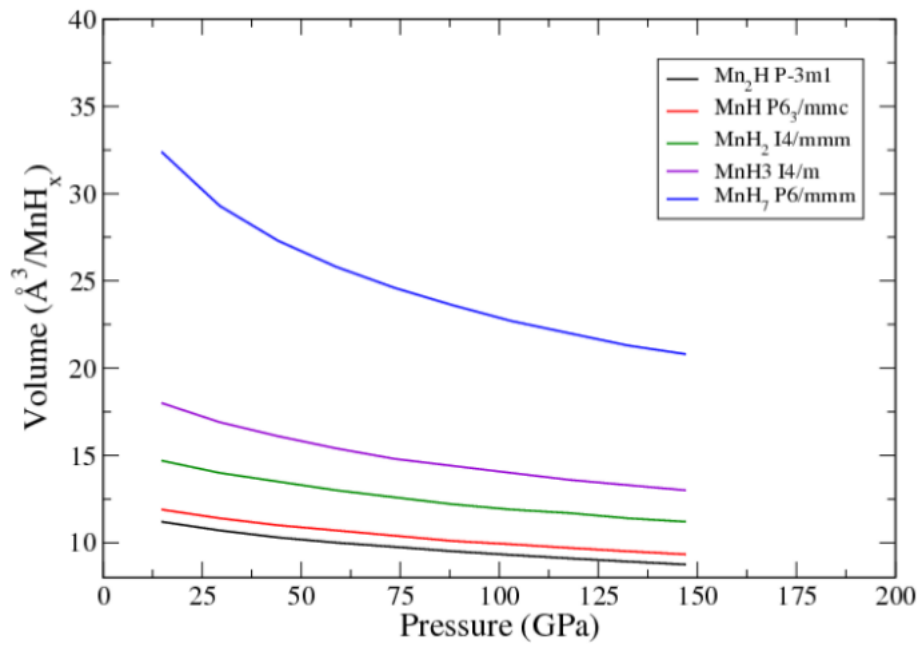


Figure E.15: Volume versus pressure for the stable Mn hydrides.

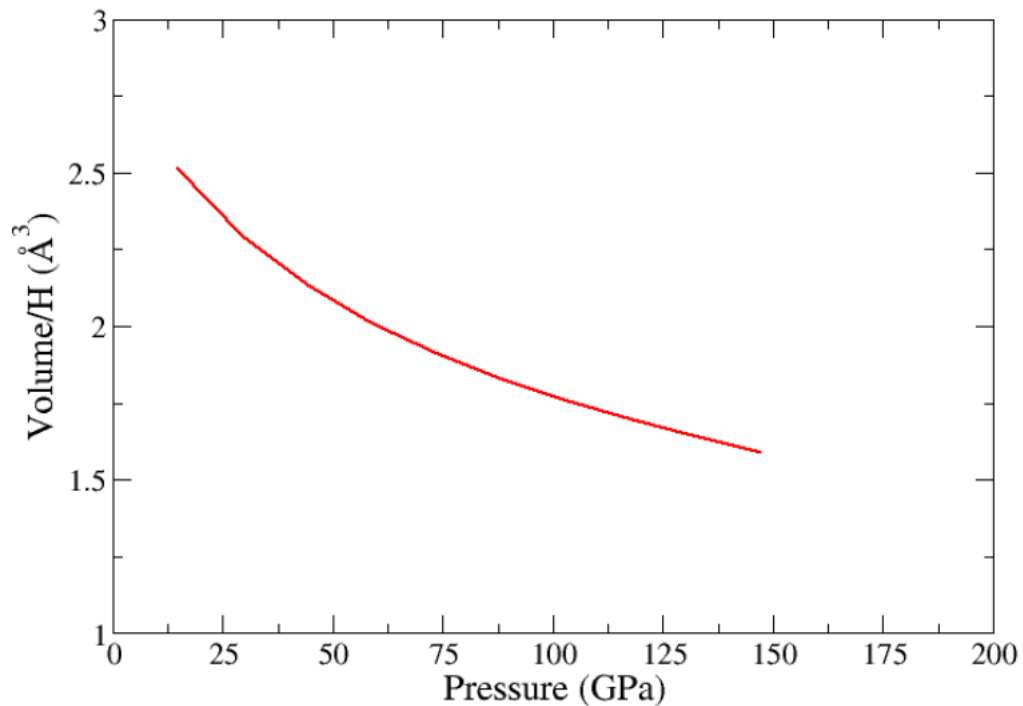


Figure E.16: Volume increment per hydrogen atom (from Mn₂H to MnH₃) as a function of pressure.

Phonon Band Structure of MnH and MnH₂: Pressure Evolution

MnH

FIG E.17 displays the phonon dispersion curves and phonon density of states of ϵ -MnH for two pressures: 50 and 150 GPa. FIG E.17 illustrates the dynamical stability of ϵ -MnH as well as the hardening of the phonon modes with pressure, especially the highest-frequency phonon modes, which correspond to atomic motions within the hydrogen sublattice.

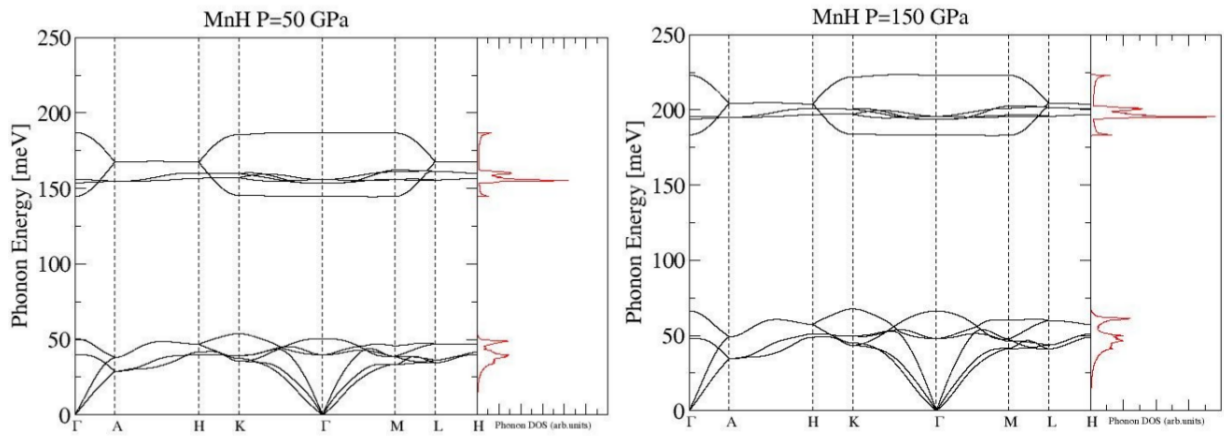


Figure E.17: Phonon dispersion curves and vibrational density of states of ϵ -MnH at 50 GPa and 150 GPa (space group P63/mmc).

MnH₂

FIG E.18 displays the phonon dispersion curves and phonon density of states of MnH₂ for pressures of 50 and 100 GPa (the one for 150 GPa is presented in the main text).

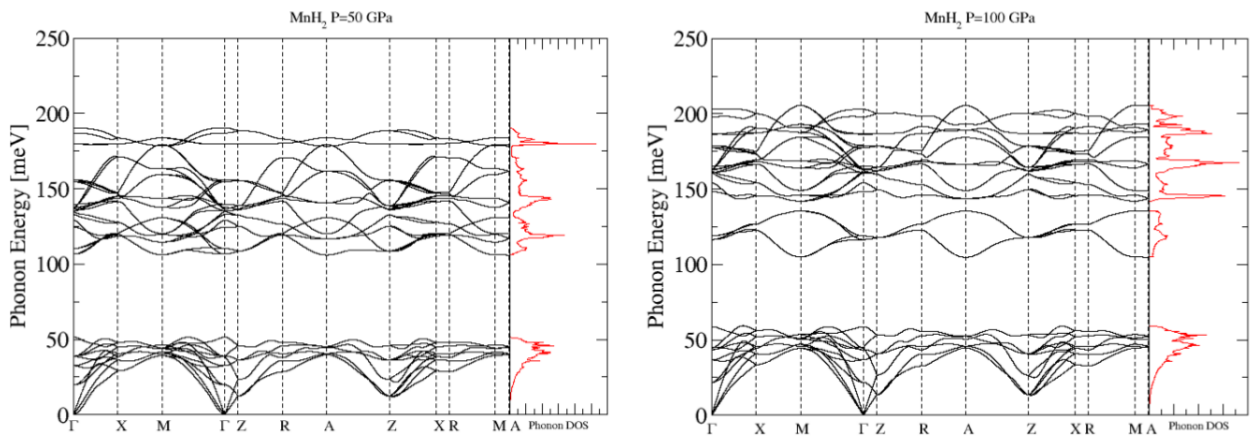


Figure E.18: Phonon dispersion curves and vibrational density of states of MnH₂ at 50 GPa, and 100 GPa (space group I4/mmm).

Phonons of Energetically Unstable Hydrides at 150 GPa

Mn_2H_9 (Space Group Immm)

The DFPT calculation is performed using the conventional unit cell (multiplicity 2). FIG E.19 displays the phonon dispersion curves of Mn_2H_9 at 150 GPa over a path in the Brillouin Zone associated with this conventional unit cell.

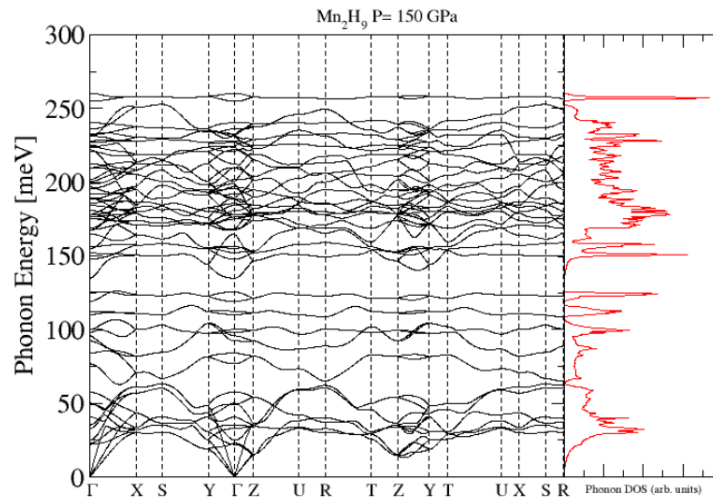


Figure E.19: Phonon dispersion curves and vibrational density of states of Mn_2H_9 at 150 GPa (space group Immm).

MnH₅ (Space Group C2/m)

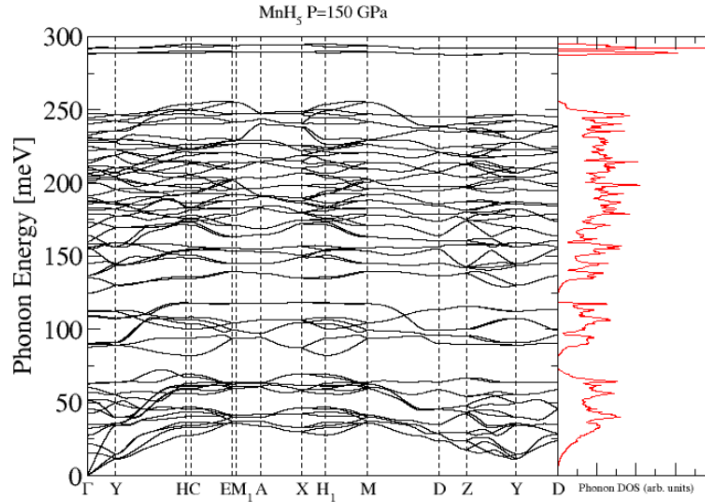


Figure E.20: Phonon dispersion curves and vibrational density of states of MnH₅ at 150 GPa (space group C2/m).

Mn₂H₁₁ (Space Group Pc)

The DFPT calculation performed on the structure identified by AIRSS (13 atoms per unit cell) exhibited unstable modes at the Zone boundary. Thus, we increased the size of our supercell by replicating once the structure along the three cartesian directions, leading to a 104-atom supercell. Then we performed a new full structural optimization within this supercell, starting from the AIRSS structure on which small displacements along the phonon eigenvector have been superimposed. This lead to a slightly more stable structure. Since the phonon instabilities were observed only at the Zone boundary of the 13-atom structure, we only performed the phonon calculation at the Γ point of the new 104-atom structure, and observed that the unstable modes have indeed disappeared. Then we ran the findsym program over this new structure, which provided a monoclinic space group Pc.

Mn₂H₁₃ (Space Group P-1)

This is a mixed hydride at 150 GPa. Phonon dispersion curves at this pressure are shown on FIG E.21.

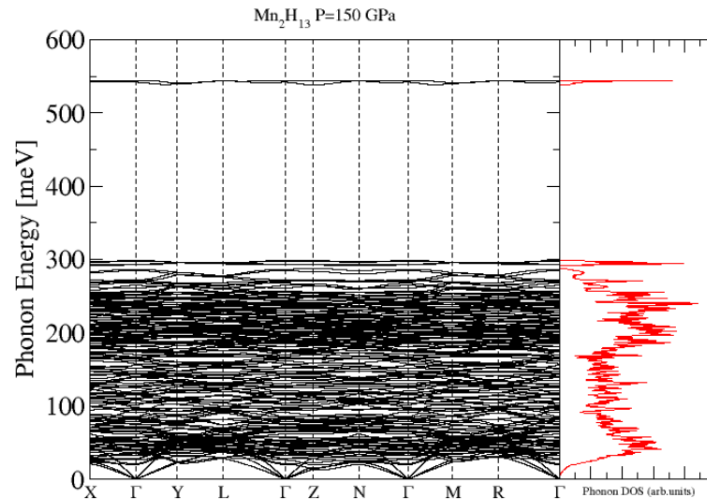


Figure E.21: Phonon dispersion curves and vibrational density of states of Mn_2H_{13} (space group P-1) at 150 GPa.

K-Point and Q-Point Meshes Used in the DFPT Calculations

The tables below E.22 present the k-point meshes and q-point meshes employed to compute the dynamical stability of the considered compounds. The k-meshes correspond to those applied to sample the Brillouin Zone in the electronic DFPT calculation. The q-point meshes are those used to compute the dynamical matrices.

Compound	MnH	MnH_2	MnH_3	MnH_7
k-point mesh	$16 \times 16 \times 10$	$18 \times 18 \times 6$	$8 \times 8 \times 12$	$8 \times 8 \times 12$
q-point mesh	$8 \times 8 \times 5$	$9 \times 9 \times 3$	$4 \times 4 \times 6$	$4 \times 4 \times 6$

Compound	MnH_4	Mn_2H_9	MnH_5	Mn_2H_{11}	Mn_2H_{13}	MnH_8
k-point mesh	$8 \times 8 \times 6$	$8 \times 8 \times 8$	$10 \times 10 \times 10$	$6 \times 4 \times 4$	$6 \times 6 \times 4$	$12 \times 6 \times 4$
q-point mesh	$4 \times 4 \times 3$	$4 \times 4 \times 4$	$5 \times 5 \times 5$	$1 \times 1 \times 1$	$3 \times 3 \times 2$	$6 \times 3 \times 2$

Figure E.22

Appendix	F
----------	----------

Machine Learning and AIRSS

Potential's Evaluation

Correlations on the Training Database

After being trained, the correlations between predictions and DFT values for energy, forces and stress on the configurations stored in the training database are given below.

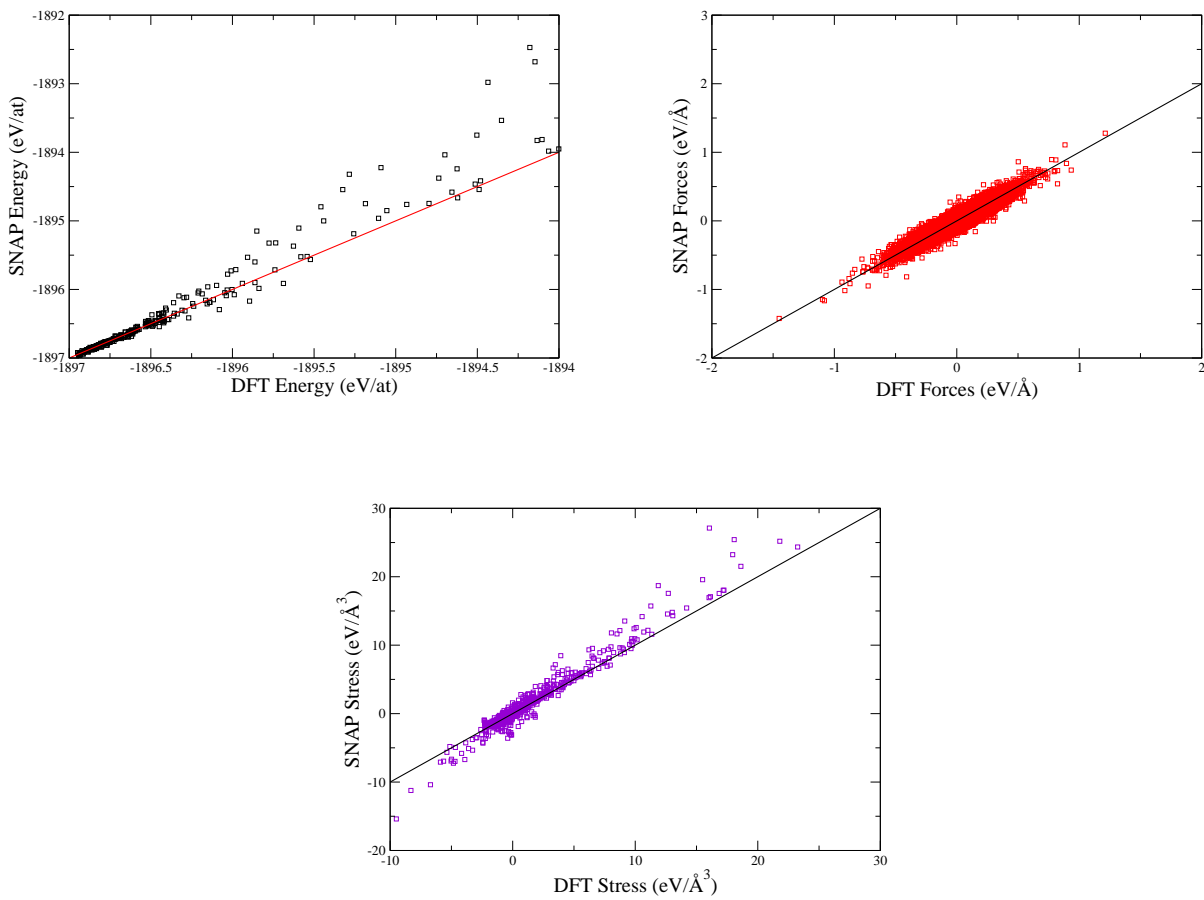


Figure F.1: Energy, forces and stress correlations on the training database

The RMSE obtained for these three physical quantities are gathered in the table F.1.

	Energy (meV/at)	Forces (meV/Å)	Stress (meV/Å ³)
RMSE (meV/at)	54.43	22.41	49.51

Table F.1: RMSE values for energy, forces and stress on the initial training database

Fig F.1 and table F.1 are confirming the good correlations regarding SNAP predictions and DFT

values. It can be inferred that the numerical potential could be able to reproduce the DFT computed physical properties of the training database in a relatively good and accurate way.

Cold Curve Predictions

An other test on the trained potential is to assess how it can reproduce the energy evolution of a crystal type configuration with the atomic density (the cold curves). This test is important because it evaluates whether the potential can reproduce accurately crystal's cold curves, and in particular for phases it has not seen before. This ability is worth because we expect the potential to predict stability for phases not necessary included in the training database.

The following figure F.2 shows the cold curves of a Bcc crystal structure. The open circle corresponds to the cold curve computed with DFT. The black squares are the values predicted by the numerical potential trained on the three main ingredients of the database (liquid, crystal, triclinic structures). The diamonds are the values computed with the numerical potential trained only with the liquid and triclinic structures.

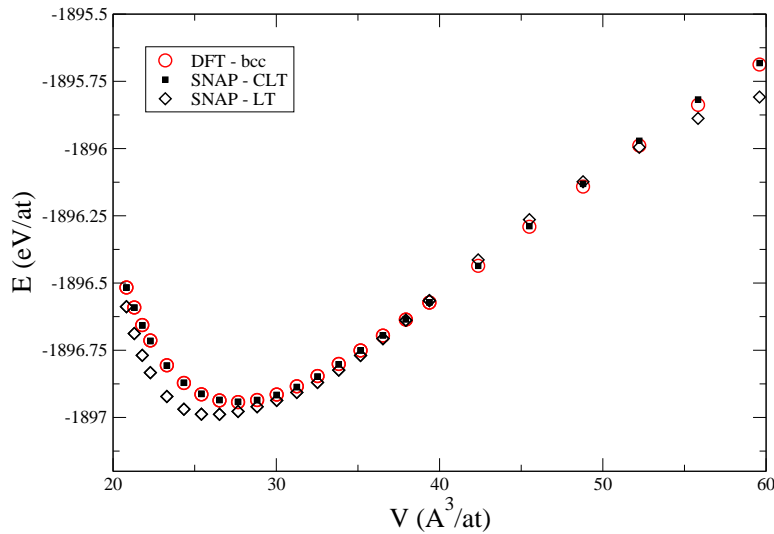


Figure F.2: Evaluation of the potential for the Bcc cold curves with and without the training on the crystal configurations.

Without any surprise, the cold curve is well predicted when the model is trained on the crystal configurations (The Bcc phase is included in this part of the training database). This curve shows moreover the ability of the model to reproduce the Bcc cold curve quite accurately even if it is not trained on the crystal structures. It is a strong indication the model can predict physical properties of structures non included in the training database during the search. Moreover, once

a new crystal structure is predicted and included in the database, it confirms the model is able to accurately estimate its physical properties.

Structure Generation

After being trained, the numerical potential is then employed to optimize randomly generated structures. This structure generation is realized throughout a determined process explained below.

Global Generation Parameters

The crystal structure searches are realized on a set of 10000 randomly generated configurations. These structures are distributed uniformly in an atomic density domain between 20 and 40 $\text{\AA}^3/\text{at}$. This domain is justified because tin's density under normal conditions and low pressure is around 30 $\text{\AA}^3/\text{at}$ (this is the first starting information taken at the beginning of this work). One thus expect a density domain approximately centered at this value and taking into account the atomic repulsion forbids to have too small atomic density (below 20 $\text{\AA}^3/\text{at}$). The choice of a such wide atomic density range is to make sure all the atomic densities of the different possible tin's phases are included in it.

So as to avoid distorted generated geometries, some global parameters are set. First of all a maximal rate of 2.5 is imposed between the values of the primitive vectors' lengths. The angle values between these vectors are moreover included in a distribution centered at 90 degrees with a width of ± 20 degrees.

Steps of the Generation Process

For a given structure to generate, the volume per atom and their number are firstly chosen. Here it is an even number chosen randomly between 2 and 12 (this is the second starting information taken at the beginning of this work). From these data the total volume of the structure is computed. The angles between the primitive cell vectors are then determined. Finally the cell's parameters so as to respect the total volume computed and the maximal rate specified between them are obtained. After generating the primitive cell's geometry, the atomic positions need to be chosen inside it. A first position set is thus generated with random atoms' coordinates. This first atomic configuration is then subject to a few steps of structural optimization with the repulsive empirical potential ZBL to avoid too close atom pairs. The figure F.3 shows the atomic density distribution of the generated structures. This generation is realized so as to have a uniform distribution between the range 20 and 40 $\text{\AA}^3/\text{at}$.

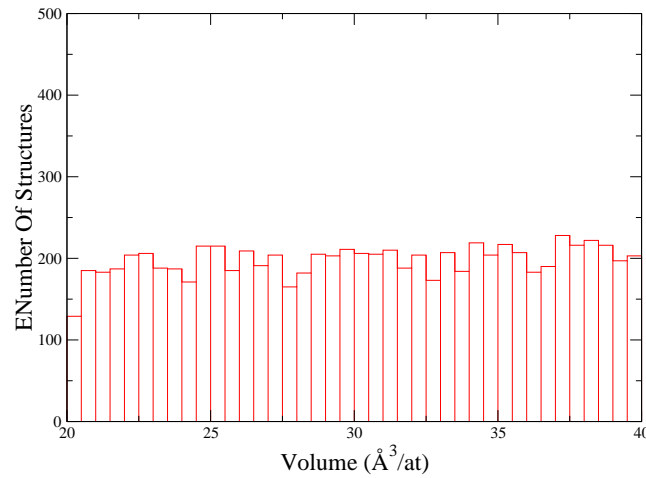


Figure F.3: Volume per atom distribution of the 10000 generated structures.

Comparison with DFT Results

The structure search using the above described active learning process was able to predict at 0 GPa, an α phase and at 20 GPa a γ phase. The first question considering the ability to predict phases making part in the tin's thermodynamical diagram is thus successfully answered. Moreover, with this method, the numerical potential acquires knowledge on very different tin's structures in comparison with those initially present in the database. In particular, the α phase which has covalent bonds and thus is far from the structures stored in the database at the beginning has been identified and learned afterwards. It's now time to challenge the second question asking whether the predicted phases are those expected to appear.

Most Stable Structure At 0 GPa

At 0 GPa, the figure F.4 here gives the enthalpy differences of the tin's phases computed by DFT (using DFTO-2 parameter set). These enthalpy differences are computed from the basis of the Bcc phase. It shows clearly the stability of the α phase at 0 GPa and 0 K.

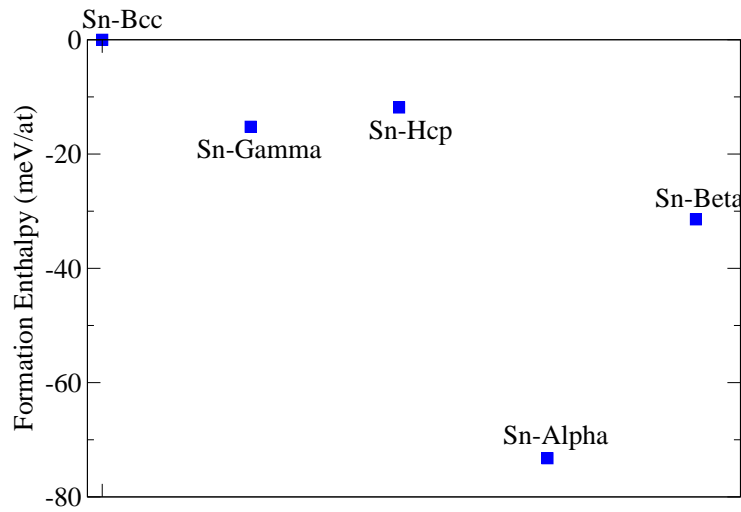


Figure F.4: Enthalpy formation of tin's phases at 0 GPa. The stabilities are compared with the Bcc phase.

Most Stable Structure At 20 GPa

Under pressure, the γ , Bcc and Hcp phases have very close relative stability between each other. The figure F.5 computes the relative enthalpy formation in comparison with the Bcc phase of these three crystal structures. It is computed on a range between 0 and 200 GPa.

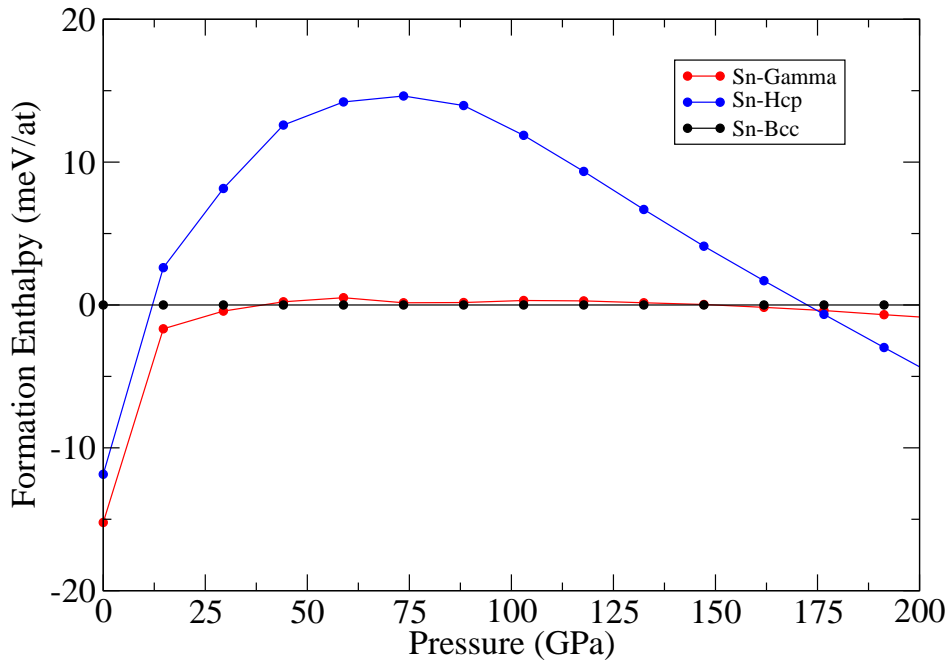


Figure F.5: Enthalpy formation of tin's phases between 0 and 200 GPa. The stabilities are compared with the BCC phase.

These results are coherent with former studies [206], [207], [208]. In particular one retrieve:

- An Hcp phase, which is expected to be stable under high pressure [206] (above 150 GPa).
- Two phases have very close formation enthalpy between 10 and 150 GPa: the Bcc and a γ phase.
- The three considered phases have very tiny formation enthalpy difference (in the range of a few meV/at) at 20 GPa. The γ phase being the most stable however.

Perspectives for Ternary Superhydride Predictions

As said, the absence or presence of a transition pressure is evaluated by a SVM model employed for a binary classification. This model is trained on all the data related to the pure alloys forming binary (super)hydrides. Electronegativity and crystal radius of the elements forming the alloys are taken as features. Each alloy is associated to a 1 label if it forms a superhydride (ie H/M ratio above five), 0 otherwise (H/M ratio strictly below five). The kernel used for SVM is a gaussian one. A scale parameter has to be fixed (noted C). It has to be determined with the γ one (see [212]). The figure G.1 presents the mean accuracy's evolution with these two parameters. For each set of (C, γ) , a cross validation process is realized. The training database is splitted twenty times with 90% of the data for training, the rest for testing. The best mean accuracy of 0.96 is reached for $C=10$ and $\gamma=14.38$.

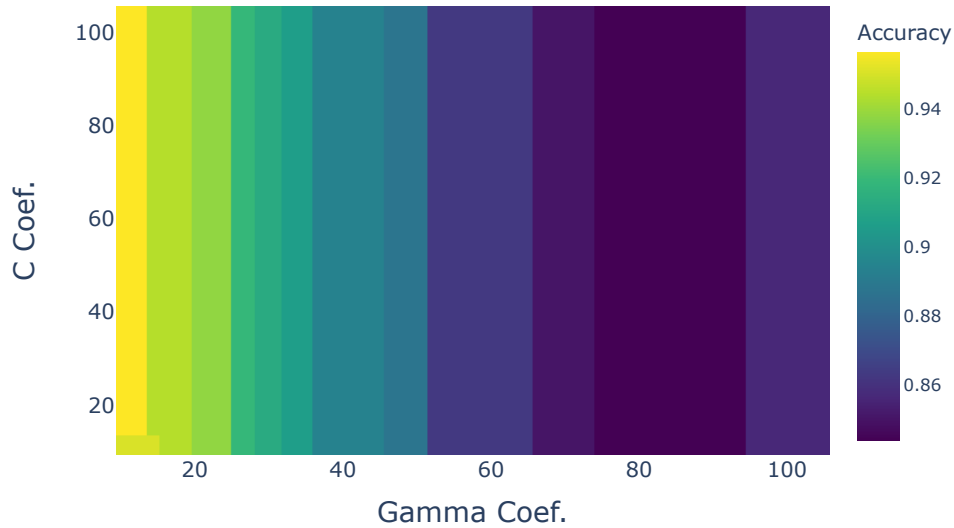


Figure G.1: Cross validation to determine SVM's hyperparameters

When a transition pressure exists (i.e. an alloy which forms a superhydride and thus associated to a label of 1 by the SVM classifier), a GPR model (called GPR_0) is employed to estimate a value of it. This GPR model also uses a gaussian type kernel characterized by a scale parameter. To realize this estimation, it is trained on the subdatabase gathering the elements having a transition pressure (i.e. forming a superhydride). Similarly as the SVM classifier, both electronegativity and crystal radius are taken as features. To assess the scale parameter entering in the kernel's definition, the distances' distribution of the data points in the plan electronegativity and crystal radius is computed. This distribution is represented in FIG G.2. In addition, it shows the integration of this histogram and a normalization of it. From this integration, a distance range is fixed so as to have between five and twenty percent of the distances. This range is the interval between 0.1

and 0.3.

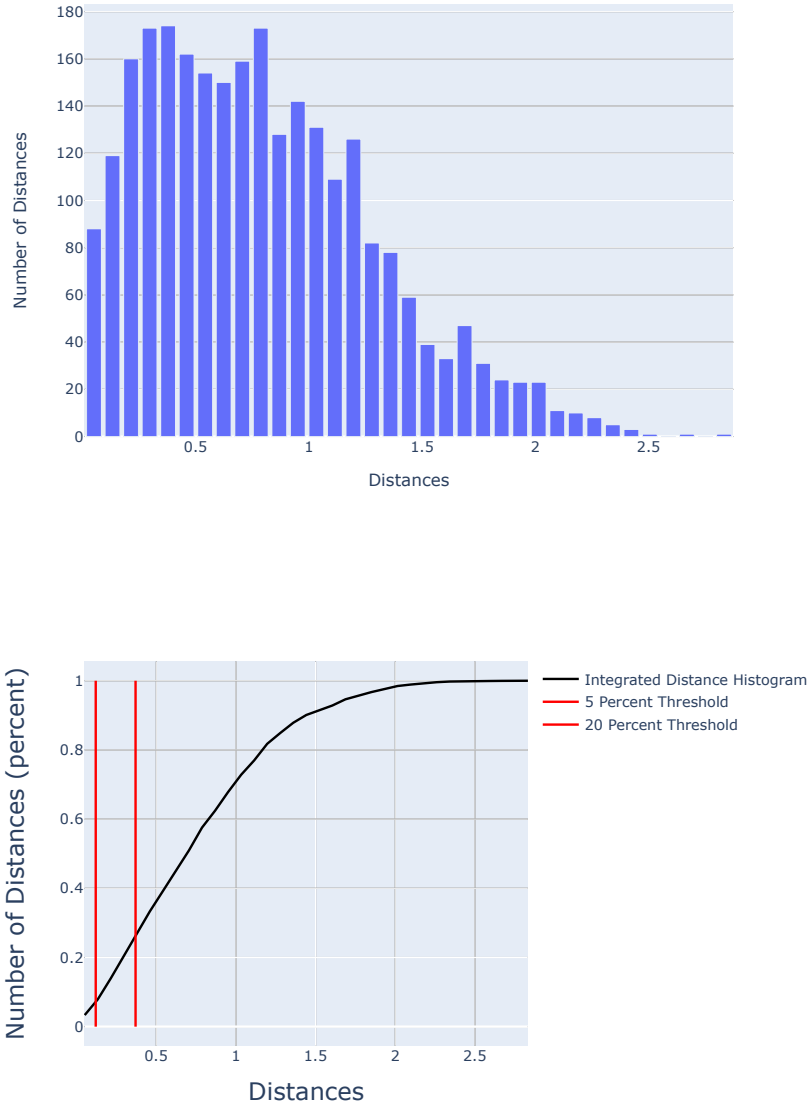


Figure G.2: Distances' distribution of the data points in the plan electronegativity and crystal radius

A cross validation process is enforced in this range to assess the best possible σ parameters. For each value of sigma, 30 datasets are randomly generated where 90 % of the data are used for training, the rest for testing. The FIG G.3 shows how the mean error values evolve on both training and testing as a function of the σ parameter. It can be seen that the σ parameter has to be fixed above 0.18.

Model	Parameters	Label	Mean Error or Accuracy	Purpose
SVM (See S.1)	$C=14.38, \gamma=10$	(0,1)	0.97	Form a superhydride or not (Classification)
GPR ₀ (See S.2)	$\sigma=0.2$	Transition Pressure	18 GPa	Transition pressure values (Regression)

Table G.1: Hyperparameters of both SVM and GPR models

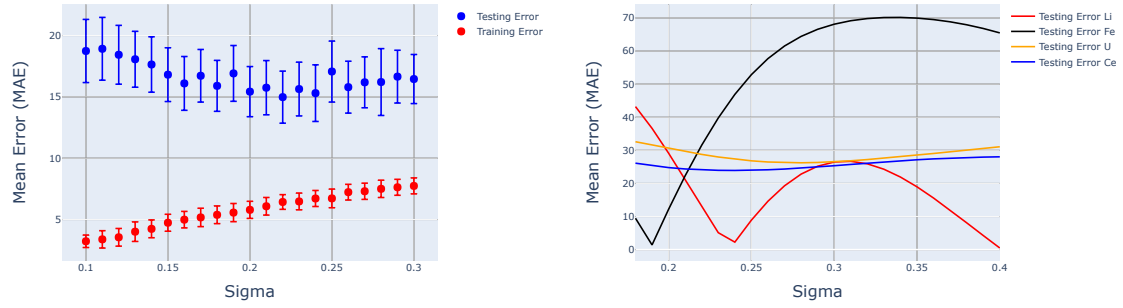


Figure G.3: Up: Mean testing (blue) and training (red) error evolution for the transition pressure prediction, as a function of σ . Down: Error evolution for the transition pressure prediction regarding U, Li, Fe and Ce systems.

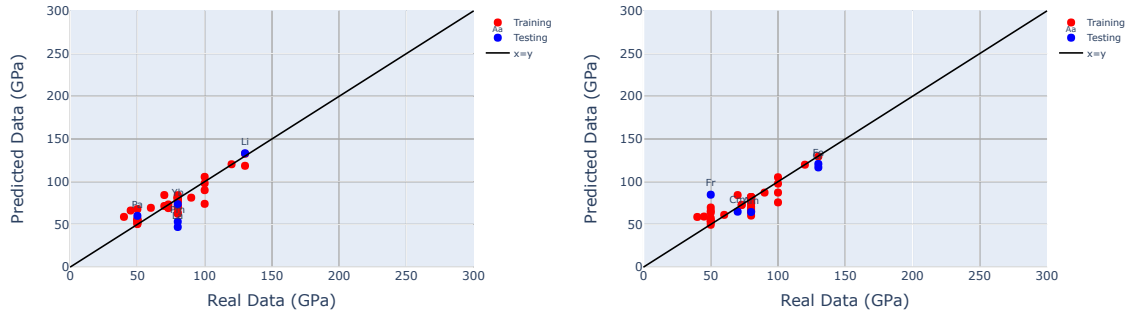


Figure G.4: Correlations between real transition pressures and those predicted by the GPR_0 model. The red points correspond to data used for training, the blue ones for testing.

To determine a σ value between 0.18 and 0.3, some known systems (Fe, Li, U, Ce) for which transition pressures were evaluated both theoretically and experimentally are used (respectively 130 GPa [12], 120 GPa [18], 36 GPa [38], 80 GPa [14]). The error made for these systems as a function of σ is computed. It can be seen that a σ value of 0.2 enables a good compromise regarding the error made of each of these compounds. With this value of σ parameter, the error for the transition pressure prediction is in the range of 14-16 GPa (see FIG G.3 Up). The FIG G.4 shows the predictions made on random training (red) and testing (blue) sets. The table G.1 presents the different models' hyperparameters finally selected.

Hydride and Superhydride Database

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
LiH	Li	Alc.	0.0	1.0	130	0.9	0.9	[18, 17, 217]
LiH ₂	Li	Alc.	150.0	2.0	130	0.9	0.9	[18, 17, 217]
LiH ₂	Li	Alc.	200.0	2.0	130	0.9	0.9	[18, 17, 217]
LiH ₆	Li	Alc.	150.0	6.0	130	0.9	0.9	[18, 17, 217]
LiH ₆	Li	Alc.	200.0	6.0	130	0.9	0.9	[18, 17, 217]
LiH ₈	Li	Alc.	200.0	8.0	130	0.9	0.9	[18, 17, 217]
BeH ₂	Be	Alc. Terr.	0.0	2.0	None	1.5	0.4	[218]
BeH ₂	Be	Alc. Terr.	200.0	2.0	None	1.5	0.4	[218]
BH	B	Metalloid	50.0	1.0	None	2.0	0.2	[219]
BH	B	Metalloid	100.0	1.0	None	2.0	0.2	[219]
BH	B	Metalloid	100.0	1.0	None	2.0	0.2	[219]
B ₁₀₀ H ₁₄	B	Metalloid	50.0	1.4	None	2.0	0.2	[220]
B ₁₀₀ H ₁₄	B	Metalloid	100.0	1.4	None	2.0	0.2	[220]
B ₄ H ₁₀	B	Metalloid	50.0	2.5	None	2.0	0.2	[221]
BH ₃	B	Metalloid	20.0	3.0	None	2.0	0.2	[219]
BH ₃	B	Metalloid	50.0	3.0	None	2.0	0.2	[219]
BH ₃	B	Metalloid	100.0	3.0	None	2.0	0.2	[219]
BH ₃	B	Metalloid	150.0	3.0	None	2.0	0.2	[219]
NaH	Na	Alc.	0.0	1.0	80	0.9	1.1	[23]
Na ₂ H ₃	Na	Alc.	80.0	1.5	80	0.9	1.1	[23]
Na ₃ H ₅	Na	Alc.	80.0	1.6	80	0.9	1.1	[23]
NaH ₃	Na	Alc.	80.0	3.0	80	0.9	1.1	[23]
NaH ₇	Na	Alc.	80.0	7.0	80	0.9	1.1	[23]
NaH ₉	Na	Alc.	80.0	9.0	80	0.9	1.1	[23]
MgH ₂	Mg	Alc. Terr.	0.0	2.0	80	1.3	0.8	[222]
MgH ₂	Mg	Alc. Terr.	50.0	2.0	80	1.3	0.8	[222]
MgH ₄	Mg	Alc. Terr.	80.0	4.0	80	1.3	0.8	[222]
MgH ₄	Mg	Alc. Terr.	150.0	4.0	80	1.3	0.8	[222]
MgH ₆	Mg	Alc. Terr.	180.0	6.0	80	1.3	0.8	[222]
AlH ₃	Al	Other	0.0	3.0	73	1.6	0.5	[223]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
AlH ₅	Al	Other	73.0	3.0	73	1.6	0.5	[223]
SiH ₄	Si	Metalloid	10.0	4.0	None	1.9	0.4	[224]
SiH ₄	Si	Metalloid	0.0	4.0	None	1.9	0.4	[224]
SiH ₄	Si	Metalloid	90.0	4.0	None	1.9	0.4	[225]
PH	P	No Metal	100.0	1.0	None	2.1	0.3	[217]
PH ₂	P	No Metal	200.0	2.0	None	2.1	0.3	[217]
PH ₃	P	No Metal	80.0	3.0	None	2.1	0.3	[217]
PH ₃	P	No Metal	100.0	3.0	None	2.1	0.3	[217]
PH ₃	P	No Metal	200.0	3.0	None	2.1	0.3	[217]
SH ₂	S	No Metal	17.0	2.0	None	2.5	0.4	[226]
SH ₂	S	No Metal	0.0	2.0	None	2.5	0.4	[226]
SH ₃	S	No Metal	150.0	3.0	None	2.5	0.4	[226]
ClH	Cl	Hal.	51.0	1.0	None	3.1	1.6	[227]
ClH	Cl	Hal.	150.0	1.0	None	3.1	1.6	[227]
ClH	Cl	Hal.	0.0	1.0	None	3.1	1.6	[227]
ClH ₂	Cl	Hal.	200.0	2.0	None	3.1	1.6	[228]
ClH ₂	Cl	Hal.	100.0	2.0	None	3.1	1.6	[228]
KH ₅	K	Alc.	100.0	5.0	100	0.8	1.5	[19]
KH ₅	K	Alc.	100.0	5.0	100	0.8	1.5	[19]
KH ₅	K	Alc.	150.0	5.0	100	0.8	1.5	[19]
KH ₆	K	Alc.	150.0	6.0	100	0.8	1.5	[229]
KH ₉	K	Alc.	150.0	9.0	100	0.8	1.5	[19]
KH ₉	K	Alc.	100.0	9.0	100	0.8	1.5	[19]
KH ₉	K	Alc.	50.0	9.0	100	0.8	1.5	[19]
KH ₁₀	K	Alc.	150.0	10.	100	0.8	1.5	[217]
KH ₁₂	K	Alc.	150.0	12.	100	0.8	1.5	[217]
CaH ₂	Ca	Alc. Terr.	0.0	2.0	90	1.0	1.1	[20]
CaH ₄	Ca	Alc. Terr.	50.0	4.0	90	1.0	1.1	[20]
CaH ₉	Ca	Alc. Terr.	100.0	9.0	90	1.0	1.1	[20]
CaH ₁₂	Ca	Alc. Terr.	100.0	12.	90	1.0	1.1	[20]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (Å)	Ref
ScH ₃	Sc	Trans.	0.0	3.0	80	1.3	0.8	[230]
ScH ₃	Sc	Trans.	100.0	4.0	80	1.3	0.8	[230]
ScH ₄	Sc	Trans.	150.0	4.0	80	1.3	0.8	[230]
ScH ₆	Sc	Trans.	150.0	6.0	80	1.3	0.8	[230]
ScH ₇	Sc	Trans.	150.0	7.0	80	1.3	0.8	[230]
ScH ₈	Sc	Trans.	150.0	8.0	80	1.3	0.8	[230]
ScH ₉	Sc	Trans.	150.0	9.0	80	1.3	0.8	[230]
ScH ₁₂	Sc	Trans.	150.0	12.	80	1.3	0.8	[230]
TiH ₂	Ti	Trans.	0.0	2.0	80	1.5	0.7	[231]
Ti ₂ H ₅	Ti	Trans.	50.0	2.5	80	1.5	0.7	[231]
Ti ₅ H ₁₄	Ti	Trans.	50.0	2.8	80	1.5	0.7	[231]
TiH ₃	Ti	Trans.	70.0	3.0	80	1.5	0.7	[231]
TiH ₁₂	Ti	Trans.	100.0	12.	80	1.5	0.7	[231]
TiH ₁₄	Ti	Trans.	150.0	14.	80	1.5	0.7	[231]
V ₂ H	V	Trans.	0.0	0.5	100	1.6	0.6	[232]
V ₂ H	V	Trans.	100.0	0.5	100	1.6	0.6	[232]
VH	V	Trans.	100.0	1.0	100	1.6	0.6	[232]
VH	V	Trans.	200.0	1.0	100	1.6	0.6	[232]
VH ₂	V	Trans.	200.0	2.0	100	1.6	0.6	[232]
VH ₂	V	Trans.	200.0	2.0	100	1.6	0.6	[232]
V ₂ H ₅	V	Trans.	100.0	2.5	100	1.6	0.6	[232]
V ₂ H ₅	V	Trans.	200.0	2.5	100	1.6	0.6	[232]
VH ₃	V	Trans.	200.0	3.0	100	1.6	0.6	[232]
VH ₅	V	Trans.	100.0	5.0	100	1.6	0.6	[232]
VH ₅	V	Trans.	200.0	5.0	100	1.6	0.6	[232]
VH ₈	V	Trans.	200.0	8.0	100	1.6	0.6	[232]
CrH	Cr	Trans.	0.0	1.0	130	1.6	0.7	[115]
Cr ₂ H ₃	Cr	Trans.	20.0	1.5	130	1.6	0.7	[115]
CrH ₂	Cr	Trans.	20.0	1.0	130	1.6	0.7	[115]
CrH ₃	Cr	Trans.	81.0	3.0	130	1.6	0.7	[115]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
CrH ₃	Cr	Trans.	90.0	3.0	130	1.6	0.7	[115]
CrH ₄	Cr	Trans.	100.0	4.0	130	1.6	0.7	[115]
CrH ₄	Cr	Trans.	150.0	4.0	130	1.6	0.7	[115]
CrH ₄	Cr	Trans.	200.0	4.0	130	1.6	0.7	[115]
CrH ₈	Cr	Trans.	132.0	8.0	130	1.6	0.7	[115]
CrH ₈	Cr	Trans.	200.0	8.0	130	1.6	0.7	[115]
Mn ₂ H	Mn	Trans.	0.0	0.5	45	1.5	0.8	[134]
Mn ₂ H	Mn	Trans.	15.0	0.5	45	1.5	0.8	[134]
MnH	Mn	Trans.	15.0	1.0	45	1.5	0.8	[134]
MnH ₂	Mn	Trans.	44.0	2.0	45	1.5	0.8	[134]
MnH ₇	Mn	Trans.	44.0	7.0	45	1.5	0.8	[134]
MnH ₂	Mn	Trans.	117.0	2.0	45	1.5	0.8	[134]
MnH ₃	Mn	Trans.	117.0	3.0	45	1.5	0.8	[134]
MnH ₇	Mn	Trans.	117.0	7.0	45	1.5	0.8	[134]
MnH ₂	Mn	Trans.	147.0	2.0	45	1.5	0.8	[134]
MnH ₃	Mn	Trans.	147.0	3.0	45	1.5	0.8	[134]
MnH ₇	Mn	Trans.	147.0	7.0	45	1.5	0.8	[134]
MnH ₈	Mn	Trans.	150.0	8.0	45	1.5	0.8	[134]
FeH	Fe	Trans.	0.0	1.0	130	1.8	0.6	[12]
FeH ₂	Fe	Trans.	67.0	2.0	130	1.8	0.6	[12]
FeH ₃	Fe	Trans.	87.0	3.0	130	1.8	0.6	[12]
FeH ₅	Fe	Trans.	130.0	5.0	130	1.8	0.6	[12]
FeH ₆	Fe	Trans.	150.0	6.0	130	1.8	0.6	[233]
CoH	Co	Trans.	10.0	1.0	None	1.8	0.5	[118]
CoH ₂	Co	Trans.	10.0	2.0	None	1.8	0.5	[118]
CoH ₃	Co	Trans.	30.0	3.0	None	1.8	0.5	[118]
CoH ₃	Co	Trans.	300.0	3.0	None	1.8	0.5	[118]
NiH	Ni	Trans.	0.0	1.0	None	1.9	0.7	[120]
Ni ₂ H ₃	Ni	Trans.	60.0	1.5	None	1.9	0.7	[120]
NiH ₂	Ni	Trans.	60.0	2.0	None	1.9	0.7	[120]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
Cu ₂ H	Cu	Trans.	0.0	0.5	Unk	1.9	0.7	[135]
CuH	Cu	Trans.	30.0	1.0	Unk	1.9	0.7	[135]
Cu ₂ H	Cu	Trans.	17.0	0.5	Unk	1.9	0.7	[135]
CuH ₂	Cu	Trans.	40.0	2.0	Unk	1.9	0.7	[135]
CuH ₃	Cu	Trans.	40.0	3.0	Unk	1.9	0.7	[135]
ZnH ₂	Zn	Trans.	0.0	2.0	Unk	1.6	0.7	[234]
GaH	Ga	Other	0.0	1.0	Unk	1.8	0.7	[235]
GaH	Ga	Other	0.0	1.0	Unk	1.8	0.7	[235]
GaH ₂	Ga	Other	0.0	2.0	Unk	1.8	0.7	[235]
GaH ₃	Ga	Other	5.0	3.0	Unk	1.8	0.7	[235]
GaH ₃	Ga	Other	160.0	3.0	Unk	1.8	0.7	[235]
GeH ₄	Ge	Metalloid	0.0	4.0	None	2.0	0.5	[236]
GeH ₄	Ge	Metalloid	72.0	4.0	None	2.0	0.5	[236]
AsH ₃	As	Metalloid	0.0	3.0	None	2.1	0.7	[237]
SeH ₂	Se	No Metal	0.0	2.0	Unk	2.5	0.5	[238]
SeH ₃	Se	No Metal	5.0	3.0	Unk	2.5	0.5	[238]
BrH	Br	Hal.	0.0	1.0	Unk	2.9	1.8	[]
RbH	Rb	Alc.	0.0	1.0	80	0.8	1.6	[239]
RbH ₃	Rb	Alc.	100.0	3.0	80	0.8	1.6	[239]
RbH ₅	Rb	Alc.	80.0	5.0	80	0.8	1.6	[239]
RbH ₉	Rb	Alc.	100.0	9.0	80	0.8	1.6	[239]
RbH ₁₄	Rb	Alc.	200.0	14.	80	0.8	1.6	[239]
SrH ₂	Sr	Alc. Terr.	0.0	2.0	50	0.9	1.3	[240]
SrH ₂	Sr	Alc. Terr.	20.0	2.0	50	0.9	1.3	[240]
SrH ₄	Sr	Alc. Terr.	50.0	4.0	50	0.9	1.3	[240]
SrH ₁₂	Sr	Alc. Terr.	50.0	12.	50	0.9	1.3	[240]
SrH ₁₂	Sr	Alc. Terr.	200.0	12.	50	0.9	1.3	[240]
YH ₂	Y	Trans.	0.0	2.0	80	1.2	1.0	[98]
YH ₂	Y	Trans.	150.0	2.0	80	1.2	1.0	[98]
YH ₃	Y	Trans.	10.0	3.0	80	1.2	1.0	[98]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
YH ₃	Y	Trans.	50.0	3.0	80	1.2	1.0	[98]
YH ₃	Y	Trans.	150.0	3.0	80	1.2	1.0	[98]
YH ₄	Y	Trans.	150.0	4.0	80	1.2	1.0	[98]
YH ₆	Y	Trans.	200.0	6.0	80	1.2	1.0	[98]
YH ₇	Y	Trans.	150.0	7.0	80	1.2	1.0	[241]
YH ₇	Y	Trans.	117.0	7.0	80	1.2	1.0	[241]
YH ₇	Y	Trans.	176.0	7.0	80	1.2	1.0	[241]
YH ₁₂	Y	Trans.	180.0	12.	80	1.2	1.0	[98]
ZrH	Zr	Trans.	0.0	1.0	100	1.3	0.8	[242]
ZrH	Zr	Trans.	50.0	1.0	100	1.3	0.8	[242]
ZrH	Zr	Trans.	100.0	1.0	100	1.3	0.8	[242]
ZrH ₂	Zr	Trans.	0.0	2.0	100	1.3	0.8	[242]
ZrH ₂	Zr	Trans.	50.0	2.0	100	1.3	0.8	[242]
ZrH ₂	Zr	Trans.	100.0	2.0	100	1.3	0.8	[242]
ZrH ₃	Zr	Trans.	50.0	3.0	100	1.3	0.8	[242]
ZrH ₃	Zr	Trans.	100.0	3.0	100	1.3	0.8	[242]
Zr ₄ H ₁₅	Zr	Trans.	50.0	3.7	100	1.3	0.8	[242]
Zr ₄ H ₁₅	Zr	Trans.	100.0	3.7	100	1.3	0.8	[242]
ZrH ₄	Zr	Trans.	110.0	4.0	100	1.3	0.8	[243]
ZrH ₆	Zr	Trans.	150.0	6.0	100	1.3	0.8	[243]
NbH ₂	Nb	Trans.	0.0	2.0	100	1.6	0.7	[244]
NbH ₃	Nb	Trans.	50.0	3.0	100	1.6	0.7	[244]
NbH ₄	Nb	Trans.	200.0	4.0	100	1.6	0.7	[244]
NbH ₆	Nb	Trans.	100.0	4.0	100	1.6	0.7	[244]
MoH	Mo	Trans.	0.0	1.0	None	2.1	0.7	[245]
MoH	Mo	Trans.	20.0	1.0	None	2.1	0.7	[245]
MoH	Mo	Trans.	100.0	1.0	None	2.1	0.7	[245]
MoH ₂	Mo	Trans.	80.0	2.0	None	2.1	0.7	[245]
MoH ₃	Mo	Trans.	100.0	3.0	None	2.1	0.7	[245]
TcH	Tc	Trans.	50.0	1.0	None	1.9	0.7	[246]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
TcH ₂	Tc	Trans.	60.0	2.0	None	1.9	0.7	[246]
TcH ₃	Tc	Trans.	80.0	3.0	None	1.9	0.7	[246]
TcH ₃	Tc	Trans.	300.0	3.0	None	1.9	0.7	[246]
RuH	Ru	Trans.	100.0	1.0	None	2.2	0.8	[247]
RuH	Ru	Trans.	7.0	1.0	None	2.2	0.8	[247]
RuH ₃	Ru	Trans.	66.0	3.0	None	2.2	0.8	[247]
RuH ₃	Ru	Trans.	100.0	3.0	None	2.2	0.8	[247]
RuH ₃	Ru	Trans.	200.0	3.0	None	2.2	0.8	[247]
RhH	Rh	Trans.	0.0	1.0	None	2.2	0.8	[248]
PdH	Pd	Trans.	0.0	1.0	None	2.2	0.7	[249]
AgH	Ag	Trans.	0.0	1.0	None	1.9	1.2	[250]
CdH	Cd	Trans.	0.0	1.0	Unk	1.6	0.9	[]
InH ₃	In	Other	0.0	3.0	120	1.7	0.9	[251]
InH ₅	In	Other	120.0	5.0	120	1.7	0.9	[251]
SnH ₄	Sn	Other	0.0	4.0	None	1.9	0.6	[252]
SnH ₄	Sn	Other	108.0	4.0	None	1.9	0.6	[252]
SbH	Sb	Other	100.0	1.0	None	2.0	0.9	[253]
SbH ₃	Sb	Other	0.0	3.0	None	2.0	0.9	[253]
SbH ₃	Sb	Other	300.0	3.0	None	2.0	0.9	[253]
SbH ₄	Sb	Other	150.0	4.0	None	2.0	0.9	[253]
TeH ₂	Te	Metalloid	0.0	2.0	None	2.1	1.1	[254]
TeH ₄	Te	Metalloid	170.0	4.0	None	2.1	1.1	[254]
IH ₂	I	Hal.	100.0	2.0	None	2.6	2.0	[255]
IH ₂	I	Hal.	150.0	2.0	None	2.6	2.0	[255]
IH ₂	I	Hal.	200.0	2.0	None	2.6	2.0	[255]
IH ₄	I	Hal.	100.0	4.0	None	2.6	2.0	[255]
IH ₄	I	Hal.	150.0	4.0	None	2.6	2.0	[255]
IH ₄	I	Hal.	200.0	4.0	None	2.6	2.0	[255]
CsH	Cs	Alc.	0.0	1.0	80	0.2	1.8	[256]
CsH ₃	Cs	Alc.	150.0	3.0	80	0.2	1.8	[256]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
CsH ₃	Cs	Alc.	30.0	3.0	80	0.2	1.8	[256]
CsH ₃	Cs	Alc.	100.0	3.0	80	0.2	1.8	[256]
CsH ₅	Cs	Alc.	150.0	5.0	80	0.2	1.8	[256]
CsH ₇	Cs	Alc.	80.0	7.0	80	0.2	1.8	[256]
CsH ₇	Cs	Alc.	150.0	7.0	80	0.2	1.8	[256]
CsH ₁₄	Cs	Alc.	150.0	14.	80	0.2	1.8	[256]
BaH ₂	Ba	Alc. Terr.	10.0	2.0	73	0.8	1.4	[257]
BaH ₁₂	Ba	Alc. Terr.	73.0	6.0	73	0.8	1.4	[257]
BaH ₁₀	Ba	Alc. Terr.	150.0	8.0	73	0.8	1.4	[257]
BaH ₁₂	Ba	Alc. Terr.	100.0	8.0	73	0.8	1.4	[257]
BaH ₁₂	Ba	Alc. Terr.	150.0	12.	73	0.8	1.4	[257]
LuH ₃	Lu	Trans.	0.0	3.0	80	1.2	1.0	[217]
LuH ₃	Lu	Trans.	50.0	3.0	80	1.2	1.0	[217]
LuH ₃	Lu	Trans.	150.0	3.0	80	1.2	1.0	[217]
LuH ₇	Lu	Trans.	150.0	7.0	80	1.2	1.0	[217]
LuH ₁₂	Lu	Trans.	150.0	12.	80	1.2	1.0	[217]
LuH ₁₃	Lu	Trans.	150.0	13.	80	1.2	1.0	[217]
HfH ₂	Hf	Trans.	0.0	1.0	70	1.3	0.8	[217]
HfH ₂	Hf	Trans.	300.0	2.0	70	1.3	0.8	[217]
Hf ₄ H ₁₅	Hf	Trans.	13.0	3.7	70	1.3	0.8	[258]
Hf ₃ H ₁₃	Hf	Trans.	100.0	4.3	70	1.3	0.8	[217]
HfH ₆	Hf	Trans.	200.0	6.0	70	1.3	0.8	[217]
Ta ₅ H	Ta	Trans.	0.0	0.2	50	1.5	0.7	[259]
Ta ₂ H	Ta	Trans.	0.0	0.5	50	1.5	0.7	[259]
TaH	Ta	Trans.	0.0001	1.0	50	1.5	0.7	[259]
TaH ₂	Ta	Trans.	0.0001	2.0	50	1.5	0.7	[259]
TaH ₂	Ta	Trans.	200.0	2.0	50	1.5	0.7	[259]
TaH ₃	Ta	Trans.	50.0	3.0	50	1.5	0.7	[259]
TaH ₄	Ta	Trans.	50.0	4.0	50	1.5	0.7	[259]
TaH ₄	Ta	Trans.	250.0	4.0	50	1.5	0.7	[259]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
TaH ₆	Ta	Trans.	300.0	6.0	50	1.5	0.7	[259]
WH	W	Trans.	100.0	1.0	None	2.3	0.7	[260]
WH ₂	W	Trans.	100.0	2.0	None	2.3	0.7	[260]
OsH	Os	Trans.	100.0	1.0	None	2.2	0.7	[261]
OsH	Os	Trans.	150.0	1.0	None	2.2	0.7	[261]
OsH ₃	Os	Trans.	150.0	3.0	None	2.2	0.7	[261]
IrH ₂	Ir	Trans.	100.0	2.0	None	2.2	0.7	[262]
IrH ₃	Ir	Trans.	50.0	3.0	None	2.2	0.7	[262]
IrH ₃	Ir	Trans.	100.0	3.0	None	2.2	0.7	[262]
Pt ₂ H	Pt	Trans.	100.0	0.5	None	2.2	0.7	[263]
PtH	Pt	Trans.	100.0	1.0	None	2.2	0.7	[263]
PtH ₂	Pt	Trans.	100.0	2.0	None	2.2	0.7	[263]
AuH	Au	Other	100.0	1.0	None	2.5	1.5	[128]
Pb	Pb	Other	0.0	0.0	None	2.3	1.4	[264]
Pb	Pb	Other	80.0	0.0	None	2.3	1.4	[264]
PbH ₄	Pb	Other	132.0	4.0	None	2.3	1.4	[264]
BiH	Bi	Other	250.0	1.0	None	2.0	1.1	[265]
BiH ₂	Bi	Other	150.0	2.0	None	2.0	1.1	[265]
BiH ₃	Bi	Other	0.0	3.0	None	2.0	1.1	[265]
BiH ₄	Bi	Other	150.0	4.0	None	2.0	1.1	[265]
PoH	Po	Metalloid	200.0	1.0	None	2.0	1.0	[266]
PoH ₂	Po	Metalloid	0.0	2.0	None	2.0	1.0	[266]
PoH ₂	Po	Metalloid	100.0	2.0	None	2.0	1.0	[266]
PoH ₂	Po	Metalloid	100.0	2.0	None	2.0	1.0	[266]
PoH ₄	Po	Metalloid	150.0	4.0	None	2.0	1.0	[266]
PoH ₄	Po	Metalloid	200.0	4.0	None	2.0	1.0	[266]
FrH	Fr	Alc.	0.0	1.0	50	0.7	1.9	[217]
FrH ₇	Fr	Alc.	50.0	7.0	50	0.7	1.9	[217]
FrH ₇	Fr	Alc.	100.0	7.0	50	0.7	1.9	[217]
FrH ₁₃	Fr	Alc.	50.0	13.	50	0.7	1.9	[217]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
FrH ₁₅	Fr	Alc.	150.0	15.	50	0.7	1.9	[217]
RaH	Ra	Alc. Terr.	0.0	1.0	50	0.8	1.6	[217]
RaH ₁₀	Ra	Alc. Terr.	50.0	10.	50	0.8	1.6	[217]
RaH ₁₀	Ra	Alc. Terr.	150.0	10.	50	0.8	1.6	[217]
RaH ₁₂	Ra	Alc. Terr.	150.0	12.	50	0.8	1.6	[217]
RaH ₁₂	Ra	Alc. Terr.	200.0	12.	50	0.8	1.6	[217]
LaH ₂	La	Lanthanide	0.0	2.0	80	1.1	1.3	[98]
LaH ₂	La	Lanthanide	50.0	2.0	80	1.1	1.3	[98]
LaH ₃	La	Lanthanide	50.0	3.0	80	1.1	1.3	[98]
LaH ₃	La	Lanthanide	150.0	3.0	80	1.1	1.3	[98]
LaH ₄	La	Lanthanide	150.0	4.0	80	1.1	1.3	[98]
LaH ₅	La	Lanthanide	150.0	5.0	80	1.1	1.3	[98]
LaH ₈	La	Lanthanide	150.0	8.0	80	1.1	1.3	[98]
LaH ₁₀	La	Lanthanide	150.0	10.	80	1.1	1.3	[98]
CeH ₃	Ce	Lanthanide	0.0	3.0	50	1.1	1.1	[14]
CeH ₃	Ce	Lanthanide	50.0	3.0	50	1.1	1.1	[14]
CeH ₄	Ce	Lanthanide	50.0	4.0	50	1.1	1.1	[14]
CeH ₆	Ce	Lanthanide	50.0	6.0	50	1.1	1.1	[14]
CeH ₄	Ce	Lanthanide	100.0	4.0	50	1.1	1.1	[14]
CeH ₄	Ce	Lanthanide	200.0	4.0	50	1.1	1.1	[14]
CeH ₈	Ce	Lanthanide	100.0	8.0	50	1.1	1.1	[14]
CeH ₄	Ce	Lanthanide	150.0	4.0	50	1.1	1.1	[14]
CeH ₉	Ce	Lanthanide	150.0	9.0	50	1.1	1.1	[14]
CeH ₉	Ce	Lanthanide	200.0	9.0	50	1.1	1.1	[14]
CeH ₁₀	Ce	Lanthanide	200.0	10.	50	1.1	1.1	[14]
PrH	Pr	Lanthanide	0.0	1.0	50	1.1	1.3	[128]
PrH	Pr	Lanthanide	20.0	1.0	50	1.1	1.3	[128]
PrH	Pr	Lanthanide	50.0	1.0	50	1.1	1.3	[128]
Pr ₃ H ₁₀	Pr	Lanthanide	50.0	3.3	50	1.1	1.3	[128]
PrH ₃	Pr	Lanthanide	50.0	3.0	50	1.1	1.3	[128]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
PrH ₄	Pr	Lanthanide	50.0	4.0	50	1.1	1.3	[128]
PrH ₇	Pr	Lanthanide	50.0	7.0	50	1.1	1.3	[128]
PrH ₇	Pr	Lanthanide	150.0	7.0	50	1.1	1.3	[128]
PrH ₈	Pr	Lanthanide	150.0	8.0	50	1.1	1.3	[128]
PrH ₉	Pr	Lanthanide	150.0	9.0	50	1.1	1.3	[128]
NdH	Nd	Lanthanide	150.0	1.0	50	1.1	1.3	[267]
Nd ₄ H ₇	Nd	Lanthanide	150.0	1.7	50	1.1	1.3	[267]
NdH ₃	Nd	Lanthanide	50.0	3.0	50	1.1	1.3	[267]
NdH ₄	Nd	Lanthanide	100.0	4.0	50	1.1	1.3	[267]
NdH ₇	Nd	Lanthanide	50.0	7.0	50	1.1	1.3	[267]
NdH ₈	Nd	Lanthanide	150.0	8.0	50	1.1	1.3	[267]
NdH ₉	Nd	Lanthanide	150.0	9.0	50	1.1	1.3	[267]
NdH ₉	Nd	Lanthanide	100.0	9.0	50	1.1	1.3	[267]
PmH ₃	Pm	Lanthanide	0.0	3.0	80	1.1	1.2	[99]
PmH ₄	Pm	Lanthanide	100.0	4.0	80	1.1	1.2	[99]
PmH ₆	Pm	Lanthanide	100.0	6.0	80	1.1	1.2	[99]
PmH ₉	Pm	Lanthanide	200.0	9.0	80	1.1	1.2	[99]
SmH ₃	Sm	Lanthanide	0.0	3.0	80	1.1	1.1	[99]
SmH ₄	Sm	Lanthanide	100.0	4.0	80	1.1	1.1	[99]
SmH ₆	Sm	Lanthanide	100.0	6.0	80	1.1	1.1	[99]
SmH ₉	Sm	Lanthanide	200.0	9.0	80	1.1	1.1	[99]
EuH ₃	Eu	Lanthanide	0.0	3.0	80	1.2	1.3	[99]
EuH ₄	Eu	Lanthanide	100.0	4.0	80	1.2	1.3	[99]
EuH ₉	Eu	Lanthanide	100.0	9.0	80	1.2	1.3	[99]
GdH ₃	Gd	Lanthanide	0.0	3.0	80	1.2	1.0	[99]
GdH ₄	Gd	Lanthanide	100.0	4.0	80	1.2	1.0	[99]
GdH ₉	Gd	Lanthanide	100.0	9.0	80	1.2	1.0	[99]
TbH ₃	Tb	Lanthanide	0.0	3.0	80	1.1	1.1	[99]
TbH ₆	Tb	Lanthanide	100.0	6.0	80	1.1	1.1	[99]
TbH ₉	Tb	Lanthanide	100.0	9.0	80	1.1	1.1	[99]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
DyH ₃	Dy	Lanthanide	0.0	3.0	80	1.2	1.0	[99]
DyH ₄	Dy	Lanthanide	100.0	4.0	80	1.2	1.0	[99]
DyH ₆	Dy	Lanthanide	100.0	6.0	80	1.2	1.0	[99]
DyH ₉	Dy	Lanthanide	100.0	9.0	80	1.2	1.0	[99]
HoH	Ho	Lanthanide	150.0	1.0	80	1.2	1.0	[217]
HoH ₃	Ho	Lanthanide	0.0	1.0	80	1.2	1.0	[217]
HoH ₃	Ho	Lanthanide	50.0	1.0	80	1.2	1.0	[217]
HoH ₃	Ho	Lanthanide	150.0	3.0	80	1.2	1.0	[217]
HoH ₄	Ho	Lanthanide	150.0	4.0	80	1.2	1.0	[217]
HoH ₇	Ho	Lanthanide	150.0	7.0	80	1.2	1.0	[217]
ErH ₃	Er	Lanthanide	0.0	3.0	80	1.2	1.0	[99]
ErH ₄	Er	Lanthanide	100.0	4.0	80	1.2	1.0	[99]
ErH ₆	Er	Lanthanide	100.0	6.0	80	1.2	1.0	[99]
ErH ₉	Er	Lanthanide	200.0	9.0	80	1.2	1.0	[99]
TmH ₃	Tm	Lanthanide	0.0	3.0	80	1.2	1.0	[99]
TmH ₄	Tm	Lanthanide	100.0	4.0	80	1.2	1.0	[99]
TmH ₆	Tm	Lanthanide	100.0	6.0	80	1.2	1.0	[99]
TmH ₉	Tm	Lanthanide	200.0	9.0	80	1.2	1.0	[99]
YbH ₃	Yb	Lanthanide	0.0	3.0	80	1.1	1.1	[99]
YbH ₄	Yb	Lanthanide	100.0	4.0	80	1.1	1.1	[99]
YbH ₆	Yb	Lanthanide	100.0	6.0	80	1.1	1.1	[99]
AcH	Ac	Actinide	1.0	1.0	50	1.1	1.2	[217]
AcH ₂	Ac	Actinide	60.0	2.0	50	1.1	1.2	[217]
Ac ₃ H ₁₀	Ac	Actinide	60.0	3.3	50	1.1	1.2	[217]
Ac ₃ H ₁₀	Ac	Actinide	200.0	3.3	50	1.1	1.2	[217]
AcH ₃	Ac	Actinide	0.0	3.0	50	1.1	1.2	[217]
AcH ₃	Ac	Actinide	70.0	3.0	50	1.1	1.2	[217]
AcH ₄	Ac	Actinide	60.0	4.0	50	1.1	1.2	[217]
AcH ₅	Ac	Actinide	50.0	5.0	50	1.1	1.2	[217]
AcH ₇	Ac	Actinide	30.0	7.0	50	1.1	1.2	[217]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
AcH ₈	Ac	Actinide	90.0	8.0	50	1.1	1.2	[217]
AcH ₁₀	Ac	Actinide	80.0	10.	50	1.1	1.2	[217]
AcH ₁₀	Ac	Actinide	200.0	10.	50	1.1	1.2	[217]
AcH ₁₂	Ac	Actinide	150.0	12.	50	1.1	1.2	[217]
AcH ₁₆	Ac	Actinide	110.0	16.	50	1.1	1.2	[217]
AcH ₁₆	Ac	Actinide	200.0	16.	50	1.1	1.2	[217]
ThH ₂	Th	Actinide	0.0	2.0	60	1.2	1.1	[37]
ThH ₂	Th	Actinide	15.0	2.0	60	1.2	1.1	[37]
Th ₃ H ₁₀	Th	Actinide	85.0	3.3	60	1.2	1.1	[37]
ThH ₃	Th	Actinide	100.0	3.0	60	1.2	1.1	[37]
ThH ₄	Th	Actinide	85.0	4.0	60	1.2	1.1	[37]
ThH ₄	Th	Actinide	300.0	4.0	60	1.2	1.1	[37]
ThH ₆	Th	Actinide	90.0	6.0	60	1.2	1.1	[37]
ThH ₇	Th	Actinide	115.0	7.0	60	1.2	1.1	[37]
Th ₄ H ₁₅	Th	Actinide	10.0	3.7	60	1.2	1.1	[37]
ThH ₁₀	Th	Actinide	100.0	10.	60	1.2	1.1	[37]
PaH ₃	Pa	Actinide	50.0	3.0	50	1.5	1.0	[217]
PaH ₅	Pa	Actinide	50.0	5.0	50	1.5	1.0	[217]
PaH ₇	Pa	Actinide	50.0	7.0	50	1.5	1.0	[217]
PaH ₄	Pa	Actinide	150.0	4.0	50	1.5	1.0	[217]
PaH ₆	Pa	Actinide	150.0	6.0	50	1.5	1.0	[217]
PaH ₈	Pa	Actinide	150.0	8.0	50	1.5	1.0	[217]
PaH ₉	Pa	Actinide	150.0	9.0	50	1.5	1.0	[217]
UH ₃	U	Actinide	0.0	3.0	40	1.3	0.8	[38]
UH ₇	U	Actinide	36.0	7.0	40	1.3	0.8	[38]
UH ₈	U	Actinide	83.0	8.0	40	1.3	0.8	[38]
UH ₉	U	Actinide	200.0	9.0	40	1.3	0.8	[217]
Am ₂ H ₃	Am	Actinide	50.0	1.5	80	1.3	1.1	[217]
AmH ₂	Am	Actinide	0.0	2.0	80	1.3	1.1	[217]
AmH ₅	Am	Actinide	80.0	5.0	80	1.3	1.1	[217]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
AmH ₈	Am	Actinide	150.0	8.0	80	1.3	1.1	[217]
CmH	Cm	Actinide	0.0	1.0	70	1.3	1.1	[217]
CmH ₃	Cm	Actinide	70.0	3.0	70	1.3	1.1	[217]
CmH ₂	Cm	Actinide	150.0	2.0	70	1.3	1.1	[217]
CmH ₈	Cm	Actinide	150.0	8.0	70	1.3	1.1	[217]
CaYH ₁₂	CaY	Ternary	180.0	6.0	nan	1.1	1.0	[83]
Li ₂ MgH ₄	Li2Mg	Ternary	0.0	1.3	nan	1.0	0.8	[84]
Li ₂ MgH ₁₆	Li2Mg	Ternary	200.0	5.3	nan	1.0	0.8	[84]
MgSiH	MgSi	Ternary	200.0	0.5	nan	1.6	0.7	[268]
MgSiH ₂	MgSi	Ternary	200.0	1.0	nan	1.6	0.7	[268]
MgSiH ₃	MgSi	Ternary	200.0	1.5	nan	1.6	0.7	[268]
MgSiH ₅	MgSi	Ternary	200.0	2.5	nan	1.6	0.7	[268]
MgSiH ₄	MgSi	Ternary	200.0	2.0	nan	1.6	0.7	[268]
MgSiH ₇	MgSi	Ternary	200.0	3.5	nan	1.6	0.7	[268]
MgSiH	MgSi	Ternary	200.0	0.5	nan	1.6	0.7	[268]
MgSiH ₂	MgSi	Ternary	200.0	1.0	nan	1.6	0.7	[268]
MgSiH ₃	MgSi	Ternary	200.0	1.5	nan	1.6	0.7	[268]
MgSiH ₅	MgSi	Ternary	200.0	2.5	nan	1.6	0.7	[268]
MgSiH	MgSi	Ternary	200.0	0.5	nan	1.6	0.7	[268]
Mg ₂ SiH ₅	Mg2Si	Ternary	200.0	1.6	nan	1.5	0.7	[268]
Mg ₂ SiH ₆	Mg2Si	Ternary	200.0	2.0	nan	1.5	0.7	[268]
Mg ₂ SiH ₅	Mg2Si	Ternary	200.0	1.6	nan	1.5	0.7	[268]
Mg ₂ SiH ₆	Mg2Si	Ternary	200.0	2.0	nan	1.5	0.7	[268]
Mg ₂ SiH ₅	Mg2Si	Ternary	200.0	1.6	nan	1.5	0.7	[268]
Mg ₂ SiH ₆	Mg2Si	Ternary	200.0	2.0	nan	1.5	0.7	[268]
MgSiH ₂	MgSi	Ternary	200.0	1.0	nan	nan	nan	[268]
TiPH	TiP	Ternary	50.0	0.5	nan	1.8	0.6	[269]
TiPH	TiP	Ternary	100.0	0.5	nan	1.8	0.6	[269]
TiPH ₄	TiP	Ternary	200.0	2.0	nan	1.8	0.6	[269]
YS ₄ H ₄	YS4	Ternary	200.0	0.8	nan	2.3	0.6	[270]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
YSH ₅	YS	Ternary	200.0	2.5	nan	1.9	0.8	[270]
YSH ₆	YS	Ternary	200.0	3.0	nan	1.9	0.8	[271]
MgCH ₄	MgC	Ternary	121.0	2.0	nan	1.9	0.6	[?]
MgGeH ₆	MgGe	Ternary	200.0	3.0	nan	1.6	0.7	[272]
LaSH ₆	LaS	Ternary	200.0	3.0	nan	1.8	1.0	[271]
YFe ₂ H ₅	YFe2	Ternary	0.0	1.6	nan	1.6	0.8	[176]
YFe ₂ H ₇	YFe2	Ternary	5.0	2.3	nan	1.6	0.8	[273]
YFe ₂ H ₁₀	YFe2	Ternary	25.0	3.3	nan	1.6	0.8	[176]
YFe ₂ H ₇	YFe2	Ternary	25.0	2.3	nan	1.6	0.8	[273]
YFe ₂ H ₆	YFe2	Ternary	25.0	2.0	nan	1.6	0.8	[176]
YFe ₂ H ₅	YFe2	Ternary	25.0	1.6	nan	1.6	0.8	[176]
YFe ₂ H ₁₀	YFe2	Ternary	55.0	3.3	nan	1.6	0.8	[176]
YFe ₂ H ₈	YFe2	Ternary	55.0	2.6	nan	1.6	0.8	[176]
YFe ₂ H ₇	YFe2	Ternary	55.0	2.3	nan	1.6	0.8	[176]
YFe ₂ H ₆	YFe2	Ternary	55.0	2.0	nan	1.6	0.8	[176]
YFe ₂ H ₅	YFe2	Ternary	55.0	1.6	nan	1.6	0.8	[176]
Li ₅ MoH ₁₁	Li5Mo	Ternary	160.0	1.8	nan	1.1	0.8	[274]
NaNiH ₃	NaNi	Ternary	0.0	1.5	nan	1.4	0.9	[275]
KAuH ₂	Kau	Ternary	0.0	1.0	nan	1.6	1.5	[272]
KAuH ₂	Kau	Ternary	30.0	1.0	nan	1.6	1.5	[272]
KAuH ₂	Kau	Ternary	120.0	1.0	nan	1.6	1.5	[272]
BaAu ₂ H ₄	BaAu2	Ternary	0.0	1.3	nan	1.7	1.5	[272]
CsAuH ₂	CsAu	Ternary	0.0	1.0	nan	1.6	1.6	[272]
SrAu ₂ H ₄	SrAu2	Ternary	0.0	1.3	nan	1.7	1.4	[272]
LiPH	LiP	Ternary	50.0	0.5	nan	1.5	0.7	[276]
LiPH ₆	LiP	Ternary	50.0	3.0	nan	1.5	0.7	[276]
LiPH ₇	LiP	Ternary	50.0	3.5	nan	1.5	0.7	[276]
LiPH	LiP	Ternary	100.0	0.5	nan	1.5	0.7	[276]
LiPH ₆	LiP	Ternary	100.0	3.0	nan	1.5	0.7	[276]
LiPH ₇	LiP	Ternary	100.0	3.5	nan	1.5	0.7	[276]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (Å)	Ref
LiPH	LiP	Ternary	150.0	0.5	nan	1.5	0.7	[276]
LiPH ₄	LiP	Ternary	150.0	2.0	nan	1.5	0.7	[276]
LiPH ₆	LiP	Ternary	150.0	3.0	nan	1.5	0.7	[276]
LiPH ₇	LiP	Ternary	150.0	3.5	nan	1.5	0.7	[276]
LiPH	LiP	Ternary	200.0	0.5	nan	1.5	0.7	[276]
LiPH ₃	LiP	Ternary	200.0	1.5	nan	1.5	0.7	[276]
LiPH ₄	LiP	Ternary	200.0	2.0	nan	1.5	0.7	[276]
LiPH ₆	LiP	Ternary	200.0	3.0	nan	1.5	0.7	[276]
LiPH ₇	LiP	Ternary	200.0	3.5	nan	1.5	0.7	[276]
LiPH	LiP	Ternary	200.0	0.5	nan	1.5	0.7	[276]
LiPH ₃	LiP	Ternary	200.0	1.5	nan	1.5	0.7	[276]
LiPH ₆	LiP	Ternary	200.0	3.0	nan	1.5	0.7	[276]
LiPH ₇	LiP	Ternary	200.0	3.5	nan	1.5	0.7	[276]
Fe ₂ SH ₃	Fe ₂ S	Ternary	100.0	1.0	nan	2.0	0.6	[277]
YCaH ₁₂	Yca	Ternary	180.0	6.0	nan	1.1	1.0	[278]
Li ₂ BH ₆	Li ₂ B	Ternary	200.0	2.0	nan	1.3	0.7	[279]
Pr ₂ AlH ₄	Pr ₂ Al	Ternary	0.0	1.3	nan	1.2	1.1	[280]
Ti ₂ AlH ₄	Ti ₂ Al	Ternary	0.0	1.3	nan	1.4	0.6	[281]
Ti ₂ AlH ₂	Ti ₂ Al	Ternary	0.0	0.6	nan	1.4	0.6	[282]
Ti ₂ CoH ₃	Ti ₂ Co	Ternary	0.0	1.0	nan	1.6	0.6	[283]
Ti ₂ CuH ₄	Ti ₂ Cu	Ternary	0.0	1.3	nan	1.6	0.7	[284]
Ti ₂ FeH	Ti ₂ Fe	Ternary	0.0	0.3	nan	1.6	0.7	[285]
Ti ₂ MnH ₂	Ti ₂ Mn	Ternary	0.0	0.6	nan	1.5	0.7	[286]
Ti ₂ NiH ₄	Ti ₂ Ni	Ternary	0.0	1.3	nan	1.6	0.7	[287]
Ti ₂ PdH ₃	Ti ₂ Pd	Ternary	0.0	1.0	nan	1.7	0.7	[288]
Y ₂ AlH ₃	Y ₂ Al	Ternary	0.0	1.0	nan	1.3	0.9	[289]
Zr ₂ CoH ₅	Zr ₂ Co	Ternary	0.0	1.6	nan	1.5	0.7	[290]
LaYH ₁₂	LaY	Ternary	180.0	6.0	nan	1.1	1.2	[]
Y ₂ MgH ₁₂	Y ₂ Mg	Ternary	196.0	4.0	nan	1.1	0.9	[]
Mg ₂ FeH ₄	Mg ₂ Fe	Ternary	0.00196	1.3	nan	1.4	0.8	[291]

Chem. Form.	Alloy	Type	Pressure	H/M (GPa)	Trans. Press. (GPa)	Electro. (Pauling)	Crystal Radius (\AA)	Ref
LiNiH ₃	LiNi	Ternary	3.0	1.5	nan	1.4	0.8	[292]
Na ₃ NiH ₅	Na3Ni	Ternary	5.0	1.2	nan	1.1	1.0	[]
Li ₅ WH ₁₁	Li5W	Ternary	5.0	1.8	nan	1.2	0.8	[293]
Li ₅ NbH ₁₁	Li5Nb	Ternary	5.0	1.8	nan	1.0	0.8	[293]
Li ₆ TaH ₁₁	Li6Ta	Ternary	5.0	1.5	nan	1.0	0.8	[293]
CaRhH ₃	CaRh	Ternary	0.0	1.5	nan	1.6	1.0	[293]
BaAu ₂ H ₄	BaAu2	Ternary	120.0	1.3	nan	1.9	1.5	[272]
SrAu ₂ H ₄	SrAu2	Ternary	120.0	1.3	nan	1.7	1.4	[272]
Na ₂ PdH ₄	Na2Pd	Ternary	0.5	1.3	nan	1.3	1.0	[294]
Li ₅ PtH ₉	Li5Pt	Ternary	0.5	1.5	nan	1.1	0.8	[294]
Na ₂ PtH ₆	Na2Pt	Ternary	0.5	2.0	nan	1.3	1.0	[294]
K ₂ PtH ₆	K2Pt	Ternary	0.5	2.0	nan	1.3	1.3	[294]
Rb ₂ PtH ₆	Rb2Pt	Ternary	0.5	2.0	nan	1.3	1.4	[294]
Cs ₂ PtH ₆	Cs2Pt	Ternary	0.5	2.0	nan	1.2	1.5	[294]
Na ₃ MnH ₅	Na3Mn	Ternary	0.5	1.2	nan	1.0	1.0	[294]
K ₃ MnH ₅	K3Mn	Ternary	0.5	1.2	nan	1.0	1.4	[294]
Rb ₃ MnH ₅	Rb3Mn	Ternary	0.5	1.2	nan	1.0	1.5	[294]
Cs ₃ MnH ₅	Cs3Mn	Ternary	0.5	1.2	nan	0.9	1.6	[294]
Na ₃ ReH ₅	Na3Re	Ternary	0.5	1.2	nan	1.1	1.0	[294]
K ₃ ReH ₁₀	K3Re	Ternary	0.5	2.5	nan	1.0	1.4	[294]
Rb ₃ ReH ₁₀	Rb3Re	Ternary	0.5	2.5	nan	1.0	1.5	[294]
Cs ₃ ReH ₁₀	Cs3Re	Ternary	0.5	2.5	nan	1.0	1.6	[294]
LiSiH ₅	LiSi	Ternary	200.0	2.5	nan	1.4	0.7	[215]
LiSiH ₆	LiSi	Ternary	200.0	2.0	nan	1.4	0.7	[215]
LiSi ₂ H ₉	LiSi	Ternary	200.0	3.0	nan	1.4	0.7	[215]
Li ₂ SiH ₆	LiSi	Ternary	200.0	2.0	nan	1.4	0.7	[215]
LiSiH ₄	LiSi	Ternary	200.0	2.0	nan	1.4	0.7	[215]
LiSiH ₈	LiSi	Ternary	200.0	4.0	nan	1.4	0.7	[215]
CSH ₇	CS	Ternary	100.0	3.5	nan	2.5	0.3	[216]

Bibliography

- [1] H2Europe
- [2] H2Cars
- [3] M. Dornheim: Thermodynamics of Metal Hydrides: Tailoring Reaction Enthalpies of Hydrogen Storage Materials, in: Thermodynamics – Interaction Studies – Solids, Liquids and Gases, J. C. Moreno-Pirajan (Ed.), InTech, Chapter 33, p. 900, (2011).
- [4] N. Eigen, C. Keller, M. Dornheim, T. Klassen, and R. Bormann, Scripta Materiala **56**, 847-851 (2007).
- [5] M. B. Ley, L. H. Jepsen, Y.-S. Lee, Y. W. Cho, J. B. von Colbe, M. Dornheim, M. Rokni, J. O. Jensen, M. Sloth, Y. Filinchuk, J. E. Jorgensen, F. Besenbacher, and T. R. Jensen, Materials Today **17**, 122-128 (2014).
- [6] J. B. von Colbe, J.-R. Ares, J. Barale, M. Baricco, C. Buckley, G. Capurso, N. Gallandat, D. M. Grant, M. N. Guzik, I. Jacob, E. H. Jensen, T. Jensen, J. Jepsen, T. Klassen, M. V. Lototsky, K. Manickam, A. Montone, J. Puszkiel, S. Sartori, D. A. Sheppard, A. Stuart, G. Walker, C. J. Webb, H. Yang, V. Yartys, A. Züttel, and M. Dornheim, Int. Journal of Hydrogen Energy **44** 7780-7808 (2019).
- [7] M. V. Lototsky, I. Tolj, M. W. Davids, Y. V. Klochko, A. Parsons, D. Swanepoel, and R. Ehlers, Int. J. Hydrogen Energy **41**, 13831-13842 (2016).
- [8] Toshiba H2OneTM "Hydrogen Based Autonomous Energy Supply System Now Providing Power to a Kyushu resort Hotel", Toshiba Press Released (2016).
- [9] NasaHydrures
- [10] A. Andreasen, "Predicting formation enthalpies of metal hydrides", Riso National Laborator, Denmark, (2004).
- [11] J.-C Crivello, B. Dam, R. V. Denys, M. Dornheim, D. M. Grant, J. Huot, T. R. Jensen, P. de Jongh, M. Latroche, C. Milanese, D. Milcius, G. S. Walker, C. J. Webb, C. Zlotea, and V. A. Yartys, Applied Physics A **122**, 97 (2016).

- [12] C. M. Pepin, G. Geneste, A. Dewaele, M. Mezouar, and P. Loubeyre, *Science* **357**, 382-385 (2017).
- [13] M. Somayazulu, M. Ahart, A. K. Mishra, Z. M. Geballe, M. Baldini, Y. Meng, V. V. Struzhkin, and R. J. Hemley, *Phys. Rev. Lett.* **122**, 027001 (2019).
- [14] N. P. Salke, M. Mahdi Davari Esfahani, Y. Zhang, I. A. Kruglov, J. Zhou, Y. Wang, E. Greenberg, V. B. Prakapenka, A. R. Oganov, and J.-F. Lin, *Nature Commun.* **10**, 4453 (2019).
- [15] N. W. Ashcroft, *Phys. Rev. Lett.* **92**, 187002 (2004).
- [16] P. Loubeyre, F. Occelli, and P. Dumas, *Nature* **577**, 631-635 (2020).
- [17] E. Zurek, R. Hoffmann, N. W. Ashcroft, A. R. Oganov, and A. O. Lyakhovc, *PNAS* **106**, 17640-17643 (2009).
- [18] C. Pepin, P. Loubeyre, F. Occelli, and P. Dumas, *PNAS* **112**, 7673-7676 (2015).
- [19] J. Hooper, P. Baettig, and E. Zurek, *J. Applied Phys.* **111**, 112611 (2012).
- [20] Z. Shao, D. Duan, Y. Ma, H. Yu, H. Song, H. Xie, D. Li, F. Tian, B. Liu, and T. Cui, *Inorg. Chem.* **58**, 2558-2564 (2019).
- [21] Y. Xie, Q. Li, A. R. Oganov, and H. Wang, *Acta Cryst. C* **70**, 104-111 (2014).
- [22] D. Zhou, X. Jin, X. Meng, G. Bao, Y. Ma, B. Liu, and T. Cui, *Phys. Rev. B* **86**, 014118 (2012).
- [23] V. V. Struzhkin, D. Y. Kim, E. Stavrou, T. Muramatsu, H.-K. Mao, C. J. Pickard, R. J. Needs, V. B. Prakapenka, and A. F. Goncharov, *Nature Commun.* **7**, 12267 (2016).
- [24] Z. Bazhanova, A. R. Oganov, and O. Gianola, *Phys.-Usp.* **55**, 489-497 (2012).
- [25] C. Pepin, A. Dewaele, G. Geneste, P. Loubeyre, and M. Mezouar, *Phys. Rev. Lett.* **113**, 265504 (2014).
- [26] A. G. Kvashnin, I. A. Kruglov, D. V. Semenok, and A. R. Oganov, *J. Phys. Chem. C* **122**, 4731-4736 (2018).
- [27] S. Zhang, J. Lin, Y. Wang, G. Yang, A. Bergara, and Y. Ma, *J. Phys. Chem. C* **122**, 12022-12028 (2018).
- [28] N. Zarifi, T. Bi, H. Liu, and E. Zurek, *J. Phys. Chem. C* **122**, 24262-24269 (2018).
- [29] A. Majumdar, J. S. Tse, M. Wu, and Y. Yao, *Phys. Rev. B*, **96**, 201107(R) (2017).
- [30] C. Heil, G. B. Bachelet, and L. Boeri, *Phys. Rev. B* **97**, 214510 (2018).

- [31] Y. Li, J. Hao, H. Liu, J. S. Tse, Y. Wang, and Y. Ma, *Sci. Rep.* **5**, 9948 (2015).
- [32] A. P. Drozdov, P. P. Kong, V. S. Minkov, S. P. Besedin, M. A. Kuzovnikov, S. Mozafari, L. Balicas, F. F. Balakirev, D. E. Graf, V. B. Prakapenka, E. Greenberg, D. A. Knyazev, M. Tkacz, and M. I. Erements, *Nature* **569**, 528-531 (2019).
- [33] E. Zurek and T. Bi, *J. Chem. Phys.* **150**, 050901 (2019).
- [34] C. Heil, S. di Cataldo, G. B. Bachelet, and L. Boeri, *Phys. Rev. B* **99**, 220502(R) (2019).
- [35] I. A. Kruglov, A. G. Kvashnin, A. Goncharov, A. R. Oganov, S. L. Lobanov, N. Holtgrewe, S. Jiang, V. B. Prakapenka, E. Greenberg, and A. V. Yanilkin, *Sci. Advances* **4**, eaat9776 (2018).
- [36] D. V. Semenok, A. G. Kvashnin, I. A. Kruglov, and A. R. Oganov, *J. Phys. Chem. Lett.* **9**, 1920-1926 (2018).
- [37] A. G. Kvashnin, D. V. Semenok, I. A. Kruglov, I. A. Wrona, and A. R. Oganov, *arXiv:1711.00278* (2018).
- [38] B. Guigue, A. Marizy, and P. Loubeyre, *Phys. Rev. B* **102**, 014107 (2020).
- [39] D. J. Siegel, C. Wolverton, and V. Ozolins, *Phys. Rev. B* **76**, 134102 (2007).
- [40] J. M. Pasini, C. Corgnale, B. A. van Hassel, T. Motyka, S. Kumar, and K. L. Simmons, *Int. J. Hydrogen Energy* **38**, 9755-9765 (2013).
- [41] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- [42] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [43] M. Torrent and J. Jollet, "The Projector Augmented-Wave Method", Presentation International School Lattice Dynamics of Solids with ABINIT, (2014).
- [44] D. R. Hamann, M. Schluter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).
- [45] D. Vanderbilt, *Phys. Rev. B* **41**, 7892(R) (1990).
- [46] P. E. Blochl, *Phys. Rev. B* **50**, 17953 (1994).
- [47] G. Kresse, and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- [48] N. A. W. Holzwarth, A. R. Tackett, and G. E. Matthews, *Comput. Phys. Comm.*, **135**, 329-347 (2001).
- [49] J. P. Perdew and K. Schimdt, *AIP Conference Proceedings* **577**, 1 (2001).

- [50] S. Lehtola, C. Steigemann, M. J. T. Oliveira, and M. A. L. Marques, *Software X* **7**, 1-5 (2018).
- [51] M. A. L. Marques, M. J. T. Oliveira, and T. Burnus, *Comput. Phys. Commun.* **183**, 2272 (2012).
- [52] M. L. Hu, Z. Yu, J. L. Yin, C. X. Zhang, and L. Z. Sun, *Comput. Mat. Science* **54**, 165-169 (2012).
- [53] A. Marini, G. Onida, and R. Del Sole, *Phys. Rev. B* **64**, 195125 (2001).
- [54] I. N. Yakovkin and P. A. Dowben, *Surface Review and Letters* **14**, 481-487 (2007).
- [55] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [56] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Phys. Rev. B* **46**, 6671 (1992).
- [57] V. Ozolins and M. Koerling, *Phys. Rev. B* **48**, 18304 (1993).
- [58] D. R. Hammann, *Phys. Rev. Lett.* **76**, 660 (1996).
- [59] J. P. Perdew and L. A. Constantin, *Phys. Rev. B* **75**, 155109 (2007).
- [60] J. Tao, J. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
- [61] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, and J. Sun, *Phys. Rev. Lett.* **103**, 026403 (2009).
- [62] B. Xiao, J. Sun, A. Ruzsinszky, J. Feng, R. Haunschild, G. E. Scuseria, and J. P. Perdew, *Phys. Rev. B* **88**, 184103 (2013).
- [63] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, *Nature Chemistry* **8**, 563-568 (2016).
- [64] Y. Yao and Y. Kanai, *J. Chem. Phys.* **146**, 224105 (2017).
- [65] J. Sun, A. Ruzinsky, and J. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).
- [66] J. Sun, M. Marsman, A. Ruzsinszky, G. Kresse, and J. P. Perdew, *Phys. Rev. B* **83**, 121410 (2011).
- [67] B. Patra, S. Jana, L. A. Constantin, and P. Samal, *Phys. Rev. B* **100**, 045147 (2019).
- [68] J. Sun, C. R. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, L. M. Klein, and J. P. Perdew, *Nature Chemistry* **8**, 831-836 (2016).

- [69] J. W. Furness, Y. Zhang, C. Lane, I. G. Buda, B. Barbiellini, R. S. Markiewicz, A. Bansil, and J. Sun, *Commun. Phys.* **1**, 11 (2018).
- [70] M. Chen, H. Y. Ko, R. C. Remsing, M. F. Calegari Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. L. Klein, and J. P. Perdew, *PNAS* **114**, 10846-10851 (2017).
- [71] A. P. Bartok and J. R. Yates, *J. Chem. Phys.* **150**, 161101 (2019).
- [72] Y. Yamamoto, A. Salcedo, C. M. Diaz, M. S. Alam, T. Baruah, and R. R. Zope, *arXiv:2004.13393*, (2020).
- [73] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, *J. Phys. Chem. Lett.* **11**, 8208-8215 (2020).
- [74] S. Grimme, A. Hansen, S. Ehlert, and J.-M. Mewes, *J. Chem. Phys.* **154**, 064103 (2021).
- [75] A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- [76] J. P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).
- [77] VASP-Hybrid
- [78] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003).
- [79] S. Jana, and P. Samal, *Phys. Chem. Chem. Phys.* **21**, 3002-30015 (2019).
- [80] S. Jana, and P. Samal, *Phys. Chem. Chem. Phys.* **20**, 8999-9005 (2018).
- [81] M. Torrent, F. Jollet, F. Bottin, G. Zerah, and X. Gonze, *Comput. Mat. Science* **42**, 337-351 (2008).
- [82] M. Torrent, N. A. W. Holzwarth, F. Jollet, D. Harris, N. Lepley, X. Xu, *Comput. Phys. Commun.* **181**, 1862-1867 (2010)
- [83] X. Liang, A. Bergara, L. Wang, B. Wen, Z. Zhao, X.-F. Zhou, J. He, G. Gao, and Y. Tian, *Phys. Rev.* **99**, 100505 (2019).
- [84] Y. Sun, J. Lv, Y. Xie, H. Liu, and Y. Ma, *Phys. Rev. B* **123**, 097001 (2019).
- [85] C. J. Pickard and R. J. Needs, *Nature Physics* **3**, 473-476 (2007).
- [86] Surface-PES.
- [87] C. J. Pickard and R. J. Needs, *J. of Phys. Cond. Matt.* **23**, 053201 (2011).
- [88] C. J. Pickard, M. Martinez-Canales, and R. J. Needs, *Phys. Rev. B* **86**, 059901 (2012).

- [89] C. J. Pickard, and R. J. Needs, *Nature Materials* **9**, 624-627 (2010).
- [90] C. J. Pickard, R. J. Needs, *Nature Mater.* **7**, 775 (2008).
- [91] C. J. Pickard and R. J. Needs, *Phys. Rev. Lett.* **97**, 045504 (2006).
- [92] C. J. Pickard and R. J. Needs, *Phys. Rev. B* **76**, 144114 (2007).
- [93] C. J. Pickard and R. J. Needs, *Phys. Rev. Lett.* **102**, 125702 (2009).
- [94] C. G. Broyden, "A new double-rank minimization algorithm", *Notices American Math. Soc.* **16**, 670 (1969)
- [95] S. T. Call, D. Y. Zubarev, and A. I. Boldyrev, *J. Comput. Chem.* **28**, 1177-1186 (2007).
- [96] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, and Q. Zhu, *Comput. Phys. Comm.* **184**, 1172-1182 (2013).
- [97] Y. Wang, J. Lv, L. Zhu, and Y. Ma, *Phys. Rev. B* **82**, 094116 (2010).
- [98] H. Liu, I. Naumov, R. Hoffmann, N. W. Ashcroft, and R. J. Hemley, *PNAS* **114**, 6990-6995 (2017).
- [99] F. Peng, Y. Sun, C. J. Pickard, R. J. Needs, Q. Wu, and Y. Ma, *Phys. Rev. Lett.* **119**, 107001 (2017).
- [100] F. Curtis, X. Li, T. Rose, A. Vazquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli, and N. Marom, *J. Chem. Theory Comput.* **14**, 2246-2264 (2018).
- [101] R. Martonak, A. Laio, and M. Parrinello, *Phys. Rev. Lett.* **90**, 075503 (2003).
- [102] A. Laio and M. Parrinello, *PNAS* **99**, 12562-12566 (2002).
- [103] D. Branduardi, G. Bussi, and M. Parrinello, *J. Chem. Theory Comput.* **8**, 2247-2254 (2012).
- [104] C. J. Pickard, *Phys. Rev. B* **99**, 054102 (2019).
- [105] Z. Lu, B. Zhu, B. W. B. Shires, D. O. Scanlon, and C. J. Pickard, *J. Chem. Phys.* **154**, 174111 (2021).
- [106] F. Curtis, T. Rose, and N. Marom, *Faraday Discussions* **211**, 61-77 (2018).
- [107] M. S. Jorgensen, M. N. Groves, and B. Hamme, *J. Chem. Theory. Comput.* **3**, 1486-1493 (2017).
- [108] V. L. Deringer, C. J. Pickard, and G. Csanyi, *Phys. Rev. Lett.* **120**, 156001 (2018).

- [109] G. Cheon, L. Yang, K. McCloskey, E. J. Reed, and E. D. Cubuk, Arxiv:2012.02920v2 (2020).
- [110] E. V. Podryabinkin, E. V. Tikhonov, A. V. Shapeev, and A. R. Oganov, Phys. Rev. B **99**, 064114 (2019).
- [111] I. A. Kruglov, A. Yanilkin, A. R. Oganov, and P. Korataev, Phys. Rev. B **100**, 174104 (2019).
- [112] I. A. Kruglov, A. V. Yanilkin, Y. Propad, and A. R. Oganov, Arxiv:2101.10153 (2021).
- [113] Q. Tong, L. Xue, J. Lv, Y. Wang, and Y. Ma, Faraday Discussions **211**, 31-43 (2018).
- [114] Q. Tong, P. Gao, H. Liu, Y. Xie, J. Lv, Y. Wang, and J. Zhao, J. Phys. Chem. Lett. **11**, 8710–8720 (2020).
- [115] S. Yu, X. Jia, G. Frapper, D. Li, A. R. Oganov, Q. Zeng, and L. Zhang, Sci. Rep. **5**, 17764 (2015).
- [116] L. Pauling, J. Am. Chem. Soc. **54**, 3570-3582 (1932).
- [117] A. L. Allred, J. Inorg. and Nuclear Chemistry **17**, 215-221 (1961).
- [118] L. Wang, D. Duan, H. Yu, H. Xie, X. Xiaoli, Y. Ma, F. Tian, D. Li, B. Liu, and T. Cui, Inorg. Chem. **57**, 181-186 (2018).
- [119] J. Binns, M.-E. Donnelly, M. Wang, A. Hermann, E. Gregoryanz, P. Dalladay-Simpson, and R. T. Howie, Phys. Rev. B **98**, 140101(R) (2018).
- [120] J. Ying, H. Liu, E. Greenberg, V. B. Prakapenka, and V. V. Struzhkin, Phys. Rev. Mater. **2**, 085409 (2018).
- [121] A. Wurtz, "On copper hydride", C. R. Hebd Acad Sci Paris **18**, p. 702 (1844)
- [122] R. Burtovyy, E. Utzig, and M. Tkacz, Thermochemica Acta **363**, 157 (2000).
- [123] N. P. Fitzsimons, W. Jones, and P. J. Herley, J. Chem. Soc., Faraday Trans. **91**, 713-718 (1995).
- [124] S. Aldridge and A. J. Downs, Chem. Rev. **101**, 3305-3366 (2001).
- [125] M. Tkacz, S. Majchrzak, B. Baranovski, Zeitschrift für Physikalische Chemie Neue Folge, Bd. **163**, S. 467-468 (1989).
- [126] M. Tkacz, S. Majchrzak, B. Baranovski, High-Pressure Research **6**, 85-90 (1989).
- [127] M. Tkacz and R. Burtovyy, Solid State Communications **132**, 37-41 (2004).

- [128] T. Bi, N. Zarifi, T. Terpstra, and E. Zurek, "Reference Module in Chemistry, Molecular Sciences and Chemical Engineering", Elsevier (2019).
- [129] R. Burtovyy and M. Tkacz, Solid State Communications **132**, 169 (2004).
- [130] J. Binns, M. Pena-Alvarez, M.-E Donnelly, E. Gregoryanz, R. T. Howie, and P. Dalladay-Simpson, Engineering **5**, (3) 505-509 (2019).
- [131] C. Donnerer, T. Scheler, and E. Gregoryanz, J. of Chem. Phys. **138**, 134507 (2013).
- [132] T. Meier, F. Trybel, G. Criniti, D. Laniel, S. Khandarkhaeva, E. Koemets, T. Fedotenko, K. Glazyrin, M. Hanfland, G. Steinle-Neumann, N. Dubrovinskaia, and L. Dubrovinsky, Phys. Rev. B **102**, 165109 (2020).
- [133] V. E. Antonov, J. Alloys and Compounds **330-332**, 110-116 (2002).
- [134] J.-B Charraud, G. Geneste, and M. Torrent, Phys. Rev. B **100**, 224102 (2019).
- [135] X.-H. Xiao, D.-F. Duan, Y.-B. Ma, H. Xie, H. Song, D. Li, F.-B. Tian, B.-B. Liu, H.-Y. Yu, and T. Cui, Front. Phys. **14**, 43601 (2019).
- [136] D. Yin, H. A. Murdoch, B. Chad Hornbuckle, E. Hernández-Rivera, and M. K. Dunstan, Electrochemistry Communications **98**, 96-100 (2019).
- [137] H. T. Stokes, D. M. Hatch, and B. J. Campbell, FINDSYM Software Suite, iso.byu.edu
- [138] H. T. Stokes and D. M. Hatch, J. Appl. Cryst. **38**, 237-238 (2005).
- [139] The routines for generating the random numbers (authors: R. Chandler and P. Northrup) have been taken from Richard Chandler homepage. See
- [140] X. Gonze, F. Jollet, F. Abreu Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J. M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. Da Pieve, M. Delaveau, M. Di Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, A. Gerossier, M. Giantomassi, Y. Gillet, D. R. Hamann, L. He, G. Jomard, S. Laflamme Janssen, J. Le Roux, S. Levitt, A. Lherbier, A. Liu, F. Lukacevic, I. Martin, A. Martins, C. Oliveira, M. J. T. Poncé, Y. Pouillon, T. Rangel, G. M. Rignanese, A. H. Romero, B. Rousseau, O. Rubel, A. A. Shukri, M. Stankovski, M. Torrent, M. J. Van Setten, B. Van Troeye, M. J. Verstraete, D. Waroquiers, J. Wiktor, B. Xu, A. Zhou, and J. W. Zwanziger, Comput. Phys. Commun. **205**, 106-131 (2016).
- [141] F. Jollet, M. Torrent, and N. Holzwarth, Comput. Phys. Commun. **185**, 1246 (2014).
- [142] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, Phys. Rev. Lett. **65**, 353 (1990).

- [143] J. Von Pezold, A. Dick, M. Friák, and Jörg Neugebauer, *Phys. Rev. B* **81**, 094203 (2010).
- [144] A. Marizy, G. Geneste, P. Loubeyre, B. Guigue, and G. Garbarino, *Phys. Rev. B* **97**, 184103 (2018).
- [145] R. W. Bader, *Chem. Rev. (Washington, D. C.)* **91**, 893-928 (1991).
- [146] C. Audouze, F. Jollet, M. Torrent and X. Gonze, *Phys. Rev. B* **73**, 235101 (2006).
- [147] C. Audouze, F. Jollet, M. Torrent and X. Gonze, *Phys. Rev. B* **78**, 035105 (2008).
- [148] H. Sugiura, S. M. Filipek, V. Paul-Boncour, I. Marchuk, R.-S. Liu, and S. Il Pyun, *Nukleonika* **51**, S73 (2006).
- [149] M. Krukowski and B. Baranowski, *Roczniki Chemii* **49**, 1183 (1975)
- [150] M. Krukowski and B. Baranowski, *J. Less-Common Metals* **49**, 385 (1976).
- [151] A. San-Martin, and F. D. Manchester, *J. Phase Equilibria* **16**, 255-262 (1995).
- [152] V. E. Antonov, T. E. Antonova, N. A. Chirin, E. G. Ponyatovsky, M. Baier, and F. E. Wagner, *Scripta Materialia* **34**, 1331-1336 (1996).
- [153] V. E. antonov, K. Cornell, B. Dorner, V. K. Fedotov, G. Grosse, A. I. Kolesnikov, F. E. Wagner, and H. Wipf, *Solid State Commun.* **113**, 569-572 (2000).
- [154] Y. Fukai and H. Ishikawa, *Zeitschrift für Physikalische Chemie Neue Folge* **163**, 479-482 (1989).
- [155] S. M. Filipek, S. Majchfeak, A. B. Sawaoka, and M. Cernansky, *High Pressure Research* **7**, 271-273 (1991).
- [156] S. M. Filipek, H. Sugiura, and A. B. Sawaoka, *High Pressure Research* **4**, 354-356 (1990).
- [157] Y. Fukai, T. Haraguchi, H. Shinoniya, and K. Mori, *Scripta Materialia* **46**, 679-684 (2002).
- [158] H. Fujihisa and K. Takemura, *Phys. Rev. B* **52**, 13257 (1995).
- [159] D. Hobbs, J. Hafner, and D. Spisak, *Phys. Rev. B* **68**, 014407 (2003).
- [160] L. Morris, U. James, J. Hales, M. L. Trudeau, P. Georgiev, J. P. Embs, J. Eckert, N. Kaltsoyannis, and D. M. Antonelli, *Energy and Environmental Science* **12**, 1580-1591 (2019).
- [161] W. L. Mao, W. Sturhahn, D. L. Heinz, H.-K. Mao, J. Shu, and R. J. Hemley, *Geophys. Res. Lett.* **31**, L15618 (2004).

- [162] T. R. Paudel, S. S. Jaswal, and E. Y. Tsymbal, *Phys. Rev. B* **85**, 104409 (2012).
- [163] CRC Handbook of Chemistry and Physics, 85 th ed., edited by D.R. Lide (CRC Press, Boca Raton, FL, 2004)
- [164] M. Wang, J. Binns, M.-E. Donnelly, M. Pena-Alvarez, P. Dalladay-Simpson, and R. T. Howie, *J. Chem. Phys.* **148**, 144310 (2018).
- [165] T. Bi, N. Zarifi, T. Terpstra, and E. zurek, "Reference Module in Chemistry, Molecular Sciences and Chemical Engineering", Elsevier (2019).
- [166] A. V. Irodova, V. P. Glazkov, V. A. Somenkov, S. Sh. Shilstein, V. E. Antonov, and E. G. Ponyatovsky, *Sov. Phys. Solid State* **29**, 1562 (1987).
- [167] V. K. Fedotov, V. E. Antonov, A. I. Kolesnikov, A. I. Beskrovnyi, G. Grosse, and F. E. Wagner, *Solid State Commun.* **107**, 787 (1998).
- [168] I. A. Troyan, D. V. Semenok, A. G. Kvashnin, A. V. Sadakov, O. A. Sobolevskiy, V. M. Pudalov, A. G. Ivanova, V. B. Prakapenka, E. Greenberg, A. G. Gavriluk, I. S. Lyubutin, V. V. Struzhkin, A. Bergara, . Errea, R. Bianco, M. Calandra, F. Mauri, L. Monacelli, R. Akashi, and A. R. Oganov, *Adv. Mater.*, **33**, 15 2006832 (2021).
- [169] L.-L. Liu, H.-J. Sun, C. Z. Wang, and W.-C. Lu, *J. Phys. Condens. Mat.* **29**, 325401 (2017).
- [170] P. P. Kong, V. S. Minkov, M. A. Kuzovnikov, S. P. Besedin, A. P. Drozdov, S. Mozaffari, L. Balicas, F. F. Balakirev, V. B. Prakapenka, E. Greenberg, D. A. Knyazev, and M. I. Erements, *arXiv:1909.10482* (2019).
- [171] E. Snider, N. Dasenbrock-Gammon, R. McBride, X. Wang, N. Meyers, K. V. Lawler, E. Zurek, A. Salamat, and R. P. Dias, *Phys. Rev. Lett.* **126**, 117003 (2021).
- [172] M. Rahm, R. Hoffmann, and N. W. Ashcroft, *J. Am. Chem. Soc.* **25**, 8740–8751 (2017).
- [173] D. Laniel, V. Svitlyk, G. Weck, and P. Loubeyre, *Phys. Chem. Chem. Phys.* **20**, 4050-4057 (2018).
- [174] H. Sugiura, I. Marchuk, V. Paul-Boncour, A. Percheron-Guégan, T. Kitazawa, and S.M. Filipek, *J. Alloy Compd.* **32**, 356-357 (2003).
- [175] V. Paul-Boncour and S. Matar, *Phys. Rev. B* **70**, 184435 (2004).
- [176] T. Leblond, V. Paul-Boncour, and A. Percheron-Guégan, *J. Alloy Compound* **419**, 446-447 (2007).

- [177] V. Paul-Boncour, S. M. Filipek, A. Percheron-Guégan, I. Marchuk, and J. Pielaszek, *J. Alloy Compd* **317-318**, 83-87 (2001).
- [178] V. Paul-Boncour, L. Guénée, M. Latroche, and A. Percheron-Guégan, *J. Alloys Compound*, **255**, 195-202 (1997).
- [179] V. Paul-Boncour, L. Guénée, M. Latroche, A. Percheron-Guégan, B. Ouladdiaf, and F. Bourée-Vigneron, *J. Solid State Chem.* **142**, 120-129 (1999).
- [180] V. Paul-Boncour, M. Guillot, G. Wiesinger, and G. André, *Phys. Rev. B* **72**, 174430 (2005).
- [181] J. Ropka, R. Cerny, V. Paul-Boncour, and T. Proffen, *J. Solid State Chem.* **182**, 1907-1912 (2009).
- [182] V. Paul-Boncour, S.M. Filipek, I. Marchuk, G. André, F. Bouré, G. Wiesinger, and A. Percheron-Guégan, *J. Phys. Condens. Matter* **15**, 4349-4359 (2003).
- [183] V. Paul-Boncour, M. Guillot, O. Isnard, B. Ouladdiaf, A. Hoser, T. Hansen, and N. Stuesser, *Journal of Solid State Chemistry* **245**, 98-109 (2017)
- [184] B. Guigue, "Structure et propriétés magnétiques de quelques super-hydrures sous pression : recherche d'un supraconducteur à Tc ambiante et de similarités avec l'hydrogène métallique", *Chimie-Physique [physics.chem-ph]*, Université Paris sciences et lettres, (2019).
- [185] A. Marizy, B. Guigue, and P. Loubeyre, private communication.
- [186] E. V. Podryabinkin, E. V. Tikhonov, A. V. Shapeev, and A. R. Oganov, *Phys. Rev. B* **99**, 064114 (2019).
- [187] A. P. Bartok, R. Kondor, and G. Csanyi, *Phys. Rev. B* **87**, 184115 (2013).
- [188] J. Schmidt, M. R. G. Marques, S. Botti and M. A. L. Marques, *npj Comput. Mater.* **5**, 83 (2019).
- [189] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *Phys. Rev. Lett.* **108**, 058301 (2012).
- [190] F. Faber, A. Lindmaa, O. A. Lilienfeld, and R. Armiento, *Int. J. Quantum Chem.* **115**, 1094–1101 (2015).
- [191] K. T. Schutt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross, *Phys. Rev. B* **89**, 205118 (2014).
- [192] J. E. Moussa, *Phys. Rev. Lett.* **109**, 059801 (2012).
- [193] A. Shapeev, *Multiscale Model. Simul.* **14**, 1153-1173 (2016).

- [194] J. Behler, J. of Chem. Phys. **134**, 074106 (2011).
- [195] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
- [196] A.P Bartok and G. Csanyi, int J. Quantum Chem **115**,1051-1057 (2015).
- [197] T. Xie and J. C. Grossman, Phys. Rev. Lett. **120**, 145301 (2018).
- [198] Y. Li, D. Tarlow, M. Brockschmidt, and R. Zemel, arXiv:1511.05493 (2015).
- [199] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, J. Chem. Phys. **148**, 241722 (2018).
- [200] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry", In Proc. 34th International Conference on Machine Learning, Proc. Machine Learning Research **70**, 1263–1272 (2017).
- [201] A. P. Bartok, M. C. Payne, R. Kondor, and G. Csanyi, Phys. Rev. Lett. **104**, 136403 (2010).
- [202] A. P Thompson, L. P Swiler, C. R Trott, S. M Foiles, and G. J Tucker, J. of Comput. Phys. **285**, 316-330 (2015).
- [203] R. Storn and K. Price, Journal of Global Optimization **11**, 341-359 (1997).
- [204] `lammps.sandia.gov`
- [205] L. Xu, Y. Bi, X. Li, and Y. Wang, Journal of Applied Physics, **115**, 164903 (2014).
- [206] A. Salamat, G. Garbarino, A. Dewaele, P. Bouvier, S. Petitgirard, C.J. Pickard, P. F. Mcmillan, and M. Mezouar, Phys. Rev. B **84**, 140104(R) (2011).
- [207] A. Salamat, R. Briggs, P. Bouvier, S. Petitgirard, A. Dewaele, M. E. Cutler, F. Cora, D. Daisenber, G. Garbarino, and P. F. Mcmillan, Phys. Rev. B **88**, 104104 (2013).
- [208] A. Aguado, Phys. Rev. B **67**, 212104 (2003).
- [209] GMM-Sklearn
- [210] Silh-Score
- [211] cryscalc
- [212] V. Vapnik and A. Lerner, Pattern Recognition using Generalized Portrait Method, Automation and Remote Control (1963).
- [213] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. MIT Press, (2006) Online

- [214] D. V. Louzguine-Luzgin, N. Chen, A. Y. Churymov, and L. V. Louzguina-Luzguina, *J. of Mater. Sc.* **50**, 1783–1793 (2015)
- [215] P. Zhang, Y. Sun, X. Li, J. Lv, and H. Liu, *Phys. Rev. B* **102**, 184103 (2020).
- [216] W. Cui, T. Bi, J. Shi, Y. Li, H. Liu, E. Zurek, and R. J. Hemley, *Phys. Rev. B* **101**, 134504 (2020).
- [217] D. V. Semenok, I. A. Kruglov, I. A. Savkin, A. G. Kvashnin, and A. R. Oganov, *Current Opinion in Solid State and Material Science* **24**, 100808 (2020).
- [218] P. Vajeeston, P. Ravindran, A. Kjekshus, and H. Fjellvag, *Appl. Phys. Lett.* **84**, 34-36 (2004).
- [219] C.-H. Hu, A. R. Oganov, Q. Zhu, G.-R. Qian, G. Frapper, A. O Lyakhov, and H.-Y Zhou, *Phys. Rev. Lett.* **110**, 165504 (2013).
- [220] S. Nakano, R. J. Hemley, E. A. Gregoryanz, A. F. Goncharov, and H.-K. Mao, *J. Phys.: Condens. Matter* **14**, 10453 (2002).
- [221] A. Suarez-Alcubilla, I. G. Gurtubay, and A. Bergara, *High Pressure Research* **34**, 59-69 (2014).
- [222] X. Feng, J. Zhang, G. Gao, H. Liu, and H. Wang, *RSC Adv.* **5**, 59292-59296 (2015).
- [223] P. Hou, X. Zhao, F. Tian, D. Li, D. Duan, Z. Zhao, B. Chu, B. Liua, and T. Cui, *RSC Adv.* **5**, 5096-5101 (2014).
- [224] O. Degtyareva, M. M. Canales, A. Bergara, X.-J. Chen, Y. Song, V. V. Struzhkin, H-K. Mao, and R. J. Hemley, *Phys. Rev. B* **76**, 064123 (2007).
- [225] Y. Yao, J. S. Tse, Y. Ma, and K. Tanaka, *Euro. Phys. Lett.* **78**, 37003 (2007).
- [226] A. P. Drozdov, M. I. Erements, I. A. Troyan, V. Ksenofontov, and S. I. Shylin, *Nature* **525**, 73-76 (2015).
- [227] K. Aoki, E. Katoh, H. Yamawaki, M. Sakashita, and H. Fujihisa, *Physica B: Condens. Mat.* **265**, 83-86 (1999).
- [228] Z. Wang, H. Wang, J. S. Tse, T. Itakad, and Y. Maac, *Chem. Sci.* **6**, 522-526 (2015).
- [229] D. Zhou, X. Jin, X. Meng, G. Bao, Y. Ma, B. Liu, and T. Cui, *Phys. Rev. B* **86**, 014118 (2012).
- [230] X. Ye, N. Zarifi, E. Zurek, R. Hoffmann, and N. W. Ashcroft, *J. Phys. Chem. C* **122**, 6298-6309 (2018).

- [231] J. Zhang, J. M. McMahon, A. R. Oganov, X. Li, X. Dong, H. Dong, and S. Wang, *Phys. Rev. B* **101**, 134108 (2020).
- [232] X. Li and F. Peng, *Inorg. Chem.* **56**, 13759–13765 (2017).
- [233] A. G. Kvashnin, I. A. Kruglov, D. V. Semenov, and A. R. Oganov, *J. Phys. Chem. C* **122**, 4731-4736 (2018).
- [234] A. Shayesteh, S. Yu, and P. F. Bernath, *Chem. Eur. J.* **11**, 4709-4712 (2005).
- [235] H.-J. Himmel, A. J. Downs, and T. M. Greene, *Am. Chem. Soc.* **124**, 4448-4457 (2002).
- [236] M. Martinez-Canales, A. Bergara, J. Feng, and W. G. Rochalade, *J. Phys. Chem. Solids.* **67**, 2095-2099 (2006).
- [237] Y. Fu, X. Du, L. Zhang, F. Peng, M. Zhang, C. J. Pickard, R. J. Needs, D. J. Singh, W. Zheng, and Y. Ma, *Chem. Mater* **28**, 1746-1755 (2016).
- [238] X. Zhang, W. Xu, Y. Wang, S. Jiang, F. A. Gorelli, E. Greenberg, V. B. Prakapenka, and A. F. Goncharov, *Phys. Rev. B* **97**, 064101 (2018).
- [239] J. Hooper and E. Zurek, *Chem. Eur. J.* **18**, 5013–5021 (2012).
- [240] J. Hooper, T. Terpstra, A. Shamp, and E. Zurek, *J. Phys. Chem. C.* **118**, 6433-6447 (2014).
- [241] J. B. Charraud, Thèse
- [242] H. Xie, W. Zhang, D. Duan, X. Huang, Y. Huang, H. Song, X. Feng, Y. Yao, C. J. Pickard, and T. Cui, *J. Phys. Chem. Lett.* **11**, 646-651 (2020).
- [243] K. Abe, *Phys. Rev. B* **98**, 134103 (2018).
- [244] G. Gao, R. Hoffmann, N. W. Ashcroft, H. Liu, A. Bergara, and Y. Ma, *Phys. Rev. B* **88**, 184104 (2013).
- [245] X. Feng, J. Zhang, H. Liu, T. Iitaka, K. Yina, and H. Wang, *Solid State Commun.* **239**, 14-19 (2016).
- [246] X. Li, H. Liu, and F. Peng, *Chem. Phys.* **18**, 28791–287916 (2016).
- [247] Y. Liu, D. Duan, F. Tian, C. Wang, Y. Ma, D. Li, X. Huang, B. Liu, and T. Cui, *Phys. Chem. Chem. Phys.* **18**, 1516-1520 (2016).
- [248] G. Sudhapriyanga, *Acta Phys. Pol. A* **125**, 29–35 (2014).
- [249] Y. Fukai and N. Okuma, *Phys. Rev. Lett.* **73**, 1640 (1994).

- [250] C. Donnerer, T. Scheler, and E. Gregoryanz, *J. Chem. Phys.* **138**, 134507 (2013).
- [251] Y. Liu, D. Duan, F. Tian, H. Liu, C. Wang, X. Huang, D. Li, Y. Ma, B. Liu, and T. Cui, *Inorg. Chem.* **54**, 9924-9928 (2015).
- [252] M. M. D. Esfahani, Z. Wang, A. R. Oganov, H. Dong, Q. Zhu, S. Wang, M. S. Rakitin, and X.-F. Zhou, *Sc. Rep.* **6**, 22873 (2016).
- [253] Y. Ma, D. Duan, D. Li, Y. Liu, F. Tian, X. Huang, Z. Zhao, H. Yu, B. Liu, and T. Cui, *arXiv:1506.03889*, (2015).
- [254] X. Zhong, H. Wang, J. Zhang, H. Liu, S. Zhang, H.-F. Song, G. Yang, L. Zhang, and Y. Ma, *Phys. Rev. Lett.* **116**, 057002 (2016).
- [255] A. Shamp and E. Zurek, *J. Phys. Chem. Lett.* **6**, 4067-4072 (2015).
- [256] A. Shamp, J. Hooper, and E. Zurek, *Inorg. Chem.* **51**, 9333-9342 (2012).
- [257] W. Chen, D. V. Semenov, A. G. Kvashnin, I. A. Kruglov, M. Galasso, H. Song, X. Huang, D. Duan, A. F. Goncharov, V. B. Prakapenka, A. R. Oganov, and T. Cui, *arXiv:2004.12294*, (2021).
- [258] M. A. Kuzovnikov, and M. Tkacz, *J. Phys. Chem. C* **123**, 30059-30066 (2019).
- [259] Q. Zhuang, X. Jin, T. Cui, Y. Ma, Q. Lv, Y. Li, H. Zhang, X. Meng, and K. Bao, *Inorg. Chem.* **56**, 3901-3908 (2017).
- [260] S. Zheng, S. Zhang, Y. Sun, J. Zhang, J. Lin, G. Yang, and A. Bergara, *Front. Phys.* **6**, 101 (2018).
- [261] Y. Liu, D. Duan, X. Huang, F. Tian, D. Li, X. Sha, C. Wang, H. Zhang, T. Yang, B. Liu, and T. Cui, *J. Phys. Chem. C* **119**, 15905-15911 (2015).
- [262] T. Scheler, M. Marqués, Z. Konopkova, C. L. Guillaume, R. T. Howie, and E. Gregoryanz, *Phys. Rev. Lett.* **111**, 215503 (2013).
- [263] X.-F. Zhou, A. R. Oganov, X. Dong, L. Zhang, Y. Tian, and H.-T. Wang, *Phys. Rev. B* **84**, 054543 (2011).
- [264] Y. Cheng, C. Zhang, T. Wang, G. Zhong, C. Yang, X.-J. Chen, and H.-Q. Lin, *Sc. Rep.* **5**, 16475 (2015).
- [265] Y. Ma, D. Duan, D. Li, Y. Liu, F. Tian, H. Yu, C. Xu, Z. Shao, B. Liu, and T. Cui, *arXiv:1511.05291*, (2015).
- [266] Y. Liu, D. Duan, F. Tian, C. Wang, G. Wu, Y. Ma, H. Yu, D. Li, B. Liu, and T. Cui, *RSC Adv.* **5**, 103445-103450 (2015). *

- [267] D. Zhou, D. V. Semenov, H. Xie, X. Huang, D. Duan, A. Aperis, P. M. Oppeneer, M. Galasso, A. I. Kartsev, A. G. Kvashnin, A. R. Oganov, and T. Cui, *J. Am. Chem. Soc.* **142**, 2803–2811 (2020).
- [268] Y. Ma, D. Duan, Z. Shao, H. Yu, H. Liu, F. Tian, X. Huang, D. Li, B. Liu, and T. Cui, *Phys. Rev. B* **96**, 144518 (2017).
- [269] X. Guo, R.-L. Wang, H.-L. Chen, W.-C. Lu, K. M. Ho, and C. Z. Wang, *Physics Letters A* **384**, 126189 (2019).
- [270] J. Chen, W. Cui, J. Shi, M. Xu, J. Hao, A. P. Durajski, and Y. Li, *ACS Omega* **4**, 14317-14323 (2019).
- [271] X. Liang, S. Zhao, C. Shao, A. Bergara, H. Liu, L. Wang, R. Sun, Y. Zhang, Y. Gao, Z. Zhao, X.-F. Zhou, J. He, D. Yu, G. Gao, and Y. Tian, *Phys. Rev. B* **100**, 184502 (2019).
- [272] M. Rahm, R. Hoffmann, and N. W. Ashcroft, *J. Am. Chem. Soc.* **139**, 8740-8751 (2017).
- [273] J.B Charraud, Thèse
- [274] D. Meng, M. Sakata, K. Shimizu, Y. Iijima, H. Saitoh, T. Sato, S. Takagi, and S.-I. Orimo, *Phys. Rev. B* **99**, 024508 (2019).
- [275] K. Spektor, W. A. Crichton, S. Filippov, J. Klarbring, S. I. Simak, A. Fischer, and U. Haussermann, *ACS Omega* **5**, 8730–8743 (2020).
- [276] Z. Shao, D. Duan, Y. Ma, H. Yu, H. Song, H. Xie, D. Li, F. Tian, B. Liu, and T. Cui, *npj Computational Materials* **5**, 104 (2019).
- [277] S. Zhang, L. Zhu, H. Liu, and G. Yang, *Inorg. Chem.* **55**, 11434-11439 (2016).
- [278] H. Xie, D. Duan, Z. Shao, H. Song, Y. Wang, X. Xiao, D. Li, F. Tian, B. Liu, and T. Cui, *J. Phys.: Condens. Matter* **31**, 245404 (2019).
- [279] C. Kokail, W. von der Linden, and L. Boeri, *Phys. Rev. Materials* **1**, 074803 (2017).
- [280] Hydpark database Semenenko. 1985
- [281] Hydpark database Sorby. 2000
- [282] Hydpark database Semenenko. 1982
- [283] Hydpark database Verbetsky. 1986
- [284] Hydpark database Beck. 1962

- [285] Hydpark database Beck. 1962
- [286] Hydpark database Beck. 1962
- [287] Hydpark database Yamanaka. 1975
- [288] Hydpark database Kadel. 1978
- [289] Hydpark database Semenenko. 1985
- [290] Hydpark database van Essen. 1979
- [291] J. J. Didisheim, P. Zolliker, K. Yvon, P. Fischer, J. Schefer, M. Gubelmann, and A. F. Williams, *Inor. Chem.* **23**, 1953 (1984).
- [292] R. Sato, H. Saitoh, N. Endo, S. Takagi, M. Matsuo, K. Aoki, and S.-I. Orimo, *Appl. Phys. lett.* **102**, 091901 (2013).
- [293] S. Takagi, Y. Iijima, T. Sato, H. Saitoh, K. Ikeda, T. Otomo, K. Miwa, T. Ikeshoji, and S.-I. Orimo, *Sc. Rep.* **7**, 44253 (2017).
- [294] W. Bronger and G. Auffermann, *Chem. Mater.* **10**, 2723-2732 (1998).