THE UNIVERSITY *of* EDINBURGH

Theoretical Investigation of Solid Hydrogen and Deuterium

Ioan-Bogdan Magdău

Doctor of Philosophy
The University of Edinburgh
September 2016

Abstract

Solid hydrogen forms at extreme conditions, under high pressures. Although the hydrogen atom is easy to understand theoretically, when interacting in the solid state it becomes complicated. Up to now, five different solid phases have been confirmed experimentally and theory has predicted numerous competing crystal candidates. The goal is to obtain solid metallic hydrogen which has been predicted theoretically eighty years ago and has since been considered the holy grail of high pressure science. In nature, this form of matter is believed to exist at the core of large planets like Jupiter and Saturn, being responsible for the planets' large magnetic fields. Understanding the different phases of hydrogen is a test for our most advanced theories of quantum mechanics in condensed matter and it is fundamentally important for both planetary and material science.

Recently discovered solid phase IV is stabilized by entropy and therefore only exists at relatively high temperatures. Using molecular dynamics (MD) I studied the room temperature behavior of phase IV starting with the ground state candidate structures reported in the literature. Additionally, I devised a velocity projection method for extracting Raman spectra from MD in light of direct comparison to experiment. My results helped establish the true nature of phase IV and validated the structure against experimental data. Applying the same method to the previously proposed $C2/c$ crystal structure, I obtained results that confirm this structure is the best candidate for phase III.

Within the last year, a new phase V of solid hydrogen was discovered in Raman experiments. While attempting to identify the crystal structure associated with this new phase, I discovered a manifestation of solid hydrogen in the form of long polymeric chains that could be stabilized by a charge density wave. Here I discuss the possibility of such a state of matter as an intermediate on the path to molecular dissociation of hydrogen. *Chains* could, however, be a spurious structure - the effect of a subtle non-convergence problem in the MD, which could

indicate serious issues with many previous studies reported in the literature. A far more likely candidate for phase V is a structure similar to that of phase IV with a subtle dynamical modification. I will present Raman and phonon results from both static and dynamic calculations to support this claim. I conclude my work on pure solid hydrogen with an instructive model that could explain the entire phase diagram based on simple thermodynamic considerations. All of the assumptions were extracted from our previous *ab initio* studies through analysis and observations. This model encodes a comprehensive summary of the current understanding of solid hydrogen at high pressures.

Raman and infrared spectroscopy have been the methods of choice in most hydrogen studies. Another way to look at the problem is to analyze the behavior of isotopic mixtures: hydrogen-deuterium binary alloys. Using isotopic substitutions, I revealed a textbook effect in hydrogen: phonon localization by mass disorder. The effect might be unique to this element, owing to the large mass ratio between hydrogen and deuterium. Phonon localization explains the complicated Raman spectra obtained experimentally in hydrogen-deuterium mixtures at various concentrations. More recent experimental results claim an unexpected phase transition in mixtures at low temperatures based on splittings in the infrared spectra. Here I will show that the infrared splitting seen experimentally could be induced by mass disorder in phase III and does not necessarily indicate a structural transformation.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in [1–9].

(*Ioan-Bogdan Magdău, September 2016*)

Acknowledgements

First, I would like to thank my supervisor Prof. Graeme Ackland, who has been the best mentor I could have hoped for. Many of the ideas presented here were born from long discussions and brainstorming with Graeme. He has offered me excellent guidance and support and formed me into a young scientist.

I thank Prof. Eugene Gregoryanz and collaborators: Dr. Phillip Dalladay-Simpson and Dr. Ross Howie who were very supportive and helpful all along, giving me insight into the experimental side of the story. Thanks to Lecturer Andreas Herman and Reader Ingo Loa for the meaningful discussions during our weekly group meetings.

I thank all the summer students I worked with: Bálint Borgulya, Benjamin Tyson, Floris Balm, Apurva Dhingra and Francesco Sarandrea. They have contributed to my work, some of which was included in the second part of chapter 3, where I will give them appropriate recognition.

Thanks to my office mates and true friends Giovanni Brandani and Li Tao for interesting discussions and good company.

Last but not least, my eternal gratitude to my family and friends for surrounding me with their love and care.

Contents

Chapter 1

Introduction

1.1 Short History

It was late in the 18^{th} century when the chemist Smithson Tennant burned pieces of diamond and recognized they transformed into regular charcoal [10]. This crucial observation spawned a global race for achieving the inverse process, namely obtaining diamonds out of ordinary carbon products. It took more than a hundred years of development until diamond could be produced in the lab by compressing carbon at elevated temperatures. These initial struggles gave birth to the new field of high pressure science.

Soon, scientists around the world realized they could pressurize a range of materials, from pure elements to complex mixtures of elements. In 1946, Percy W. Bridgman received the Nobel Prize for his achievements in high pressure physics. He had built an apparatus to produce pressures in excess of 10 GPa [11]. He studied the elements, plastics, glasses, minerals, alloys and many other compounds, monitoring their compressibility, thermal conductivity, electrical resistivity etc; and found intriguing properties [12, 13] that spawned great interest in the field.

High pressure science became a tool for creating exotic materials but also a means for testing our understanding and developing new theories regarding the structure of matter. With slight refinements to Bridgman's apparatus, the diamond anvil cell (DAC) was invented in 1959 [14], pushing up the pressure scale about two orders of magnitude. This device, which is still being refined today, can now

reach static pressures comparable to the ones found in the center of the Earth (400 GPa). In the light of new technologies, all elements from the periodic table have been compressed to extreme conditions, giving rise to unexpected behavior: oxygen was transformed into a superconductor at pressures around 100 GPa [15], sodium changed from a metallic conductor to a transparent insulator around 200 GPa [16] and so on. Very interesting, and still the topic of much debate is the unusual case of hydrogen.

The most abundant atom in the Universe, hydrogen is found naturally at all pressures scales: from the intergalactic vacuum ($10^{-18}Pa$) to the core of exoplanets and brown dwarfs ($10^{15}Pa$) [17, 18]. Hydrogen is the lightest of all elements in the periodic table and has the simplest structure. It comprises of one proton and a single electron which can be analytically described using quantum mechanical theory. Although well understood at normal conditions, the behavior of hydrogen at elevated pressures continues to surprise the scientific community.

Wigner and Huntington discussed in 1935 the possibility of hydrogen becoming metallic at pressures in excess of 25 GPa [19] and later on, in 1968 Ashcroft predicted it might also become a superconductor at high temperatures [20]. Since then hydrogen has been in the spotlight of high pressure science and considerable effort has been made towards achieving its metalization. In the following, I will shortly review the most important experimental and theoretical advancements in the race to metalize hydrogen. I will also outline how this thesis contributes to extending our knowledge in this field.

1.2 Overview

When Wigner and Huntington studied the possibility of hydrogen metalization back in 1935 [19], they assumed at high pressures hydrogen would adopt the simple body centered cubic (bcc) configuration. They used the simple Coulomb potential to model the ions and numerically solved the Schrödinger equation for the electrons in the bcc structure. Upon adding three important corrections to the electron energy: exchange, correlation and Madelung; they found that atomic bcc hydrogen is much less stable than the regular molecular crystal at normal conditions and the bcc structure could only be achieved at considerably higher densities, at pressures in excess of 25 GPa, which seemed out of reach at the time. Perhaps more intriguing is that they mentioned the possibility that "layered-like

lattices" might become metallic faster than bcc upon increasing pressure. As I will show later, modern theoretical calculations show they had the right intuition.

Wigner's suggestion aroused a lot of interest in solid hydrogen and motivated more theoretical studies with exotic predictions. Ashcroft applied the Bardeen-Cooper-Schrieffer (BCS) theory to the proposed metallic state. He calculated the transition temperature to a superconducting state to be considerably higher than that of other alkali metals [20]. Ashcroft also outlined that understanding solid hydrogen could bear significance for planetary science. He pointed out that conditions at the core of the largest planets in our solar system might be sufficient to sustain hydrogen in a superconductive state which could explain the strong magnetic fields surrounding some of these celestial bodies.

The exciting predictions regarding solid hydrogen do not stop at superconductivity. At high density, it is expected that the electron will migrate from the intramolecular to the intermolecular region. In combination with the high zero point energy of the light protons, this could result in more intriguing physical phenomena. For instance, using *ab initio* methods Bonev *et al.* predicted that hydrogen might become a metallic liquid at pressures of around 400 GPa and low temperatures close to zero kelvin [21]. They also noted that the existence of a ground state liquid requires a turn-over in the slope of the melting curve. Interestingly, the turn-over has been recently confirmed experimentally by Howie *et al.* [22], but the existence of a ground state liquid is still debated. Simultaneously with Bonev, Babaev *et al.* applied a Ginzburg-Landau model to liquid metallic hydrogen, accounting for the possible interaction between electron-electron and proton-proton Cooper pairs [23]. Their study predicts a unique superconductivity to superfluidity transition, switchable with an external magnetic field.

Enthusiasm in the field of high pressure hydrogen was ignited by theoretical predictions which, in turn, promoted technological advancements, making experiments at high pressure possible. The most notable advancement was the development of the diamond anvil cell (DAC), which increased the pressures achievable in experiments by orders of magnitude. In a DAC, two diamonds are centered on a hole in a metal gasket and pressed towards one another as shown in figure 1.1. There are many types of DACs, but they are all based on the same principle: the high pressure is achieved by focusing the applied stress on a very small region where the sample is positioned [25]. This clever technique presents an important disadvantage in that only small samples, of the order of a few μm

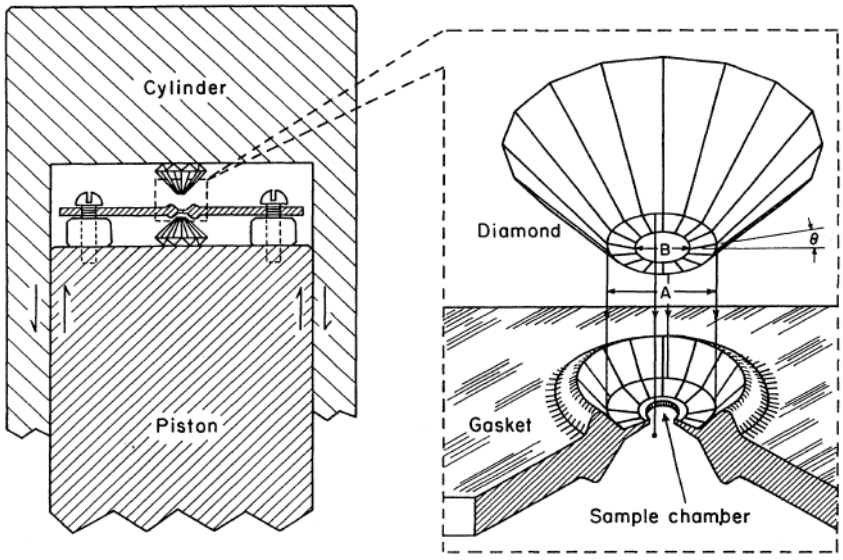

Figure 1.1 *The figure reproduced from ref [24] shows the schematics of a diamond anvil cell: simple experimental setup (left), zoom on the actual diamond and important parameters (right)*

can be examined. Additionally, the stress-strain fields are non-uniform, making it difficult to obtain homogeneous conditions of pressure and temperature.

The crystal structure of a material is the most important aspect needed for understanding the nature and characteristics of a given phase. Knowing the crystal structure allows one to perform powerful theoretical calculations and predict the physical properties of any material. Unfortunately, particularly in the case of hydrogen, the limited sample size combined with the much stronger scattering from the diamond anvils makes X-ray and neutron diffraction exceedingly difficult. These methods are critical for obtaining reliable crystallographic information and since they do not apply for hydrogen at high pressure, there is little hope for obtaining the precise crystal structure of the different phases of hydrogen. Optical spectroscopy does, however, provide a workaround and, to date, offers the most reliable connection between theory and experiment.

Apart from being the strongest naturally occurring material, diamond offers another important advantage to high pressure experiments: it is transparent to visible and infrared light up to very high pressures. This makes it possible to obtain important information about the tiny hydrogen samples, using mainly Raman and infrared absorption techniques, alongside visible absorption and reflectivity observation. Nevertheless, the crystallographic details cannot be obtained by these means alone and it is here where the theoretical models come

into play. A joint effort between experiment and theory has been proven very successful in understanding the various hydrogen phases.

On the theoretical side, most studies rely on a range of *ab intio* methods which I will review in the next chapter. The goal is to calculate the most stable from a range of tryout structures. Raman and infrared calculations are then performed for the energetically competitive lattices and the results are compared to the experiment. This general approach informs on the possible structure candidates for the actual hydrogen phases found in the lab. As we shall see later, there are many imperfections associated with this approach and these shortcomings generate many debates in the field.

In the following chapter, I will clarify the various theoretical methods that have been used throughout the literature and I will summarize our current understanding of the hydrogen phase diagram at extreme conditions. I will then follow up with three chapters of results to report on my own work in the field and finally conclude with a chapter of remarks and comments on the future outlook for high pressure hydrogen research.

Chapter 2

Background

2.1 Review of Theoretical Methods

I will begin with a short summary of the most important computational methods that are being used to study hydrogen at extreme conditions.

2.1.1 Structure Searching and Energy Calculations

Density Functional Theory (DFT) is at the center of most algorithms in condensed matter, and it has been the driving force for advancing the field of solid hydrogen as well. More recently, quantum Monte Carlo methods have been used to substitute for DFT, but since all the results I will present here were obtained with DFT, I will focus on reviewing this method.

The general Hamiltonian of a condensed matter system with N_e electrons and N_n nuclei can be expressed as:

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{ne} \tag{2.1}$$

which comprises of the kinetic energy of the nuclei, the kinetic energy of the electrons, the nuclear-nuclear interactions, the electron-electron interactions and the nuclear-electron interactions. Solving the Schrödinger equation for this general Hamiltonian is a difficult problem, but a solution can be obtained within the DFT formalism by introducing a series of clever approximations. Since the

mass of the nuclei is considerably larger than that of the electrons, the general problem can be decoupled in two parts: one dealing with the electronic degrees of freedom and one with the nuclear ones. This is called the Born-Oppenheimer approximation.

The interesting Hamiltonian is the one associated with the electrons and it can be written as:

$$\hat{H} = -\sum_{i=1}^{N_e} \frac{\hbar^2}{2m_i} \nabla_i^2 + \frac{1}{4\pi\varepsilon_0} \sum_{i=1}^{N_e} \sum_{j=i+1}^{N_e} \frac{e}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \frac{1}{4\pi\varepsilon_0} \sum_{i=1}^{N_e} \sum_{k=1}^{N_n} \frac{-Q_k}{|\boldsymbol{r}_i - \boldsymbol{R}_k|} \tag{2.2}$$

where r_i and R_k are the positions in space of the electrons and nuclei, respectively, m_i is the mass of the electrons and Q_k is the charge of the nuclei. The first two terms are universal, whereas the last term depends on the crystal structure.

Having defined the Hamiltonian, the ground state energy can generally be found by solving for the many-body wave function of the system. This is, however, almost intractable for a setup with many electrons. DFT provides an alternative where instead of describing the system based on the complicated wave function, the problem can be reformulated in terms of a scalar field, called the electron charge density $n(\boldsymbol{r})$. The first of the Hohenberg-Kohn theorems states that the Hamiltonian is entirely determined by $n(\boldsymbol{r})$ and each of the energy terms is a functional of the electron density:

$$E[n(\boldsymbol{r})] = T_e[n(\boldsymbol{r})] + V_{ee}[n(\boldsymbol{r})] + V_{ne}[n(\boldsymbol{r})] \tag{2.3}$$

The second theorem states that the minimum energy of the system is obtained if and only if the electron density is the ground state density. As a result the ground state can be computed starting from an initial guess for the charge density, and advancing towards the minimum energy in a self-consistent fashion.

The last term in equation 2.3 is simply:

$$V_{ne}[n(\boldsymbol{r})] = \int_V n(\boldsymbol{r}) V_{ne}(\boldsymbol{r}) d\boldsymbol{r} \tag{2.4}$$

The challenging problem is evaluating the first two terms in equation 2.3, in particular the kinetic one. The kinetic functional of the electron density is not known exactly and it is a major drawback in standard DFT. This problem is solved in Kohn-Sham DFT, where the system of interacting electrons is reduced to an equivalent system of noninteracting particles described by the pseudo-orbitals

ψ_i. The corresponding fictitious fermions are chosen to reproduce the charge density of the original system:

$$n(\boldsymbol{r}) = \sum_{i}^{N_e} |\psi_i(\boldsymbol{r})|^2 \tag{2.5}$$

The kinetic energy can then be easily computed from:

$$T_{KS} = -\sum_{i=1}^{N_e} \frac{\hbar^2}{2m_i} \int_V \psi_i^*(\boldsymbol{r}) \nabla_i^2 \psi_i(\boldsymbol{r}) d\boldsymbol{r} \tag{2.6}$$

However, this is just an approximation to the true kinetic energy T_e. The electron-electron interaction V_{ee}, comprises of a classical Coulomb repulsion that gives the Hartree energy:

$$V_H = \frac{e}{4\pi\varepsilon_0} \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \int_V \int_V \frac{|\psi_i(\boldsymbol{r})|^2 |\psi_j(\boldsymbol{r'})|^2}{|\boldsymbol{r}_i - \boldsymbol{r}'_j|} d\boldsymbol{r} d\boldsymbol{r'} \tag{2.7}$$

and a quantum interaction that includes the Pauli exclusion and is part of a contribution called the exchange-correlation energy V_{XC}. In KS-DFT, the exchange-correlation also includes the kinetic correction $T_e - T_{KS}$. A common problem with equation 2.7 is that it overestimates the energy by including self-interactions, where an electron falsely ends up interacting with itself.

The equations that determine the pseudo-orbitals can be found by applying the variational principle to the total energy:

$$\left[-\frac{\hbar^2}{2m_i} \nabla_i^2 + \frac{e}{4\pi\varepsilon_0} \sum_{j=1}^{N_e} \int_V \frac{|\psi_j(\boldsymbol{r'})|^2}{|\boldsymbol{r}_i - \boldsymbol{r}'_j|} d\boldsymbol{r} d\boldsymbol{r'} + \frac{\delta V_{XC}}{\delta n(\boldsymbol{r})} + V_{ne}(\boldsymbol{r}) \right] \psi_i = \varepsilon_i \psi_i \tag{2.8}$$

where these are called the Kohn-Sham equations and are equivalent to the Schrödinger equation for the fictitious system [26, 27]. In this theoretical framework, each pseudo-particle interacts with the other particles via a mean field potential as shown above.

The remaining problem with Kohn-Sham DFT is the evaluation of the exchange-correlation energy. The functional for exchange-correlation is unknown and it can only be approximated, usually based on the energy of a homogeneous electron gas. The simplest approximation is the local density approximation (LDA) which estimates the exchange-correlation energy based on the value of the charge density

at every location. More advanced methods estimate the exchange-correlation energy not only based on the point value of the density but also the gradient. These methods are called general gradient approximations (GGA), the most used of which is the Perdew Burke and Ernzerhof (PBE) method [28]. Quantum Monte Carlo solves the exchange-correlation problem by statistically sampling the real many body wave function. My results, which I will present in the later chapters, were all based on DFT.

In KS-DFT, the electron density is normally expended in an orthonormal basis set, usually as a set of gaussians (local basis set) or a set of plane waves (non-local basis set). The latter is more practical for use in periodic crystals, therefore most results in solid hydrogen are based on plane wave calculations. An additional complication is given by the divergent nature of the Coulomb interaction. Since nuclei are treated as massive point charges, in hydrogen the value of the potential at the center of the atom is infinite and the wave function has a cusp. To circumvent this problem a range of pseudo-potentials are used in practice to approximate the interaction near the center of the nucleus. These pseudo-potentials are identical to the Coulomb potential outside a cut-off radius and vary smoothly near the center. Pseudo-potentials are further divided into norm-conserving and ultra-soft depending on whether they conserve the norm of the wave function inside the cutoff radius or not. For each value of the electron angular momentum, a different non-local pseudo-potential is required. This sums up in a very concise manner the main challenges with DFT.

DFT can be seen as the fundamental building unit in algorithmic condensed matter. Going a level up, more complex routines like geometry optimization are constructed upon DFT. The goal of geometry optimization is to find the equilibrium atomic configuration of a crystal at zero temperature. The algorithm proceeds as follows: it starts with an initial crystal structure, calculates the total potential energy using DFT, then evaluates the forces on each atom i as a gradient of the energy surface: $\vec{F_i} = -\nabla_{\vec{r}_i} U(\vec{r})$, and finally employs a minimization method to update the atomic positions and minimize the total strain. The routine is repeated until the force on each of the atoms is zero. The geometry optimization technique is crucial and most algorithms employ it to relax the initial structures, which then become the basis of subsequent calculations.

Yet one level up in complexity we find the recently proposed Ab Initio Random Structure Searching (AIRSS) algorithm. AIRSS was first used by Pickard and Needs to successfully solve two of the high pressure phases of SiH_4 [29]. AIRSS

involves a rather simple, yet effective method. The routine generates a series of random crystals parametrized by the lattice constants: cell lengths (a, b, c) and cell angles (α, β, γ); together with a number N of atomic positions (r_i, $i = 1, 2, ...N$). The variables are constrained by some reasonable conditions; for example, the crystal should be a homogeneous continuum, with nearest neighbor distances ranging from 0.75 to 3 Å[30]. Each structure is then geometrically optimized. At the same time, the enthalpy is calculated for each candidate arrangement and then compared to decide the most energetically favorable crystal conformation. For a given number of atoms in the unit cell, the procedure is repeated until many configurations are generated more than once, ensuring an exhaustive search.

As we will see in the literature review section, the AIRSS approach has been essential for the understanding of the high pressure phases of solid hydrogen. In the case of hydrogen, there is a complication to the energy evaluation, in that some phases are stabilized by entropy and zero point energy. The usual geometry optimization drives the structure in the ground state, at zero temperature, and therefore evaluates the static energy E_{tot}. To correct for dynamical contributions a thermodynamics calculation is usually performed, which evaluates the vibrational and zero point energies. The latter is found to be substantial in the case of atoms with light nuclei (e.g. hydrogen and its isotopes). In the following, the steps for the thermodynamical correction are given as implemented in the code CASTEP [31].

At first the phonons are calculated from Density Functional Perturbation Theory (DFPT) [32] which I will return to in section 2.1.3. As such the phonon density of states $f(\omega)$ can be evaluated. The zero point energy is thus given by:

$$E_{zp} = \int_0^\infty \frac{\hbar\omega}{2} f(\omega) d\omega \tag{2.9}$$

Each phonon is a boson that in the harmonic approximation stores an energy quantum $\hbar\omega$. The energy distribution over the collection of phonons follows Bose-Einstein statistics, thus at a given finite temperature T, the energy stored in the vibrational motion is:

$$E_{vib} = \int_0^\infty \frac{\hbar\omega f(\omega)}{\exp(\frac{\hbar\omega}{kT}) - 1} d\omega \tag{2.10}$$

while the total Helmholtz free energy at fixed volume, in the harmonic approximation, is given by:

$$F(V,T) = E_{tot} + E_{zp} + kT \int_0^\infty f(\omega) \ln \left[1 - \exp\left(-\frac{\hbar\omega}{kT} \right) \right] d\omega \tag{2.11}$$

The entropy can also be evaluated from:

$$S = \frac{1}{T}(E - F) = \frac{1}{T}(E_{tot} + E_{zp} + E_{vib} - F(T)) \tag{2.12}$$

where $S = S(V,T)$.

Although these calculations give a good measure of the dynamical effects, they are still an approximation in the harmonic regime. In some materials, like various phases of hydrogen, the anharmonic behavior is pronounced. Therefore, for a better theoretical description one can employ proper dynamical calculations.

2.1.2 The Role of Temperature: Dynamical Simulation

Classical molecular dynamics (MD)[33] is an algorithm that simulates the motion of atoms at finite temperature. For a given interaction potential and a set of initial conditions and boundaries, the MD algorithm estimates, at every given step, the force on each particle i as a sum of contributions from all other particles $j \neq i$. Subsequently, the routine employs an integration scheme (e.g. Runge-Kutta, Leap Frog, Verlet) and estimates both the velocities and the new positions for each of the particles. Combined with DFT, the method can be used to simulate real crystal structures at finite temperatures and pressures. There are two main variants of *ab initio* MD: Born-Oppenheimer (BO) and Car-Parrinello (CP) [34]. In BO-MD, the electronic forces on each atom are extracted at every step from a DFT calculation. When DFT is performed, the nuclei are fixed and later the atomic positions are updated following an integration scheme (like in classical MD). On the other hand, in CP-MD the electron degrees of freedom are associated small fictitious masses and they are included in the dynamics. The results I will show later were derived from BO-MD.

Ab initio MD can be performed in a range of ensembles depending on the boundary conditions. The most natural for simulations is the NVE ensemble

where the number of particles, volume and total energy are kept fixed. NVE is appropriate when the system under study is insulated and does not exchange energy and momentum with the outside world. In experiments, however, the sample interacts with the setup and the system is more appropriately described by the Gibbs free energy. Calculations can mimic these conditions by using the NPT ensemble, where the number of particles, pressure and temperature are kept constant. In NPT, temperature is kept fixed by a thermostat, usually Langevin or Nose-Hoover and pressure is fixed by a barostat, typically Andersen-Hoover or Parrinello-Rahman [31]. The Langevin thermostat adjusts the velocities by introducing a set of friction forces to simulate a viscous behavior, whereas the Nose-Hoover introduces an additional fictitious particle acting as a heat bath into the dynamics. The Anderson barostat allows volume to vary by introducing three additional degrees of freedom corresponding to the lengths of the box, whereas the Parrinello barostat also allows the angles to change.

Many dynamical properties can be extracted from MD. Since the routine simulates the actual motion of the atoms, both the potential and the kinetic energy are readily estimated from the dynamics. However, MD treats the nuclei classically, therefore the zero point motion is not accounted for in the calculations. A modification of the MD algorithm called PIMD, is based on the path integral formalism and can achieve proper quantum mechanical treatment of light nuclei, such as the single protons in the case of hydrogen.

According to the minimum action principle in classical mechanics, from all the possible paths a particle can follow, the particle will always choose the path that requires the smallest possible action. In the path integral formalism of quantum mechanics, all possible paths must be considered. The final path of the particle is the sum over all possible paths weighted by a complex exponential factor that depends on the action of each path. As such, the quantum system can be simulated with an infinite number of classical systems that explore the different paths.

PIMD exploits precisely this idea. The nucleus is simulated by a closed chain of beads, where each bead samples a different path in imaginary time, and all the beads are connected to one another by harmonic oscillators as shown in figure 2.1. In practice, this means that PIMD essentially comprises of a number of separate *ab initio* MD trajectories that each simulate one of the beads from the closed chain that describes one quantum nucleus. If the number of beads is large enough, this algorithm asymptotically approximates the real wave function of the

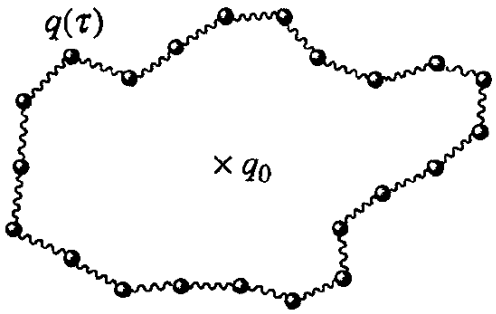

Figure 2.1 *A nucleus modeled as a polymeric ring of fictitious classical nuclei, living in imaginary time, with periodic boundary conditions and connected by an harmonic potential [35, 36].*

nucleus and the associated zero point energy. In the case of hydrogen which has a light nucleus, the zero point energy is important, so PIMD is more appropriate than MD to study the dynamics. However, since PIMD typically consists of a large number of beads, these simulations are substantially more expensive than MD.

Returning to structure searching, MD simulations can be used to solve some entropy stabilized structures. While in the AIRSS approach, the most stable crystal at a given pressure and temperature is chosen from a collection of randomly generated candidates, in MD and PIMD the energetically favored form should be achieved naturally, especially in ensembles that allow the cell to vary as shown by Parrinello *et al.* [37]. Given a large and long enough simulation, independent of the starting atomic configuration, MD and PIMD should drive the crystal through various phase transitions to the most stable structure.

In practice, however, the energy barrier of many phase transformations might be much higher than the available thermal energy and thus the probability of such an event happening within reasonable simulation time is well reduced. Additionally, the transition can only take place if the number of atoms in the unit cell of the initial structure is a multiple of the number of atoms of the final structure. These factors can lead to nonphysical metastability.

On the other hand, small simulation cells can promote transitions, but these are spurious, owing to large energy fluctuations compared to the size of the system. In practice, MD simulation with small systems tend to go back and forth between all the energetically competitive crystal structures. These are called finite size effects [1] and can be a big problem in hydrogen simulations. Intuitively, in

smaller cells, fewer atoms need to rearrange, thus increasing the probability of a bogus phase transition.

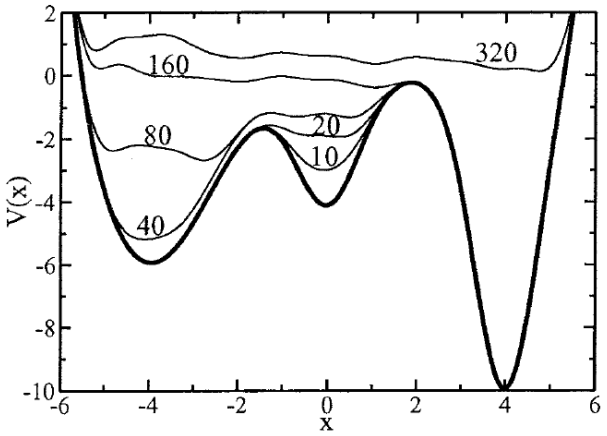

Figure 2.2 *A one dimensional representation demonstrating how metadynamics efficiently scans the potential energy surface (PES). The structure starts in a local minimum but it is encouraged to migrate towards the lowest energy point. The figure was reproduced from ref [38].*

A solution to the metastability problem in dynamical structure searching is offered by the metadynamics algorithm. The method was introduced by Laio *et al.* [38] and later refined for pressure induced phase transformations [39]. In one metadynamics step, several MD simulation replicas are run simultaneously, independent of each other, starting from the same set of initial coordinates s_{0i}. A finite number of MD steps ensures that a local region on the potential energy surface (PES) has been scanned around s_{0i}. A collection of forces is then obtained as derivatives of PES with respect to s_i. The forces obtained as such also contain a history term that slowly fills up the PES local minima, which encourages the system to move out of equilibrium and sample other regions as shown in figure 2.2. Filling local minima one by one, the method eventually drives the system towards a globally stable configuration.

There are many other algorithms that attempt to gauge dynamical properties and find the most stable structures at given conditions of pressure and temperature. Notably, Monte Carlo molecular modeling (MCMM) and Path Integral MCMM simulations which do not generate real dynamics, but can sample dynamical properties according to proper quantum mechanical statistics. Details of those algorithms are, however, beyond the purpose of this review.

2.1.3 Phonons: the Connection to Experiment

Finally, there are the spectroscopy simulation methods which are crucial for advancing the field of solid hydrogen. Predictions based on energetics alone are not always informative for the experiment. The errors involved in the different *ab initio* approximations such as the choice of exchange-correlation functional, combined with the experimental errors for measuring pressure and temperature, make it difficult to identify which crystal structure corresponds to which phase. The way forward is to consider all competitive structures as candidates and make a decision based on spectroscopic data.

Theory of lattice dynamics (LD) [40] is a well established framework that prescribes the recipe for phonon calculations. The basic steps in LD are given in the following [32]. Consider a solid with equilibrium atomic positions $r_{0i\alpha}$, where the indices i, j label the atom number and α, β, γ one of the x, y, z Cartesian coordinates. Call $V(r)$ the effective potential felt by the nuclei in the electronic environment. If all the atoms experience small displacements $u_{i\alpha}$ around the equilibrium r_0, $V(r)$ can be approximated as:

$$V(r_0+u) = V(r_0) + \sum_{i\alpha} \frac{\partial V}{\partial u_{i,\alpha}} u_{i\alpha} + \frac{1}{2} \sum_{i,j,\alpha,\beta} u_{i\alpha} \frac{\partial^2 V}{\partial u_{i\alpha} \partial u_{j\beta}} u_{j\beta} + ... \tag{2.13}$$

However, since the perturbation is added to the equilibrium state, the first derivatives cancel. Ignoring the first term which is just a constant and the higher order terms (i.e. anharmonic contributions), the new potential becomes:

$$V(r) = \frac{1}{2} \sum_{i,j,\alpha,\beta} u_{i\alpha} \phi_{i,j}^{\alpha,\beta} u_{j\beta}, \quad \phi_{i,j}^{\alpha,\beta} = \frac{\partial^2 V}{\partial u_{i\alpha} \partial u_{j\beta}} \tag{2.14}$$

where $\phi_{i,j}^{\alpha,\beta}$ is called the *force constant matrix* or *Hessian matrix*. The equation of motion for the above potential takes the form (left) with Ansatz (right):

$$m_i \ddot{u}_{i\alpha} = -\sum_{j\beta} \phi_{i,j}^{\alpha,\beta} u_{j\beta}, \quad u_{i\alpha} = \sqrt{m}_i \epsilon_{i\alpha q} e^{i(q r_{i\alpha} - \omega_q t)} \tag{2.15}$$

with q labeling the phonons' wave vector, ω_q their frequency and $\epsilon_{i\alpha q}$ the normal

mode. After substituting the solution, an eigenvalue problem is obtained:

$$D_{i,j}^{\alpha\beta}(q) \cdot \epsilon_{i\alpha q} = \omega_q \epsilon_{i\alpha q}, \quad D_{i,j}^{\alpha\beta}(q) = \frac{1}{\sqrt{m_i m_j}} \sum_{\gamma} \phi_{i,j}^{\alpha,\beta} e^{-iqr_\gamma} \tag{2.16}$$

where $D_{i,j}^{\alpha\beta}(q)$ is just the Fourier Transform of the Hessian matrix [31]. For any q there are exactly 3N eigenvalues (i.e. frequencies and eigenmodes), with N being the number of atoms in the unit cell. The gamma point phonons can be calculated by setting $q = 0$.

In practice, the problem lies in constructing the Hessian, for which two independent methods exist: Finite Displacement (FD) and Density Functional Perturbation Theory (DFPT). In FD the atoms in the unit cell are displaced one at a time along x, y and z directions. For each perturbation, the force contribution to all other atoms is calculated, which in turn allows the evaluation of the Hessian matrix. The main limitation of this algorithm is that, in reality, contributions from mirror images, located in neighboring cells, are wrongly included in the values of the forces. Therefore this procedure only works well when the atomic interactions are short range and do not extend beyond the unit cell.

The DFPT algorithm calculates the Hessian from linear response theory using Hellmann-Feynman theorem, which can be derived as [32, 41]:

$$V(r) = \langle\psi| H |\psi\rangle = \int \psi^* H \psi dr \tag{2.17}$$

$$\frac{dV}{dr} = \left\langle \frac{d\psi}{dr} \middle| H |\psi\rangle + \langle\psi| H \middle| \frac{d\psi}{dr} \right\rangle + \langle\psi| \frac{dH}{dr} |\psi\rangle \tag{2.18}$$

$$\frac{dV}{dr} = V \frac{d}{dr} \langle\psi| \psi\rangle + \langle\psi| \frac{dH}{dr} |\psi\rangle = \int \psi^* \frac{dH}{dr} \psi \tag{2.19}$$

Taking second derivatives, which are needed for the Hessian, one finally obtains:

$$\frac{d^2V}{dr^2} = \left\langle \frac{d\psi}{dr} \middle| \frac{dH}{dr} |\psi\rangle + \langle\psi| \frac{dH}{dr} \middle| \frac{d\psi}{dr} \right\rangle - \langle\psi| \frac{d^2H}{dr^2} |\psi\rangle \tag{2.20}$$

The problem is finally solved by calculating the response terms (i.e. derivatives of the wave functions) in a self-consistent fashion by employing DFPT. Once the

phonons are calculated, one can proceed by evaluating which of the vibrations are infrared or Raman active, thus making the ultimate bridge to experiment.

Infrared absorption results from the transition of a vibrational mode at the gamma point (i.e. $q = 0$) to an excited level by absorbing part of the energy from an external field. When light is passing through a crystal, the infrared active modes will absorb at their natural frequencies, leaving gaps in the transmitted light spectrum. The intensity of the absorption is influenced by the change in the electric dipole moment occurring in a given oscillation mode. The intensity of a transition from state m to state n is given by [42]:

$$I_{nm} \propto \sum_{\alpha} \langle \psi_m | \, \mu_\alpha \, | \psi_n \rangle^2 = \sum_{\alpha} \left(\int \psi_m^* \mu_\alpha \psi_n \right)^2 \tag{2.21}$$

with α, β, γ being one of the Cartesian components x, y or z, as before and μ_α the dipole moment. Labeling the normal modes with k and normal vector $\epsilon_{k\alpha}$, the dipole moment can be developed as a series in the basis of the modes:

$$\mu_\alpha = \mu_\alpha^0 + \sum_k \frac{\partial \mu_\alpha}{\partial \epsilon_{k\alpha}} \epsilon_{k\alpha} \tag{2.22}$$

The intensity of absorption, thus, becomes:

$$I_{nm} \propto \sum_{\alpha} (\langle \psi_m | \, \mu_\alpha^0 \, | \psi_n \rangle + \sum_k \langle \psi_m | \, \frac{\partial \mu_\alpha}{\partial \epsilon_{k\alpha}} \epsilon_{k\alpha} \, | \psi_n \rangle)^2 \tag{2.23}$$

Since the states ψ_m, ψ_n are orthogonal to one another, the term $\langle \psi_m | \, \mu_\alpha^0 \, | \psi_n \rangle$ does not give any contributions. As such, the only contribution to the intensity arises from the change of dipole moment. Note that the integral $\langle \psi_m | \, (\partial_{\epsilon_{k\alpha}} \mu_\alpha) \epsilon_{k\alpha} \, | \psi_n \rangle$ is non-zero if the derivative is non-zero [42]. In simulations, the change in dipole moment with respect to the normal modes can be calculated as a linear response with DFPT or by numerical methods with finite displacement [43]:

$$\frac{\partial \mu_\alpha}{\partial \epsilon_{k\beta}} = -\frac{\partial^2 E(r)}{\partial A_\alpha \partial \epsilon_{k\beta}} = \frac{\partial F_k}{\partial A_\alpha} \tag{2.24}$$

where A_α is an external field and F_k are the forces corresponding to each mode k. Within the finite displacement method, for each value of the external field, a new

relaxation can be performed and the force contributions recalculated. As such, the infrared activity can be easily computed since only first order derivatives of the forces are required.

The Raman process results from the inelastic scattering of the photons by phonons. When an external field $A_\alpha = A^0_\alpha \cos(\omega t)$ that is not in resonance with some eigenfrequency of the lattice interacts with the electron density, it induces a dipole moment proportional to the polarizability tensor $\tilde{a}(\omega)$ [42]:

$$\mu_\alpha = \tilde{a}(\omega) A^0_\alpha \cos(\omega t) \tag{2.25}$$

For a normal mode labeled by angular frequency ω_k, the polarizability tensor can be written as:

$$\tilde{a}(\omega) = \tilde{a}(\omega_0) + \frac{\partial \tilde{a}}{\partial \epsilon_{k\alpha}} d\epsilon_{k\alpha} = \tilde{a}(\omega_0) + \frac{\partial \tilde{a}}{\partial \epsilon_{k\alpha}} \epsilon_{k\alpha} \cos(\omega_k t) \tag{2.26}$$

Combining the two equations yields:

$$\mu_\alpha = \tilde{a}(\omega_0) A^0_\alpha \cos(\omega t) + \frac{\partial \tilde{a}}{\partial \epsilon_{k\alpha}} A^0_\alpha \epsilon_{k\alpha} \cos([\omega - \omega_k]t) + \frac{\partial \tilde{a}}{\partial \epsilon_{k\alpha}} A^0_\alpha \epsilon_{k\alpha} \cos([\omega + \omega_k]t) \tag{2.27}$$

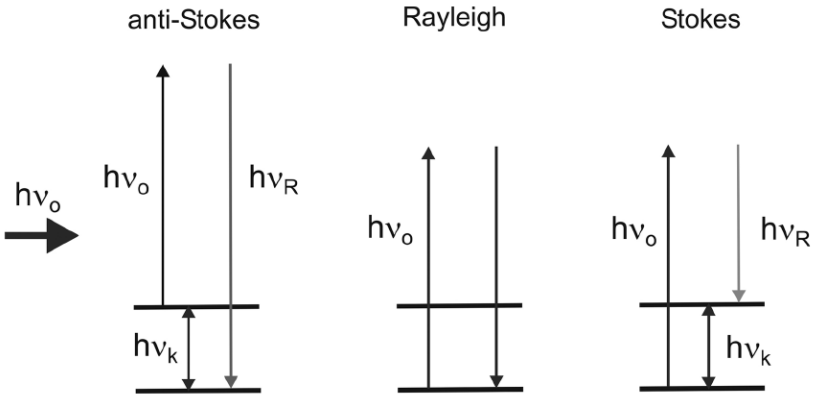

Figure 2.3 *This figure reproduced from ref [42] illustrates the Raman processes.*

This model offers a simple classical description of a Raman process. When an external field interacts with the crystal, it will excite the system to a virtual state. If the system was initially in the ground state, it can relax either back to the ground state, conserving the energy (i.e. elastic photon scattering or Rayleigh process) or to a real excited state, producing a photon with lower energy (i.e.

Stokes process). If the system was found in an excited state, the external field will add to the energy, and finally, the configuration will relax to the ground state, emitting a photon with higher energy (i.e. anti-Stokes process). All three mechanisms (i.e. Rayleigh, Stokes and anti-Stokes) are illustrated in figure 2.3 and they are present in equation 2.27 as the tree terms of the summation. In reality, there is a much smaller probability of finding the system in the excited state, so the cross-section of the anti-Stokes scattering is much smaller than that of Stokes, however, this is not explained by the simple classical model. Experimentally, by looking at the frequency difference between the Rayleigh and Stokes radiation, one can determine the frequencies of the Raman active modes. The temperature of the sample can also be deduced from the ratio between the Stokes and anti-Stokes Raman intensities.

The intensity of the active Raman modes depends on the derivative of the polarizability tensor. In practice this derivative can be evaluated within DFPT from [43, 44]:

$$\frac{\partial \tilde{a}}{\partial \epsilon_{k\alpha}} = -\frac{\partial^3 E(r)}{\partial A_\alpha \partial A_\beta \partial \epsilon_{k\gamma}} = \frac{\partial^2 F_k}{\partial A_\alpha \partial A_\beta} \tag{2.28}$$

In contrast with infrared, the Raman intensities require second order force derivatives to be evaluated as a response to different external fields. The calculation is, therefore, much more expensive.

The entire theory presented above works in the harmonic approximation. As we shall see later, hydrogen does not always subscribe to this special case. A standard method to account for anharmonicity is to extract the phonon frequencies from MD simulations, by taking the Fourier Transform of the velocity auto-correlation function (VACF) [45]. When the simulation is long enough and the structure remains well behaved, this method can extract reliable frequencies even in the anharmonic case. However, VACF gives the entire phonon spectrum, and there is no way to distinguish the infrared and Raman active modes. In chapter 3 I will show the method we developed to map the phonons onto the atomic trajectories, which allows us to calculate the Raman activity at finite temperature for anharmonic crystal structures [1, 3, 4].

Having summarized the most important theoretical techniques involved in the search and characterization of the high pressure hydrogen phases, we shall turn our attention to the progress made in the field up to date.

2.2 The Phase Diagram: Theory and Experiment

In this section, I will review our current understanding of the phase diagram of hydrogen under pressure, see figure 2.4. Starting with a short revision of the liquid phase, I will then progress through all the solid phases from I to the newly discovered phase V. This survey will cover the most important experimental and theoretical studies in the field in the past decades. For more comprehensive reviews, the reader can consult refs [18, 24, 46–49].

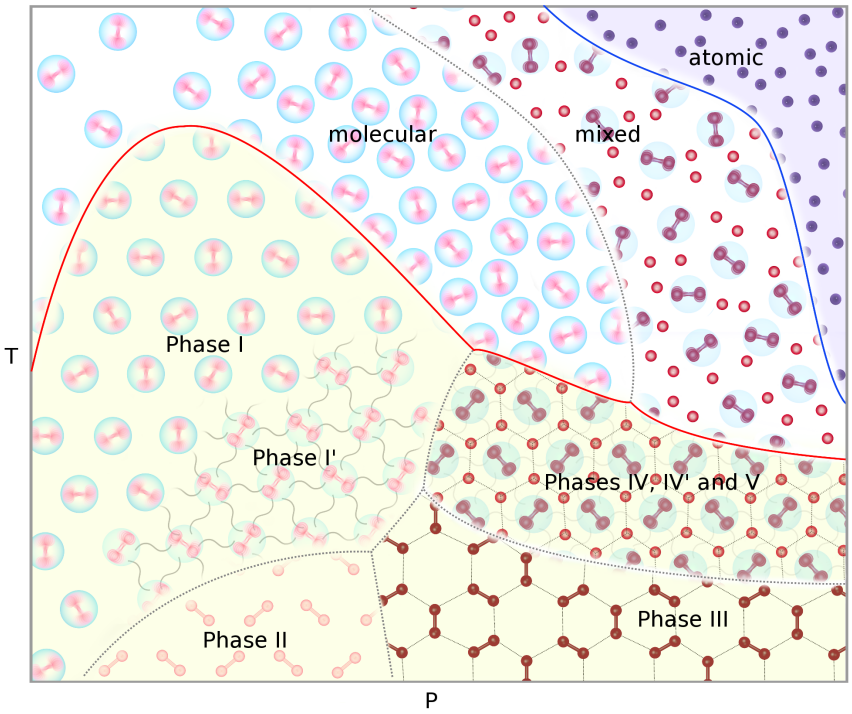

Figure 2.4 *Schematic illustration of the hydrogen phase diagram as we currently understand it. I show the solid phases in light yellow, separated from the liquid phases by the melting curve, depicted in red. In blue I show the line of molecular dissociation in the liquid phase. From ref [50].*

In figure 2.4 I show our current view on the phase diagram. The liquid is stable at high temperatures. At low pressure, the liquid is molecular and insulating, but at high pressure, it becomes atomic and metallic. A mixed atomic-molecular liquid is expected at intermediate pressures. In the second half of chapter 4, I will present an intuitive thermodynamical model that explains this intermediate state as a mixture following Boltzmann statistics. The melting line has a positive slope early in phase I and then switches sign when the liquid becomes denser than the solid. Higher up in pressure, the line flattens out as the solid phases become

more entropic.

There are five distinct solid phases, none of which is metallic. At low pressures, there are phases I and II. Phase I is a hexagonal close packing of quantum rotors, while phase II is a broken symmetry phase, where molecules become classical and assume a fixed orientation. Within phase I, there could be a variation - phase I', where the quantum rotors become correlated. At higher pressures, phases III, IV and V are stackings of molecular layers. While phase III comprises of one type of layer, phases IV, V comprise of two or more. Phases IV, IV' and V are very similar to one another and the subtle differences between them will be the topic of chapter 3. Further up in pressure, the solid could become atomic and then the liquid could become the ground state. In chapter 4 I will present an additional intermediate metallic structure that could become stable before the atomic phase.

2.2.1 Liquid Phases

Hydrogen metalization can be achieved following two different paths: by compressing in the solid phase or in the liquid phase [51]. Since the former remains out of reach, a lot of effort has been made to achieve metalization in the liquid state. The liquid is stable at high temperatures, where static compression experiments in DACs are limited by hydrogen diffusion and subsequent embrittlement. The alternative is to use shock compression where a pressure pulse compresses the sample along an equation of state called Hugoniot.

The first experimental confirmation of metallic hydrogen came in 1996, when Weir *et al.* [52] used shock compression to measure the electric conductivity of small hydrogen samples at temperatures of 2200-4400 K and pressures of 93-190 GPa. They found a significant increase in conductivity up to 140 GPa, above which the conductivity remained constant. They interpreted the results as a continuous transition from a semi-metallic to a metallic fluid at 140 GPa and 3000 K. This was the first account of metallic hydrogen. A more recent shock wave experiment [53] achieved the metalization of deuterium at around 300 GPa as illustrated in figure 2.5.

An experimental alternative used to achieve high temperatures in the liquid state is the laser heating technique [54, 55]. The laser heating experiments are performed in DACs where a laser is used to heat up only a small region

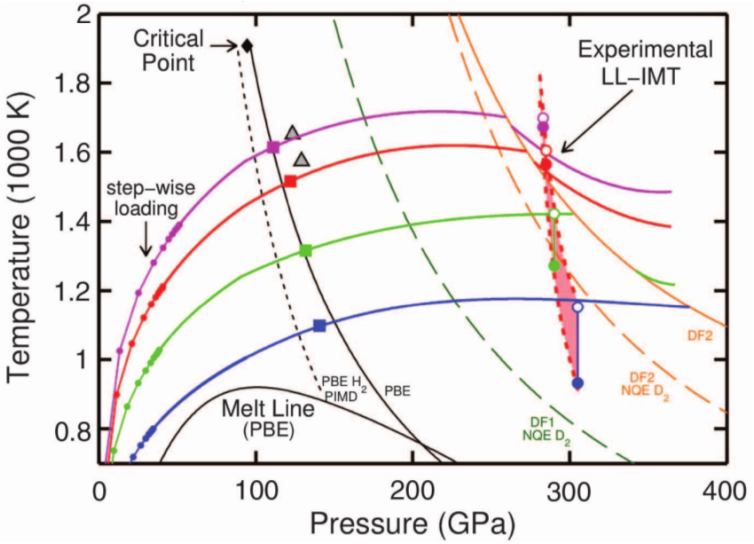

Figure 2.5 *The figure taken from ref [53] shows the PPT line in liquid deuterium at 300 GPa. The experimental result is illustrated in dashed red line. Solid purple, red, light green and blue are the compression paths. The black, dark green and orange lines are theoretical predictions within various approximations.*

of the sample in the vicinity of a strong absorber. These studies claimed to have identified a plasma phase transition (PPT) in liquid hydrogen at pressures much lower than the shock experiment. Concerns about the loss of sample and uneven temperatures make these results the subject of heated debates. The changes in the optical properties of the samples could, however, be related to a first order transition to a mixed atomic-molecular phase [56].

Unlike shock studies, laser heating experiments can also access lower temperature regimes, closer to the melting line [57, 58]. The transition from solid to liquid has captured a lot of attention mainly because of the prediction that the melting line could change slope with pressure as the liquid is gradually becoming more stable. More traditional DAC experiments where the samples are uniformly heated using a thermostat [59, 60] and are thus more reliable than laser heating studies, have attempted to find the maximum in the melting curve. The breakthrough came last year when Howie *et al.* mapped the melting line to a record of 250 GPa [22] by observing changes in the Raman spectra as shown in figure 2.6. For the first time, they showed clear evidence that the melting line has a maximum above which the melting temperature is decreasing upon increasing pressure. They also showed that the Raman vibron persists in the liquid, which brings counter-evidence to the liquid being atomic.

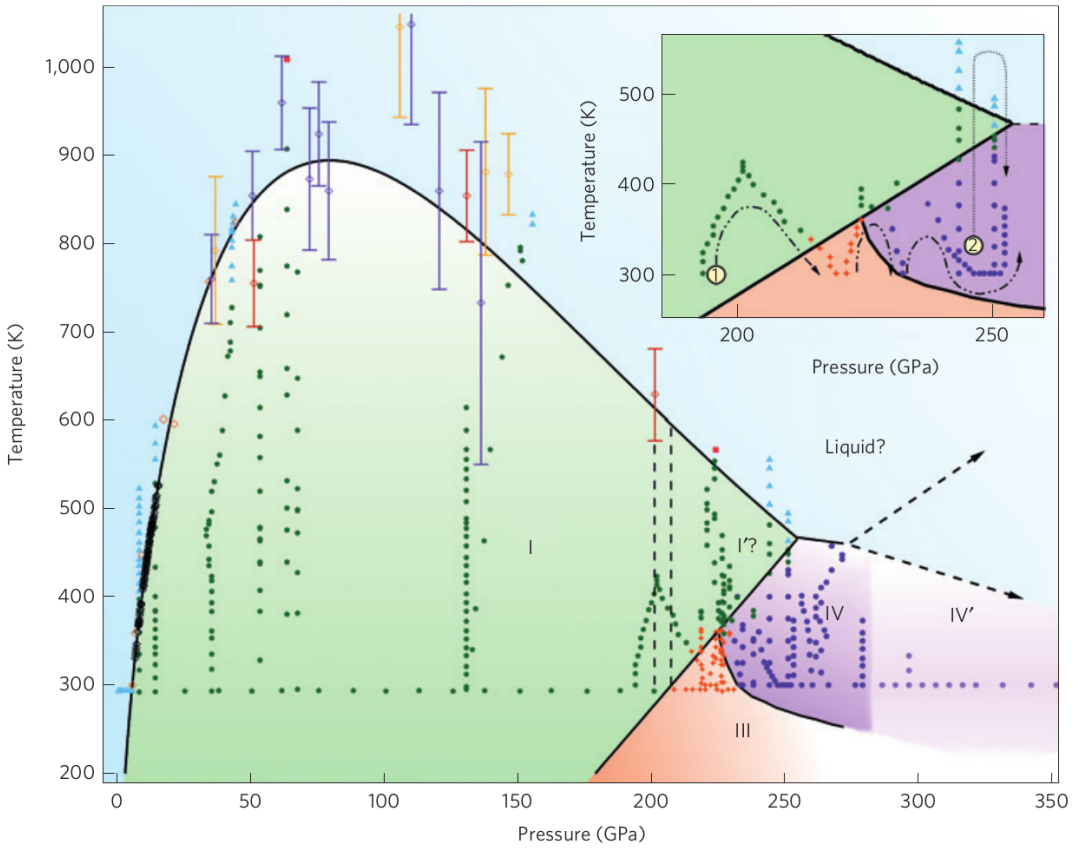

Figure 2.6 *The figure taken from ref [22] shows the most recent data on the melting line with clear evidence of a maximum and a change of slope sign with pressure. The data below 150 GPa is from previous studies, while the red diamonds with associated errors are from theory [21]. The data shown with filled symbols and without error bars is from the latest experiment [22].*

On the theory side, the first study attempting to find the PPT line was performed by Scandalo in 2003, who used *ab initio* MD to simulate liquid hydrogen in the pressure range 75-175 GPa and temperatures around 1500 K [61]. He found a first order transition from a molecular to an atomic fluid, indicated by an abrupt change in the pair distribution function. Moreover, the transition with little hysteresis was also accompanied by a substantial change in volume that made the liquid denser than the solid. Based on Clausius-Clapeyron relation this result indicated for the first time a possible downturn in the melting line.

One year later, Bonev *et al.* extended this previous theoretical study. They used the phase coexistence method to map out the melting line in the pressure range 50-200 GPa [21]. In their calculation, the PPT line was found at 200 GPa, 900-1000 K, in agreement with the former result. They also found a crossover

between the PPT and the melting line, yielding a liquid-liquid-solid triple point at 300 GPa, 400 K. Furthermore, their results indicated the existence of a ground state metallic fluid and supported the idea of a maximum in the melting line.

These initial results were based on DFT-MD. Nevertheless, they predicted the maximum in the melting curve long before the experimental confirmation. Recently, more advanced studies have used quantum Monte Carlo methods (QMC) to calculate the forces and run the dynamics [56, 62, 63]. QMC provides an exact numerical approximation for the exchange-correlation interactions. However, QMC is more expensive than DFT so, generally, it is performed on smaller systems, with fewer k-points. As I will show later, this can lead to finite size effects which, in some cases, could be more important than errors in the exchange-correlation energy.

Morales *et al.* compared DFT with QMC and found that QMC-MD places the liquid-liquid transition 50 GPa higher than DFT. In a more recent study [64] they found that the zero point energy of the nuclei and the van der Waals corrections can shift the PPT higher and lower in pressure by more than 100 GPa. Last year, Mazzola *et al.* showed that there might be in fact two different transitions in the liquid which changes gradually from molecular to mixed atomic-molecular and finally atomic [56]. This result inspired our simple thermodynamical model which I will present in the second half of chapter 4 and which could reconcile the experimental observations from laser heating and shock waves.

The prediction of a liquid ground state has intrigued the theoretical community. Recent studies have extended the calculations of the melting line to higher pressures to investigate the possibility of a ground state liquid. Liu *et al.* used the Z-method and found that after going through a maximum at around 100 GPa, the melt line drops with pressure but eventually flattens and the solid remains stable at room temperature up to pressures higher than 600 GPa [65]. Chen *et al.* obtained a similar result - they found that the stable phase between 500 GPa and 800 GPa is an atomic solid [66]. However, they also found that including the zero point energy in the dynamics, using PIMD was crucial for the calculations. With PIMD, the liquid did become the ground state above 800 GPa, while with classical MD the solid remained stable. These pressures remain too high for experiments at low temperature, so the experimental confirmation will have to wait. However, all these studies agree on a low temperature melting line, which could be related to a class of relatively unstable solids such as the one I will present in chapter 4.

2.2.2 Phases I and II

Hydrogen was first solidified in 1899 [67], and for almost a century only one solid phase of hydrogen was known - phase I. Later on, a transition to a new phase was observed in para-hydrogen, ortho-deuterium [68] and in hydrogen deuteride [69], by means of Raman spectroscopy, at low temperature and high pressure.

These initial Raman results followed by subsequent X-ray diffraction experiments [70] led to the conclusion that phase I is a quantum solid consisting of free quantum rotors arranged in a hexagonal closed packed (hcp) structure. Since the transformation to phase II only happened in spherically symmetric molecules ($J = 0$), it was classified as a broken symmetry orientational transition [24].

A series of crystal structures (i.e. $P2_1/c$, $Pca21$, $P6_3/m$) have been proposed in early theoretical studies to account for phase II. Moraldi *et al.* approximated the quantum rotational energies in $Pca21$, $Cmc21$, $P2_1/c$ and $P6_3/m$, finding $Pca21$ to be the best candidate for phase II of hydrogen [71]. More recently, Li *et al.* performed a series of more sophisticated, PIMD simulations starting in both $P6_3/m$ and $P2_1/c-24$ [72]. With the inclusion of zero point energy, which is implicit in PIMD, both $P6_3/m$ and $P2_1/c-24$ [73] were identified as good candidate structures for phase II. The calculated X-ray scattering, Raman and infrared spectra for $P2_1/c-24$ were found to be in good agreement with the experimental observations [74, 75] as shown in figure 2.8.

Classically, the phase I to II transition can be viewed as an ordering by quadrupole-quadrupole interactions as we showed in a recent MD study [7]. This is the perspective we adopt in our simple thermodynamic picture discussed in the second half of chapter 4.

2.2.3 Phase III

Hemley and Mao were the first who managed to compress pure hydrogen in DACs up to 200 GPa at low temperatures (77 K) [76]. Around 150 GPa, they found a new phase transformation, indicated by a considerable discontinuity in the Raman vibron. Their results were first interpreted as an orientational phase transition, similar to the phase I-II transformation. Lorenzana *et al.* repeated the experiments and carefully analyzed the data [77]. They claimed the transformation was structural and they named the new solid phase A, later

known as phase III.

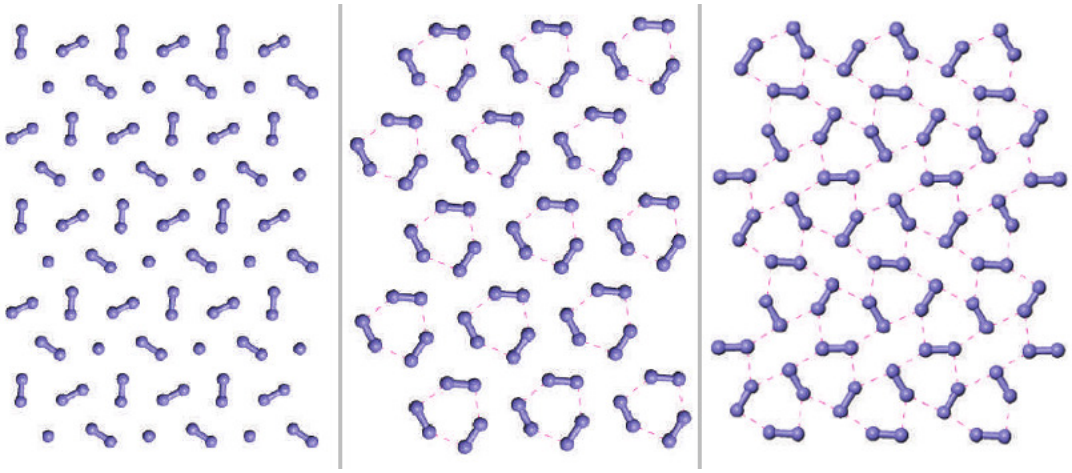

Figure 2.7 *From left to right: $P6_3/m$, $C2/c$ and $Cmca-12$, candidates for: phase II, III. The figure was adapted from ref [78].*

In a systematic theoretical study [78], Pickard and Needs employed AIRSS to find the most energetically favorable zero-temperature crystal configurations up to 500 GPa. Comparing the enthalpies, they found the most stable structures as follows: $P6_3/m$ (at pressures lower than 105 GPa), $C2/c$ (105-270 GPa), $Cmca-12$ (270-385 GPa) and $Cmca-4$ (385-490 GPa). Other predictions, similar in energy with $Cmca-12$ were a new class of mixed layer structures: $Pbcn$, $C2$ and $Ibam$. The atomic crystal structure $I4_1/amd$ was found to be similar in energy to $Cmca-4$. Including the zero point energy (ZPE) through the quasi-harmonic approximation changed slightly the regions of stability such that $C2/c$ became stable below 240 GPa and $Cmca-12$ above. Figure 2.7 shows the crystal structures of the likely candidates.

Their study found $C2/c$ to be a good match for phase III and predicted its metalization above 350 GPa. For $C2/c$ both Raman and infrared data were found to agree with earlier experiments [79, 80], as shown by Tse *et al.* [81].

A series of more recent experiments were reported by Akahama *et al.* who performed Raman measurements up to 296 GPa at low temperatures (90 K and 100K) [82]. Darkening of the samples was observed above 270 GPa in agreement with previous results [79, 80], indicating the gradual closing of the band gap.

Akahama *et al.* also measured the X-ray powder diffraction for hydrogen up to 183 GPa, 100 K [75]. As shown in figure 2.8, they found two diffraction peaks 100 and 101 to be continuous across the II-III phase transition. They interpreted this as evidence that the transition to phase III is non-reconstructive, being another

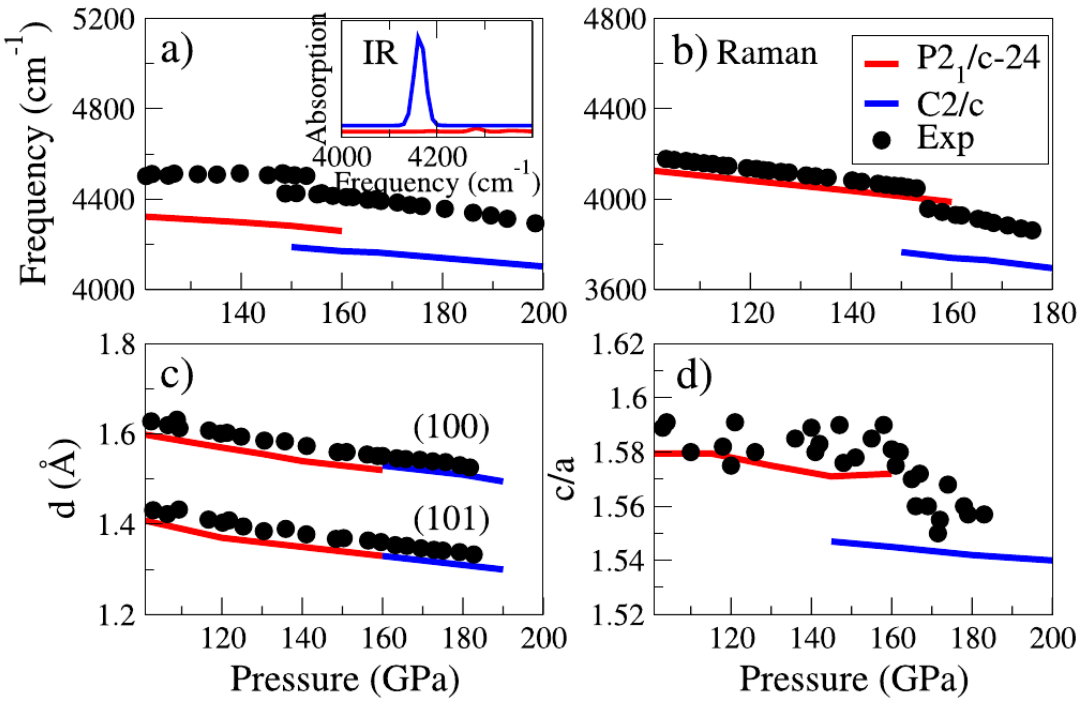

Figure 2.8 *Comparison between theoretical predictions for phases II (solid red line) and phase III (solid blue line) and experimental observations (black dots) of: a) infrared, b) Raman and c), d) crystallographic information. This figure is from ref [72].*

broken symmetry, in agreement with some theoretical studies [83]. Therefore, they conclude that the phase transformation from II to III could be orientational in nature, similar to the I-II transformation, and $C2/c$ cannot be a good match for phase III.

Akahamas's conclusions were carefully analyzed in a more recent theoretical study by Li *et al.* They demonstrate that the X-ray diffraction data is compatible with a structural transformation [72]. They compared the Raman, infrared vibron and X-ray diffraction parameters of phases II and III with their analog theoretical predictions for the structures $P2_1/c-24$ and $C2/c$, respectively; and found reasonable agreement as shown in figure 2.8.

Further optical measurements at low temperatures and much higher pressures in phase III were obtained by Zha *et al.* [84]. They measured infrared and optical absorption as illustrated in figure 2.9, claiming to have achieved 360 GPa. Their findings dismiss the possibility for a phase transition or metalization at those pressures since their samples remain transparent and no discontinuity is found in the infrared spectrum. These results are at odds with Pickard's prediction that a

phase transformation should occur at around 250 GPa and at low temperatures.

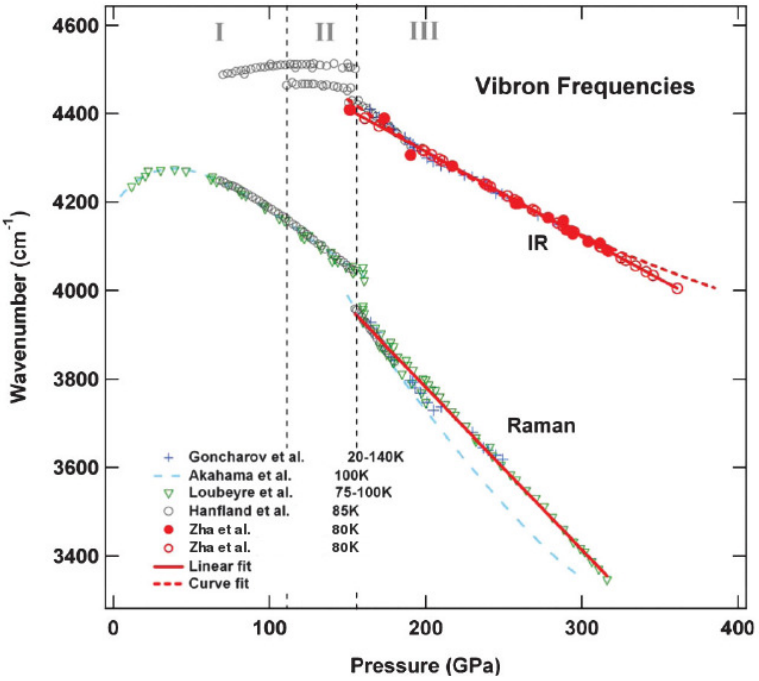

Figure 2.9 *Experimental infrared and Raman spectra collected from various sources. Figure adapted from [84].*

In their recent work, Li *et al.* [72] used an improved DFT functional that accounts for dispersion and found a shift in the stability range of $C2/c$ from the earlier prediction of 120-240 GPa [78] to 150-300 GPa, in better agreement with the experiment. Furthermore, including contributions from ZPE shifts the transition to the metallic $Cmca-4$ at higher pressures, which reconciles with the experimental observations that could not find the metallic state at similar conditions.

The use of different exchange-correlation functionals [85] and more sophisticated methods like quantum Monte Carlo [86] can drastically change the stability of the different candidate structures. However, the ultimate confirmation comes from the comparison to experimental data. Pressure calibration studies have shown that the experimentally measured pressures are also subject to substantial uncertainty. Errors from the DFT exchange-correlation functionals and anharmonic effects of the free energies can result in large variations of the region of stability of different structure candidates. It is no wonder that so many theoretical and experimental studies obtain conflicting results. The approach in my work was to understand the connection between the various experimental phases and theoretical crystal structure predictions through spectroscopic data

rather than energetics. The $C2/c$ structure remains the best candidate for phase III - to which I will bring supportive evidence in chapters 3 and 5.

2.2.4 Phase IV

Eremets *et al.* claimed the discovery of the metallic phase [87]. In their experiments, they compressed small hydrogen and deuterium samples in DACs to pressures as high as 300 GPa at room temperature. They measured Raman, optical absorption and photo-induced conductance. Following a continuous darkening in the interval 220-260 GPa and an abrupt drop in resistance at 270 GPa, they claimed their samples became metallic. Later, their experimental findings were found to be highly controversial. Silvera *et al.* investigated their work [88], and attributed the drop in resistance to leakage currents and chemical reactions of the sample with the epoxy gasket.

Howie *et al.* conducted similar experiments with different outcomes [89]. They manage to obtain clean Raman spectra as shown in figure 2.10, of both the vibron and the lower frequency modes along the room temperature isotherm up to 310 GPa. At around 210 GPa a transition from phase I to phase III is indicated by the abrupt shift in the intensity of the Raman vibron and the appearance of a broad low energy phonon band. At slightly higher pressures, at around 220 GPa, a new, sharper vibron ν_2 appears, the original vibron ν_1 softens and rapidly broadens, indicating the transition to a new phase IV. The new phase also yields three strong lower energy modes and pronounced optical absorption. A further, more subtle modification occurs at 270 GPa, where ν_1 changes slope and a new low frequency mode appears. These latter modifications were noticed by Howie *et al.* and were associated with a slightly different phase IV' [90].

The changes in the Raman signature are also present in the results by Eremets *et al.*, however, they focused more on the drop of resistivity. Howie *et al.* were the first to suggest that the mixed *Pbcn* structure proposed by Pickard *et al.* [78] as a candidate for phase III, could, instead, account for changes observed in the new phase IV. *Pbcn* is a molecular, layered structure, but unlike the other molecular candidates $C2/c$, $Cmca-12$ or $Cmca-4$; *Pbcn* comprises of two different molecular environments, which could, in turn, explain the two vibrons of very different frequencies. The structure consists of bromine-like layers, called B layers and graphene like layers, called G layers. On the one hand, the molecules in the B layer are shorter and form stronger bonds, generating a high frequency

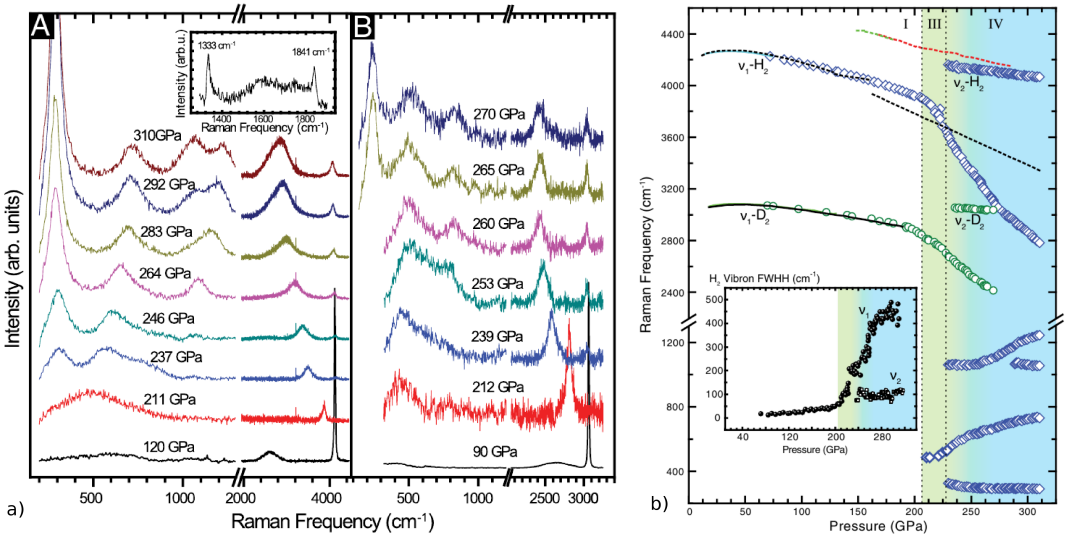

Figure 2.10 *a) Representative Raman of hydrogen (A) and deuterium (B) at room temperature. The diamond response (inset) is eliminated from the Raman spectra (the gap around 1300 - 2000); b) The dependency of the Raman spectra with pressure accounting for phases I, III and IV at room temperature. Inset: the dependence of the FWHM with pressure, which also changes at the phase transitions. The figures were reproduced from ref [89].*

vibron ν_2. On the other hand, the molecules in the G layer are weaker and longer, giving rise to a softer vibron ν_1 [91].

The initial Raman spectroscopy experiments were followed by a series of infrared studies that also found phase IV at room temperature [92–94]. Infrared as depicted in figure 2.11 shows behavior similar to the Raman spectrum.

Within a given band of vibron frequencies, the Raman active vibrons are symmetric and occupy the lower frequency of the band, while infrared ones are anti-symmetric and usually live in the higher frequency range of the same band. Except for a constant offset between Raman and infrared, the same behavior is observed upon entering phase IV: the vibron splits in half, one of which softens rapidly with pressure, the other remains pressure independent. This result supports the model of a mixed molecular structure like *Pbcn*.

Although *Pbcn* appears to be a good candidate for phase IV, in phonon calculations it gives negative frequencies. The existence of imaginary phonons usually means that the structure is unstable, however, in the case of hydrogen, the large zero point motion could stabilize the structure.

Following the discovery of phase IV, Pickard *et al.* extended their initial AIRSS

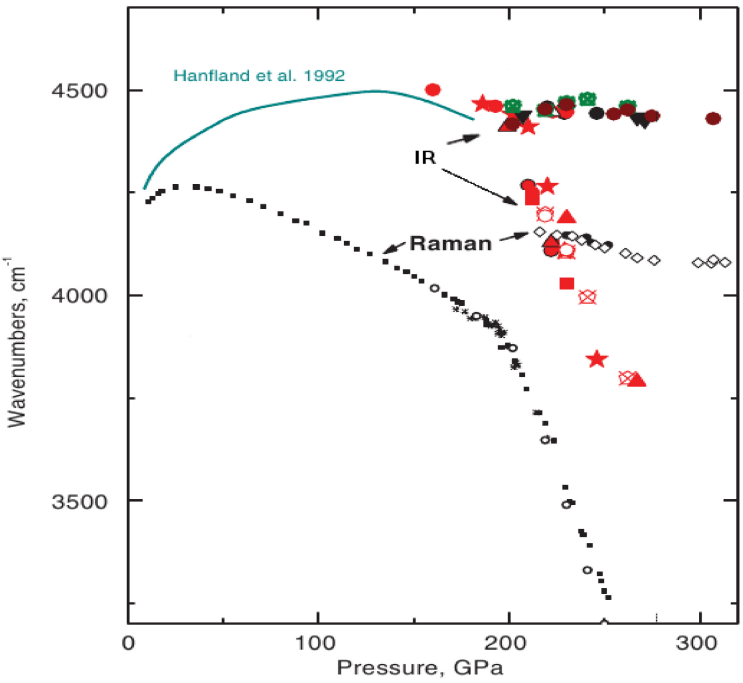

Figure 2.11 *Infrared measurements in phase IV of solid hydrogen at room temperature in comparison with Raman measurements at the same conditions. The figure is from ref [92].*

study based on PBE-DFT and found a new structure $Pc-48$ as shown in figure 2.12. The new $Pc-48$ is similar to $Pbcn$ but with more distorted G layers [91]. In phonon calculations $Pc-48$ gives negative frequencies as well, however, a larger unit cell, $Pc-96$ eliminates those. The study also estimated the energies including zero point motion and vibrational energies in the harmonic approximation for $C2/c$, $Pbcn$, $Pc-48$, $Pc-96$, $Cmca-12$ and $Cmca-4$. Additionally, the phonon densities of states were calculated alongside Raman and infrared activity at the gamma point. The extended version of Pc was found to be the best candidate for phase IV, accounting for both energetics and Raman spectra. An erratum published later by the same authors found the metallic $Cmca-4$ more stable than Pc on the entire pressure range [95] which further extended the metallic/non-metallic controversy started by the conflicting experimental results of Eremets and Howie. As shown by more recent studies [86], the increased stability of the molecular $Cmca-4$, could be an anomaly due to the PBE functional.

Ignoring the energetics, Pc partially explains the experimental results obtained in phase IV. On closer inspection, the two molecular environments in Pc behave very differently with increasing pressure. While the bond lengths in the B layer change very little, the G layer molecules strongly elongate and the intramolecular

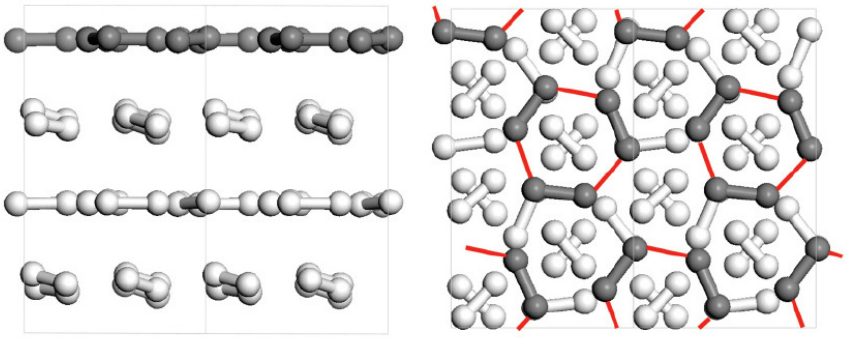

Figure 2.12 *An illustration of the extended Pc structure (96 atoms) at 250 GPa. Left: stacking; Right: top view and layer configuration. Figure taken from ref [91].*

distances get closer to the intermolecular ones as the pressure increases. The structure seems to continuously transform towards the more symmetric *Ibam*, where the G layers are entirely atomic.

The Raman spectra computed in the harmonic approximation for *Pc*, however, do not entirely agree with the experimental result for phase IV. Although ν_2 is well explained by the *Pc* crystal, the calculated ν_1 is softer in DFPT calculations than in experiment. In our work [1, 3] we showed that phase IV, like phase I, is stabilized by entropy and a better description comes from molecular dynamics (MD) simulations. In chapter 3, I will explore a method to extract the Raman vibrations from MD [4], which narrows the discrepancy between theory and experiment.

Other studies have also investigated phase IV using dynamical methods. Liu *et al.* employed metadynamics calculations to map the phases of hydrogen within the pressure range 150-300 GPa at room temperature [96]. They found a new candidate for phase IV (*Cc*), which is energetically degenerate with *Pc*. According to their calculations, *Cc* does not contain imaginary phonons, which makes it more stable than *Pc*. The molecules in the B layers are found to be rotating freely in *Cc* which brings evidence that phase IV is inherently dynamical and cannot be entirely understood with any of the 0 K, ground state symmetries.

In a different article, Liu *et al.* extended their calculations attempting to give a better dynamical description of their *Cc* structure [97]. In their *ab initio* MD simulations, they report an interesting proton diffusion effect which they associate with the rapid softening and broadening of the ν_1 vibron. This diffusion phenomenon comprises of two distinct effects: spontaneous rotation of

the molecular trimmers in the G layers, combined with the change of trimmer identity through complete layer reconstruction.

Seeking to understand the intricate nature of phase IV, Goncharov *et al.* also used MD to simulate the *Pbcn* crystal. They found layer structural fluctuations where B and G switch identity and molecular "hopping" [98]. We later show their small cell simulations were affected by finite size effects [1]. They also computed the phonon density of states extracted from auto-correlation function and showed that the soft vibron could be highly anharmonic.

The results I will present in chapter 3 summarize our own work intended to understand these issues. In agreement with Liu's work [96, 97] we found that the nature of phase IV is dynamical with free rotation of molecules in the B layers and rotations of molecular trimmers in the G layers [1, 3]. In our simulations we found little diffusion in the region of stability for phase IV and we believe this could not affect the vibron frequency since it takes place at a much longer time scale.

In parallel with dynamical studies, other groups continued to investigate ground state candidates for phase IV in static calculations. They argue that the apparent shortcomings of the *Pc* structure regarding energetic stability [95] stem from the problems with PBE-DFT [86, 99], which is a poor approximation in hydrogen. In quantum Monte Carlo (QCM) calculations, Drummond *et al.* found a more stable *Pc* which exists in the region of stability of phase IV, in good agreement with the experiment.

We believe that phase IV has a dynamical nature with free rotation in the B and G layers, therefore MD is a better approach to investigate this phase [1, 3, 4]. The usual *ab initio* MD does not include quantum effects on the nuclei, which were shown to be important, so we do not expect the pressure and temperature transitions to match experimental observations perfectly. Nevertheless, as shown in the results chapter 3, MD simulations can help us understand this phase better and explain the experimental findings.

2.2.5 Phase V and Beyond

This year, three new experimental studies have claimed new phases of solid hydrogen [100–102]. In particular, Dalladay-Simpson *et al.* showed a further

modification at room temperature that occurs in hydrogen and hydrogen-deuteride, but not in deuterium [100]. At 325 GPa, the soft vibron ν_1 in solid hydrogen becomes almost pressure independent and two out of the four low frequency modes gradually disappear. This phase is new and little is know about it. In the second half of chapter 3, section 3.3, I will show our preliminary computational results on phase V. There, I will also discuss the experimental observations in more detail. I believe phase V is a subtle dynamical modification from phase IV, where the main change takes place in one of the layers.

It is clear that the new phase reported by Dalladay-Simpson *et al.* is not an atomic phase, because the Raman vibrons persist at those conditions. It is also clear that the new phase comprises of at least two molecular environments just like phase IV, evident by the two vibrons of different origins. All these issues will be discussed in more detail in the second half of chapter 3.

Theory predicts that at pressures higher than 800 GPa, the liquid becomes stable close to zero kelvin [66]. The molecular phases persist up to almost 400 GPa as shown by the recent experimental results [100].

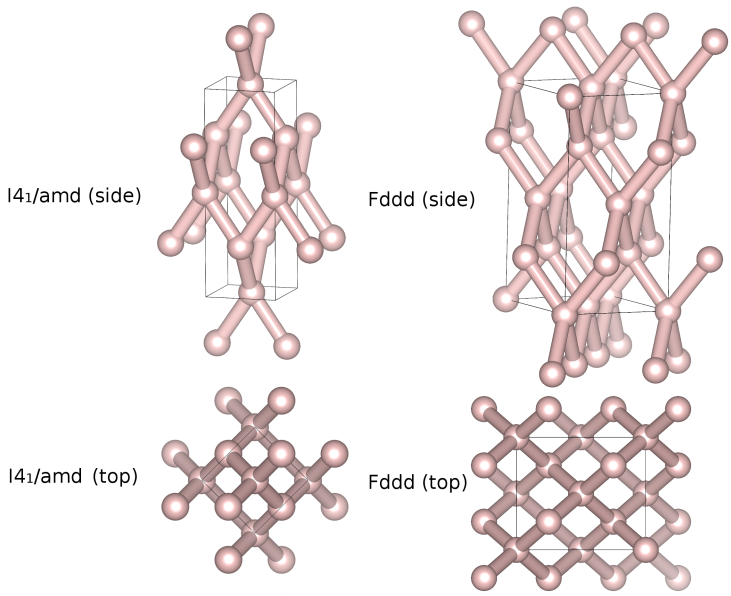

Figure 2.13 *The figure shows the best candidate structures for the atomic solid phase [103, 104]. On the left is $I4_1/amd$ and on the right $Fddd$, a slightly distorted version of $I4_1/amd$. The distortion is visible in the top-down view.*

The region between these pressures is of great interest to both theory and experiment. A series of DFT studies [103–105] have shown that at these pressures, the most likely candidate is an atomic phase with a distorted diamond structure.

This crystal candidate as shown in figure 2.13, has $I4_1/amd$ symmetry and is also the structure of cesium IV ($Cs-IV$). A similar structure, almost degenerate in energy [106], is $Fddd$ which is a small distortion from $I4_1/amd$ as illustrated in figure 2.13. These initial results have been validated by quantum Monte Carlo studies [107, 108] which also found these structures to be stable at high pressures.

In chapter 4, I investigated $I4_1/amd$ using molecular dynamics simulations at pressures around 400 GPa which might soon become accessible in experiments. I also compared this structure against $Cmca-4$ which has been found to be the most stable from the molecular structures and remains a competitive candidate for solid hydrogen above 400 GPa.

2.2.6 Hydrogen-Deuterium Mixtures

Deuterium is an isotope of hydrogen, with twice the nuclear weight. Since deuterium has the same electronic structure as hydrogen, there is no difference between the two isotopes in the Born-Oppenheimer approximation, where the nuclear and electronic degrees of freedom are decoupled. Owing to its larger mass, the deuterium atom has, however, lower zero point energy than hydrogen. Studying the phase diagram of solid deuterium in comparison to that of hydrogen can help us assess the contribution of zero point energy in stabilizing the different phases [109]. The heavier mass of the deuterium isotope can also be exploited as an experimental tool to investigate the strength of intramolecular and intermolecular interactions. For instance, at the same pressures and temperatures, the ratio between the HH vibron and the DD vibron in the harmonic approximation should be $\sqrt{2}$. The deviation from this result can inform on the weakening of the bond with increasing pressure [89]. Similarly, the ratio between the free rotors of the two isotopes is expected to be 2, absent any intermolecular interactions. Upon increasing pressure, the strength of the intermolecular forces can be estimated from changes in this ratio.

Even more interesting than pure hydrogen and deuterium solids are the hydrogen-deuterium mixtures. As I have extensively discussed above, many studies have investigated pure solid hydrogen and deuterium across the accessible pressure and temperature scales. However, hydrogen-deuterium mixtures have received less attention, partially because they are much more complicated to deal with. In the following, I will review the most notable studies in mixtures.

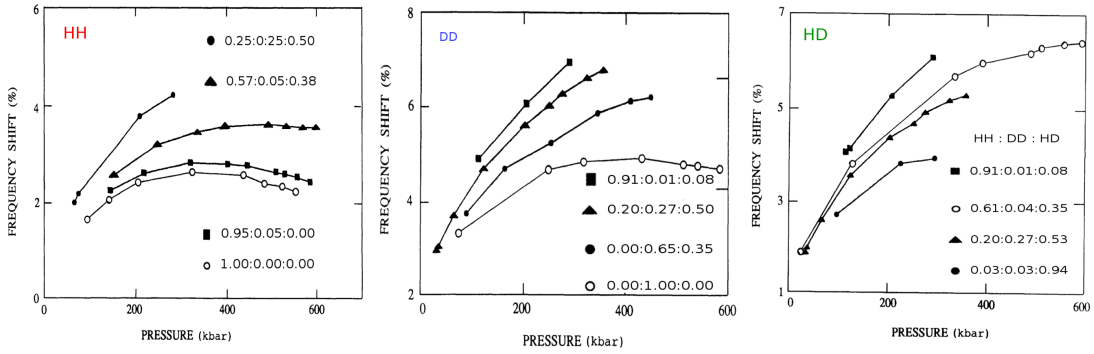

Figure 2.14 *The shift of the HH (left), DD (middle) and HD (right) Raman vibrons as a function of mixture composition and pressure, at room temperature. From ref [110].*

Brown and Daniels performed a pioneering experiment where they compressed hydrogen and deuterium mixtures up to 60 GPa at room temperatures [110]. As shown in figure 2.14, they observed that, in mixtures, the vibron frequencies corresponding to each molecular species are higher than those obtained in the corresponding pure solids. The lower the concentration of a particular species and the higher the pressure, the larger its frequency shift from the pure solid. They explained this phenomenon with a model of coupled oscillators [111]. In a pure lattice, the molecules can couple and oscillate all in phase. This vibration pattern has a lower Raman frequency than that of noninteracting free molecules in the gas phase. In a mixture, the mass disorder breaks the coupling, increasing the frequency of the minority species. Upon increasing pressure, the difference between coupled and uncoupled molecules is enhanced as the strength of the intermolecular interaction increases. This phenomenon has been observed in other mixtures as well, where a small concentration of hydrogen molecules is embedded in a matrix of heavier inert atoms like helium, neon and argon [112].

Our collaborators have recently managed to extend the experimental study of hydrogen-deuterium mixtures into phase IV [2]. In the first half of chapter 5, I will present my computational work on phase IV mixtures that explains the experimental observations through a phenomenon called phonon localization. Basically, this phenomenon results from breaking the coupling between molecules through mass induced disorder as previously noticed by Brown and Daniels.

Other experimental studies have employed infrared spectroscopy to investigate pure hydrogen-deuteride (HD). Figure 2.15 shows the result of Chijioke and Silvera in pure HD, where they mapped the phase boundaries of phase II with phase I and III [113]. Upon increasing pressure, at around 4-8 K the sample

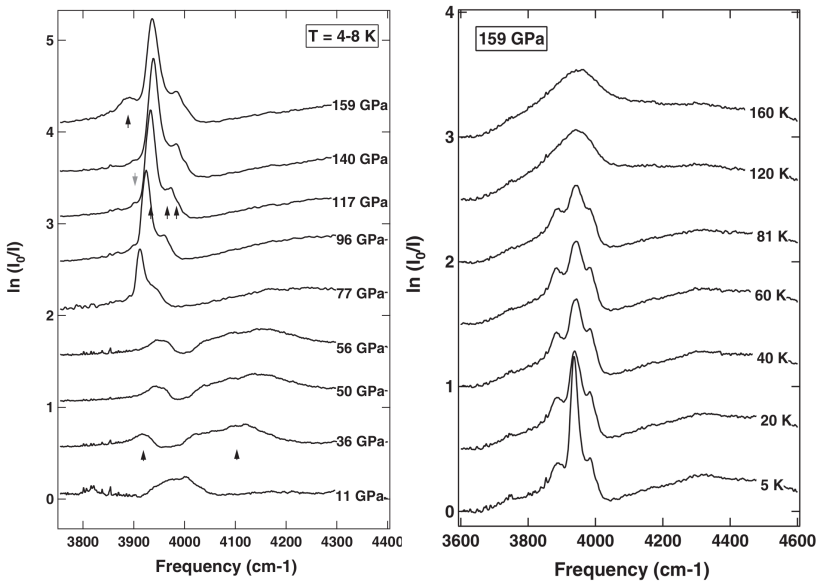

Figure 2.15 *This figure reproduced from ref [113] shows the infrared spectra in pure hydrogen-deuteride (HD) upon pressurizing, at around 4-8 K (left), and upon heating at 159 GPa (right).*

transformed on the path I-II-III. In phase III, the HD peak has two shoulders and there is no evidence of HH or DD peaks.

In a recent study this year, Dias *et al.* extended the original work at low temperatures in HD up to pressures of 350 GPa [114]. Based on infrared splitting they claim two new phases. I will dedicate the second half of chapter 5 to analyzing their result using *ab initio* calculations, starting from the best known structural candidates for phase III, $C2/c$ and $Cmca-12$.

Chapter 3

Solid Hydrogen Phases III, IV and V

3.1 Overview

In the first half of this chapter, I will present my initial work on phases III and IV of solid hydrogen. The experimental discovery of phase IV [87, 89] revived the field of hydrogen. The main indicator of phase III to phase IV transition in experiment is the splitting of the Raman active vibron into two individual vibrons: a higher frequency one which is pressure independent and a lower frequency one that varies strongly with pressure. This is surprising for a molecular crystal, especially for an element as simple as hydrogen, and it proves the existence of two distinct molecular environments.

The study I report in the second half of the chapter was aimed at understanding the origin of the newly discovered phase V [100]. My hypothesis is that this new phase is a subtle modification from phase IV. The new structure maintains the mixed molecular environments as evident by the continuity of the two different vibrons from phase IV.

3.2 Phases III-IV

3.2.1 Introduction

A substantial body of work, both theoretical [1, 3, 72, 96–98] and experimental [75, 84, 90, 92–94] has been dedicated to understanding the crystal structures and the physical properties of phases III and IV. The breakthrough came from the AIRSS studies [78, 91] which proposed two crystal candidates based on energy stability: $C2/c$ for phase III and Pc for phase IV. Raman computed with density functional perturbation theory (DFPT) yielded the correct number of active vibrons in both cases and at roughly the correct frequencies, with one notable exception: the lower frequency vibron in Pc is much softer than the experimentally measured vibron in phase IV.

The important observation here is that phase IV does not extend to low temperatures, indeed experiments that started in phase IV found a transition back to phase III upon decreasing the temperature [89]. This indicates that phase IV is stabilized by entropy and it is a high temperature phase. Therefore, the true nature of this phase cannot be understood in zero kelvin (0 K) calculations but can be studied using molecular dynamics (MD) simulations.

In this section I will present a range of MD simulations starting from the best known candidates for phases III and IV at 0 K. The calculations were extended in the pressure interval 180 to 350 GPa and at temperatures ranging from 150 to 500 K. Apart from investigating the dynamic behavior of these structures, I also devised a simple technique to extract the Raman active vibrons directly from the MD trajectories. The results presented in this section were also published in two peer review papers [1, 3].

3.2.2 Methods

Simulation Details

All calculations presented here are molecular dynamics (MD) simulations performed with CASTEP [31, 115]. For the density functional theory (DFT) step of the calculations, the electron density was expanded in plane waves with an

energy cut-off of 1200 eV. The density was treated as spin neutral (unpolarized).

Since the focus here was on phases III and IV, mostly at pressures below 300 GPa, all structures were modeled as insulators by fixing the occupancy of the electronic bands. In the later parts of the chapter, section 3.3, the calculations were mainly focused on the recently discovered phase V and those simulations have variable occupancy to allow for possible metallic modifications.

Here I employed the Perdew Burke Ernzerhof PBE functional [28] to evaluate the exchange-correlation energy. For electron minimization I used the Broyden density mixing (dm) scheme and the total energy was converged within 10^{-11} eV at every iteration step. With these settings, I performed a series of MD simulations as summarized by table 3.1.

To avoid the cusp at the center of the hydrogen nucleus, CASTEP employs pseudo-potentials. In simulations 1-25, table 3.1, I used the ultra-soft pseudo-potential generated on the fly by CASTEP from the input string "H 1|0.8|0.8|0.6|2|6|8|10(qc = 6)" [116], while for simulations 26-29, I used the norm-conserving pseudo-potential generated by "H 0|0.7|2|6|8|10L(qc = 10)" [116].

In all the MD runs, the time step was fixed to 0.5 fs which is the setting I used throughout the thesis. The highest frequency likely to be present in the system is the vibration of the hydrogen molecule at around 4200 cm^{-1}, which is equivalent to approximately 8 fs. With a time step of 0.5 fs, the simulation samples the highest frequency period at roughly 16 points, which should be sufficient to conserve the energy of the system.

In order to understand the true nature of phases III and IV, but also to gauge the contributions of spurious effects, I employed a variety of MD settings: pressures, temperatures, ensembles, different cell sizes, k-point MP grids; all reported in table 3.1. The structural candidate I used for phase III was $C2/c$ [78], while the candidates for phase IV were Pc [91] and Cc [96]. In the later parts of this chapter, I will return to phase IV and consider more structural candidates. Note that here all cells with 288 atoms comprise of four wider layers, while later in section 3.3 the 288 atoms cells have eight narrower layers.

All the MD simulation here were performed in one of two thermodynamic ensembles: constant volume and energy (NVE) or constant pressure and temperature (NPT), both of which conserve the number of particles. In the

Table 3.1 *Summary of all MD calculations discussed in this section. Simulations 1-25 employ ultra-soft pseudo-potentials, while simulations 26-29 use norm-conserving pseudo-potentials. Shaded in blue are the simulations I used to extract Raman information from the dynamics.*

Id	Start	Atoms	Length/Ensemble	MP grid	Pres.	Temp.
1	Pc	48	4.5ps NVE	5x3x3	250 GPa	60 K
2	Pc	48	3.0ps NVE	5x3x3	250 GPa	145 K
3	Pc	48	3.0ps NVE	5x3x3	250 GPa	215 K
4	Pc	48	3.0ps NVE	5x3x3	250 GPa	285 K
5	Pc	48	3.0ps NVE	5x3x3	250 GPa	360 K
6	Pc	48	3.0ps NVE	5x3x3	250 GPa	430 K
7	Pc	48	3.0ps NVE	5x3x3	250 GPa	500 K
8	Pc	288	1.5ps NVE	1x1x2	250 GPa	145 K
9	Pc	288	1.5ps NVE	1x1x2	250 GPa	215 K
10	Pc	288	1.5ps NVE	1x1x2	250 GPa	285 K
11	Pc	288	1.5ps NVE	1x1x2	250 GPa	360 K
12	Pc	288	1.5ps NVE	1x1x2	250 GPa	430 K
13	Pc	288	1.5ps NVE	1x1x2	250 GPa	500 K
14	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	180 GPa	220 K
15	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	200 GPa	220 K
16	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	220 GPa	220 K
17	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	250 GPa	220 K
18	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	270 GPa	220 K
19	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	300 GPa	220 K
20	Pc	288	0.25ps NPT+0.75ps NVE	1x1x2	350 GPa	220 K
21	C2/c	288	0.25ps NPT+1.5ps NVE	1x1x2	200 GPa	220 K
22	C2/c	288	0.25ps NPT+1.5ps NVE	1x1x2	220 GPa	220 K
23	C2/c	288	0.25ps NPT+1.5ps NVE	1x1x2	250 GPa	220 K
24	C2/c	288	0.6ps NPT	1x1x2	250 GPa	200→400 K
25	C2/c	288	0.4ps NPT+1.0ps NVE	1x1x2	250 GPa	300 K
26	Pc	768	0.25ps NPT+1.5ps NVE	1x1x1	270 GPa	300 K
27	Pc	768	0.25ps NPT	1x1x1	340 GPa	300 K
28	Pc	288	1.5ps NPT+1.5ps NVE	2x2x4	270 GPa	300 K
29	Cc	384	0.3ps NPT+2.0ps NVE	1x1x2	270 GPa	300 K

NPT ensemble, the temperature was kept constant using the Nose-Hoover chain thermostat with 5 particles and the default characteristic ionic time of 5 fs. The pressure was kept fixed using either the Andersen-Hoover (simulations 14-20) or the Parrinello-Rahman (simulations 21-29), in both cases with characteristic cell time of 50 fs.

NVE keeps the simulation box fixed and usually does not allow for structural phase transitions. NPT-Andersen assigns three degrees of freedom corresponding to the lengths of the box, while NPT-Parrinello assigns six degrees of freedom corresponding to the three lengths and three angles of the box, which are then

allowed to change in time. Therefore, NPT was generally employed here to equilibrate the simulations to the desired conditions, while NVE was employed to study the properties of an equilibrated structure. Some of the trajectories were used to analyze the behavior, like the average crystal structure or the dynamics, while others were used to extract information about Raman activity. These latter simulations were colored light blue in table 3.1.

Structure and Dynamics

The room temperature behavior of these crystal structures was analyzed through time average trajectories and mean square displacements. To obtain the average crystal structure in the case of NPT, I first calculated the average lattice parameters:

$$\widetilde{M}_{\alpha,\beta} = \frac{1}{t_{max}} \sum_{t=1}^{t_{max}} M_{\alpha,\beta}(t) \tag{3.1}$$

where α and β label Cartesian coordinates, $M_{\alpha,\beta}$ is the transpose lattice matrix, the tilde denotes time averaging, while t_{max} denotes the length of the simulations.

Since CASTEP allows atoms to cross the periodic boundary during the MD simulation, in both NVE and NPT, the average positions can simply be obtained from:

$$\tilde{x}_l^\alpha = \frac{1}{t_{max}} \sum_{t=1}^{t_{max}} x_l^\alpha(t) \tag{3.2}$$

where $l = 1, N$ labels the atoms and x_l^α denotes the absolute atomic positions.

Finally, the average crystal structures presented in the results section were obtained by reducing the average absolute positions to the main average cell:

$$\hat{x}_l^\alpha = \sum_{\beta=1}^{3} \widetilde{M}_{\alpha,\beta} \left[\left(\sum_{\gamma=1}^{3} \widetilde{M}_{\beta,\gamma}^{-1} \tilde{x}_l^\gamma \right) \mod 1 \right] \tag{3.3}$$

here, the hat represents time average and reduction to unit cell, as the equation shows.

I investigated diffusion and molecular rotations through mean displacement:

$$MSD(t) = \frac{1}{N}\sum_{l=1}^{N}\sqrt{\sum_{\alpha=1}^{3}\left[x_l^\alpha(t) - x_l^\alpha(0)\right]^2} \tag{3.4}$$

Raman from MD

The idea here was to extract Raman from MD and account for the anharmonicity of the modes in hydrogen, phases III and IV. The method presented below was successfully applied before to classical MD simulations [117]. Here the method was extended to *ab initio* calculations.

Given the trajectory of an equilibrium quasi-harmonic structure with associated eigenmodes $\varepsilon_{l,k}^\alpha$ (where k labels the $3N$ frequencies), one can expand the atomic positions of this structure at every time step in terms of its normal modes [4]:

$$x_l^\alpha(t) = \tilde{x}_l^\alpha + \sum_{k=1}^{3N} c_k(t)\varepsilon_{l,k}^\alpha \tag{3.5}$$

where c_k is the projection of the trajectory onto each normal mode k.

To eliminate the complication introduced by the average positions, one can expand the atomic velocities v_l^α instead. Since the center of mass is kept fixed in MD, the velocities average to zero and the equation becomes:

$$v_l^\alpha(t) = \sum_{k=1}^{3N} \dot{c}_k(t)\varepsilon_{l,k}^\alpha \tag{3.6}$$

Since $\varepsilon_{l,k}^\alpha$ span an orthonormal space, the contributions $\dot{c}_k$ to each mode k can be obtained reversely by projecting the velocity onto each normal mode:

$$\dot{c}_k(t) = \sum_{l=1}^{N}\sum_{\alpha=1}^{3} v_l^\alpha(t)\varepsilon_{l,k}^\alpha \tag{3.7}$$

All $\dot{c}_k(t)$ are oscillating in time. Furthermore, the frequency of this oscillation is the temperature corrected frequency f_k of each mode k.

Now, if r_k^{0K} is the Raman intensity of mode k at zero kelvin (0K), then the finite temperature Raman distribution $r_k^T(f)$ can be extracted from the discrete Fourier Transform of $\dot{c}_k(t)$:

$$r_k^T(f) = \frac{r_k^{0K}}{t_{max}} \left| \sum_{t=1}^{t_{max}} \dot{c}_k(t) \exp(-2\pi i f t) \right| \tag{3.8}$$

At this point, $r_k^T(f)$ is a gaussian distribution centered around a frequency f_k which is the temperature corrected frequency of mode k. The final spectrum can then be obtained as a sum over all k:

$$r^T(f) = \sum_{k=1}^{3N} r_k^T(f) \tag{3.9}$$

Unfortunately, in the case of hydrogen, this method fails because, as shown later in the results section, the hydrogen structures are not harmonic and they include free rotations of molecules and trimers. Figure 3.1 shows how such a rotation can render the method unusable.

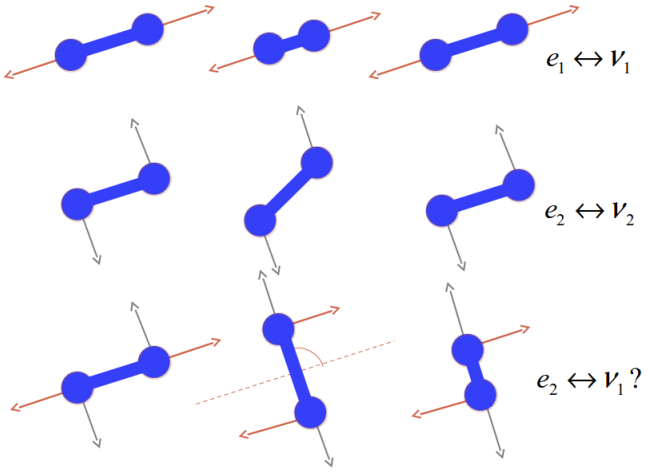

Figure 3.1 *The figure shows why the projection method onto fixed eigenmodes fails when there are free rotations in the system. Top, mode e_1 is a molecular vibration with frequency ν_1. Middle, mode e_2 is a molecular libration with frequency ν_2. Bottom, the molecule rotates freely by 90°, the vibrational motion ν_1 moves along the librational eigenmode e_2 and the method fails.*

In conclusion, the low frequency modes are nearly impossible to project. As shown in the later parts of the chapter, section 3.3.3, they comprise of complicated

oscillation motions. However, it turns out that the high frequency Raman vibrons are simple, in phase oscillations of molecules in the B and G layers, respectively. Therefore, here, I calculated the high frequencies Raman spectra by simply projecting the velocities onto molecular axes a_l^α, which are bound to and rotating together with the molecules:

$$r^T(f) = \frac{1}{t_{max}} \left| \sum_{t=1}^{t_{max}} \sum_{l=1}^{N} \sum_{\alpha=1}^{3} v_l^\alpha(t) a_l(t)^\alpha \exp(-2\pi i f t) \right| \tag{3.10}$$

To calculate the molecular axes, one needs to identify the molecules first. For each atom, I find the closest atom pair that is not already part of a molecule. This becomes problematic when molecules are no longer well defined, in which case the order of the atomic labels matters. For instance, if both atoms 1 and 2 are close to an atom 3, but it is not obvious which of the pairs 1-3 and 2-3 is the molecule, than the algorithm will automatically assign the molecule to the pair 1-3 and atom 2 will be left out. These cases are actually difficult to deal with and there is no best solution to this problem.

Generally, I use two approaches for finding molecules. In the first approach, the molecules are only identified once at the start of the simulations and then the labels are kept throughout. However, in some cases, molecules are short lived and dissociation and rebonding become frequent. Therefore, in the second approach, the molecules are recalculated at every time step. The results from the two methods are different only when dissociation is important.

3.2.3 Results

In figure 3.2 I show evidence that phase IV only exists at high temperatures and it is stabilized by entropy. The $C2/c$ structure is the best candidate for phase III and it consists of a series of distorted G layers (GGGG stacking). MD simulations started in $C2/c$ and performed at 220 K are stable for almost 2 ps. However, upon heating to 300 K, $C2/c$ transforms into a stacking of BGBG layers which is characteristic to phase IV.

Interestingly, the transformation is accompanied by a complete reconstruction of the entire cell. Whereas the original layers in $C2/c$ are perpendicular to the z-axis, after the transition, the layers in BGBG reform in a direction perpendicular to the

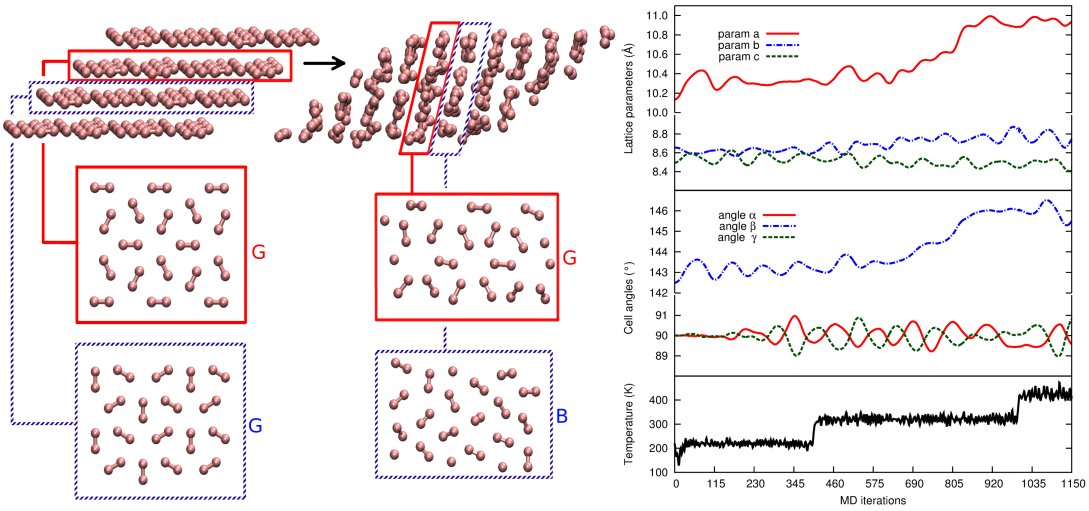

Figure 3.2 *Results of the MD simulation performed in NPT (number 24 from table 3.1) which was started in a 288 atom supercell of $C2/c$ and which transformed to a series of BGBG layers upon heating at constant 250 GPa. Left, the relaxed structure at the start of the simulation and a snapshot of the BGBG structure after the transition occurred. Alongside the 3D cell I also show the pattern of the different layers as insets. The bromine-like layers are labeled B, while the graphene-like layers are labeled G. Right, the temperature as function of time, as well as the evolution of the lattice parameters across the transformation.*

x-axis. This new structure is similar to the ground state structure $C2$ discovered by AIRSS [78], with the distinction that at room temperature in BGBG, the molecules in the B layer appear to be rotating almost freely. The rotating B molecules generate large entropy and therefore decrease the free energy of this configuration, making BGBG more stable than the GGGG stacking at the same pressure and higher temperature. In the next chapter, I will discuss in more detail the thermodynamics behind these arrangements.

The structure of phase IV has been thoroughly studied in ground state calculations [91], metadynamics [96] and molecular dynamics [1, 3, 97, 98, 118], being a unique molecular crystal with two different kinds of molecular environments. Some of the earliest dynamics calculations employed simulations boxes as small as the unit cell (48 atoms). In table 3.1 I summarized my own simulations (numbers 1-7) with such small cells. Although fewer atoms allow for longer trajectories which are important in MD, these simulations are bugged with a range of severe finite size effects. For instance, layers appear to often switch identity, with B layers becoming G and vice versa. To eliminate these spurious effects, I extended the cells used in MD to 288 atoms or more and these were found to be stable.

As I will discuss later in chapter 4 the k-point sampling is also important, in particular for metals. Here, however, the structures were treated as insulators which are less affected by the number of k points.

To understand the true nature of phase IV, I performed a series of MD simulations with 288 atoms and four layers, spanning the pressure range 180 to 350 GPa. Interestingly, the results show a distinction between the behavior at pressures below 260 GPa and above, see figure 3.3.

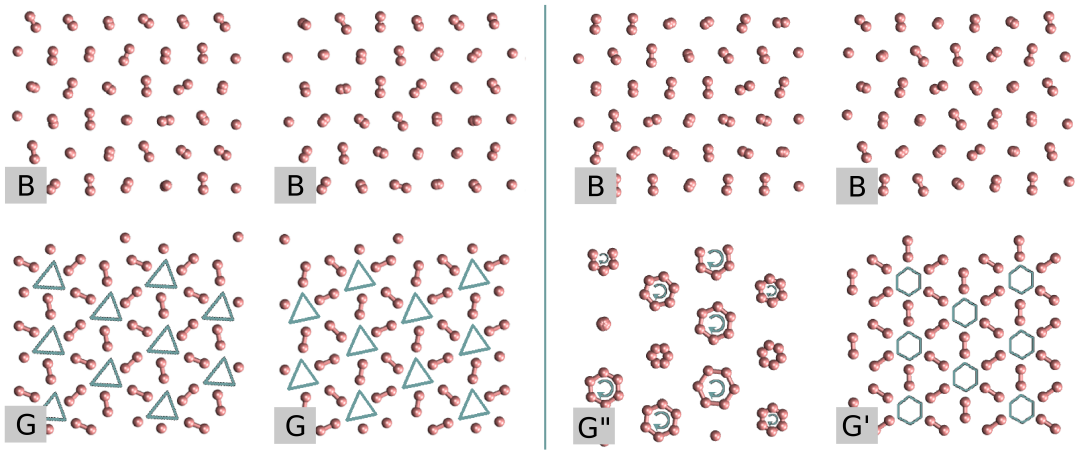

Figure 3.3 *Atomic positions averaged over 1 ps of MD simulation with 288 atoms, originally started with Pc. Left, average structure at 220 GPa and 220 K (simulation 16 from table 3.1). Right, average structure at 270 GPa and 220 K (simulation 18 from table 3.1). In both cases the structures comprised of four layers which were unfolded here for illustration purposes. Upon increasing pressure the G layers differentiate into layers that allow trimer rotation G" and layers that comprise of symmetric molecular trimers G'.*

At the lower pressures, the simulations started in *Pc* remain close to the original symmetry in the G layer, while in the B layer, the molecules are almost freely rotating. The rotation is evident in the average trajectory, where the mean position of the two atoms of the B molecule coincides with the center of mass of the molecule. As a result, the B layer acquires hexagonal symmetry on average. We call this stacking BGBG. This region is, however, outside the experimental stability regime for phase IV.

At the higher pressures, above 260 GPa, where the experiment finds phase IV, the B layers remain unchanged, but the G layers differentiate into two distinct types: G' and G" [1]. In the G" layers the molecular trimers rotate around their axis of symmetry, independently of one another. The rotation is indicated by the time average of the G" layers where the mean position of six atoms belonging to

each trimer coincides with the center of mass of that trimer after long enough time. On the other hand, the G' layer becomes more symmetric and appears to be more stable against rotation. This stacking was called BG'BG" and it is possibly the true nature of phase IV. The rotation of the trimers in the G" layer generates additional entropy which further stabilizes the BG'BG" structure at finite temperature.

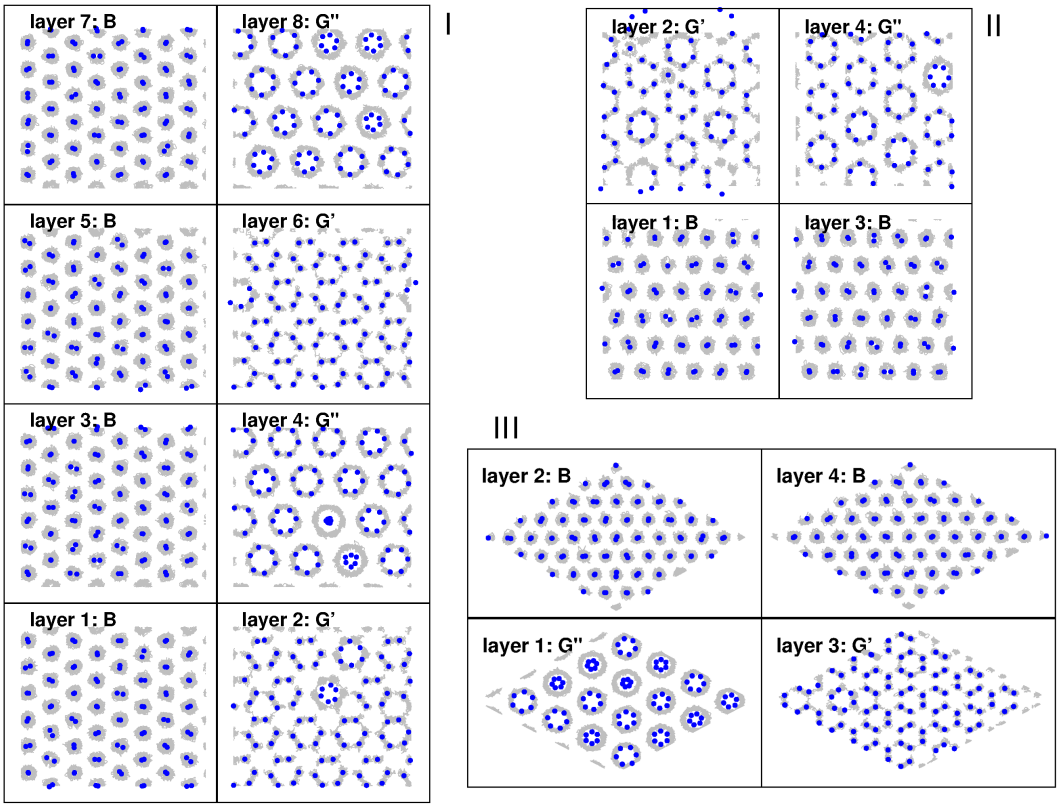

Figure 3.4 *Full trajectories in grey and the average atomic positions with blue dots, extracted from MD simulations at 270 GPa and 300 K. Panel I, the result of the 768 atoms simulation started in Pc with 1x1x1 k−point MP grid (number 26 from table 3.1). Panel II, the 288 atoms Pc simulation with 2x2x4 k−point grid (number 28). Panel III, the 384 atoms simulation started in Cc [96] with 1x1x2 k−point MP grid (number 29). When averaged over a long enough period to allow full rotations, the two blue dots in the B layer molecule / six blue dots in the G″ trimer will coincide with the center of molecule / trimer.*

In order to check if this dynamic behavior is robust against finite size effects I repeated the simulations with larger unit cells, consisting of 768 atoms and double the number of layers (eight). I also performed a simulation starting from a hexagonal setting of the BG pattern, having Cc symmetry at 0 K [96] and a simulation starting with Pc but with denser k-point sampling.

The results of these calculations are shown in figure 3.4 and are consistent with the previous observations. At room temperature, the eight layer cell with 768 atoms displays a BG'BG"BG'BG" stacking which is just a repeat of BG'BG" in the direction perpendicular to the layers. Similarly, the *Cc* structure displays the same behavior with two types of G layers. The simulation performed with a denser k-point grid (panel III in figure 3.4) is slightly different because the rotation of the trimers took longer to initiate and therefore the distinction between G' and G" layers is less clear.

The studies presented in refs [97, 118] found substantial diffusion in the G layers owing to a two step process: 1) rotation of molecular trimers and 2) complete layer reconstruction such that an atom initially belonging to one trimer can switch over to the neighbor one. While the results I show here, indicate clear evidence of trimer rotation, at these pressures and temperatures I found no sign of layer reconstructions. Subsequently, the diffusion cannot take place under these circumstances.

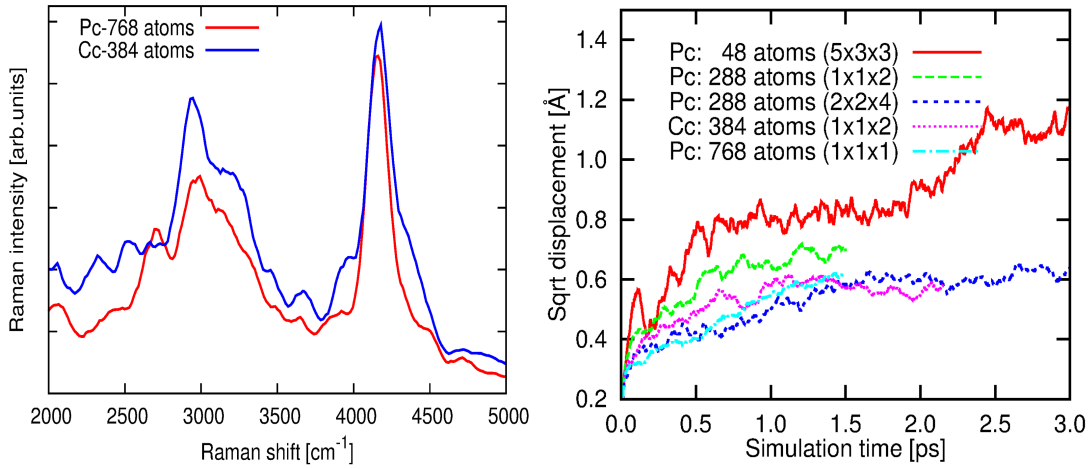

Figure 3.5 *On the left, I show the Raman vibrons computed with the molecular projection method for 768 atoms Pc in red (simulation number 26) and 384 atoms Cc in blue (simulation number 29). On the right, I show the mean displacement (MSD) calculated for phase IV at 270 GPa with different simulation settings: starting cell, the number of atoms and k-point grid as reported in the legend and table 3.1.*

Figure 3.5 shows the MSD calculated for phase IV at 270 GPa from the various MD trajectories presented in table 3.1. The displacement is unusually large owing to molecular rotation in the B layers and trimer rotation in the G layers, however, it does eventually reach an upper limit. The only simulation that appears to sustain diffusion is the 48 atom cell. In this case, the increased diffusion is a

result of finite size effects.

The study in ref [118] argued that the lack of diffusion in our calculation is mainly the result of insufficiently long simulations. In the second half of this chapter in section 3.3.3, I will present new simulations extended to higher pressures and longer time scale that reconcile the disagreement between our studies [1, 3, 97, 118]. As we will see later, the structure that undergoes true diffusion is slightly different than BG'BG" and could be related to the newly discovered phase V.

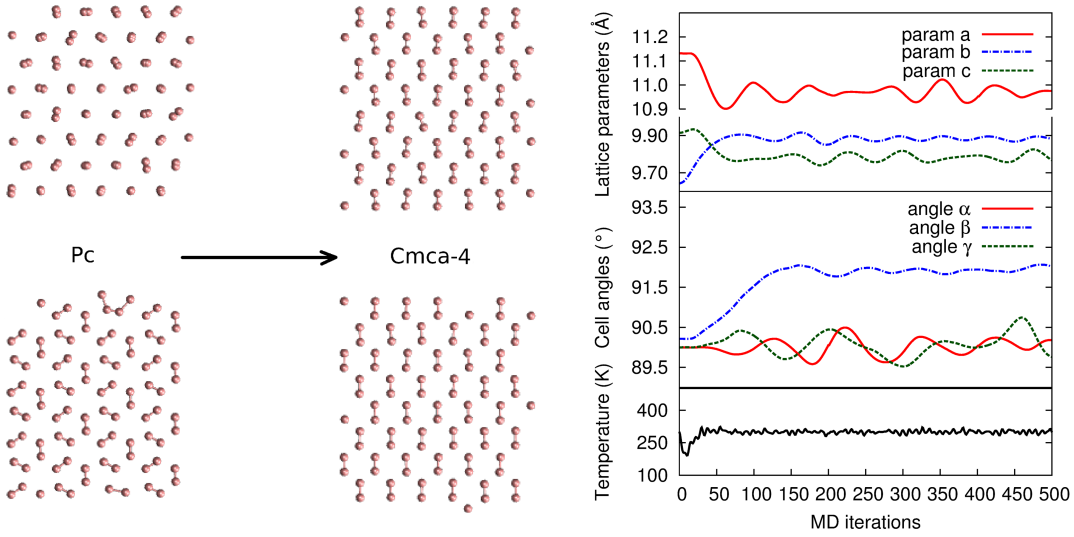

Figure 3.6 *Here I report the result of a short NPT MD simulation at 340 GPa and room temperature (number 27 from table 3.1) which was started in a 768 atom supercell of Pc and immediately transformed to Cmca−4. The left figure shows the average atomic positions in four of the eight layers. The right figure shows the temperature as well as the evolution of the lattice parameters across the transformation.*

Finally, at a pressure of 340 GPa, the simulation started in Pc immediately transformed to $Cmca-4$ which is a molecular structure with a single type of molecule (see figure 3.6). The stabilization of $Cmca-4$ at higher pressures was found in both static [91] and dynamic simulations [1, 96, 98], however recent studies have shown this result is highly dependent on the exchange-correlation functional used.

When Raman is calculated in Pc from perturbation theory, the hard vibron fits the experiment, but the soft vibron does not. This is expected since phase IV is an entropically stabilized structure and at finite temperature, it has a higher symmetry than Pc. This disagreement was the main motivation for performing MD in the first place and for devising a method to calculate the vibrons from

MD trajectories at finite temperature.

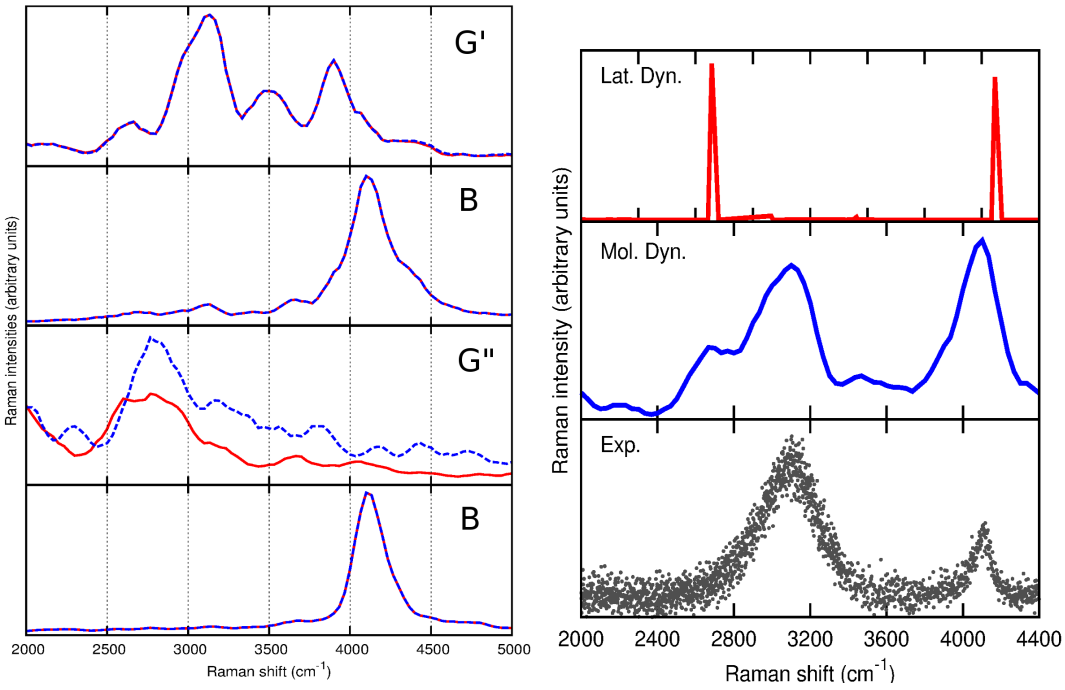

Figure 3.7 *This figure reports the Raman vibrons calculated with the projection method (equation 3.10) at 270 GPa and 220 K for phase IV (simulation 18). On the left, I show the vibrons calculated separately for each layer using two different approaches: 1) the molecules are identified at the start of the simulation and the labels are kept throughout (continuous red line), 2) the molecules are recalculated at each step (dotted blue line). On the right, I compare the Raman extracted from MD at 270 GPa and 220 K (blue) with the result calculated with DFPT at 0 K (red) [1] and the experimental data from ref [89] (dark gray).*

In figure 3.7 I compare the Raman active vibrons extracted from MD at 270 GPa with the experimental data [89] and the calculation obtained from perturbation theory at 0 K [1, 91]. In this particular case, the vibrons extracted from the MD agree well with the experiment and correct for the earlier discrepancy. Interestingly, the simulation also predicts a shoulder on the softer vibron, which was identified in more recent experiments [119], however, higher up in frequency. By 270 GPa, the simulation is a BG'BG" stacking and the agreement of the Raman data with the experiment brings new evidence that the true nature of phase IV is dynamic and phase IV cannot be described well by any of the 0 K symmetries found in AIRSS.

In figure 3.7, I also show the Raman spectra computed by projection separately for each layer. It is clear that the high frequency vibron originates from the

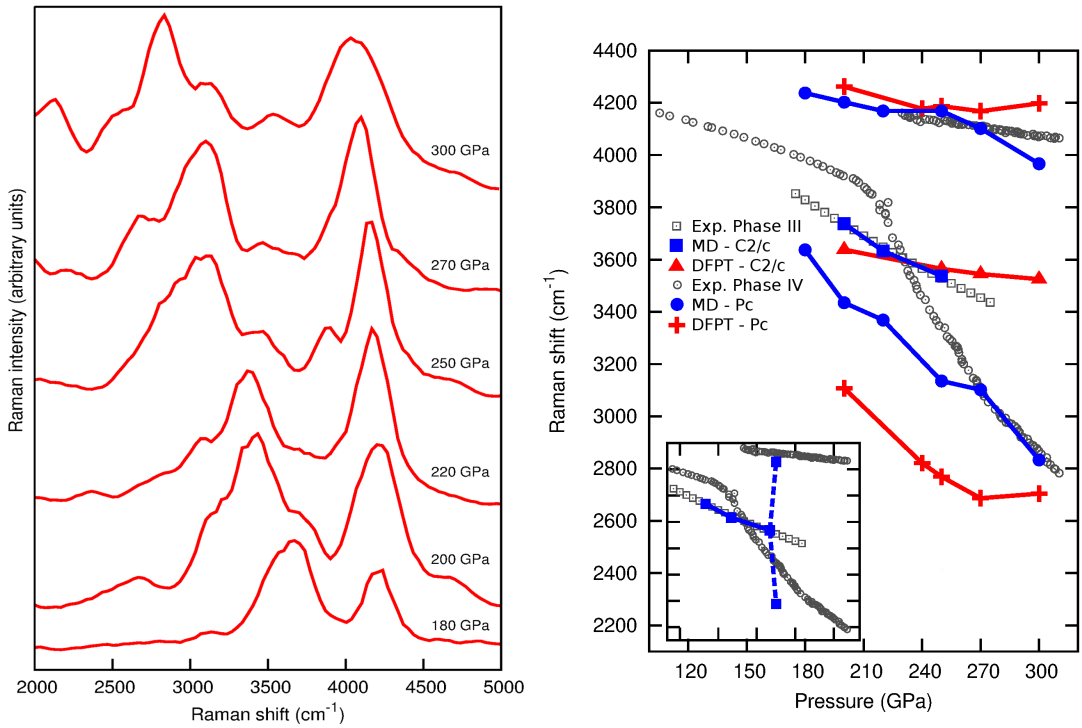

Figure 3.8 *In this figure, I illustrate the pressure dependence of the Raman vibrons computed by molecular projection. On the left I show the spectra calculated for phase IV (starting in Pc) at 220 K and pressures ranging from 180 GPa (bottom) to 300 GPa (top) (corresponding to simulations 14-19 in table 3.1). On the right I compare the Raman obtained from simulations (solid) with the Raman measured experimentally (open symbols). Experimental data for phase IV was obtained at 300 K in ref [89], while for phase III at 90 K in ref [82]. DFPT data was calculated at 0 K in refs [1, 78] while my MD results were obtained at 220 K (numbers 14-23 of table 3.1). The inset shows the Raman peaks extracted from the BGBG structure after C2/c reconstructed at 250 GPa and 300 K. Notice how the vibrons split after the transformation.*

B layer while the softer vibron originates from the G layer. As I explained in the methods section 3.2.2 the molecules were determined with two approaches: 1) identified in the beginning and then kept throughout the simulation and 2) recomputed at every step. The two methods give the same result in layers B and G', but a slightly different result in layer G". This difference indicates that in layer G", molecules are shorter lived and undergo continuous breaking and rebonding.

The final test for the projection method and for the validity of the BG'BG" model as a candidate for phase IV comes from the pressure dependence of the vibrons presented in figure 3.8.

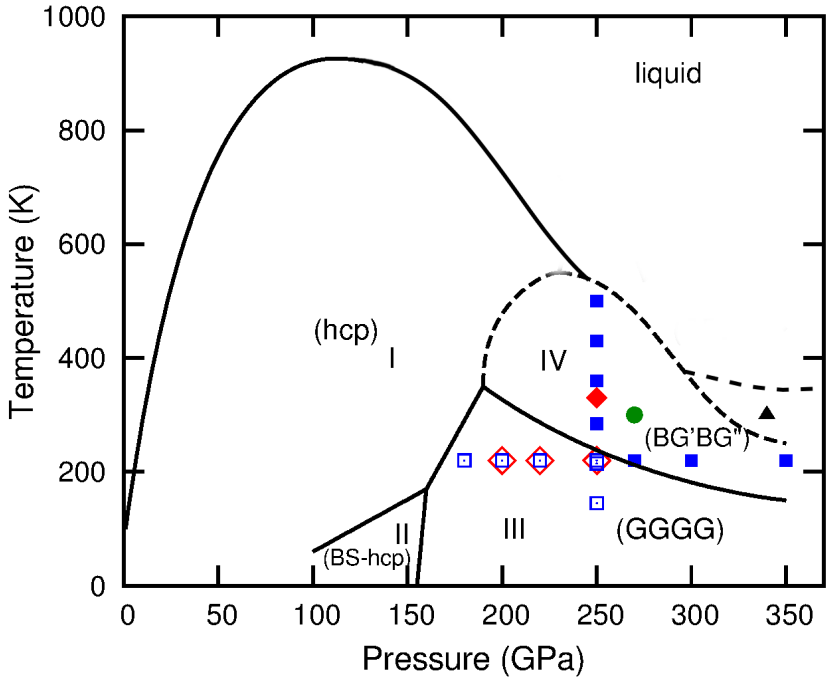

Figure 3.9 *This figure summarizes on the P-T phase diagram [90, 96] the simulations performed in this section: MD started $C2/c$ which remained unchanged (open red rhombi); $C2/c$ which transformed to a BGBG stacking upon heating (solid red rhombus); simulations started in Pc which stayed in a BGBG stacking and are probably metastable (open blue squares); Pc that transformed to a BG'BG" stacking (solid blue squares and solid green circle) and finally MD initiated in Pc which transformed to $Cmca-4$ (solid black triangle). Note that both the solid green circle and the solid black triangle label simulations containing 768 atoms, while all the other labels are simulations with 288 atoms.*

For both phases III and IV, I compare the vibrons extracted from MD with those computed by perturbation theory (DFPT) and the experimentally measured ones. The result obtained from MD for $C2/c$ matches the experimental data for phase III better than the DFPT result. This brings additional evidence to support that $C2/c$ is a good candidate structure for phase III [78], despite the small discrepancy between DFPT and experiment.

In phase IV, the vibrons obtained with the new projection method from the MD started in Pc agree well with the experiment, especially at pressures above 260 GPa where the stacking becomes BG'BG". This result solves a major disagreement between theory and experiment and indicates that the high temperature nature of phase IV is the BG'BG" model.

Finally, in figure 3.9, I summarized all the MD simulation on the phase diagram

[1]. The calculations agree with the experiment regarding the stability regions of each phase. In the region of phase III, simulations started in $C2/c$ remain unchanged, while simulations started in Pc form a metastable BGBG stacking. On the other hand, in the region of phase IV, the MD simulation started in $C2/c$ and heated up transformed to a BGBG type stacking. Additionally, in the same regime, the Pc simulations changed to a BG'BG" stacking, which is the high temperature manifestation of phase IV. Further in the chapter I will present additional simulations in the pressure range 250 to 500 GPa that indicate more subtle modifications beyond the BG'BG" model.

3.2.4 Summary

This section was aimed at understanding the dynamic behavior of phases III and IV of solid hydrogen. For doing so, I performed a series of MD simulations in the pressure range 180-350 GPa and at temperatures from 150 to 500 K. Additionally, I developed a new technique for extracting the Raman active vibrons from the dynamics in light of direct comparison with the experiment.

The Raman vibron computed from the dynamics of $C2/c$ agreed well with experiment even better than the previous DFPT result. Two conclusions can be drawn: 1) the $C2/c$ vibron which originates from distorted G layers is slightly anharmonic and the new method corrects for it and 2) $C2/c$ is a good candidate for phase III.

Upon heating, the $C2/c$ structure transforms to a BGBG type stacking accompanied by complete layer reconstruction. This transformation agrees with the experimental result that phase IV is stabilized by entropy and only exists at finite temperatures above phase III.

MD is crucial for understanding the dynamic nature of phase IV. The finite temperature behavior is the stacking model BG'BG", where molecules are almost freely rotating in the B layer and the G layers differentiate into two types: one which allows trimer rotation (G") and one which does not (G'). The Raman vibrons calculated with this model fit the experimental data well, bringing evidence that BG'BG" is a good candidate for phase IV.

At the pressures and temperatures studied here, I found no diffusion in the BG layered models, but as I will show in the next section 3.3, diffusion becomes

possible after a slight change of symmetry which takes place at higher pressures.

3.3 Phases IV-V

3.3.1 Introduction

The new studies I will present here were prompted by a very recent experiment which claims the discovery of a yet new solid phase V [100] based on modifications recorded in the Raman spectra as shown in figure 3.10. The study also brings new evidence for the phase IV to IV' transition, claimed in previous work [90]. These findings were confirmed by an independent experimental study [101].

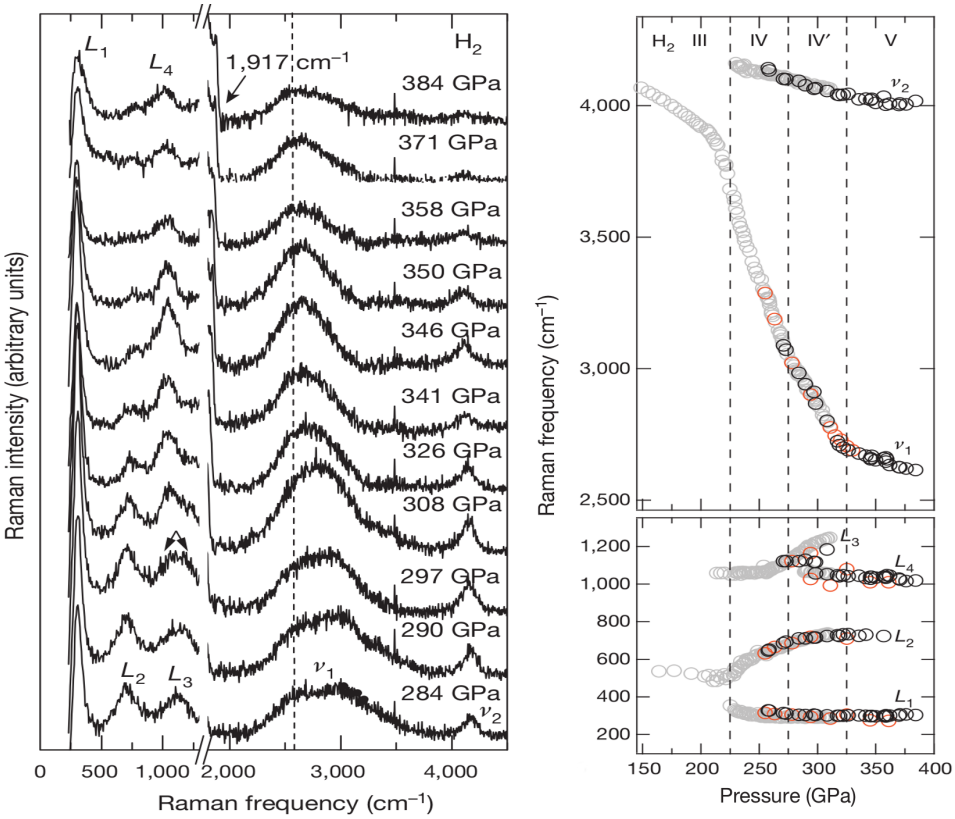

Figure 3.10 *This figure is taken from ref [100] and summarizes their experimental results in hydrogen where the new phase V is identified. On the left is the change of Raman spectra with increasing pressure, while on the right is the change in the position of the vibrons ν_2, ν_1 and lattice modes L_{1-4}.*

The experimental signature of these phases can be summarized as follows. The softer vibron originating from the G layers which varies strongly with pressure

in phase IV, becomes less pressure dependent in phase IV' and finally almost pressure independent in phase V. Whereas phase IV generates three independent low frequency lattice modes, labeled L_1, L_2 and L_3, in phase IV' a new mode appears to split from L_3 and is labeled L_4. Upon transition into phase V, modes L_2 and L_3 vanish. Additionally, L_1 and L_4 increase in relative intensity, while the softer vibron ν_1 decreases. Finally, the full width at half maximum (FWHM) of L_1 starts to increase rapidly at the IV'-V transition.

My hypothesis is that these modifications could be caused by a gradual symmetrization of the G layer, where the intermolecular distances become similar to the intramolecular ones. Since both vibrons persist in phase V, the new structure is likely a variation of the bromine-graphene (BG) layering. The softer vibron becomes pressure independent in phase V which indicates that while the bond elongates rapidly with pressure in phases IV/IV', upon transition to V, it asymptotically reaches a physical limit. There are no abrupt changes in the low frequency modes. Two of them gradually increase in intensity, while two gradually decrease in intensity until they can no longer be detected in the experiment. This again indicates subtle dynamical changes in one of the B or G layers.

To investigate these modifications, I analyzed a series of phonon calculations, Raman and infrared alongside molecular dynamics in the pressure range 200 to 500 GPa. As shown previously in section 3.2.2 low frequency modes cannot be extracted from MD in the case of phase IV. Therefore to understand the nature of these modes we relied on the ground state structural candidates. Good candidates for phase IV,IV' and V at 0 K are the crystal structures *Pc*, *Pbcn* and *Ibam* as discovered based on AIRSS [78, 91]. This structure sequence also represents a continuous symmetrization from the least symmetric *Pc* to the most symmetric *Ibam* which could be related to the transition into phase V.

I start from these candidates to investigate changes with pressure in both ground state and at room temperatures, attempting to understand the origin of phase V. The 0 K transformation were studied using energy calculations, bond analysis and phonon calculations. On the other hand, the room temperature behavior which reflects the true nature of the experimental phases, was investigated using molecular dynamics MD alongside tools I developed to analyze the trajectories in terms of structure and optical properties. Note that when I discuss MD calculations, the space groups only denote the start of the simulations and do not refer to the actual average symmetry at finite temperature. On the other hand, in static calculations, the symmetries are precisely as named. The results

of this section are soon to be published in a journal article [8].

3.3.2 Methods

The CASTEP simulations presented in this second half of the chapter were performed by two summer students Benjamin Tyson and Bálint Borgulya under my guidance. Ben performed the phonon, Raman and infrared calculations, while Balint performed all the molecular dynamics (MD) calculations. I did the analysis and interpretation of the results.

DFT Details

All computations presented in this section were performed using the plane wave method as implemented by CASTEP [31, 115]. The plane wave energy cutoff was set to 1200 eV, which is necessary to achieve good convergence. For approximating the exchange-correlation energy the PBE functional [28] was employed. In all simulations, including molecular dynamics (MD) we used the norm conserving pseudo-potential generated "on the fly" by the string "H 0|0.7|2|6|8|10L(qc = 10)" [116]. The total energy in the electronic minimization was converged within 10^{-11} eV and the charge density was treated as spin unpolarized. In some calculations, the systems were considered metals while in others they were treated as insulators as explained in each of the subsections. Density mixing (dm) was the method used in the self-consistent minimization loop.

Raman and IR Spectra

Metals do not generate Raman and IR spectra. This is mainly because in conductive materials, the free electrons can easily rearrange to screen any external electromagnetic field, such that inside the bulk the field is zero. Here, the electron occupancy was set fixed and the Raman/IR calculation was only performed at 250 GPa where it is safe to assume that all crystals are still insulators [91]. For geometry convergence we used the following thresholds: 10^{-9} eV/atom for energy, 10^{-6} eV/Å for force, 10^{-5} GPa for stress and 10^{-7} Å for the displacement. These stringent constraints are required to obtain accurate spectra. Table 3.2 summarizes the cell setups of the three crystal structure candidates we considered:

Pc, *Pbcn* and *Ibam* [78, 91]. These structures have 48, 48 and 16 atoms in their standard unit cells, respectively. Here these cells were extended to ensure sufficient convergence of long range interactions. Table 3.2 reports the optimized geometries at 250 GPa with the settings detailed above.

Table 3.2 *The table shows the cell setups. The last column summarizes the cell geometries as optimized at 250 GPa.*

structure	cell	atoms	k MP grid	optimized geometry (a,b,c,α,β,γ)
Pc	2 1 1	96	6 8 6	5.799, 5.043, 5.174, 90.00, 90.16, 90.00
Pbcn	2 1 1	96	6 8 6	5.794, 5.041, 5.179, 90.00, 90.00, 90.00
Ibam	2 2 1	64	4 12 4	5.731, 3.328, 5.246, 90.00, 90.00, 90.00

Raman calculations were performed with linear response, using a phonon convergence tolerance of 10^{-5}eV/Å^2 and FINE method interpolate.

Phonon Pressure Dependence

All candidate structures become metals at high pressure [91]. To allow the metallic modifications, we set the electron occupancy to variable. With this setting, phonon calculations were performed on all three structures, every 25 GPa, in the pressure range 250-450 GPa. The occupancy was smeared using the "fermidirac" scheme with smearing width 0.1 eV. The cell setups were the same as for the Raman calculation at 250 GPa, shown in table 3.2. Additionally, 24 extra bands were added in all *Pc* and *Pbcn* calculations and 8 extra bands in all *Ibam*.

As one may notice in figure 3.15 from section 3.3.3, *Pc* is the least symmetric structure from the three, followed by *Pbcn* and finally *Ibam*, which is the most symmetric one. During the geometry optimization, only phase transformations that increase symmetry were allowed.

At the end of each geometry optimization CASTEP automatically computes charge populations and bond statistics, which were used to analyze the pressure evolution of the different distance scales in the three structures.

For each optimized structure, phonons were calculated at the gamma point only. The calculations were performed using the finite displacement method, with the super cell technique for *Ibam* and *Pbcn*. The super cell matrices chosen for the two structures were diagonal [2, 4, 2] and diagonal [2, 2, 2], respectively. *Pc*

created problems due to some software issues, so, in this case, linear response with FINE method interpolate was used instead.

I devised a technique to track the phonons of each crystal structure along pressure, based on the oscillation pattern. At the end of every calculation we obtained a list of frequencies f_j and associated eigenmodes $\varepsilon^{\alpha}_{l,j}$, where $l = 1 : N$ labels the atom, $j = 1 : 3N$ labels the mode and α the Cartesian component. The index j is not unique and it often points to different modes at different pressure points. If $\varepsilon^{\alpha}_{l,j}(p_1)$ and $\varepsilon^{\alpha}_{l,k}(p_2)$ are the eigenvectors for modes j and k at pressures p_1 and p_2, respectively, the challenge is to find the mapping between the j's and k's labeling the same mode patterns. Every mode j at a pressure p_1 is projected onto every mode k at all pressures p_2, by calculating the sums:

$$S_{j,k}(p_1, p_2) = \sum_{l=1}^{N} \sum_{\alpha=1}^{3} \varepsilon^{\alpha}_{l,j}(p_1) \varepsilon^{\alpha}_{l,k}(p_2) \tag{3.11}$$

For every mode j at p_1, I identify the label k at p_2 that gives the maximum sum. Note that the eigenmode completeness, requires that for every phonon j:

$$T_j(p_1, p_2) = \sum_{k=1}^{3N} S_{j,k}(p_1, p_2)^2 = 1 \tag{3.12}$$

In other words, the square projections of a given mode j at p_1 onto each mode k at p_2 sum to unity. If the square of the maximum sum $S^2_{j,k}$ is less than half of that the total sum T^2_j, the mode j from pressure p_1 does not continue to pressure p_2 and it is likely a hybridization of several k modes at p_2. Therefore one can say the initial mode j ceases to exist at pressure p_2 if there is no k for which:

$$S_{j,k}(p_1, p_2) \geq \sqrt{\frac{1}{2}} \approx 0.7 \tag{3.13}$$

This procedure allows us to compute the frequency dependence of a certain mode with pressure. The Raman and infrared calculations performed at 250 GPa for nonmetals help identify the active modes. One can then track those modes along the pressure axis by projecting them onto the new eigenmodes calculated for possible metallic systems at each pressure point, without performing the otherwise expensive Raman calculations. The method assumes that for a given structure, the Raman and infrared activities do not change significantly with

pressure provided the oscillation pattern of the mode remains the same. This might not always be the case.

Molecular Dynamics Simulation

I continued my previous study on the dynamics of phase IV presented in the first half of this chapter, section 3.2, and extended the calculation to longer time scales and higher pressures into the newly discovered phase V [100]. We employed molecular dynamics (MD) to study the room temperature behavior of the three structure candidates *Pc*, *Pbcn* and *Ibam*. Simulations with small cells were shown to cause serious problems with finite size effects. Therefore for this section we used large cell as reported in table 3.3. Note that here all cells comprised of eight layers, where each layer contained 18 molecules. We allow for variable occupation of the electron degrees of freedom which is the correct treatment for possible metallic systems. The table also shows the number of atoms, k-point grid and percentage of extra bands we used for each structure. In chapter 4, I will show that even for large simulation boxes, sampling the Brillouin zone at the gamma point only is not sufficient, especially when the structures could become metallic, hence the dense k-point mesh.

Table 3.3 *The table shows the cell setups used for molecular dynamics.*

structure	cell repeat	number of atoms	k-point MP grid	extra bands
Pc	3 2 2	576	2 2 2	30%
Pbcn	3 1 2	288	2 4 2	100%
Ibam	3 3 2	288	2 4 2	100%

The calculations were started with geometry optimizations on the large cells at a number of pressure points as shown later in table 3.6. MD does not require the same level of convergence needed for phonon calculations so in this case the convergence thresholds for the geometry optimization (GO) and subsequent MD, were relaxed to the CASTEP default values: 2×10^{-5} eV/atom for energy, 5×10^{-2} eV/Å for force, 10^{-1} GPa for stress and 10^{-3} Å for atom displacement.

For the MD runs one additional DFT setting was changed after the GO, namely the threshold for the total energy calculation in the self-consistent loop was set to 10^{-9}eV in light of faster convergence. The MD simulations were initiated with $t_{max} = 500$ steps of equilibration in the NPT ensemble. The time step was 0.5 fs giving a total of 0.25 ps of equilibration. The target temperature was set to

300 K and kept by a 5-chain Nose-Hoover thermostat. The pressures for each structure were chosen as in table 3.6 and stabilized with the Parrinello-Rahman barostat using the default characteristic cell time of 50 fs.

These short trajectories were analyzed using radial and angular distribution functions (RDF and ADF). Throughout the simulations, the atoms remained clustered in layers, except where the transition to $Cmca-4$ occurred. I first separated the trajectories into B layers and G layers, four of each type, and then calculated a 2D, in layer, RDF from each layer type. At each time step t, I created the histogram of atom to atom distances $n_t(r_i)$ averaged over atoms from which I could then derive the time averaged RDF as:

$$RDF(r_i) = \frac{1}{t_{max}} \sum_{t=1}^{t_{max}} \frac{4S_t}{\frac{N}{2}} \frac{n_t(r_i)}{2\pi r_i dr} \tag{3.14}$$

where r_i is the space discretization, t labels the times steps running from 1 to maximum time t_{max}. $4S_t$ is the total surface area at time t (there are four layers of each type: 4B, 4G), $N/2$ the number of atoms (four out of eight layers gives half the total number of atoms), and dr the bin width. I normalized with respect to the 2D surface element, which only makes sense to the point where neighboring layers of the same type start to contribute to $n_t(r_i)$ at around 2.5 Å. I also calculated the coordination number with respect to distance R_i as the sum over the time averaged $n_t(r_i)$:

$$c(R_i) = \frac{1}{t_{max}} \sum_{r_i}^{r_i \leq R_i} \sum_{t=1}^{t_{max}} n_t(r_i) \tag{3.15}$$

This allows us to study the average crystal structures in a somewhat standard way. To gain more insight and understand the local order better, I also investigated the angular distribution function (ADF). ADF is basically a histogram of angles θ spanned by two distances r and q from a reference point (e.g atom or molecular center) to two distinct neighboring atoms, respectively. I defined ADF differently for the B layers and G layers because one layer is molecular in nature and the other gradually becomes atomic upon increasing pressure.

For the B layer, I start from a reference atom, then limit the first distance to the molecular length from the condition $c_B(r) \leq 1$. I calculate the second distance q' from the center of the identified molecule (instead of reference atom) and limit

it to the next twelve neighboring atoms by setting $c_B(q') \leq 13$. In the G layer, the first distance was limited by the condition $c_G(r) \leq 3$, since the molecule is not always well defined. For instance, by 400 GPa the first RDF peak integrates to three, whereas if the molecule was still intact, it would integrate to one. The second distance q is limited by $c_G(r) \leq 12$ which includes the next 9 neighbors, three hexagonal trimers in total.

Note that the two neighboring atoms spanning the angle θ at the reference atom are always distinct, but while in the case of B layer only one atom falls behind the first limit $r_{c_B(r)\leq 1}$, in the case of the G layer both atoms can be closer than the first limiting distance $r_{c_G(r)\leq 3}$. Finally, at each time step, I calculate the histogram $m_t(q_i, \theta_j(r,q))$ with respect to the second reference distance q_i and the angle.

$$ADF_2(q_i, \theta_j) = \frac{1}{I}\frac{1}{t_{max}}\sum_{t=1}^{t_{max}} m_t(q_i, \theta_j(r,q)) \tag{3.16}$$

where I is the volume integral under the averaged histogram, used for normalization purpose.

Additionally, I also calculate a one variable ADF_1 as a partial integral of ADF_2 over the second reference distance q:

$$ADF_1(\theta_j) = \frac{1}{I}\frac{1}{t_{max}}\sum_{q_i=0}^{q_i<q_{max}}\sum_{t=1}^{t_{max}} m_t(q_i, \theta_j(r,q))\frac{dq}{q_i^2} \tag{3.17}$$

where dq is the bin size and q_i^2 is a factor used to emphasize the local order over long range order.

As I will discuss in the results section, the simulations started in *Pc* and *Ibam* gave the same dynamic behavior as *Pbcn* even after a very short NPT stage, therefore we continued with *Pbcn* only. The *Pbcn* trajectories were extended for 5 ps in the NVE ensemble at 300 K. At 200 GPa and 500 GPa we only performed 2.5 ps of NVE simulation. All other settings that apply were kept unchanged from the previous NPT stage.

For analyzing the average crystal structure and behavior as function of pressure at room temperature I calculated RDF and ADF_2 as explained above. I also inspected the mean square displacement separately for the two layer types.

Whereas in the first half of the chapter, section 3.2, I looked at the average crystal structures to determine the dynamic behavior, here I computed iso-surfaces of proton density (PDS). These were obtained by cumulating the whole trajectories into density fields followed by surface interpolation and plotting using VMD [120].

The trajectories were also used to extract information about phonon density of states by the standard Fourier Transform of the velocity auto-correlation function (VACF). Raman vibrons were calculated with the molecular projection method, introduced in the first half of the chapter, section 3.2.2. To remind the reader: I project the velocities at each time step onto molecular stretches and then take the Fourier Transform. The difference to the previous work in section 3.2.3, is that here the molecules are always recalculated at each step because above 300 GPa, at room temperature, molecules are short-lived and they can break and rebond during the course of the simulation. The positions and widths of the peaks were obtained by fitting gaussian functions with a constant background.

Finally, I calculated a free rotor frequency in the B layer, by using the fitted RDF peak (molecular length b) to compute the moment of inertia I_{H_2}:

$$E_J = \frac{\hbar^2}{2I_{H_2}} J(J+1),\ \ \Delta J = 2,\ \ \nu = \frac{\Delta E}{h} = \frac{3h}{2\pi^2 m_H b^2} \tag{3.18}$$

The approximation for the rotor linewidth in this case was also calculated as propagated from the uncertainty in b, which was obtained by fitting the RDF peaks:

$$\Delta\nu = \left|\frac{\partial\nu}{\partial b}\right| \Delta b = \frac{2\nu}{b}\Delta b \tag{3.19}$$

3.3.3 Results

Figure 3.11 and table 3.4 summarize the Raman and infrared result for *Pc*, *Pbcn* and *Ibam*, as calculated with linear response at 250 GPa. For consistency, I labeled the most intense Raman modes based on their similarity to the experimental modes with the same name. The prime symbol denotes modes that are very close in frequency, almost degenerate. In ref [100], the high frequency modes (vibrons) are labeled ν while the lower frequency modes are labeled L.

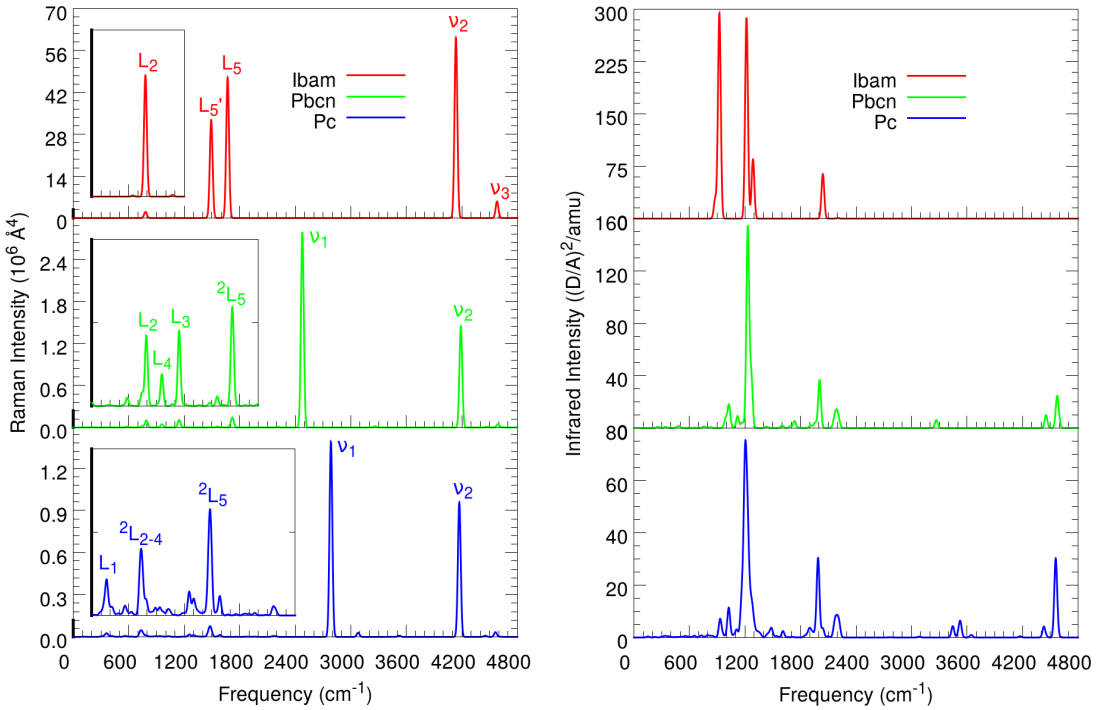

Figure 3.11 *Raman (left) and IR spectra (right) calculated with linear response theory for Ibam, Pbcn and Pc, respectively, at 250 GPa. Insets show the lower intensity, lower frequency Raman modes zoomed in. The scale of the zoom is illustrated on the y-axis in bold. Notice that the Raman intensities in Ibam are much larger than in Pc and Pbcn. The modes are labeled based on their proximity to the experimental peaks with the same names [100]: L_i for low frequency and ν_i for vibrons. The left superscript "2" indicates a doublet. All peaks were smoothed with gaussians of height unity and full width a half maximum of* $35\ cm^{-1}$.

As we have seen before, structures *Pc* and *Pbcn* are molecular so they generate two Raman vibrons, one from each layer: ν_1 from the G layer and ν_2 from the B layer. On the other hand, *Ibam* is atomic in the G layer so it loses the softer vibron ν_1. *Ibam* generates an additional low intensity vibron which I labeled ν_3 in the B layer. This vibron is also present in the other structures but it is much weaker compared to the other vibrons. Notice that Raman peaks in the *Ibam* structure are two orders of magnitude more intense than those in the *Pc* and *Pbcn* structures.

In the low frequency regime, I found three clear Raman modes in *Ibam*, four in *Pbcn* and three in *Pc*, some of which are degenerate in frequency. Modes from different structures that are labeled the same, have very similar patterns. Ref [91] shows four different low frequency modes in *Pc*, but here I only found three.

Table 3.4 *The table summarizes the most relevant Raman peaks for each structure (Pc, Pbcn and Ibam) at 250 GPa.*

struc.	data	L_1	L_2	L_4	L_3	L_5	L_5'	ν_1	ν_2	ν_3
Pc	$freq.\,(cm^{-1})$	361	725, 746		—	1469	1488	2785	4173	4564
	$int.\,(10^4 \mathring{A}^4)$	2.4	3.2, 2.9		—	4.5	4.5	140	96	3.4
Pbcn	$freq.\,(cm^{-1})$	—	790	959	1145	1711	1724	2477	4189	4590
	$int.\,(10^4 \mathring{A}^4)$	—	10	4.6	11	5.6	10	280	140	4.7
Ibam	$freq.\,(cm^{-1})$	—	781	—	—	1669	1491	—	4136	4581
	$int.\,(10^6 \mathring{A}^4)$	—	2.0	—	—	33	47	—	60	5.5

In figure 3.11 I also show the infrared spectra calculated at 250 GPa. *Pbcn* and *Pc* have almost identical infrared activity and they both generate two vibrons, one from the B layers and one from the G layers, similar to the Raman spectra. The result I obtained here is in good agreement with ref [91]. Unlike, the molecular structures, *Ibam*, shows no infrared vibron at all and an extra intense peak at lower frequencies, around 900 cm^{-1}. This is important because these differences are clear experimental signatures and could help identify easily an *Ibam*-like structure in the experiment. However, to date, there is no infrared data for phase V.

The frequencies and intensities of Raman and infrared are important, but one can learn much more by inspecting the patterns of the modes. In figures 3.12 and 3.13 I show the eigenvectors for the most important Raman oscillations. Interestingly, many of the modes are common to all three structures and that is the reason why I chose to assign them the same label. For instance variations of the modes L_5 and L_5' from *Ibam* can also be found in *Pc* and *Pbcn*, albeit with different frequencies (see table 3.4).

As I discussed previously, in the first half of this chapter, the Raman active vibrons are in phase vibrations of the molecules in the B and G layer. Figure 3.12 illustrates the eigenmodes of ν_1 and ν_2, supporting this assumption which made it possible to extract Raman signals from molecular dynamics. ν_3 is an exception, but its Raman activity is relatively weak.

The lowest frequency Raman mode is L_1 and, as shown in figure 3.12, it consists of an in-plane libration of the hexagonal trimers in the G layer, with no motion in the B layer [98]. The low energy of this mode explains why in MD at room temperature trimers can rotate. Interestingly, this mode is not Raman active in *Pbcn* and *Ibam*, so one could expect the mode to vanish upon symmetrization of the G layer.

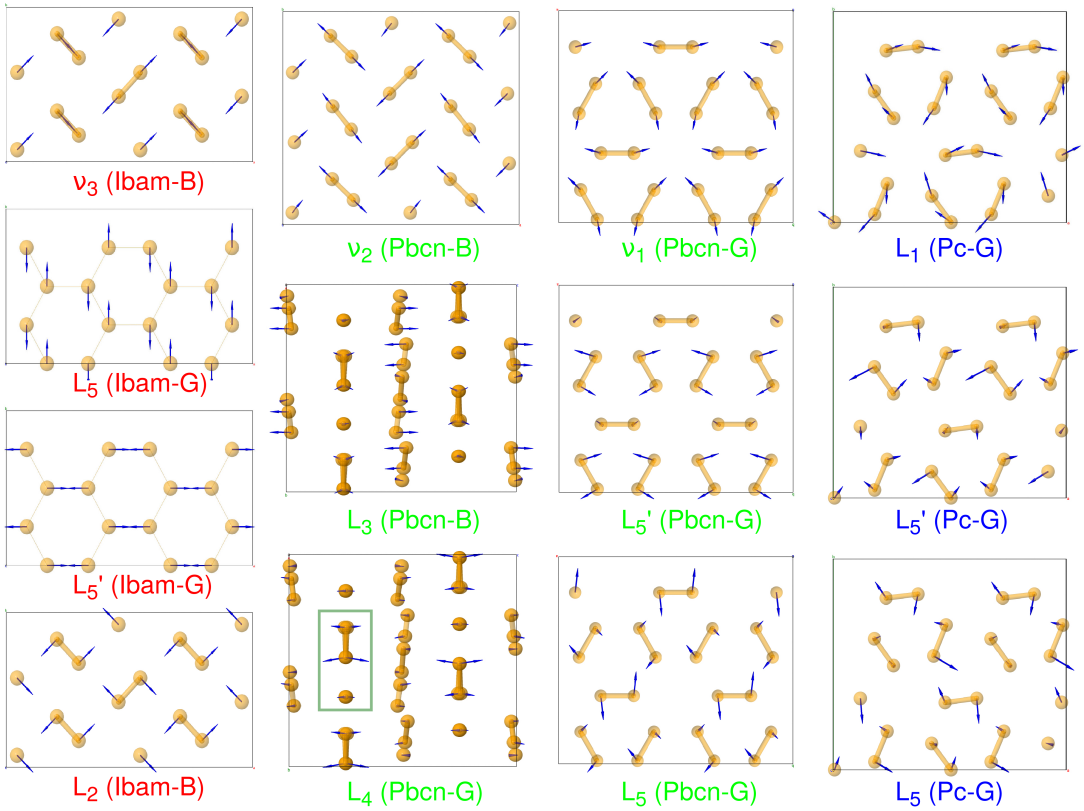

Figure 3.12 *Here I illustrate the most important Raman active modes as calculated with linear response at 250 GPa. Their frequencies and intensities are plotted in figure 3.11. Modes labeled the same in different structures have similar frequencies and oscillation patterns, check for instance the L_5, L_5' doublet in Pc and Pbcn. All the modes shown here involve atom oscillations in one layer type only, which is indicated in brackets next to the phonon label. All panels are top views (perpendicular to the layers), except the ones illustrating modes L_3 and L_4 which are side views. For reference, the green rectangle in L_4 marks a hexagonal trimer seen from the side.*

Further up in frequency is the mode L_2, which can be clearly identified in both *Pbcn* and *Ibam*. It consists of an in-plane, in-phase libration of all the molecule in the B layer, with no motion in the G layer. At room temperature, in MD, this mode becomes poorly defined, because as seen before in section 3.2.3 the molecules can rotate almost freely in the B layer.

Modes L_3 and L_4 are out of plane oscillations and seem the be specific to *Pbcn*. L_3 is a translation of B layer molecules, while L_4 is an out of plane libration of the molecules from the G layer.

Although L_4 and L_2 do not appear in *Pc* with the same clear pattern as in *Pbcn*,

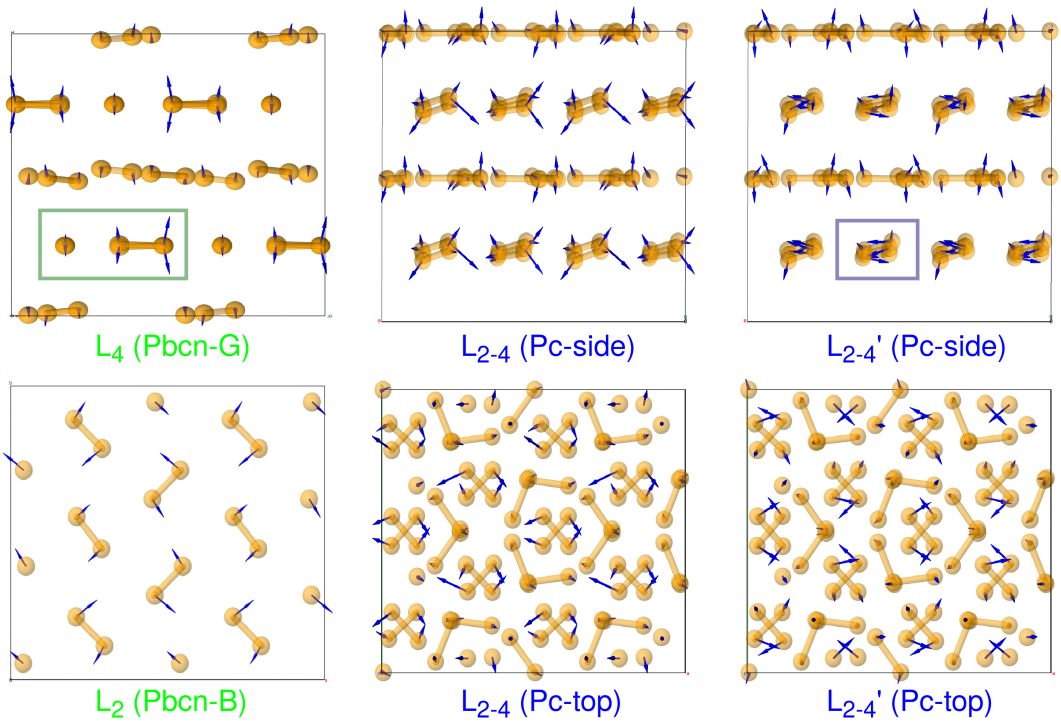

Figure 3.13 *This figure is supplementary to figure 3.12. Here I compare the pure modes L_2 and L_4 in Pbcn with the degenerate mixed modes L_{2-4} and L'_{2-4} in Pc. The green rectangle in L_4 (Pbcn−G) marks a hexagonal trimer in the G layer, seen from the side, while the blue rectangle in L'_{2-4} (Pbcn − G) outlines a molecule in the B layer. When comparing the modes notice that B layers in Pc are aligned with the G layers in Pbcn. All the figures were compiled with JMOL [121].*

the degenerate modes L_{2-4} and L'_{2-4} in Pc seem to be a combination of L_4 and L_2, being especially similar to the latter (see figure 3.13). This explains why L_{2-4} and L'_{2-4} have similar pressure dependence with L_2.

Finally, the modes L_5 and L'_5 are specific to the G layers and can be found in all of three structural candidates. These modes involve in plane librations of the molecules in the trimers. In Pc and $Pbcn$ the mode pair is almost degenerate, but upon symmetrization of the G layer, the two modes split in frequency (check figure 3.12). Interestingly, even though in $Ibam$ the molecules are completely dissociated, the two modes maintain a similar oscillation pattern.

Let us summarize all the important Raman active modes found in these BG layered structures at 0 K:

- ν_2: in plane vibration of the molecules in the B layer

- ν_1: in plane vibration of the molecules in the G layer
- L_5, L_5': in plane librations of the molecules in the G layer
- L_3: out of plane translations of the molecules in the B layer
- L_4: out of plane librations of the molecules in the G layer
- L_2: in plane libration of the molecules in the B layer
- L_1: in plane libration of the hexagonal trimer in the G layer

An important observation here is that all Raman vibrons and low frequency modes originate from one layer type only. This implies that there is little coupling between layers, but it also helps explain what happens to the layers at phase transitions. For instance, as we have seen in section 3.2.3, at the transition III to IV, half of the G layers transform slightly, while the other half change entirely into B layers. Since all Raman modes originate from one layer type only, one expects that both the vibrons and low frequency modes should split in half at the transition. This is consistent with the experimental findings [89]. Furthermore, if the assumption that phase V is a symmetrization of the G layer with no significant change in the B layer is correct, then only half of the low frequency modes and one out of the two vibrons should change. This is what is actually observed in the experiment, where ν_1 becomes almost pressure independent and L_2, L_3 vanish upon transition to phase V. To understand this better, I also investigated the pressure dependence of these Raman modes, as I will show in the following.

To study the pressure dependence I first performed geometry optimizations on all three structural candidates in the pressure range 250-450 GPa as shown in table 3.5. Upon increasing pressure, both Pc and $Pbcn$ eventually transform into the more symmetric $Ibam$, but not into one another. This result supports the assumption that phase V could be a symmetrization of the G layer from phase IV. Pc further transforms into $Cmca-4$, in agreement with the previous molecular dynamics simulations presented in section 3.2.3.

In figure 3.14 I compare the energies and geometries of the three candidates upon increasing pressure. Consistently with what is shown in table 3.5 $Pbcn$ and Pc transform into $Ibam$ at 350 GPa and 375 GPa, respectively. This is evident from the energies, volumes and lattice parameters. The more symmetric structure $Ibam$ has a higher internal energy because the molecules in the G layer are dissociated into atoms, however, it has a smaller volume with better packing so it wins at higher pressures with a smaller enthalpy. I will discuss

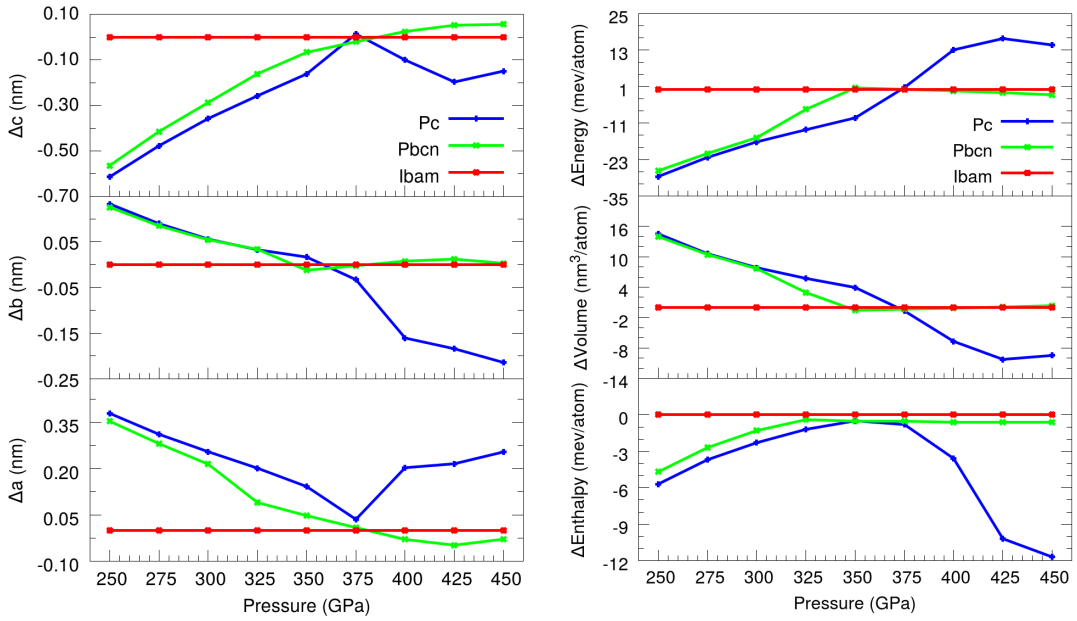

Figure 3.14 *On the left, I show the unit cell vectors of the three candidates. Pc and Pbcn are broken symmetry Ibam and their unit cells can be mapped onto a [1 3 1] multiplications of the Ibam unit cell. To directly compare the geometries of Pc and Pbcn to Ibam, the b length of the former was divided by 3. On the right, I compare the energies, the volumes and the enthalpies of the three 0 K structure candidates for phase IV and V across the pressure range. The values are normalized with respect to the number of atoms and Ibam is used as the reference structure.*

Table 3.5 *Here I summarize the results of the geometry optimizations of the small cells that were later used for phonon calculations. The leftmost column shows the starting geometry, while the rest of the table shows the final symmetry after optimization at each pressure. The top row shows pressure in GPa.*

Start	250	275	300	325	350	375	400	425	450
Ibam	Ibam	Ibam	Ibam	Ibam	Ibam	Ibam	Ibam	Ibam	Ibam
Pbcn	Pbcn	Pbcn	Pbcn	Pbcn	Ibam	Ibam	Ibam	Ibam	Ibam
Pc	Pc	Pc	Pc	Pc	Pc	Ibam	—	Cmca	Cmca

the thermodynamic characteristics of these structures in more detail in the next chapter, section 4.3. Including zero point energy (ZPE) would change the energy landscape entirely, but here I am more interested in understanding the physics of symmetrization upon increasing pressure, which could be related to phase V. For instance in figure 3.15 I show the results of the bond analysis in the three structural candidates.

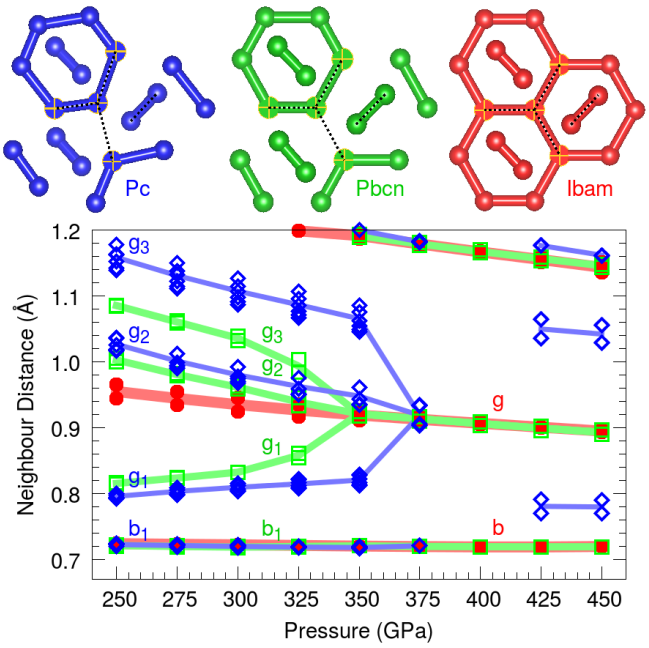

Figure 3.15 *This figure shows the pressure dependence of the characteristic lengths for each of the candidate structures. The top panels illustrate the defining motifs with the characteristic dimensions in dotted lines. The bottom panel shows the pressure evolution of all these dimensions: b_1 is the molecular length in the B layers, while g_i are distances in the G layers with g_1 the molecule length.*

Ibam is the most symmetric structure with two important dimensions: the length of the molecule in the B layer (labeled b) and size of the hexagon in the G layer (labeled g). The hexagon has atoms at the vertices and it describes equal angles spanning 120°. *Pbcn* breaks the symmetry in the G layer, and while angles remain close to 120°, the atoms join into molecules breaking the g distances into three categories: g_1 is the distance between atoms within the molecule, g_2 is the distance between molecules within the molecular trimer (hexagon) and g_3 is the distance between neighboring hexagons. An important observation here is that *Pbcn* comprise of three types of hexagons in the G layer: a small one with sides g_1 and g_2, a medium one with sides g_1 and g_3 and a large one with sides g_2 and g_3. The dynamic transition between phase IV and phase V could be related to changes of these hexagons. *Pc* further breaks the symmetry of the angles at the hexagonal vertices, which are no longer 120°. Understanding these characteristics of the G layer will help later when I discuss the high temperature results in MD which simulates the true nature of these phases.

Following the geometry optimizations, phonons were calculated at all pressure points. Using the phonon tracking technique explained in the methods section,

I investigated the pressure dependence of both Raman (figure 3.16) and infrared (figure 3.17) active modes in all three structural candidates. I also compared the Raman modes with data from the recent experiments [100, 101] and data extracted from the MD with the projection method (see later figure 3.22).

At high frequencies, the harder vibron ν_2 calculated with lattice dynamics (LD) fits the experimental data well, regardless of the structure, implying that the main difference between the candidates is in the G layers and B layer is essentially the same. The ν_2 vibron extracted from MD fits experiment even better, bringing new evidence that the true structure of phases IV and also V is dynamical and should be described with MD.

The main problem here is with the softer vibron ν_1. Lattice Dynamics gives a value too low in frequency, which is related to anharmonicity. Once the structures transform to *Ibam*, the vibron starts to harden with pressure. This is again consistent with the idea that symmetrization of the G layer might be responsible for the pressure independence of ν_1 discovered in the experiment. However, keep in mind that this vibron is not Raman active in *Ibam*.

My previous MD calculations with ultra-soft pseudo-potentials corrected for the difference ν_1 frequency between experiment and LD up to 300 GPa (see figures 3.7 and 3.8). With the new settings used here (norm conserving potentials, better k-point sampling and longer trajectories) I find a slightly different result. MD generates a ν_1 vibron which is closer to the LD value at lower pressures, below 325 GPa, but at higher pressures it does move closer to the experiment, following the same trend where the frequency flattens with pressures. Since MD behaves similar to the experiment, it means that the change in slope of ν_1 indicates a dynamic transition which was successfully captured by MD. This will be discussed in more detail later in the chapter.

Recent studies have shown that the discrepancy between the vibrons measured experimentally and calculated could be a consequence of the functional [85] and also the contribution from anharmonicity [122]. MD is important because not only it can correct for anharmonicity [1, 123], but it also captures the high temperature behavior, which in this case is crucial.

Interestingly, the vibrons measured experimentally in deuterium (scaled accordingly) do not match the hydrogen data. This could be a result of pressure calibration, isotopic phase boundary shift or the vibrons are less harmonic than previously believed.

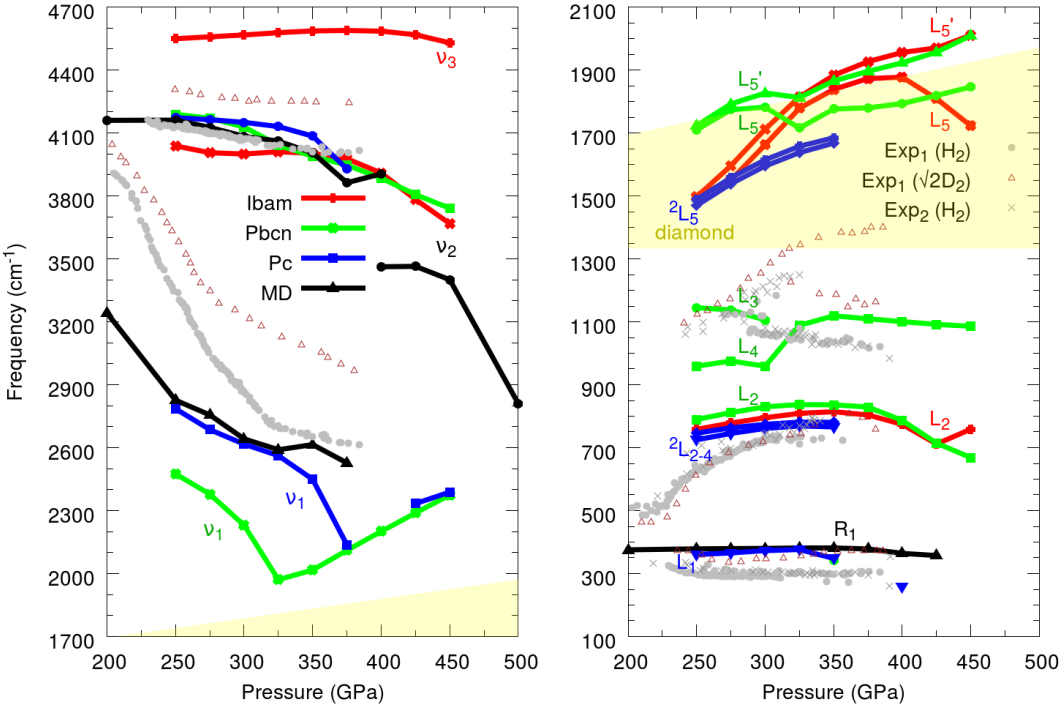

Figure 3.16 *I illustrate the pressure evolution of the high frequency (left) and low frequency (right) Raman active phonons calculated from LD in the candidate structures: Pc (blue), Pbcn (green) and Ibam (red). The Raman activity was only computed at 250 GPa (see figure 3.11) and the trends were obtained by mode tracking. A phonon disappears when the value of the projection falls below* 70% *as explained in the methods section. In black I present the Raman vibrons fitted from MD (see figure 3.22) and I approximate the B layer rotor R_1 from the RDF fit (see methods). In gray I reproduced the hydrogen and deuterium data (scaled accordingly) from the newest experiments: Exp$_1$ [100] and Exp$_2$ [101]. Shaded in yellow is the region of frequency occupied by the diamond spectrum in the experiment.*

In the low frequency Raman regime (right of figure 3.16), the L_5 mode falls under the diamond so it is not accessible in the experiment. The lowest frequency L_1, R_1 and L_2 fit the experiment well, however, we do find a discrepancy. The LD calculations suggest that L_1 is a libron in the G layer, while L_2 is a libron in the B layer. Yet, looking at the experimental data, L_1 is pressure independent like the B vibron ν_2, making it more likely to be a B layer mode. On the other hand, L_2 varies stronger with pressure and then it flattens, just like ν_1 and probably originates from the G layer. These phases are dynamical, so probably MD would give a better description of the low frequency modes. Yet, as I discussed before, the projection method in MD fails for these modes. Interestingly, the free rotor R_1, approximated based on RDF (see figure 3.16), originates from the B layer

and fits the lowest energy experimental phonon L_1.

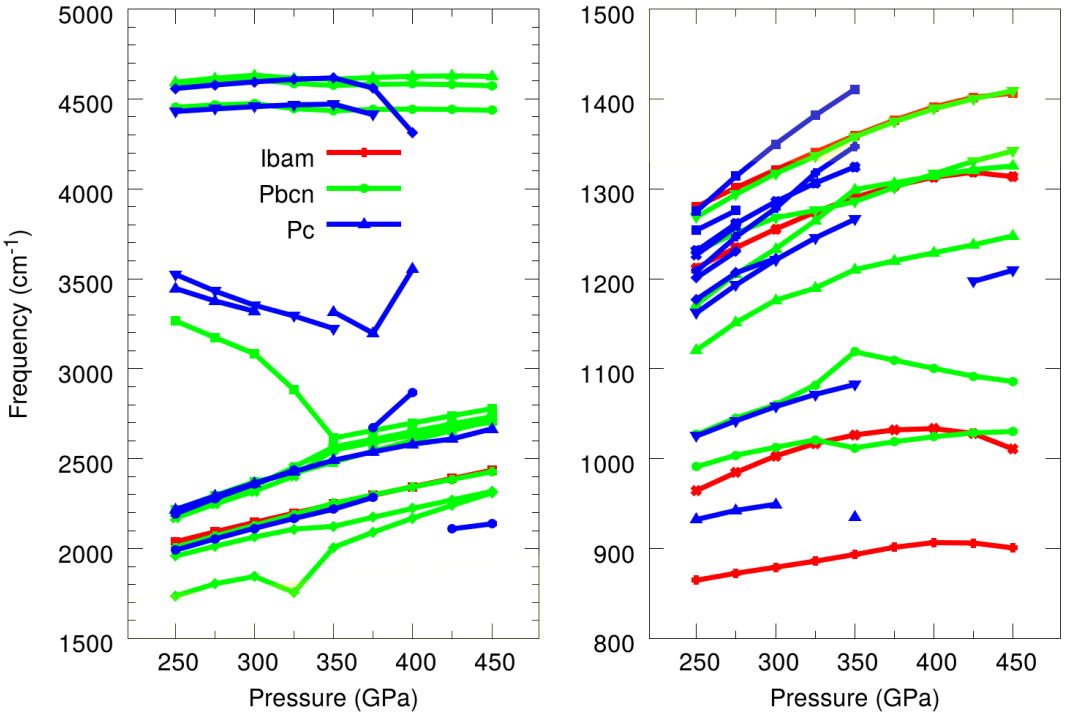

Figure 3.17 *This figure reports the pressure evolution of the infrared spectra for the three candidate structures, as labeled. As explained in the methods section 3.3.2, the actual infrared was only performed at 250 GPa (see figure 3.11), while the evolution with pressure was evaluated by tracking the modes using the eigenmode projection method. High frequency modes are illustrated on the left panel, while the low frequency phonons are on the right.*

A similar story holds true for the L_3 and L_4. Whereas the pressure trends imply that L_3 originates from the G layer and L_4 from B, the calculations indicate otherwise. The assignment to layers according to the experimental observations makes sense. It explains why at the transition to phase IV, the pairs L_3, L_4 and L_2, L_1 appear to split from one single mode each, upon differentiation of the layers into B and G. This observation also supports the idea that phase V is a G layer symmetrization, where the phonons originating from the G layer change (ν_1) or disappear gradually (L_2 and L_3). This is also consistent with different changes at different time scales as I will discuss later (see figure 3.20). The high and low frequency regimes sample different dynamical processes.

In figure 3.17 I show the pressure evolution of the infrared modes, computed with the same phonon tracking technique. In the high frequency regime, the infrared spectra behave similarly to the Raman spectra. The quickly softening vibron in *Pc* and *Pbcn* continues with slowly hardening modes in *Ibam*. Although

these new modes would not be infrared active in *Ibam*, the dynamic transition from lower to higher symmetry in MD at room temperature (later figure 3.20) is likely softer than indicated by the 0 K calculations. Therefore this sharp change in trend could manifest as a gradual change of slope in the experiment if the symmetrization of the G layer hypothesis is correct.

LD allows us to study the origin of the different low frequency modes but does not give an accurate account of the room temperature behavior. To study this, we performed a new series of molecular dynamics calculations at higher pressures, longer time scales and using better quality settings than other studies in literature [96–98].

Table 3.6 *Here I summarize the results of the geometry optimizations of the large cells that were later used for MD simulations. The left most column shows the starting geometry, while the rest of the table shows the final symmetry after optimization at each pressure. The top row shows pressure in GPa. Note that calculations with Ibam were run every 50 GPa only. Compare to table 3.5.*

Start	200	250	275	300	325	350	375	400	425	450	500
Ibam	Ibam	Ibam	-/-	Ibam	-/-	Ibam	-/-	Ibam	-/-	Ibam	Ibam
Pbcn	Pbcn	Pbcn	Pbcn	Pbcn	Pbcn	Pbcn	Pbcn	Ibam	Ibam	Ibam	Ibam
Pc	Pc	Pc	Pc	Pc	Pc	Pc	Pc	Pc	—	Cmca	Cmca

In table 3.6 I present the geometry optimizations of the large unit cells (eight layers each) ahead of the MD simulations. The results are consistent with those reported in table 3.5: *Pbcn* transforms into *Ibam* upon increasing pressure.

We continued with short MD simulations in NPT to equilibrate the structures. After only 0.25 ps of simulation, all structures converged to the same dynamic behavior BG regardless of the starting configuration: *Pc*, *Pbcn* or *Ibam*.

In figure 3.18 I compare the RDF and ADF_1 extracted from the NPT trajectories with the RDF and ADF_1 extracted from the structures at 0 K. Notice that at 0 K all three structural candidates show distinct features, whereas the MD averages out the small differences and at room temperature, all structures exhibit the same behavior. Even *Ibam* which is noticeably different in the ground state, creates molecules and becomes BG like. The thermal energy is sufficient to blur out the small enthalpy differences. We, therefore, decided to only extend the trajectories started in *Pbcn*. The subsequent simulations were run at all the pressure points for 5 ps in NVE.

Figure 3.19 shows the ADF_2 and RDF extracted from these long trajectories.

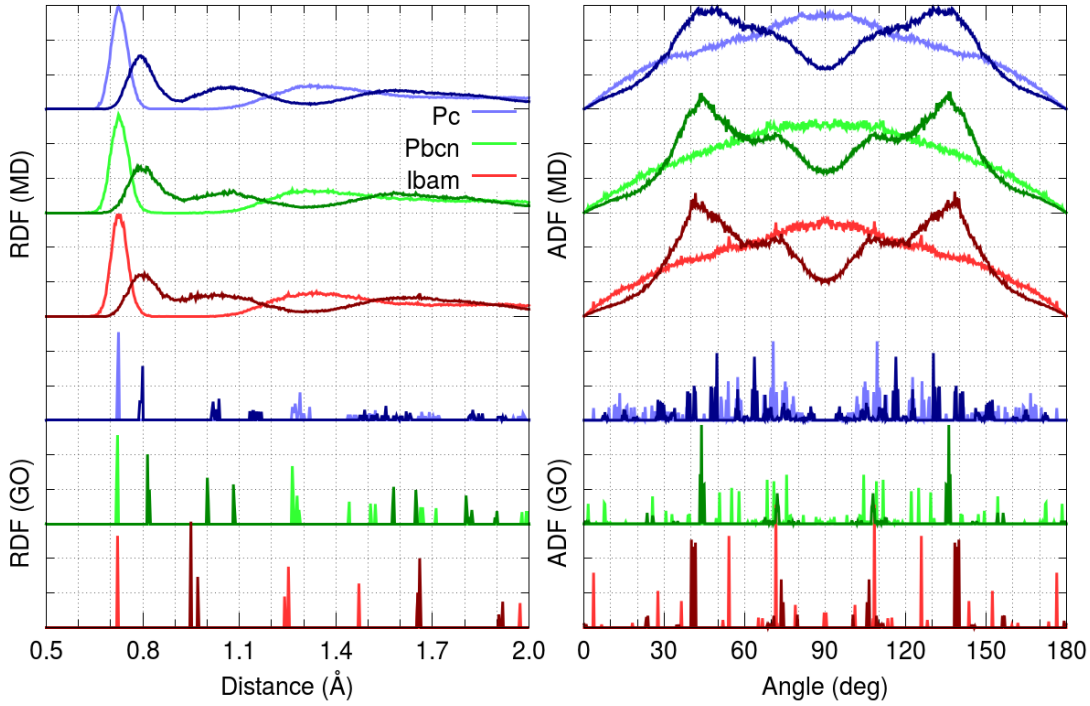

Figure 3.18 *Here, the finite temperature 0.25 ps NPT trajectories of the three candidate structures (top) were compared to the 0 K optimized geometries (bottom) using RDF and ADF_1. On the left, I show RDF, while on the right I show ADF_1, both computed from the trajectories (curves) / geometries (bars) at 250 GPa. Dark colors show the results from the G layers, while light colors from the B layers.*

In the B layer, the RDF shows one well defined peak at around 0.75 Å, which integrates to one. This peak is weakly dependent on pressure and corresponds to the molecular length. The next two peaks, integrating to 12, correspond to the six neighboring molecules also shown in the proton density surface (figure 3.19). These peaks are very broad because the molecules are almost freely rotating in the B layer. The rotation is clearer in the ADF_2 which shows the preferred angle between an atom at distance q from the reference molecular center and the molecular axis itself. Although there appears to be almost free rotation in the B layer, at a closer look there is some weak correlation. At low pressures (250 GPa) the molecules like to stay perpendicular to one another, while at high pressures the molecules prefer angles of 45° and 135°. At intermediate pressures, there is a mixture of these preferred correlations.

In the G layer the calculated RDF and ADF_2 support the assumption of symmetrization with pressure. At low pressures, the RDF gives one peak between 0.75 and 0.9 Å, that integrates to one and corresponds to the molecular length

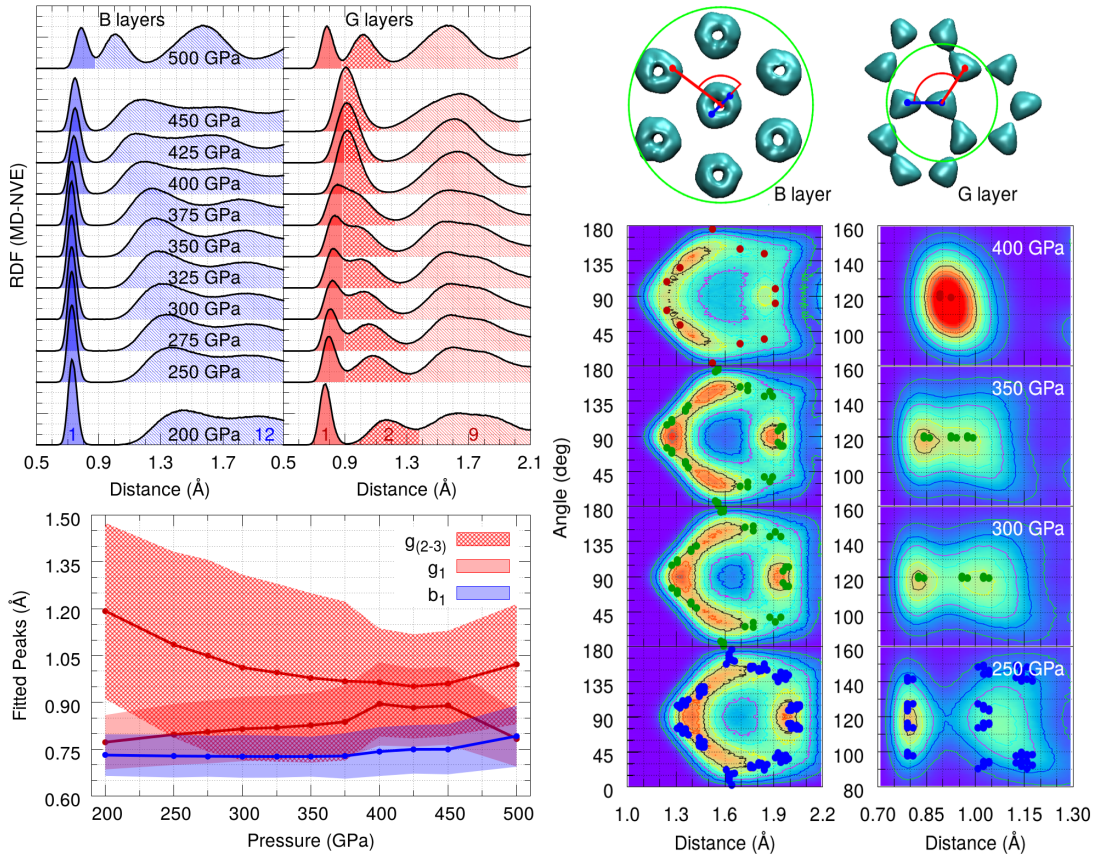

Figure 3.19 *RDF (left) and ADF_2 (right) in the B and G layers, respectively. The analysis was performed on the NVE simulations at 300 K, as explained in detail in the methods section 3.3.2. The RDF figure reports the integrals (value on the bottom curve) under the distributions, as shaded regions at specific lengths. Gaussians were fitted on the RDF peaks: one gaussian in the B layer and two in the G layer, corresponding to distances b_1, g_1 and $g_{(2-3)}$ from the static calculations (figure 3.15). The results of the fits are shown on the bottom panel with the shaded area corresponding to the full width at half maximum. In the ADF figure, the continuous color maps on the eight bottom panels represent ADF_2 extracted from MD, while the dots are ADF_2 calculated for the 0 K relaxed structures at corresponding pressures: Pc (blue), Pbcn (green) and Ibam (red). The top panels in this figure illustrate the short distance scales as iso-surfaces of the proton density (PDS) at 350 GPa. The two length scales shown in red and blue correspond to the reference distances in the ADF_2: r and q.*

(i.e. g_1 from the static calculations). This peak varies strongly with pressure and relates to the soft Raman vibron ν_1. The second important peak in the RDF integrates to two and it is a combination between the distance scales g_2 and g_3, which are no longer distinguishable at room temperatures. This implies that in

MD the G layer no longer supports three hexagonal motives but at most two in every layer: g_1-g_2 and g_2-g_2. At 400 GPa, the two peaks merge into one when the molecule in the G layer dissociates completely and the layer becomes atomic. The ν_1 vibron disappears at this pressure.

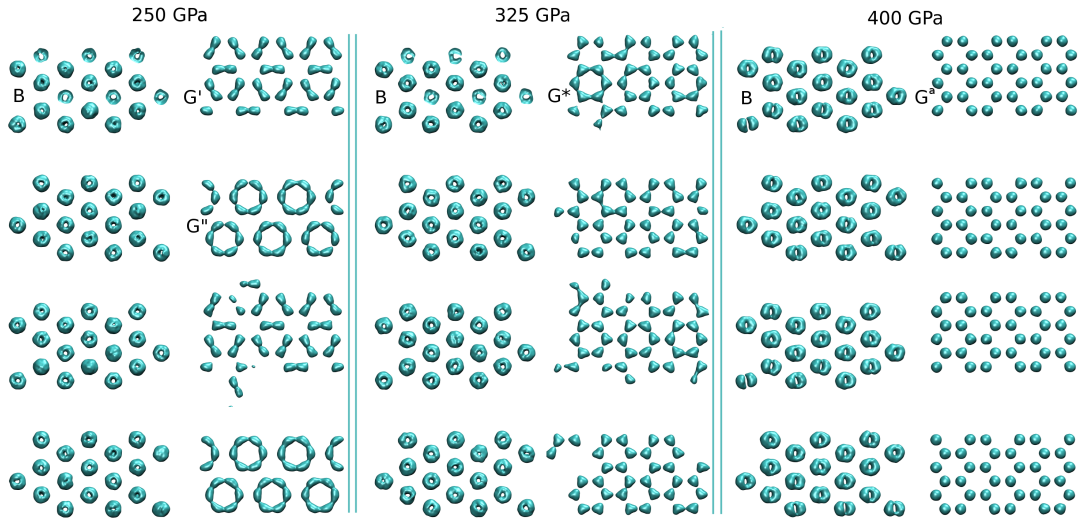

Figure 3.20 *In this figure I show the representative proton density iso-surfaces (PDS) extracted from 5 ps of NVE simulations at 300 K and different pressures as labeled on the top. Each simulation box contained 8 layers which are shown separately: B on the left and G on the right. Four different kinds of G layers gradually emerge upon increasing pressure as illustrated (see labels).*

The ADF_2 in the G layer shows the distribution of the angles around an atom (see PDS in figure 3.19). At 250 GPa, the molecule is well defined, but half of the trimers rotate (G" layers) and half are static (G' layer), shown by the double peak distribution in the ADF_2 at around 1.1 Å. So far this is consistent with the previous MD simulations in phase IV, section 3.2. At higher pressures, ADF_2 becomes more symmetric and the distinction between the G', G" layers disappears. By 350 GPa, the distinction between the intra-molecular and inter-molecular scales is greatly reduced and at 400 GPa the layer becomes atomic.

These observations are better substantiated by the proton density surfaces (PDS) shown in figure 3.20. Not shown here is the MD at 200 GPa which has four similar G layers, with distorted trimers, similar to the ones in *Pc*. By 250 GPa (figure 3.20), half of the trimers start to rotate and the other half become more symmetric. This is now phase IV, consistent with the study presented in the first half of the chapter. With slow rebonding in the G" layers and no rebonding in G', this model does not support the diffusion (see figure 3.21) reported in ref [118].

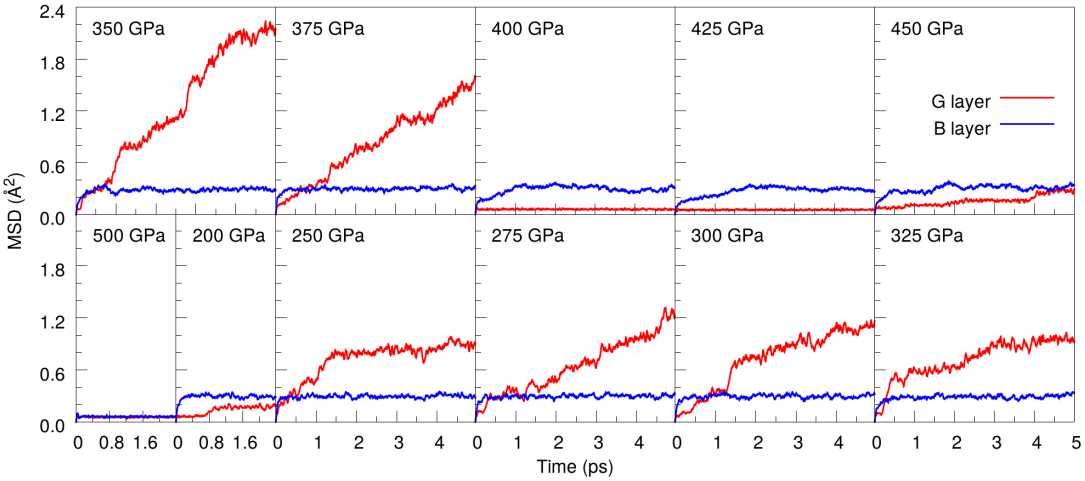

Figure 3.21 *Here I show the mean square displacement (MSD) calculated separately for the B layers (blue) and G layers (red) from the 5 ps NVE simulations. Each panel reports the result at a certain pressure as labeled. Notice that the trajectories at 200 GPa and 500 GPa are shorter and only span half the time: 2.5 ps. Compare to figure 3.5 from previous section 3.2.3.*

At higher pressure, 325 GPa, the G layers are all the same and there is no more G', G" distinction. Although RDF indicates the bond is still well defined, on average in G* the hexagons look all symmetric. The trimers rotate in both layers and the molecules are breaking and rebonding frequently. At the short time scale, there is a distinction between the two types of hexagonal patterns in the G* layer (i.e. g_1-g_2 and g_2-g_2) but through layer reconstruction, these distinctions vanish and the layer looks symmetric on average. This is important because instantaneously the layer contains molecules which give a ν_1 Raman vibron, but at long time averages the layer looks honeycomb symmetric and the low frequency modes arising from the G layer are suppressed. These changes are consistent with phase V discovered experimentally [100]. As seen in figure 3.21, above 325 GPa, this model supports the two step diffusion through trimer rotation and layer reconstruction and agrees with the studies presented in refs [97]. Finally, at 400 GPa, the G layers become atomic and they no longer support trimer rotation or diffusion.

The result of the Raman vibrons obtained through the projection method is reported in figures 3.16 and 3.22. The harder vibron ν_2 is almost constant with respect to pressure and has a width of about 150 cm^{-1}. The softer vibron ν_1 varies strongly with pressure in BG'BG" but then it becomes pressure independent when the G layer becomes symmetric in BG*. The width of ν_1 increases with pressure as the breaking and rebonding of molecules intensifies.

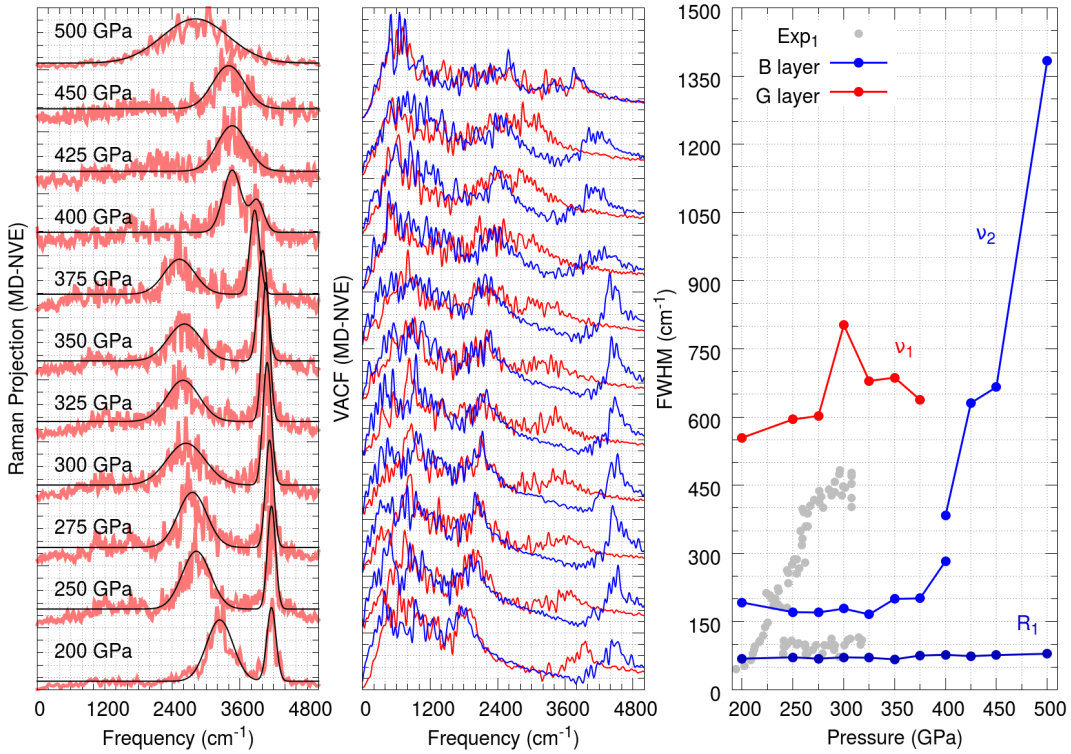

Figure 3.22 *The two panels on the left show the Raman signal extracted by projection from the 5 ps NVE and the Fourier Transform of the velocity auto-correlation function (VACF), respectively. VACF was computed separately for the B layers (blue) and for the G layers (red). I extracted the positions of the Raman peaks and their widths at different pressure by fitting a sum of two gaussians. The positions of the peaks are reported in figure 3.16, while the full widths at half maximum (FWHM= $2\sigma\sqrt{2\ln 2}$) are shown here on the right most panel. Labeled R_1 is the FWHM of the free rotor approximation, calculated from RDF as explained in methods section 3.3.2. In gray is the experimental FWHM of the vibrons, digitized from ref [89] using an online tool [124]. Interestingly, the difference between theory and experiment can be corrected by a factor of $\sqrt{2}$.*

To summarize, here I find five kinds of G layers in MD, first three are consistent with previous simulations presented in section 3.2.3 while the last two are new:

- G: at 200 GPa, distorted molecular trimers similar to those found in static *Pc*, not rebonding, not rotating, not diffusing
- G': at 250 GPa, regular molecular trimers similar to those found in static *Pbcn*, not rebonding, not rotating, not diffusing
- G": at 250 GPa, molecular trimers, slowly rebonding, rotating, slowly diffusing

- G*: at 325 GPa, molecular trimers at the short time scale, regular hexagons at the long time scale, rebonding, rotating, diffusing
- G^a: at 400 GPa, regular hexagons similar to those found in static *Ibam*, atomic, not rotating, not diffusing

The model BG'BG" is consistent with phase IV as discussed in detail in the first half of the chapter. The new model BG* is a good candidate for phase V: the softer Raman vibron becomes pressure independent and the low frequency modes originating from the G layer disappear as the layer becomes more symmetric at long time scales.

3.3.4 Conclusion

Motivated by the recent experimental claim of a new solid hydrogen phase (phase V), I extended my previous analysis from phase III and IV to higher pressures using a series of ground state calculations and molecular dynamics simulations in the pressure range 200-500 GPa, starting from the 0 K candidates *Pc*, *Pbcn* and *Ibam*.

With increasing pressure, the more symmetric structures become more favorable in static calculations. In particular, while the B layer remains largely unchanged, the G layer transforms gradually from distorted molecular trimers to regular trimers and finally atomic hexagons. In molecular dynamics, I found a few additional more subtle transformation owing to dynamical effects.

One the one hand, in MD, B layers are almost independent of pressure. While the static structures have well defined symmetry of the B layers, in MD molecules are spinning around almost freely and on average the B layers are hexagonal at all pressures. This is consistent with the previous findings discussed in the first half of the chapter and refs [1, 3, 97, 98]. Instantaneously, however, there is a weak correlation between the neighbor molecules which prefer to remain perpendicular to one another at low pressures and at angles of 45° at higher pressures.

On the other hand, the G layers experience changes on two different time scales. At short time scales the molecules elongate with increasing pressure and then start breaking and rebonding in phase V. As the molecules become longer, the associated vibron frequency decreases sharply and when rebonding is important the vibron becomes pressure independent, while the FWHM increases.

At the longer time scale, lower pressures, the G layer is of two types: G' symmetric and static and G" comprising of rotating hexagons. This behavior agrees with my previous result and it is consistent with phase IV. Higher in pressure, all G layers become the same, they can rotate, reconstruct and therefore diffuse [97]. Although, instantaneously the atoms are grouped into molecules, on average this G* layer appears symmetric. This means that while the vibron persists, the low frequency modes originating in this layer are likely to disappear.

At yet higher pressures, the G layer becomes atomic G^a, where both the vibron and the low frequency modes are lost. The BG^a structure is similar to $Ibam$ with the distinction that the B molecules are still rotating at room temperature. It turns out that this symmetry of type $P6/mmm$ is the best close packing that can be achieved with atoms and pure rotors. I will discuss this in more detail in chapter 4.

It is clear that the experiment has not reached the BG^a structure, however the other structures BG'BG", BG^* are a good fit for the phases IV/IV' and V. The BG'BG" model explains the two vibrons (ν_1 and ν_2) and four low frequency modes (L_{1-4}), of which one vibron (ν_1) and two low frequency modes (L_2, L_3) are changing strongly with pressure due to the rapidly elongating molecule in the G layers. The BG^* explains why the softer modes (ν_1, L_2, L_3) which likely originate in the G layers, become pressure independent in phase V and then gradually disappear as the G layers become more symmetric.

I conclude that the experimental observations recorded in ref [100] are consistent with a gradual symmetrization of the G layer in the BG type structures. This analysis predicts that upon increasing pressure the soft vibron will eventually disappear upon entering a new phase consistent with the BG^a model.

3.4 Summary

In this chapter, I covered my work on phases III, IV and V of solid hydrogen. All our calculations were based on the simple PBE-DFT theory which might be lacking when it comes to precise energetics, but which allows us to perform long MD trajectories on large systems that are crucial for understanding the entropically stabilized phases.

The best candidate for phase III remains $C2/c$. This phase contains a stacking

of G layers, which are known to be anharmonic as a result of the intermolecular interactions. Here I presented a method for extracting Raman from MD which corrects for this anharmonicity.

Finally, phases IV and V are high temperature phases and best described by the $BG'BG$" and BG^* models rather than the ground-state symmetries Pc, $Pbcn$ and $Ibam$.

Chapter 4

Solid Hydrogen Phase Diagram: The Missing Pieces

4.1 Overview

The objective of this chapter is two-fold. On the one hand, I will explore the region of the phase diagram around 400 GPa close to molecular dissociation, beyond phase V. Recent experiments [100–102] managed to compress hydrogen at comparable pressures, so this region of the diagram is relevant for future experimental work. A series of theoretical studies [91, 103–105] investigated the possible 0 K structure candidates at these pressures and proposed a collection of molecular and atomic crystals. Here I will focus on understanding the finite temperature regime from 0 K up to the melting temperature, searching for any structures stabilized by entropy.

On the other hand, I will present a simple but powerful thermodynamic model that captures the essential physics of the phase diagram. Many recent theoretical studies focused on increasingly more advance methods to better describe all the stable phases of hydrogen and the precise P-T boundaries in the phase diagram. Initial work was performed within the PBE approximation of the DFT. Gradually, different studies improved the resolution of the calculations in incremental steps:

- inclusion of nuclear quantum effects within PBE-DFT, quasi-harmonic approximation

- calculation of free energy based on PBE-DFT, quasi-harmonic approximation
- gauging the effect of different exchange-correlation functionals: local LDA, general gradient PBE, semi-local BLYP
- inclusion of anharmonic effect and finite temperature through DFT MD
- inclusion of nuclear quantum effect in the dynamics using path integral molecular dynamics (PIMD)
- using exact methods to calculate exchange energy - quantum Monte Carlo
- generating dynamics with quantum Monte Carlo

This work is crucial for understanding hydrogen and growing our knowledge. However, here, in section 4.3 we will take a step back and attempt to see how it all fits into the bigger picture.

4.2 Charge Density Wave: Chains

4.2.1 Introduction

The first prediction of hydrogen metalization was based on the assumption that under pressure solid hydrogen would eventually become atomic BCC [19]. It was proposed that under sufficient compressing energy PV, the electrons would be squeezed out of the covalent bond and the solid would become both atomic and metallic. Recent studies have shown that even the molecular phases can metalized under sufficient pressure [91, 125]. Although the metallic liquid was confirmed in shock experiments [52, 53], thus far, the solid metallic state remains elusive in DAC experiments [100].

Even though all solid phases identified in experiments are still molecular, there is mounting evidence that the electrons are slowly pushed out the bond. For instance, the soft vibron in phase IV is rapidly decreasing in frequency upon increasing pressure. This brings evidence of the rapid weakening of the bond and migration of the electron density elsewhere. Surprisingly, even phase V [100], where the soft bond has reached a steady length, being close to dissociation, is still a semiconductor and the band gap is yet to close.

As the electrons are leaving the bond, the solid can remain a poor conductor in two scenarios. First, the freed electric charge could localize in real space and create pockets of electrons that act as ions and form an electride structure. This is predicted in the case of the alkali metals potassium, sodium and lithium [126–128]. Second, if the charge that escapes the bond is delocalized, then it can form a charge density standing wave to screen the ion-ion interaction. In this latter case, instead of forming a close packing, the solid would rearrange in a pattern that scatters the electron wave functions at twice the Fermi vector ($2k_F$). It is possible that the high pressure solid phases III and IV are organized as layers for precisely this reason.

Here, I will present a series of MD simulations at 400 GPa and various temperatures that indicate to a candidate structure for hydrogen in the form of long polymeric chains of atoms, that has been observed by others before [104, 129–132]. This structure, although metallic, has a pseudo-gap, opened by a possible charge density wave. This is evident from the XRD pattern of *Chains*, which has a strong peak at $2k_F$. On the other hand, *Chains* could be a spurious effect of sparse k-point sampling. I will compare and discuss both possibilities in this section. The results presented below are soon to be published in a conference proceeding [6].

4.2.2 Methods

Simulation Details

The goal here was to investigate possible candidate structures for solid hydrogen beyond phase V at pressures larger than 400 GPa. AIRSS has been used successfully in the literature to identify 0 K candidates for high pressure phases [78, 91, 103]. By employing MD, I searched for candidates, beyond those presented in chapter 3, that are possible dynamically stabilized and only exist at high temperatures.

In this section, I will present a series of MD and PIMD simulations performed with CASTEP at 400 GPa and various temperatures as reported in table 4.1. The simulation settings were almost identical with those chosen in the previous section 3.3.2 on phase V: electron density expanded in plane waves with energy cut-off 1200 eV, norm conserving pseudo-potential generated on the fly by the string

Table 4.1 *The table here summarizes the calculations I performed in this section: O_i are a series of short MD runs at constant temperature and pressure starting with various cells [78, 103–105]. In A the simulations start at 0 K and are slowly heated up until melting. Blocks C and D show simulations of phase I, IV and beyond candidates at room temperature, while block B shows simulations at temperatures other than 300 K. I compared the new Chains structure against $Cmca-4$ in a series of additional simulations (B and D). The values for pressure and temperature are expressed in GPa and K, respectively.*

Id	Start	Atoms	Length	Simulation	k-grid	Pres.	Temp.	Final
O1	$I4_1/amd\,[c>a]$	128	2.75ps	MD-PBE	2x2x2	280	300	*Chains*
O2	$I4_1/amd\,[c>a]$	128	2.75ps	MD-PBE	2x2x2	350	300	*Chains*
O4	$I4_1/amd\,[c>a]$	128	2.75ps	MD-PBE	2x2x2	500	300	*undeterm*
O5	$I4_1/amd\,[c<a]$	192	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O6	$P6_3/mmc$	150	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O7	$R3m$	192	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O8	$Immm1$	180	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O9	$Immm2$	180	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O10	$Pmmn$	216	1.5ps	MD-PBE	2x2x2	400	300	*undeterm*
O11	$I4/mmm$	200	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O12	$Cmca-12$	192	1.5ps	MD-PBE	2x2x2	400	300	$Cmca-12$
O13	Pc	192	1.5ps	MD-PBE	2x2x2	400	300	*undeterm*
O14	$mC24$	144	1.5ps	MD-PBE	2x2x2	400	300	$mC24$
O15	$oC12$	192	1.5ps	MD-PBE	2x2x2	400	300	*undeterm*
O16	$Pnma$	192	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
O17	$CI16$	128	1.5ps	MD-PBE	2x2x2	400	300	*Chains*
A1	$I4_1/amd\,[c>a]$	128	3.5ps	MD-PBE	2x2x2	400	$50\rightarrow700$	*liquid*
A2	$I4_1/amd\,[c>a]$	128	4.5ps	MD-LDA	2x2x2	400	$50\rightarrow700$	*liquid*
A3	$Cmca-4$	128	3.5ps	MD-PBE	2x2x2	400	$50\rightarrow700$	*liquid*
A4	$Cmca-4$	128	4.5ps	MD-LDA	2x2x2	400	$50\rightarrow700$	*liquid*
A5	$I4_1/amd\,[c>a]$	128	1.5ps	MD-PBE	4x4x4	400	$300\rightarrow650$	*liquid*
A6	*Chains*	128	1.7ps	MD-PBE	4x4x4	400	$300\rightarrow650$	*liquid*
A7	*Chains*	128	1.2ps	MD-LDA	4x4x4	400	$300\rightarrow650$	*undeterm*
B1	*Chains*	128	3.0ps	MD-PBE	2x2x2	400	400	*Chains*
B2	*Chains*	128	2.5ps	PIMD-PBE	2x2x2	400	200	*Chains*
B3	$I4_1amd\,[c>a]$	128	2.5ps	MD-PBE	2x2x2	400	500	*liquid*
B4	$Cmca-4$	128	2.5ps	MD-PBE	2x2x2	400	500	$Cmca-4$
C1	*mol. hcp*	288	3.75ps	MD-PBE	2x2x2	50	300	*mol. hcp*
C2	Pc	288	1.5ps	MD-PBE	2x2x2	275	300	Pc
D1	*Chains*	128	5.0ps	MD-PBE	2x2x2	400	300	*Chains*
D2	*Chains*	128	2.5ps	PIMD-PBE	2x2x2	400	300	*Chains*
D3	$Cmca-4$	128	1.5ps	MD-PBE	2x2x2	400	300	$Cmca-4$
D4	$Cmca-4$	128	1.7ps	PIMD-PBE	2x2x2	400	300	$Cmca-4$

"H 0|0.7|2|6|8|10L(qc = 10)" [116] and exchange-correlation functional LDA or PBE as reported in table 4.1. The energy in the self-consistent loop was converged within 10^{-11} eV, the electron minimization was performed with the default density mixing scheme and the spin density was unpolarized. The structures were treated as metals by setting fix occupancy to false and using 100% extra bands. Almost all simulations were carried out in the NPT ensemble, using the Parrinello-Raman barostat and the Nose-Hoover thermostat. The time step was set to 0.5 fs.

The initial round of simulations labeled O_i in table 4.1, were initialized with various known crystal candidates and run at constant pressure and temperature.

As expected, most simulations were unstable at those conditions, but surprisingly they transformed to the same kind of structure which we called *Chains* [6] and which I will discuss in detail in the results section 4.2.2.

Two competitive 0 K candidates at 400 GPa are the molecular $Cmca-4$ [78] and the atomic $I4_1/amd$ ($c > a$, also structure of $Cs-IV$) [72, 103, 104]. In the next round of simulations from table 4.1, labeled A_i, these structures were slowly heated up by increasing the temperature 30 K every 0.15 ps (300 iterations) until melting occurred. Upon heating, some of the simulations resulted in *Chains*. In order to study the validity of the new result, I repeated the heating simulations with two different functionals (PBE and LDA) and with different k-point densities. For the simulations with a denser k-point grid (A_5, A_6 and A_7), I reduced the self consistent energy threshold to 10^{-8} eV for speed up.

Simulations D_1 to D_4 further extend the study of *Chains* and $Cmca-4$ at 300 K, while simulations B_1 to B_4 extend the study to other temperatures. To investigate the contribution of ZPE, I used path integral molecular dynamics (PIMD) with 16 beads, where the temperature was kept constant with a Langevin thermostat.

Finally, simulations C_1 and C_2, at smaller pressures, were run in the NVE ensemble, with fix occupancy (i.e. treated as insulators) and with a self consistency energy tolerance of 10^{-9} eV.

Trajectory Analysis

I analyzed the trajectories employing a series of methods some of which were introduced in the previous chapter. In section 3.3.2 I used a radial distribution function (RDF) reduced to 2D because it was suitable to study layered structures. Here I use the standard RDF, instead:

$$RDF(r_i) = \frac{1}{t_{max}} \sum_{t=1}^{t_{max}} \frac{V_t}{N} \frac{n_t(r_i)}{4\pi r_i^2 dr} \tag{4.1}$$

where r_i is the space discretization, t labels the times steps running from 1 to maximum time t_{max}. V_t is the volume at time t, N the number of atoms and dr the bin width.

To gain insight into the molecular orientation, I used the angular distribution function (ADF_1) as defined in section 3.3.2. Additionally, here, I also normalized

the ADF_1 with $\propto 1/\sin\theta$ to account for the probability distribution in spherical space.

The *Chains* structure can diffuse, particularly in the direction of the chains. Now, identifying this direction is not trivial because it can be different for each simulation instance. I devised an algorithm that computes the direction of the chains by averaging the orientation of the shortest distances in the system. In order to study the diffusion mechanism I used the standard mean square displacement (MSD_{3D}) but also a 1D displacement MSD_{1D} and a 2D displacement MSD_{2D} which made it possible to assess the diffusion along the chains and perpendicular to the chains. The displacements were also normalized according to dimensionality:

$$MSD_{3D}(t) = \frac{1}{3N}\sum_{l=1}^{N}\sum_{\alpha=1}^{3}\left[x_l^\alpha(t) - x_l^\alpha(0)\right]^2 \tag{4.2}$$

where l running from 1 to N labels the atoms, α the Cartesian components and $x_l^\alpha(t)$ the atomic position at time t. On the other hand, MSD_{1D} was obtained by projecting the displacements along the direction of the chains p^α:

$$MSD_{1D}(t) = \frac{1}{N}\sum_{l=1}^{N}\sum_{\alpha=1}^{3}\left[p^\alpha(x_l^\alpha(t) - x_l^\alpha(0))\right]^2 \tag{4.3}$$

and finally the 2D MSD as the difference between the two:

$$MSD_{2D}(t) = \frac{1}{2}\left[3MSD_{3D}(t) - MSD_{1D}(t)\right] \tag{4.4}$$

Electronic Structure

Chains appears to be a good candidate for a charge density wave. To investigate this possibility, I calculated the density of electronic states (DoS) and the XRD patterns from the MD trajectories.

For *mol. hcp*, *Pc* (BG layers) and *Chains*, which are only stable at finite temperature, the band structure calculations were performed at selected snapshots extracted from the MD runs as detailed in table 4.2. In each of these calculations, I increased the MP k-point grid to 8x8x8 ensuring better convergence. For

Table 4.2 *Here are the simulations I used for sampling the band structure and XRD patterns. Runs D1 and D3 were also used for computing the PDoS and project the phonons.*

Id	Start	Phase	First Sample	Last Sample	Step
C1	*mol. hcp*	I	0.25ps	3.75ps	0.175ps
C2	*Pc*	IV	0.25ps	1.5ps	0.0625ps
D1	*Chains*	beyond V	0.5ps	5.5ps	0.25ps
D2	*Chains*	beyond V	0.25ps	2.50ps	0.1125ps

comparison, I also computed the bands for the 0 K structures $Cmca-4$ and $I4_1/amd$ using their unit cells (4 and 2 atoms, respectively), and k-point meshes up to 21x21x21.

The actual Density of States (DoS) was obtained using a post-processing code offered with the CASTEP package: band2dos, which does a simple gaussian interpolation of the band structure with 0.2 eV smearing. In the results section, I will show the DoS for each snapshot shifted to the origin and the average Fermi energy which was obtained by integrating the DoS. For PIMD, I averaged the centroid positions over the beads and calculated the DoS at each snapshot for the centroid structure.

The ideal free electron gas curves and ideal Fermi Energy were calculated for comparison, using the standard formulae:

$$N(E) = \frac{V}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2}\sqrt{E} \tag{4.5}$$

$$E_F = \frac{\hbar^2}{2m_e}\left(\frac{3\pi^2 N_e}{V}\right)^{2/3} \tag{4.6}$$

where m_e is the mass of the electron, N_e is the number of electrons in each cell and V is the volume averaged over the snapshots.

XRD calculations were performed in a similar way at selected snapshots as shown in table 4.2. For each calculation I used GDIS [133] to compute the XRD pattern at a wavelength of $\lambda = 0.7$ Å, and then averaged the spectra to obtain the final result. The standard spectrum is a function of the 2θ scattering angle, but here the angle was expressed as an energy by the following argument. In Bragg's diffraction law, the scattering angle is related to λ and a family of planes at

distance d:

$$d = \frac{\lambda}{2\sin(\theta)} \tag{4.7}$$

Assuming the structures can host charge density waves, then the family of atomic planes spaced by distance d can scatter an electron wavelength q of size:

$$q = \frac{1}{2}\frac{2\pi}{d} = \frac{2\pi\sin(\theta)}{\lambda} \tag{4.8}$$

where the additional factor 1/2 relates the periodicity of electron wave function ($\propto \cos(qx)$) to the periodicity of the probability density ($\propto \cos^2(qx) \propto \cos(2qx)$). Finally, using the energy dispersion relation of the free electron gas and combining with equation 4.8:

$$E = \frac{\hbar^2 q^2}{2m_e} = \frac{h^2\sin^2(\theta)}{2m_e\lambda^2} \tag{4.9}$$

This allows us to express the diffraction pattern averaged over all the snapshots as a function of energy $XRD(E(\theta))$ and therefore compare it with the DoS.

After an electronic structure calculation, CASTEP can print the electron charge density. In the results section, I will show such a density field for an instance of the *Chains* structure. For the same sample, I also calculated the electron localization function (ELF) using the code VASP [134].

Finite Temperature Phonons

Finally, I investigated the vibrational properties of *Chains* and the other structures by calculating the phonon density of states PDoS from the Fourier Transform of the velocity auto-correlation function (VACF).

Since most of these structures are stabilized by entropy, extracting the 0 K Raman spectra is difficult. Here, I used the projection method defined in the previous chapter, section 3.2.2, to extract the frequency of the vibrons and the most symmetric lattice modes from the MD trajectories. For *Chains* the most symmetric lattice modes involve plains of chains oscillating along the three Cartesian directions.

4.2.3 Results

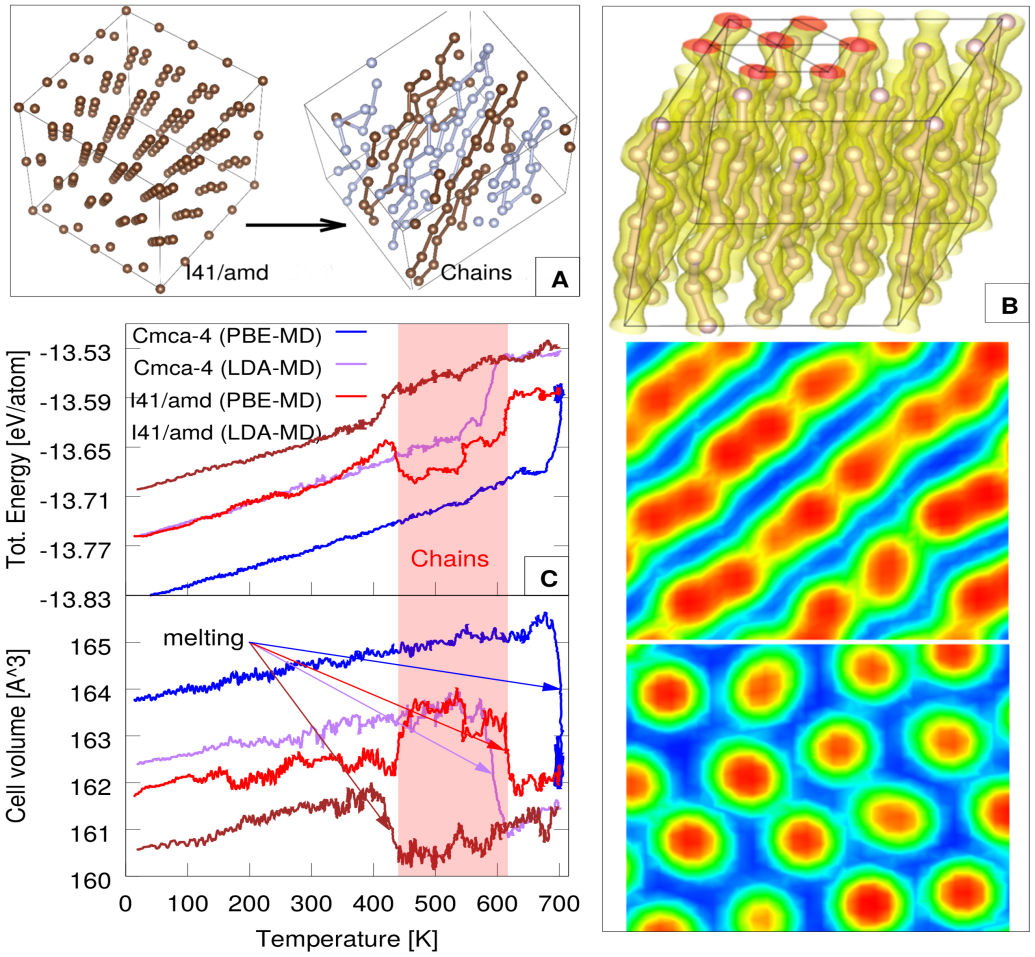

Figure 4.1 *In panel A, I show the transformation of $I4_1/amd$ into 2 distinct chains wrapped around the periodic boundaries. Panel B illustrates a snapshot from MD containing chains: associated charge density isosurface (top) and the electron localization function (ELF) [135] along and across the chains (bottom). Panel C reports the cell volume and total energy as a function of temperature upon heating at 400 GPa (simulations A_1-A_4 in table 4.1). Both quantities are moving averages over a window of 0.1 ps. Shaded in red is the region where $I4_1/amd$ transformed to Chains.*

One of the best candidates based on atomic hydrogen beyond phase V at high pressures is the distorted diamond structure $I4_1/amd$ [72, 103, 104], whereas the best candidate based on molecular hydrogen is $Cmca-4$ [78]. In this section, I performed a range of MD simulations at 400 GPa with these and other candidate structures. Surprisingly, upon increasing temperature, most of the atomic candidates including $I4_1/amd$ were unstable and transformed to the same structure comprising of long polymeric chains wrapped around the simulation box

as shown in figure 4.1.

Interestingly, every chain can slide up and down almost freely, with little interaction with neighboring chains. At the same time, within the chains, atoms can easily switch positions and diffuse along the tube of charge. This is clear from the magnitude of MSD_{1D} shown in figure 4.2. The diffusion of atoms from one chain to the other is smaller (compare MSD_{2D} to MSD_{1D}) implying the chains are well defined. The structure appears to be almost melted along the direction of the polymers, while clearly solid forming a hexagonal pattern in the transversal plane.

On closer inspection, the long chains appear to be parallel to one another, organized in a hexagonal pattern when viewed transversally. The electron charge density and the electron localization function (ELF) show that the molecular bond is no longer well defined. However, most of the charge remains near the chains, forming long uninterrupted tubes.

In figure 4.1 I compare two MD-PBE simulations started in $I4_1/amd$ and $Cmca-4$, respectively. The two runs were slowly heated up until melting occurred. At around 430 K, $I4_1/amd$ transformed to $Chains$ which has a lower total energy and a larger volume than the initial structure. Compared to $Cmca-4$, $Chains$ has higher energy at the same pressure but a smaller volume. $Chains$ has also likely larger entropy as a result of its diffusive nature.

At 400 GPa, $Cmca-4$ appears to be more stable than $Chains$ since it melts at a higher temperature, however, this could also be a result of hysteresis. Even if $Cmca-4$ is more stable, $Chains$ cannot transform into $Cmca-4$ spontaneously because it would require all atoms in the chains to simultaneously pair up to form molecules. Such a transition is very unlikely to take place in a short MD simulation. The smaller volume and larger entropy could stabilize the $Chains$ over $Cmca-4$ at higher pressures and temperatures, where $Chains$ could be the stable structure at high temperatures between the 0 K $I4_1/amd$ structure and the liquid phase.

I repeated the heating simulation using the local density approximation LDA. With the latter exchange-correlation functional, the transition to $Chains$ does not occur and both $I4_1/amd$ and $Cmca-4$ remain stable until they melt. LDA shifts down the melting temperatures by 100 K for $Cmca-4$ and 200 K for $I4_1/amd$. Consistently with PBE, $Cmca-4$ melts at higher temperatures than $I4_1/amd$ at the same pressure - 400 GPa. Although generally LDA is less reliable

than PBE, this result suggests *Chains* could be a spurious structure, albeit one worth studying because of its interesting physics.

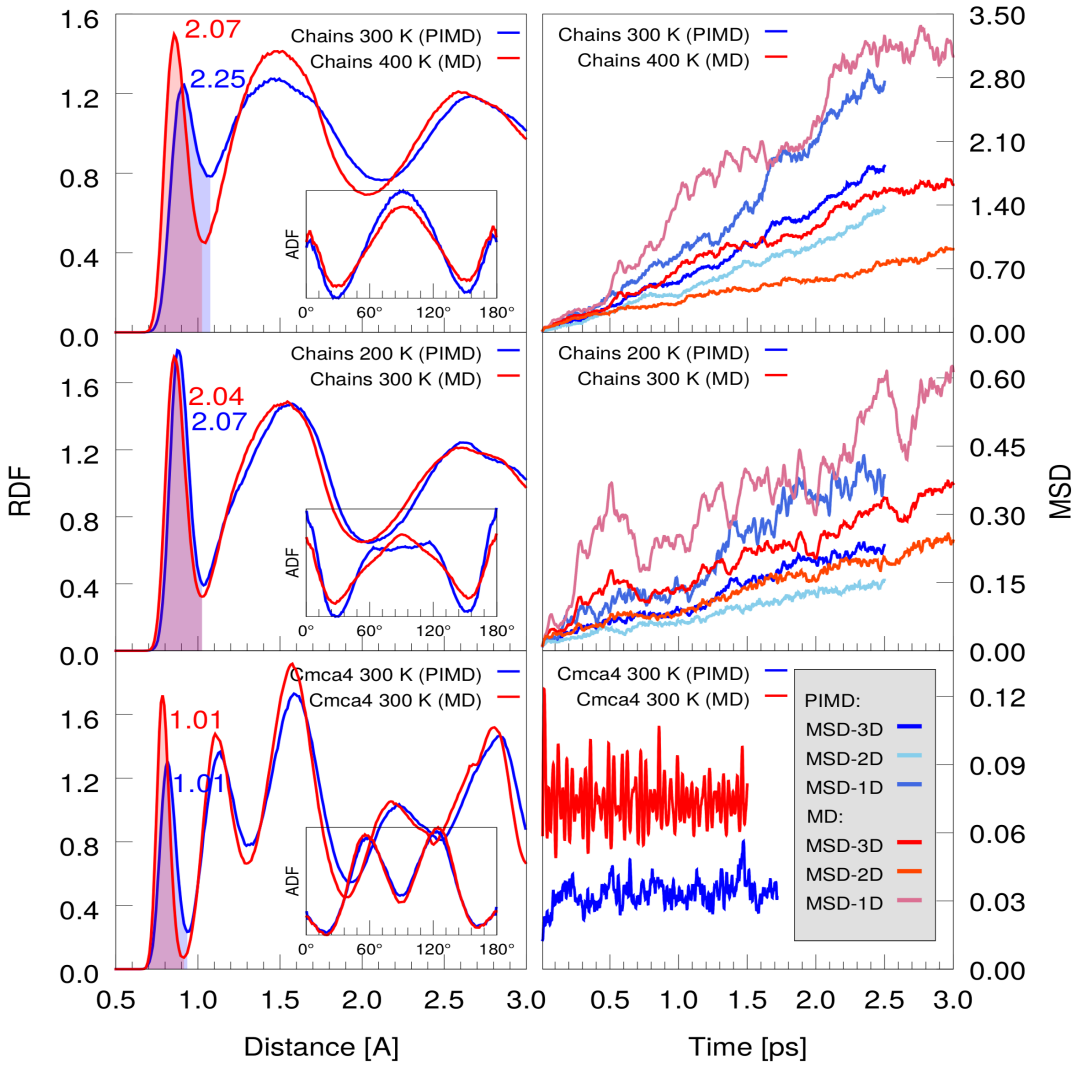

Figure 4.2 *The left panels show the RDF and ADF (insets) for Chains and Cmca − 4 calculated from both MD and PIMD at 400 GPa and temperatures as labeled (simulations B_1, B_2, D_1-D_4). Shaded is the integrated area under the first peak of the RDF which gives the coordination number. The right panels depict the corresponding MSD calculated as total displacement (3D), displacement along chains (1D) and between chains (2D) for the same simulations; normalized according to dimensionality.*

In figure 4.2, I also compare the RDF and ADF for *Chains* and *Cmca* − 4 in both MD and PIMD. Both classically and quantum mechanically, *Cmca* − 4 forms well defined molecules as indicated by the clear peak in the RDF, which integrates to one. As we have seen before in chapter 3, these molecules are weaker than the B layer molecules in the BG structures, but stronger than the G layer molecules.

On the other hand, the *Chains* structure generates a clear peak which integrates to two. This implies that: 1) the chains are well defined and remain in tubular shapes throughout the course of the simulations and 2) atoms do not pair up into molecules and the electron has been pushed out of the bond. The ADF further illustrates the difference between *Chains* and $Cmca-4$. In *Chains* the short-lived pseudo-molecules point towards one another, while in $Cmca-4$ the molecules point at angles of 60° to one another and the orientation remains fixed during the simulation. Including quantum effects in PIMD has a similar effect as increasing the temperature by about 100 K in *Chains*.

The MSD shows that *Chains* can diffuse and supports the idea of a 1D melt. This freedom gives the structure high entropy and makes it more stable than $I4_1/amd$ at high temperature. On the other hand, $Cmca-4$ shows no sign of diffusion.

More interesting than the dynamics of *Chains* is its electronic structure. In figure 4.3 I show the electronic density of states DoS and the XRD pattern computed from both MD and PIMD as explained in the methods section. Surprisingly, the DoS looks similar to that of a free electron gas, except for a small pseudo-gap near the Fermi energy.

The XRD gives one pronounced peak near twice the Fermi vector ($2k_F$), corresponding to the spacing between the different chains in the transversal plane. The position of the peak on the energy axis is such that below it, the calculated DoS is larger than that of the ideal gas and above it, smaller. This mechanism creates extra quantum states at lower energies, allowing the electrons to rearrange and decreases the overall energy of the structure. As a result, a pseudo-gap appears near the Fermi vector. These observations are consistent with a charge density wave created by the special periodicity of the chains.

Surprisingly, *Chains* is not the only solid hydrogen structure with this pattern. The previous candidate for phases IV and V (the BG layers), the 0 K structures $I4_1/amd$ and $Cmca-4$ show similar behavior. In the case of the layered structures, the distance between layers is the defining scale that scatters the charge density near $2k_F$. This could be the reason why most of the candidate structure at these high pressures appear to be organized in layers: $C2/c$, Pc, $Pbcn$, $Ibam$ etc.

Interestingly, even in *mol. hcp*, the DoS resembles the free electron gas at low electronic energies. This is surprising, because at 50 GPa, *mol. hcp* is expected to

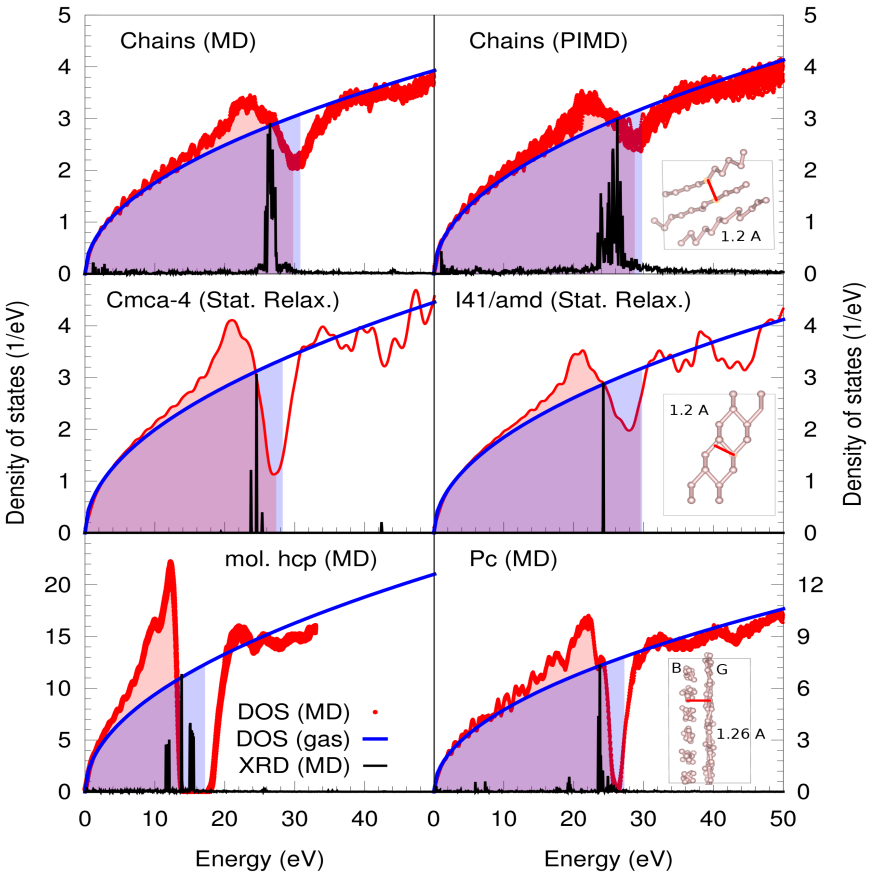

Figure 4.3 *In this figure I show in red the density of electronic states (DoS) from multiple snapshots in the MD/PIMD at 300 K and from static relaxations (middle panels) for comparison. The pressure is different for each phase candidate: mol. hcp (50 GPa), Pc (275 GPa), $Cmca-4$, $I4_1-amd$ and Chains (all at 400 GPa). Blue lines are the analytic DoS for a free electron gas at the same density. Shaded regions depict the occupied states for both calculated and ideal DoS in red and blue, respectively. In solid black I show the simulated XRD powder pattern (calculated with GDIS [133]) averaged over the MD/PIMD snapshots. Only the relative intensities are meaningful, and the peak positions are plotted in units of energy as explained in the methods section 4.2.2. The insets show the length scale that scatters the most intense XRD peaks.*

be a molecular crystal with a well defined energy of the covalent bond, whereas here the electrons have an energy spread of up to 15 eV. This could be an effect of the Kohn-Sham approximation, where electrons are often treated with plane waves which are more appropriate to describe a free electron gas rather than a molecular crystal, where all the electrons are localized on the bonds.

If *Chains* do form in the experiment, they will be hard to identify. Since the molecule ceases to exist and so does the Raman vibron. In figure 4.4 I show the

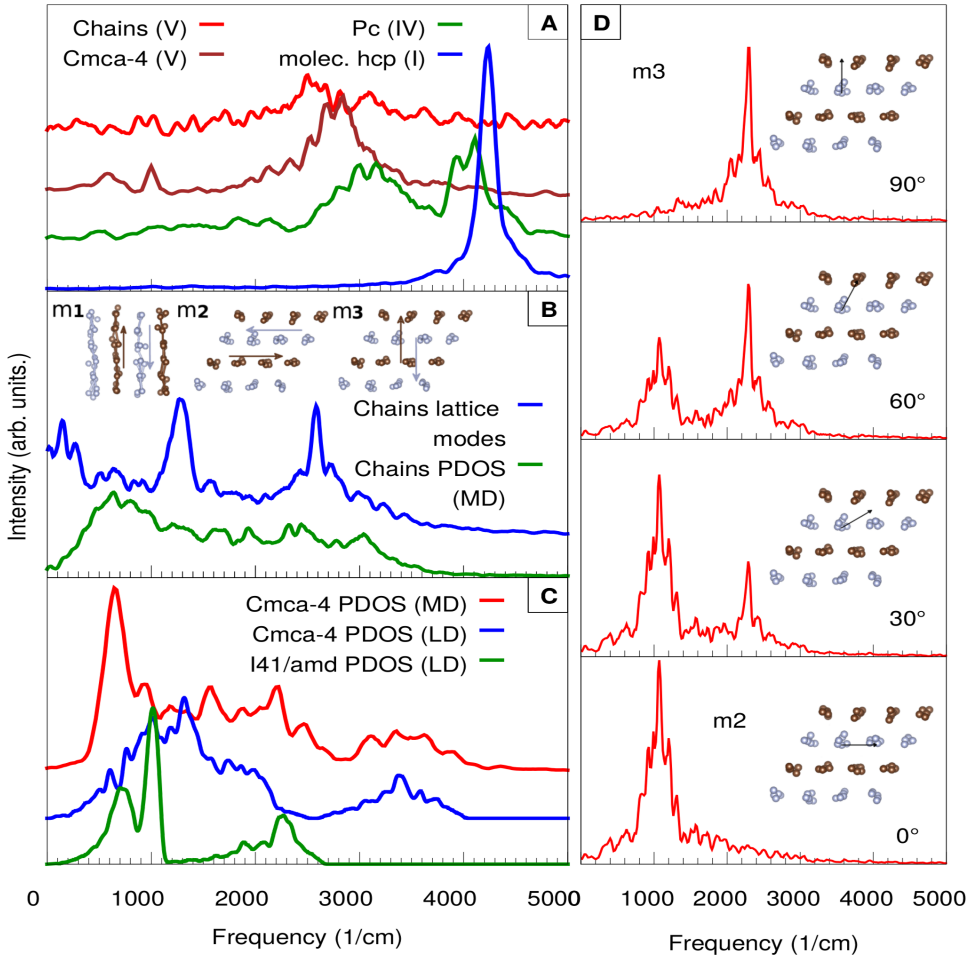

Figure 4.4 *Here, panel A illustrates the Raman vibron computed at 300 K from the MD as a Fourier Transform of the molecular projection for the candidate structures of phases I, IV and beyond. Panel B shows the phonon density of states (PDoS) extracted from VACF together with projections onto the most symmetric lattice modes generated by Chains. The eigenvector patterns are depicted in the insets. Panel C shows PDoS calculated from MD (300 K) and LD (0 K) [5] for* $Cmca{-}4$ *and* $I4_1/amd$*. Panels D illustrate how the most symmetric modes emerge in the MD projection as the reference eigenvector is rotated from* $0°$ *to* $90°$ *.*

result of the molecular projection performed on the MD trajectories of different hydrogen structural candidates. For phase I, at 50 GPa, the *mol. hcp* generates one well defined vibron at around 4300 cm^{-1}. At 275 GPa, for phase IV, the BG layers give two vibrons from the two layers, respectively. At even higher pressure, 400 GPa, $Cmca-4$ gives one vibron again, albeit at a lower frequency, comparable to the softer vibron in BG. At these conditions, *Chains* does not have a well defined vibron, but a broad range of frequencies around 2500 cm^{-1}

which corresponds to some collective oscillation of the atoms within the chains.

The phonon density of states PDoS in *Chains* is different than both PDoS in $Cmca-4$ and $I4_1/amd$. In $Cmca-4$ there is a clear band arising from molecular vibrations and one arising from low frequency lattice modes. Whereas $I4_1/amd$ gives three well defined, low frequency peaks and no vibrons. In contrast, the PDoS of *Chains* is featureless and comprises of one wide band spanning 3000 cm^{-1}. However, by projecting the velocities onto a series of possible eigenvectors, I identified the three most symmetric modes that involve oscillations of planes of chains in the three Cartesian directions as shown in figure 4.4. Some of this low frequency modes could be Raman or infrared active and facilitate possible comparison to experiment.

Finally, in figure 4.5 I show one last convergence test. It is common practice to run MD simulations with less dense k-point grids, because it is generally believed that forces need not be perfectly converged to obtain well behaved trajectories. Here, I identified a possibly spurious effect resulting from sparse k-point sampling. The atoms appear to rearrange themselves in the MD to minimize the energy on the particular k-point grid at which the simulation is carried out. For instance in panel A of figure 4.5 I show this effect in the MD-PBE simulation with a 2x2x2 k-point grid, that started in $I4_1/amd$ and was gradually heated up (simulation A_1 from table 4.1). I chose two snapshots from the trajectory, one before the transition, containing $I4_1/amd$ and one after the transition containing *Chains*. For each sample, I recalculated the energy on different k-point grids. Notice how after the transition to *Chains*, the structure reorganized itself to minimize the energy on a 2x2x2 grid. In this case, even at convergence, *Chains* is still lower in energy than $I4_1/amd$, albeit the energy difference is much smaller.

Starting from the selected snapshots, I replicated the heating simulation on a denser grid (e.g. 4x4x4). The *Chains* eventually transformed to a BG type structure, which is expected at 400 GPa, as shown in the previous chapter. $I4_1/amd$ remained stable until melting at around 450 K. Furthermore, when $I4_1/amd$ was simulated at constant 400 GPa and constant 500 K (in contrast to being slowly heated up) the structure melted immediately. As shown in panel D from figure 4.1 the resulting MSD is specific to the liquid phase. All these observations imply that *Chains* might be a spurious effect.

Only experiment can eventually conclude whether *Chains* is a real candidate for a higher pressure phase or just a simulation spurious effect. Nevertheless, *Chains*

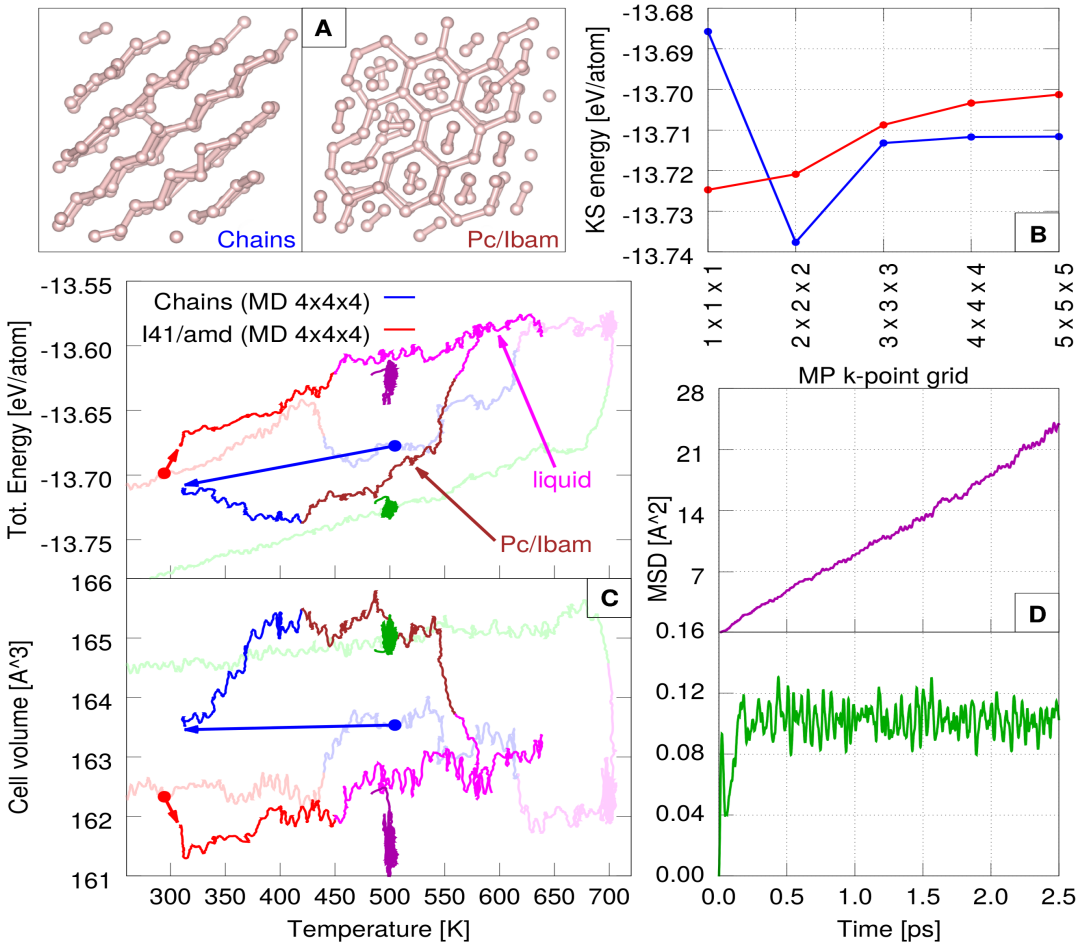

Figure 4.5 *Panel A illustrates the transitions from Chains to BG through snapshots from the simulations in Panel C. Panel B shows single point calculations (SP) illustrating the convergence of energy with k-point sampling for Chains and $I4_1/amd$. Panel C reports the results from a series of MD runs that replicate the heating simulation at 400 GPa with denser k-points (4x4x4) (simulations A_5 and A_6 from table 4.1). These calculations were started with snapshots from the original simulation (see figure 4.1, here faded lines). The colors code the different structures: $Cmca-4$ (green), Chains (blue), BG layers (brown), liquid (purple). Panel D shows the MSD extracted from the MD runs at 500 K, started in $I4_1/amd$ and $Cmca-4$, respectively (simulations B_3 and B_4).*

is an interesting structure which could be hosting a charge density wave. This study tells us a cautionary story that k-point convergence is actually important in MD, especially in treating metallic systems. If the charge density wave is a real phenomenon that stabilizes the various phases of hydrogen, it can only be studied well by sampling the reciprocal space at the location where the pseudo-gap is opened. However, this is not known a priori, and therefore we can only

choose a dense enough k-point grid hoping to sample this position in the reciprocal space. On the other hand the MD could reorganize itself to open the pseudo-gap at one of the available k-points instead, introducing spurious effects.

4.2.4 Conclusions

The experiments are moving closer to dissociation and metalization of solid hydrogen as evident by the strong softening of the molecular vibrons and darkening of the samples [80, 89, 100]. Structure searching studies have proposed a range of weak molecular and atomic candidates at high pressures [91]. As we have seen in the case of phase IV and V, the structures at those conditions are dynamic in nature, so perhaps molecular dynamics (MD) is more suitable to search for candidate structures.

In this section, I presented a series of MD simulations performed at 400 GPa and various temperatures. The calculations were initialized with a range of structures identified in 0 K studies [103–105]. Surprisingly the unstable atomic cells transformed to *Chains* [130, 131], a structure comprising of long tubes of charge hosting hydrogen atoms that are no longer bound in molecules. The structure exists at finite temperature, where it is stabilized by entropy. With large diffusion along the direction of the chains and smaller diffusion between chains, this candidate behaves like a 1D melt.

Upon inspecting the density of electronic states, it is clear that *Chains* is metallic. The calculated DoS resembles that of a free electron gas except for a pseudo-gap opened by a charge density wave. The crystal organizes itself as to allow the electrons to form standing waves extended throughout the structure. This effect is illustrated by a clustering of diffraction peaks near $2k_F$. Other solid hydrogen candidates, in particular, the layered ones present a similar pattern: the spacing of the layers creates a periodicity that could allow the formation of a charge density wave. This observation could explain why solid hydrogen continues to avoid metalization at high pressures.

Finally, more detailed calculations using denser k-point sampling and a different functional (e.g. LDA) for comparison, indicate that *Chains* could be a spurious effect. In particular, I found that structures which appear to be k-point converged at the start of the MD, can reorganize during the simulation in order to minimize the total energy on the chosen k-point grid. The subsequent structures (in this

case, *Chains*) appear to be no longer k-point converged. To my knowledge, this MD shortcoming has not been discussed in the literature before. Despite conventional wisdom, k-point convergence appears to be important in MD especially when modeling metallic systems.

Whether *Chains* can be realized in DACs or it is just a bogus effect will be decided by future experiments. As with all the other structural candidates, the experiment will be the final arbiter.

4.3 Phase Diagram

Having explored the various high pressure structures of solid hydrogen, in this section I will integrate all the essential observations and findings in the context of a simple model of the hydrogen phase diagram. This model is not a quantitative result, but qualitative, constructed around sensible assumptions derived from previous studies and the basic laws of thermodynamics. Despite its simplicity, the model gives a nice overview of hydrogen at high pressures from a simple and instructive perspective and summarizes the essential physics needed to explain the phase diagram. This study will be published in a conference proceeding [7].

The main idea here is that each phase can be assigned a simple free energy function that captures the important characteristics of the phase. In reality, these functions are very complicated and cannot be solved analytically, only calculated in DFT or quantum Monte Carlo (QMC) based simulations.

Before I explain the origin of each free energy function in this model, let us summarize all the known hydrogen phases and their characteristics. Phase I is the first solid phase, it is stable at finite temperature and it consists of molecular rotors that are hexagonally close packed [70]. Classically, the rotors can be viewed as spherical objects with high entropy. Phase II is a low temperature phase, where the rotors are arrested by quadrupole interactions [24] and the molecules assume fixed orientations. The exact structure of phase II is not known but some of the best candidates are $P6_3/m$ and $P2_1/c-24$ [72, 78]. At higher pressure above 150 GPa, phase III is another low temperature phase. As discussed previously the best candidate for this phase is the structure $C2/c$ [72, 78].

In chapter 5 on hydrogen-deuterium mixtures, I will show new evidence that $C2/c$ is a good candidate for phase III. $C2/c$ comprises of molecules which are

organized in distorted graphene layers (G). The molecules are elongated but the structure is better packed since it becomes stable at higher pressures.

At higher temperature, above the region of phase III, is the solid phase IV. Phase IV comprises of a stacking of atomic and molecular layers, where the molecular layers are freely rotating [1, 97, 98] and the atomic ones describe a honeycomb pattern. In chapter 3 I presented a detailed study that explores the subtle differences between phase IV and V and their structural candidates $Pc/Pbcn/Ibam$ at 0 K and $BG'BG"/BG^*/BG^a$ at finite temperature. These variations are negligible within this simplistic model and here phase IV is viewed as a packing of atoms and rotors (spherical objects), essentially a mixture between phases I and III. Although, to date, all experimentally discovered phases are molecular, at high enough pressures, hydrogen will probably become atomic eventually. Here I model this hypothetical atomic phase as a simple *hcp* packing of atoms, although, in reality, the atomic phase will probably adopt the $I4_1/amd$ symmetry.

The liquid phases are also complex and they undergo a transition from molecular to atomic [52, 53, 61, 63]. Recent simulations have shown that a mixed atomic and molecular fluid [56] could also exist at the transition between the former two.

The important observation is that all these phases, solid and liquid, comprise of the same underlying building blocks: free rotors that act as spherical objects (S), molecules that can be viewed as rods (R) and atoms that are basically smaller spheres (A) [136]. Table 4.3 summarizes the thermodynamic properties of these elements in our model. Both spheres and rods are essentially molecules and they were assigned similar internal energies ($U_S = U_R$). This energy is lower than that of individual atoms (U_A) since electrons gain energy when forming molecules. The volume assigned to spheres (V_S) is larger than the volume assigned to rods (V_R) which in turn is larger than that of two individual atoms (V_A). Finally, in contrast to rods and atoms, the spherical objects were given an entropy related to their orientational disorder. All the values chosen here are arbitrary and they only have qualitative meaning in relation to each other.

Starting from these assumptions, one can now define the free energies of each phase in the model. Phase I is a hexagonal close packing (*hcp*) of spherical

Table 4.3 *This table summarizes all the thermodynamic parameters used in the model. The columns report, in order, the parameters used for the basic building blocks (first three), each of the five solid phases (next five) and the liquid phase (last column). On the last two lines, C refers to packing efficiency and f to the change of packing efficiency upon melting to a liquid comprising of one of the building blocks (S,R or A). The zero point energy was introduced as a temperature offset* $T_{ZP} = 0.4$.

Quant.	S	R	A	1	2	3	4	5	liq.
U	-1.5	-1.5	0.0	0.0	-0.2	0.2	0.1	-0.2	2.0
V	3.5	2.2	1.6						
S	1.0	0.0	0.0						2.0
C				0.74	0.57	0.73	0.92	0.74	
f	0.2	0.4	0.9						

objects, so the free energy can be modeled as follows:

$$G_1(P,T) = U_1 + U_S + \frac{P \times V_S}{C_1} - (T - T_{ZP}) \times S_S \tag{4.10}$$

where U_S, V_S and S_S are the internal energy, volume and entropy of the spherical objects and C_1 is the packing efficiency of *hcp*. U_1 is the binding energy between the rotors in phase I, here set to zero (see table 4.3). The PIMD results I presented at the start of the chapter, suggest that the zero point contribution can be viewed as a temperature offset, here modeled as T_{ZP}.

In phase II, the spherical rotor (S) transforms to a simple rod (R) and the free energy can be written as:

$$G_2(P,T) = U_2 + U_R + \frac{P \times V_R}{C_2} - (T - T_{ZP}) \times S_R \tag{4.11}$$

where U_2 is slightly negative, symbolizing the attractive quadrupole-quadrupole interaction between rods which is responsible for the orientational order. In phase II, the packing C_2 is less efficient than that of *hcp*.

At higher pressures, phase III loses internal energy as the bond is elongated, but gains better packing efficiency as the rods are now organized in layers:

$$G_3(P,T) = U_3 + U_R + \frac{P \times V_R}{C_3} - (T - T_{ZP}) \times S_R \tag{4.12}$$

here, $U_3 > 0$ models the increase in energy as the electron is pushed out of the bond, while $C_3 > C_2$ shows that phase III wins over phase II at high pressures.

Above phase III, at high temperatures, we find phase IV which, in a simple picture, is a combination of spheres (S) in the B layers and atoms (A) in the G layers. Surprisingly, the BG^a model, where large spheres (S) are sandwiched between a hexagonal pattern of smaller spheres (A) is a known theoretical arrangement of high packing efficiency, even higher than *hcp* [137]. All these observations are included in the expression of the free energy:

$$G_4(P,T) = U_4 + \frac{U_A + U_S}{2} + \frac{P \times (V_A + V_S)}{C_4 \times 2} - \frac{(T - T_{ZP}) \times (S_S + S_A)}{2} \quad (4.13)$$

where U_4 is half of U_3 and represents an energy penalty for elongating molecules in the G layers (here simplified as atoms).

Finally, when all molecules have dissociated, one can expect an atomic phase with free energy:

$$G_5(P,T) = U_5 + U_A + \frac{P \times V_A}{C_5} - (T - T_{ZP}) \times S_A \quad (4.14)$$

here, U_5 is the internal energy of the atomic solid and $C_5 = C_1$ is the packing of *hcp*.

In the following, I will continue with the liquid phase. Interestingly, the melting curve has a maximum above phase I, such that at pressures below the maximum the solid is denser than the liquid, while at pressures above the maximum, less dense [21, 22]. Notice that in our simple model, the volume of a phase is considered independent of pressures. This is a big simplification, a better model would include equations of state.

However, the maximum in melting can still be achieved within this approximation, if the overall liquid is modeled as a mixture of three different liquid types, where each type of liquid comprises of one of the three different basic building units. In this model, every liquid of a certain building unit (i.e. S, R or A) is stable at higher temperatures, above the solid comprising of the same respective unit, having a larger entropy but smaller packing efficiency than the corresponding solid. The loss in packing efficiency is modeled by the parameter f, different for

each type of building unit (see table 4.3). Surprisingly, at high enough pressure, the sum liquid can achieve lower volume than the competing solids, reproducing the maximum feature of the melting line.

Below are the free energy expressions for each type of liquid:

$$G_{lS}(P,T) = U_{liq} + U_S + \frac{P \times V_S}{f_S \times C_1} - (T - T_{ZP}) \times S_S - (T - T_{ZP}) \times S_{liq} \quad (4.15)$$

$$G_{lR}(P,T) = U_{liq} + U_R + \frac{P \times V_R}{f_R \times C_3} - (T - T_{ZP}) \times S_R - (T - T_{ZP}) \times S_{liq} \quad (4.16)$$

$$G_{lA}(P,T) = U_{liq} + U_A + \frac{P \times V_A}{f_A \times C_5} - (T - T_{ZP}) \times S_A - (T - T_{ZP}) \times S_{liq} \quad (4.17)$$

here, U_{liq} is the internal energy and S_{liq} is the entropy characteristic of the liquid, regardless of the building unit. Finally, the liquid, overall, can be modeled as Boltzmann mixture of the different liquid types depending on their individual free energies:

$$G_{liq}(P,T) = \frac{G_{lS} \times \exp\left(\frac{-G_{lS}}{T}\right) + G_{lR} \times \exp\left(\frac{-G_{lR}}{T}\right) + G_{lA} \times \exp\left(\frac{-G_{lA}}{T}\right)}{\exp\left(\frac{-G_{lS}}{T}\right) + \exp\left(\frac{-G_{lR}}{T}\right) + \exp\left(\frac{-G_{lA}}{T}\right)} \quad (4.18)$$

This model for the liquid can reproduce well the melting curve measured experimentally [22]. In figure 4.6 I show the phase diagram generated using the free energies laid out above and the parameters reported in table 4.3. First I computed all the free energies on an evenly spaced pressure-temperature grid and compared them to one another to find the regions of stability for each phase.

Although this thermodynamic model is extremely simple, it reproduces the actual hydrogen phase diagram quite well. The exact values for pressure and temperature are irrelevant and can be easily tweaked by changing the parameters. However all the phases are in the right regions of the pressure-temperature (PT) space, the transition lines have plausible slopes, while the melting line has the

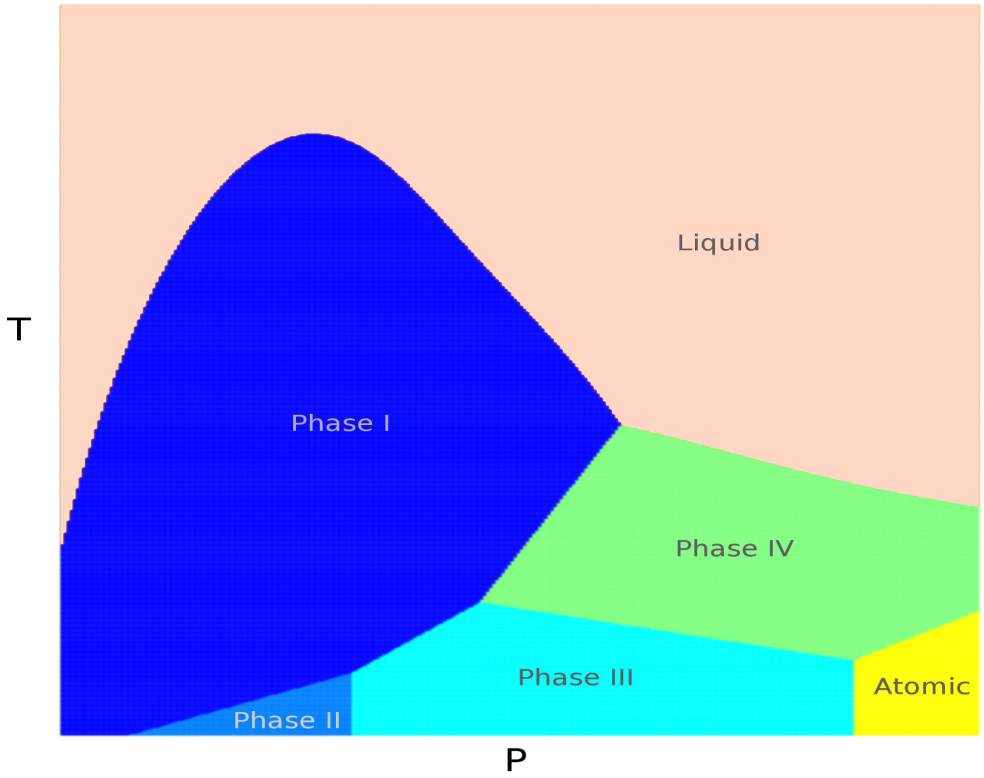

Figure 4.6 *The figure shows an example phase diagram obtained with the simple packing model and the parameters from table 4.3. The free energies were calculated on an evenly distributed P-T grid using a code I developed in GNU Octave [138].*

expected behavior.

Overall, the model is surprisingly successful in summarizing the essential physics of hydrogen at high pressure. Phase I has high entropy which makes it stable at high temperatures, whereas phase II has lower internal energy as a result of quadrupole-quadrupole interactions and is, therefore, more stable at lower temperatures. Phase III has higher internal energy as a result of the electron moving out of the bond, however, at the same time, it is a better packing so it wins at higher pressures. Phase IV, is an excellent packing of spheres (S) and atoms (A) and at the same time, it has high entropy owing to the spherical objects in the B layers. As a result, phase IV wins over phase III at higher temperatures and over phase I at higher pressures. In other words phase IV is a mixture between phases I and III, more entropic than the latter and better packed than the former. The liquid above phase I comprises predominantly of spherical objects, being less dense than the solid below. Upon increasing pressure, the liquid of rods gradually contributes more and increases the density of the fluid, which eventually becomes denser than the solid below. This gives an excellent

account of the phase diagram as it is known today.

4.4 Summary

In this chapter, I presented two important results. First I showed that in molecular dynamics, at finite temperature hydrogen can organize itself in an interesting structure comprising of long polymeric chains of atoms that could be stabilized by a charge density wave. This phenomenon could also be responsible for the increased stability of the layered molecular structures as shown by the grouping of diffraction peaks near $2k_F$ in the DoS. Alternatively, *Chains*, which has been found before in the literature [130, 131] could be a spurious effect of insufficient k-point sampling. Interestingly, k-point sampling might be much more important for molecular dynamics than previously believed. This could have far reaching implications for the more advanced studies as well. For instance, while quantum Monte Carlo molecular dynamics is praised for solving the exchange-correlation problem, it could be lacking in other respects which might be far more important.

Second, I showed that the hydrogen phase diagram can be understood from a simple thermodynamics perspective. Currently, the workflow in the field is: the theory predicts the "final product" like the pressure and temperature of transition for phases III to IV, and the experiment validates the "final product", or not. I believe that much can be gained by bridging the gap from the experimental side as well. For instance starting from the simple model that I presented here, one could further reduce the number of parameters by synthesizing more physics into the model. The equations could then be numerically fitted to the experimental phase diagram to obtain an "intermediate product" like entropy or internal energy differences, which would be more readily available as a benchmark for the theory. A model that enforces the basic physics could also be helpful for experiments. For instance, all melting experiments so far have checked their results against the Kechin fit [139], which helps to draw the line but also extrapolate the results. One could develop a simple routine that provided with the list of the pressures and temperatures, where each of the phases has been found, could output the most likely phase diagram. This way all diagrams published throughout the literature could be directly compared to one another, considerably speeding up the progress.

Chapter 5

Hydrogen-Deuterium Mixtures: Phases III and IV

5.1 Overview

The results showed so far concerned pure solid hydrogen. This chapter is centered around the two most recent experimental studies in hydrogen-deuterium mixtures: one using Raman spectroscopy in phase IV at room temperatures [2] and the other using infrared spectroscopy in the region of stability of phase III at lower temperatures [114].

The experiment in phase IV started with loading hydrogen and deuterium gas together in diamond anvil cells at room temperature. In their study [2], they claim that even pressures as low as 0.2 GPa were sufficient to induce recombination, necessary to generate hydrogen-deuteride molecules. On the other hand, the phase III experiment, which I will discuss in the second half of the chapter, started with pure hydrogen-deuteride at low temperature and claimed that dissociation and recombination (which they call DISREC) only takes place at a high pressure above 150 GPa. The probability of a dissociation event is proportional to $e^{-(H-TS)/k_BT}$ where at low temperatures, the molecular binding enthalpy H is the important energy scale. At low temperatures, dissociation has a small probability. Furthermore, for the completion of the recombination process, a second disassociation event needs to take place in the proximity of the first. This requirement decreases the probability even further. Therefore, it is expected that

mixing does not occur at low temperature but it does at high temperatures, where TS becomes comparable to the binding enthalpy H.

The concentration of the different molecular species in a hydrogen-deuterium mixture can be deduced by the ratio between their respective Raman or infrared intensities. Monitoring these ratios with time at different pressures and temperatures could serve as a direct probe for the molecular strength at those conditions. The experiment in phase III [114] already carried out such a study at one point in the phase diagram. Future experiments could extend their analysis to the rest of the phase diagram.

Recombination is just one of the interesting processes that take place in mixtures. Even more interesting are the changes in the Raman and infrared spectra owing to different hydrogen-deuterium compositions. In the first half of the chapter, I will argue that Raman changes in phase IV mixtures are driven by a textbook effect called mass induced phonon localization, which might be unique to solid molecular hydrogen. The second half will focus on a cautionary tale, namely that splitting or appearance of new infrared vibron peaks in phase III mixtures should be carefully investigated and they are not necessarily indicators of structural phase transitions. My first result was published in ref [2] and the second is being prepared for publication [9].

5.2 Mixtures in Phase IV: Phonon Localization

5.2.1 Introduction

Hydrogen-deuterium mixtures in phase IV represent a particularly interesting system to study. As I showed in chapter 3, it is generally acknowledged that phase IV comprises of at least two very different layer types B and G. Additionally, phase IV is also a molecular solid with two different scales of interactions: strong intramolecular and weak intermolecular. These different molecular environments should generate six distinct Raman active modes in pure hydrogen phase IV crystals [89]. On one hand, the two higher frequency modes called vibrons stem from the intramolecular interaction and comprise of in-phase molecular stretching. On the other hand, the four lower frequency modes are generated by the weaker intermolecular interaction. As shown in chapter 3, two of these low frequency modes are out of layer oscillations and the other two are in layer oscillations.

Similarly to vibrons, the low frequency modes are in pairs, one mode from each layer type. This complexity offers a unique playground for mixtures.

A very interesting idea was formulated in ref [78], even before the experimental discovery of phase IV [87, 89]. They proposed that owing to its unique structure, *Pbcn* could induce isotopic segregation where hydrogen and deuterium are separated in the different layer environments (i.e. B and G) to minimize the free energy. If this was true, the experiment would only measure two vibron modes one compatible with pure hydrogen, the other compatible with pure deuterium. In the low frequency range, the experiment would find four low frequency modes, two from each molecular species. This prediction motivated our collaborators on the experimental side to perform a study with hydrogen-deuterium mixtures in phase IV. I also tested the hypothesis using theoretical methods as explained later in the chapter.

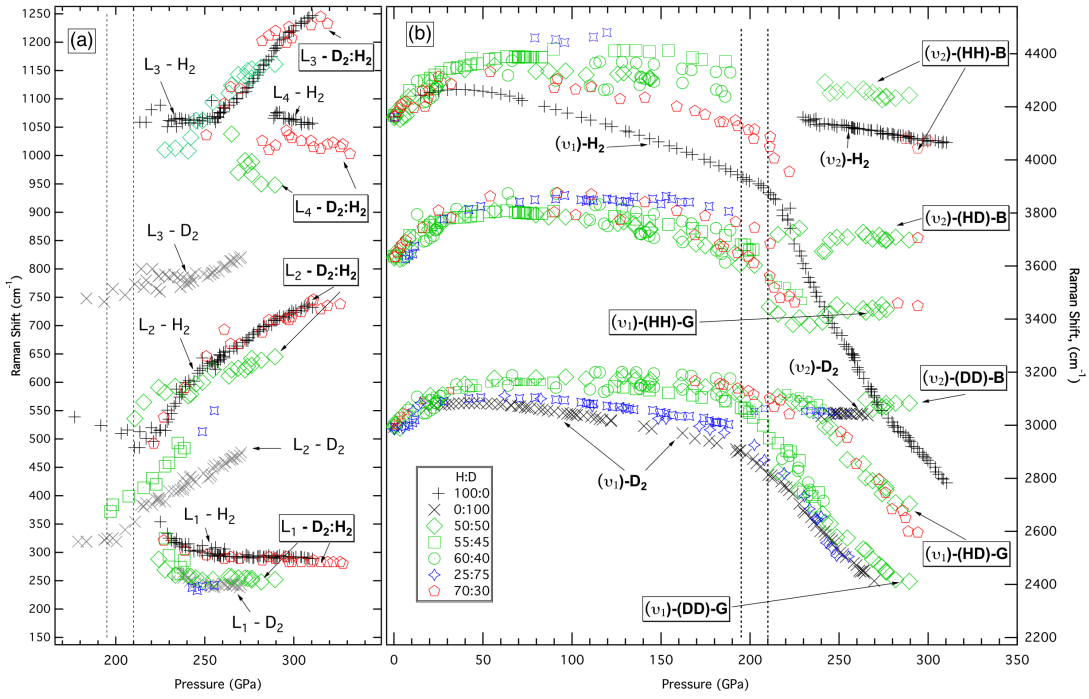

Figure 5.1 *This figure was reproduced from ref [2] and shows the experimental Raman spectra in hydrogen-deuterium mixtures as function of pressure, up to phase IV. The different colors correspond to different compositions as labeled. The convention for naming the low frequency modes and vibron modes is the same as in chapter 3.*

As shown in figure 5.1, the experiment claimed six Raman active vibrons in mixtures which implies that mixtures are randomly distributed and there is no sign of isotope segregation. Their findings were consistent with my analysis which

showed that the free energies gained by segregation would be insufficient to drive order at room temperature. Furthermore, I found a textbook effect called phonon localization, which here is purely induced by mass disorder - unique to hydrogen mixtures.

The idea of localization in random lattices was first introduced by Anderson [140] who noted that disorder can affect transport in electronic systems. Since then, localization has been thoroughly investigated in many other random systems as well. Several studies demonstrated the phenomenon: ultrasound localization in random elastic networks [141], localization of Bose-Einstein condensates [142] and even localization of light [143].

The localization of phonons means that the concept of a phonon band where all atoms across the crystal contribute to the oscillation breaks down. Instead, a small group of atoms oscillate together separated from the rest of the crystal. Hydrogen and deuterium have the same electronic structure but a large mass ratio of 1:2, so here the phonon localization is purely induced by mass disorder.

Localization itself cannot be measured directly in experiment. However, phonon localization affects the Raman spectra and it is, in turn, strongly influenced by the composition of the hydrogen-deuterium mixture. Theoretically, I investigated these dependencies and made a prediction on how Raman vibrons would vary with pressures and hydrogen-deuterium concentrations in the presence of localization. All my findings were confirmed by experimental observations, indirectly proving phonon localization in hydrogen-deuterium mixtures.

5.2.2 Methods

DFT settings

All the calculations in this chapter were performed with CASTEP [31, 115], unless otherwise specified. I used the Perdew-Burke-Ernzerhof (PBE) [28] exchange-correlation functional. The electron charge density was set to spin unpolarized and expanded in plane waves with an energy cut-off of 1200 eV. All the systems were treated as insulators by fixing the occupancy of the electronic bands, which allowed the calculation of infrared and Raman spectra. The total energy in the DFT loop was converged within 10^{-11} eV and the minimization method of choice was density mixing (dm). The pseudo-potentials were generated on the fly using

the CASTEP input string "H 0|0.7|2|6|8|10L(qc = 10)" [116].

Table 5.1 *The table shows the two settings used for geometry optimizations in this section.*

Converged Quantity	Coarse	Fine
Energy	5×10^{-6} eV/atom	1×10^{-8} eV/atom
Force	1×10^{-2} eV/Å	5×10^{-4} eV/Å
Stress	1×10^{-1} GPa	5×10^{-4} GPa
Displacement	5×10^{-4} Å	5×10^{-4} Å

Free Energy Calculations

The goal here was to test the hypothesis proposed in reference [78] according to which hydrogen and deuterium could segregate into layers to minimize the free energy in phase IV. All calculations were performed on the Pc structure, which is a good ground state candidate for phase IV. I started with a coarse geometry optimization (see table 5.1) of the Pc unit cell containing 48 atoms at 250 GPa. For sampling the reciprocal space I employed a k-point Monkhorst-Pack (MP) grid with spacing 0.04 $\mathring{A}^{-1}$.

I then continued with a phonon calculation at the gamma point using finite displacement and FINE method supercell, followed by a thermodynamics calculation as implemented in CASTEP. This procedure allowed me to evaluate the free energy at room temperature (300 K) in the quasi-harmonic approximation. For extending this calculation at other hydrogen-deuterium mixtures I used a CASTEP post-processing code called PHONONS. The software PHONONS reloads the check file written by the initial CASTEP calculation, together with the reconstructed cell file obtained by isotopic substitution to generate new phonons and thus a new free energy approximation for each of the mixture samples.

Here I focused on 50:50% mixtures containing exactly 12 hydrogen (H) and 12 deuterium (D) molecules. Generating completely random mixtures is not sufficient for this exercise, because the cases of interest where layer segregation occurs, are relatively unlikely. Therefore I used my own script to generate random samples, allowing no repetition as explained in the following.

Let n_{TL} be the number of bonds of a specific type T (i.e. HH, HD or DD) within a certain layer L (i.e. G or B). I define a class of structures as the collection of all the configurations determined by the same 6-tuple (n_{HHB}, n_{HDB}, n_{DDB}, n_{HHG},

n_{HDG}, n_{DDG}). Now, I sample each class of structures by generating a total of $N_s = 2400$ different samples uniformly in this configurational space; all subject to the following independent constraints:

$$n_{HHB} + n_{HDB} + n_{DDB} = 12 \tag{5.1}$$

$$n_{HHG} + n_{HDG} + n_{DDG} = 12 \tag{5.2}$$

$$2n_{HHB} + n_{HDB} + 2n_{HHG} + n_{HDG} = 24 \tag{5.3}$$

$$2n_{DDB} + n_{HDB} + 2n_{DDG} + n_{HDG} = 24 \tag{5.4}$$

Equations 5.1 and 5.2 reinforce that each layer type should accommodate exactly 12 molecules, while relations 5.3 and 5.4 ensure that the number of hydrogens equals the number of deuteriums and thus both equal half the total number of atoms.

Both hydrogen and deuterium have the same electronic structure and thus the enthalpy is the same for all the generated mixed samples, however, the zero point energy and the entropy vary due to changes in the phonon spectrum. For each of the 2400 different mixture samples, I performed a thermodynamic calculation and evaluated the free energy within the quasi-harmonic approximation. As shown in the results section, the free energy difference between the fully segregated and randomly mixed configurations is not sufficient to drive order at room temperature.

Phonon calculations in mixtures

As I will show later in more detail, the result derived from the thermodynamic calculation suggests that at room temperature there is no isotopic segregation in Pc hydrogen-deuterium mixtures. Therefore, for this next exercise, I doubled the unit cell of the Pc structure to 96 atoms and prepared one uniformly distributed mixture sample at 240, 270 and at 300 GPa, respectively, in which the identities of the atoms (i.e. H and D) were randomly assigned. I re-optimized the structure at all three pressure points using the fine settings reported in table 5.1. The

reciprocal space was sampled using an 8x8x8 k-point MP grid. I continued with a gamma point phonon and efield calculation at each pressure, using finite displacement and FINE method supercell. Additionally, at 270 GPa I also performed a full Raman calculation which was then used as a benchmark for the method introduced below.

Once the dynamical matrix was calculated at all three pressure points with CASTEP, I used the software PHONONS to extend the phonon calculation to other isotopic configurations. At each pressure point, I generated $N_s = 2700$ random hydrogen-deuterium configurations with hydrogen concentrations ranging from 0 to 100%. I based my method on the assumption that there is no isotopic segregation in mixtures, regardless of hydrogen concentration. Although I only tested this assumption at 50:50% mixtures, this result was demonstrated by experiment at other compositions as well.

Raman and Localization from Lattice Dynamics

The idea here was to calculate Raman spectra and phonon localization in phase IV hydrogen-deuterium mixtures at different concentrations and pressures in light of comparison to experiment [2]. Since mass disorder breaks translational symmetry in mixtures, the Raman signature for mixtures can only be obtained by sampling many different mixture configurations and averaging the spectra. Recalculating the Raman intensity for each sample is a non-trivial expensive task, so here I developed a proxy for approximating the high frequency Raman vibrons. Based on the previous observations presented in chapter 3, the Raman active vibrons in Pc comprise of in-phase stretches of the molecules. Unfortunately, the low frequency modes cannot be easily recomputed in mixtures. Here, for any random mixture sample s, I approximated the high-frequency Raman intensities by projecting the eigenmodes $\varepsilon^{\alpha}_{l,k}$ onto molecular axes:

$$r_s(f_k) = \sum_{l=1}^{N} \sum_{\alpha=1}^{3} a^{\alpha}_{l} \varepsilon^{\alpha}_{l,k} \tag{5.5}$$

where l running from 1 to N labels the atoms, k labels the modes and α the Cartesian components. The vector $\vec{a}_l = (a^1_l, a^2_l, a^3_l)$ represents the molecular axis pointing away from atom l to the nearest neighbouring atom.

Comparing the result obtained with this method with the actual Raman

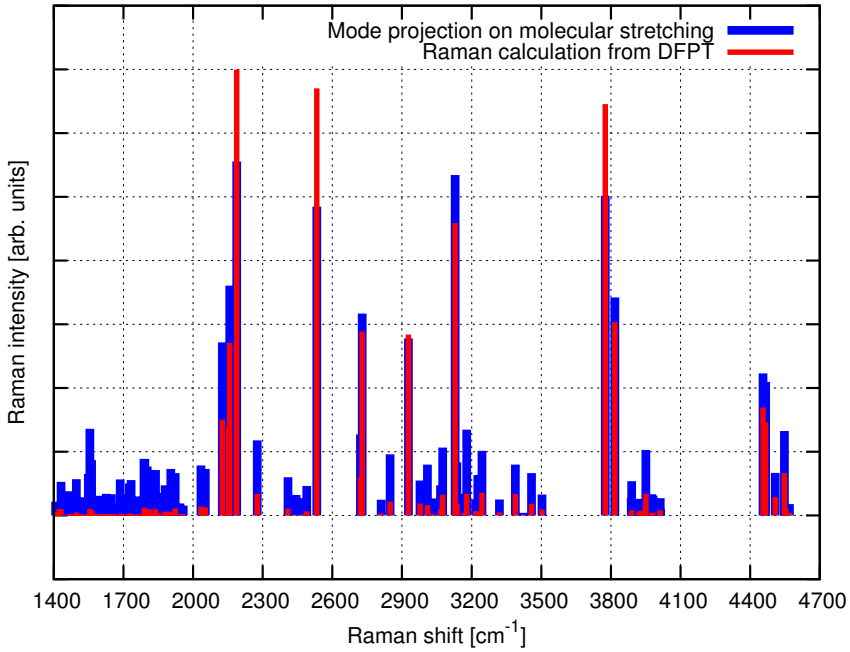

Figure 5.2 *In this figure I compare the Raman vibrons obtained in two different ways for a random mixture at 270 GPa in Pc. In red is an actual calculation with CASTEP and in blue is the proxy method I used to approximate the high frequency Raman vibrons.*

calculation performed with CASTEP for the same mixture sample at 270 GPa, I found good agreement (figure 5.2). I concluded that this projection method is sufficiently accurate to be used as a proxy in mixtures [4].

This method also allows us to calculate independently Raman contributions from HH, HD and DD, respectively, by summing selectively over those molecule types (denoted T). For instance to obtain the Raman contribution from HH molecules:

$$r_s^{HH}(f_k) = \sum_{l=1}^{N;T_l=HH} \sum_{\alpha=1}^{3} a_l^{\alpha} \varepsilon_{l,k}^{\alpha} \tag{5.6}$$

Apart from the Raman method, I also defined a measure of phonon localization which is the main finding discussed in this chapter. For a given mixture sample, the localization is calculated via the inverse participation ratio [144]:

$$pl_s(f_k) = \sum_{l=1}^{N} \left[\sum_{\alpha=1}^{3} \varepsilon_{l,k}^{\alpha} \varepsilon_{l,k}^{\alpha} \right]^2, \text{ subject to: } \sum_{l=1}^{N} \sum_{\alpha=1}^{3} \varepsilon_{l,k}^{\alpha} \varepsilon_{l,k}^{\alpha} := 1 \tag{5.7}$$

I also calculated integrated localizations over the low frequency range (LF), high frequency range (HF) and over the entire frequency range (EF). These give a measure of how localized are the vibrons, the low frequency modes and all phonons, overall.

$$pl_s^{LF} = \sum_{f_k}^{f_k<=f_0} pl_s(f_k); \quad pl_s^{HF} = \sum_{f_k}^{f_k>f_0} pl_s(f_k); \quad pl_s^{EF} = \sum_{f_k} pl_s(f_k) \tag{5.8}$$

where $f_0 = 1800\, cm^{-1}$ is an arbitrary frequency that separates the high and low frequency range.

Localization is a measure for the number of molecules participating in a phonon mode and it is related to the inverse participation ratio [144]. To get a feeling for the kind of values one would expect for this quantity, let us look at some specific examples in Pc. For instance, a completely delocalized vibrational mode would induce equal displacements in all the molecules in one layer type (B or G). Using the normalization condition, the localization of such a mode in a 96 atom Pc cell is approximately:

$$pl_{min}(f_k) \approx \frac{N}{2}\varepsilon^4 = \frac{N}{2}\frac{2^2}{N^2} = \frac{2}{N} = \frac{1}{48} \tag{5.9}$$

where ε is the displacement along each eigenvector, here assumed equal for all the atoms. On the other hand, a completely localised vibrational mode, would induce vibrations in one molecule only, so:

$$pl_{max}(f_k) \approx 2\varepsilon^4 = 2\frac{1}{2^2} = \frac{1}{2} \tag{5.10}$$

The final Raman activity and localization at one given pressure and concentration were obtained by binning up the result from many randomized hydrogen-deuterium mixture samples s. For every configuration, the eigenvectors were recomputed with PHONONS as explained in section 5.2.2. In total, I sampled 2700 different mixture samples at each of the pressures 240, 270 and 300 GPa, which was sufficient to ensure convergence. For this study, I will not delve into details on how the histograms were binned up and normalized. These specifications will become important in the next study of infrared spectra in phase III, presented later in section 5.3.2.

Molecular Dynamics Calculation

The LD method is useful for sampling many molecular environments and thus, ensuring good convergence of Raman spectra. Nevertheless, as I extensively discussed in the chapter 3, at room temperature, the crystal structure of phase IV is not exactly Pc and static methods cannot reproduce the dynamic behavior. LD calculations, alone, cannot account for phonon anharmonicity, which is especially problematic in the G layers of Pc. Here I performed Molecular Dynamics (MD) calculations which help account for the dynamical effects.

MD is expensive so here I only performed calculations with pure isotopes H (100% hydrogen), pure D (0% hydrogen) and with one random 50:50% mixture sample, no other intermediate concentrations. As a starting point, for the pure isotopes, I chose Pc cells with 288 atoms and 1x1x2 k-point MP grids. For the mixture I chose a 768 atom cell and I only sampled the reciprocal space at the gamma point. I used a larger cell for the mixtures to better represent a random mass distribution and minimize the effect of the implicit translational symmetry.

Table 5.2 *Here I present the summary of MD calculations in hydrogen-deuterium mixtures.*

Cell	H %	Atoms	Ensemble	Iterations	Length	Pressure	Temp.
pure H	100%	288	NVE	3000	1.5 ps	272 GPa	294 K
pure D	0%	288	NVE	3000	1.5 ps	272 GPa	307 K
rand mix	50%	768	NPT	1000	0.5 ps	268 GPa	219 K
rand mix	50%	768	NVE	3000	1.5 ps	267 GPa	296 K

Table 5.2 summarizes the details of the MD simulations. The cells were geometry optimized at the start using the coarse settings reported in table 5.1. The pressure and temperature reported in table 5.2 are averages over the values printed by CASTEP at each step. The simulations meant for calculating Raman, were performed in the NVE ensemble. For mixtures I used a long equilibration stage in NPT to allow the large cell to reach equilibrium first. The pressure was fixed with the Parrinello-Raman barostat and the temperature with the Nose-Hoover chain thermostat using CASTEP default settings. The time step in all the MD runs was set to 0.5 fs, consistent with the simulations presented in the previous chapters.

Raman and Localization from Molecular Dynamics

I devoted the entire chapter 3 to the method of projecting Raman from MD in molecular hydrogen. Here I will shortly summarize the basic procedure for the benefit of the reader. Instead of projecting the normal modes like in LD, in the case of MD, I project the velocities onto in-phase molecular stretching and then take the Fourier Transform of the result:

$$r(f) = FT\left[\sum_{l=1}^{N}\sum_{\alpha=1}^{3} a_l^{\alpha} v_l^{\alpha}(t)\right] \tag{5.11}$$

where $v_l^{\alpha}(t)$ is the velocity of the atom labeled l at time t.

Same as in the case of LD, in MD I can also separately compute Raman for HH, HD and DD, by performing the projection selectively:

$$r^{HH}(f) = FT\left[\sum_{l=1}^{N;T_l=HH}\sum_{\alpha=1}^{3} a_l^{\alpha} v_l^{\alpha}(t)\right] \tag{5.12}$$

The projection method was thoroughly tested in chapter 3 and refs [1, 3, 4] for pure hydrogen. Mixtures are a more difficult case than pure isotopes, so here I performed a new test. I applied the projection on short, consecutive simulation windows of 0.25 ps each as reported in table 5.2.

Table 5.3 *Peak positions with their standard deviations computed from MD in 50:50% mixtures using the window method.*

Time Interval	HH-G	HH-B	HD-G	HD-B	DD-G	DD-B
0.00 - 0.25 ps	3466	4379	2915	3731	2268	3080
0.25 - 0.50 ps	3223	4380	2916	3718	2401	3046
0.50 - 0.75 ps	3350	4393	2944	3710	2281	2943
0.75 - 1.00 ps	3470	4387	2665	3606	2010	3047
1.00 - 1.25 ps	3457	4409	2553	3737	2266	2945
1.25 - 1.50 ps	3330	4408	2911	3607	2275	2909
average	3383	4393	2817	3685	2250	2995
std. dev.	111	15	185	69	144	79

Figure 5.3 illustrates the vibrons obtained from the 50:50% mixture simulation where projections are performed selectively onto each molecular species. Table 5.3 summarizes the positions of the peaks as obtained from each window of the

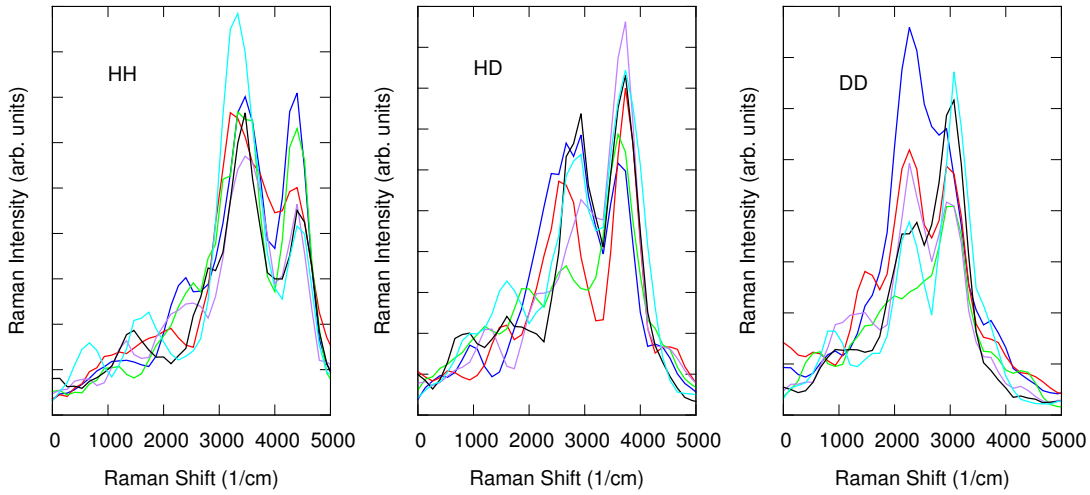

Figure 5.3 *Left, middle and right panels show selective molecular projections onto HH, HD and DD molecules as calculated from MD. Each curve is a spectrum obtained from a short 0.25 ps window of simulation. The positions of the peaks which were summarized in table 5.3 are consistent across the simulation, proving the method is robust.*

simulation. With the exception of a few outliers, the results are robust and even simulations as short as 0.25 ps are sufficient to obtain reliable Raman vibrons, once the cells have equilibrated. Note that a frequency of 2000 cm^{-1} corresponds to 0.017 ps, meaning that a 0.25 ps window contains about 15 molecular vibrations for each of the molecules. On the other hand, this windows would be too short to gain any useful information about the low frequency modes. All the vibrons presented in the results section were obtained from 1.5 ps long simulations and concern the vibrons only.

I also developed a method to approximate phonon localization from MD as I explain in the following. Denote t_{max} the total simulation time and dt the time step. One can naturally construct a set of frequencies with $f_{min} = 1/t_{max}$, $f_{max} = 1/dt$, sampled at a frequency step $df = f_{min}$, which avoids aliasing. One can, then, compute a set of displacement vectors $\delta^{\alpha}_{l,k}$ as Fourier Transforms of the atomic positions $x^{\alpha}_{l}(t)$:

$$\delta^{\alpha}_{l,j} = A_j^{-1} \cdot FT\left[x^{\alpha}_{l}(t)\right], \text{ where: } A_j^2 = \sum_{l=1}^{N}\sum_{\alpha=1}^{3}(\delta^{\alpha}_{l,j})^2 \tag{5.13}$$

where l labels the atoms and j labels the frequencies running from f_{min} to f_{max}. Note that the set of displacements $\delta^{\alpha}_{l,j}$ is an approximation to the normal modes,

but it is over-complete and encodes anharmonicity as well. Finally, using the calculated displacements, the phonon localization can be computed from the MD as:

$$pl(f_j) = \sum_{l=1}^{N} \left[\sum_{\alpha=1}^{3} (\delta_{l,j}^{\alpha} \delta_{l,j}^{\alpha}) \right]^2 \tag{5.14}$$

where as in the static case, *pl* is inverse proportional to the number of atoms that participate to the oscillation.

5.2.3 Results

Let us explore the idea proposed in ref [78] according to which phase IV mixtures could exhibit an unusual case of isotopic segregation into layers. In pure hydrogen, the B layer Raman vibron is around $f_{HHB} = 4000$ cm^{-1}, while the G layer vibron has a value of $f_{HHG} = 3000$ cm^{-1} [100]. These are just rough approximations since the frequencies change strongly with pressure. Assuming that the zero point energy (ZPE) from the vibron gives the most important contribution to ZPE, a segregated configuration where all deuteriums are in the B layers and all hydrogens are in the G layers would yield lower energies than a random mixture. In the simple harmonic approximation the frequency of the vibrons is:

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}}, \text{ with the reduced mass: } \mu = \frac{m_1 m_2}{m1 + m2} \tag{5.15}$$

For the three different molecular species, equation 5.15 yields:

$$f_{DD} = \frac{1}{\sqrt{2}} f_{HH} \text{ and } f_{HD} = \frac{\sqrt{3}}{2} f_{HH} \tag{5.16}$$

In the case of segregation, the ZPE contribution, per molecule, from the vibrons in the two layers (B and G) is:

$$ZPE_{seg} = \frac{h}{4}(f_{DDB} + f_{HHG}) = \frac{h}{4}\left(\frac{1}{\sqrt{2}} f_{HHB} + f_{HHG}\right) \tag{5.17}$$

In perfectly random 50:50% hydrogen-deuterium mixtures, both HH molecules

and DD molecules form with probability 25%, whereas HD molecules being equivalent with DH molecules are twice as likely. Therefore the contribution from the vibrons to the ZPE in random mixtures is:

$$ZPE_{mix} = \frac{h}{4}\left[\frac{1}{4}(f_{HHB} + 2f_{DDB} + f_{DDB}) + \frac{1}{4}(f_{HHG} + 2f_{DDG} + f_{DDG})\right] \quad (5.18)$$

And combining with equation 5.15:

$$ZPE_{mix} = \frac{h}{4}\frac{\sqrt{4} + 2\sqrt{3} + \sqrt{2}}{8}(f_{HHB} + f_{HHG}) \quad (5.19)$$

The difference in ZPE introduced by segregation is finally:

$$\Delta ZPE = ZPE_{seg} - ZPE_{mix} = \frac{h}{4}(-0.1527 f_{HHB} + 0.1402 f_{HHG}) \quad (5.20)$$

Placing the heavier DD molecules in the B layer stabilizes the lattice. The extra energy gained, however, is counterbalanced by the energy lost from placing all the remaining light HH molecules in the G layer. Using the experimental values for the vibron frequency, the ZPE difference per atom calculated from equation 5.20 is roughly $3\,meV/atom$.

I continued this analysis further by performing thermodynamic calculations with CASTEP on 2400 configurations at 50:50% hydrogen-deuterium concentrations. The results are shown in figure 5.4 and are consistent with my theoretical analysis. The structure with all the DD molecules in the B layer, all HH molecules in the G layer and hosting no HD molecules at all, has the lowest free energy in mixtures. One of the highest energies is attained in the inverted structure with all DD in the G layer and all HH in the B layer. However, when all modes are accounted for (i.e. not only the two vibrons as in the previous approximation) the energy differences are even smaller. In dynamics, the entropy will further help to stabilize a randomly distributed mixture. The difference between the segregated structure and the completely random is less than 1.2 meV/atom in LD, which is relatively small compared to $k_B T \sim 26\,meV$. I, therefore, conclude that at room temperature hydrogen-deuterium mixtures are randomly distributed and the energy differences found here are insufficient to drive segregation.

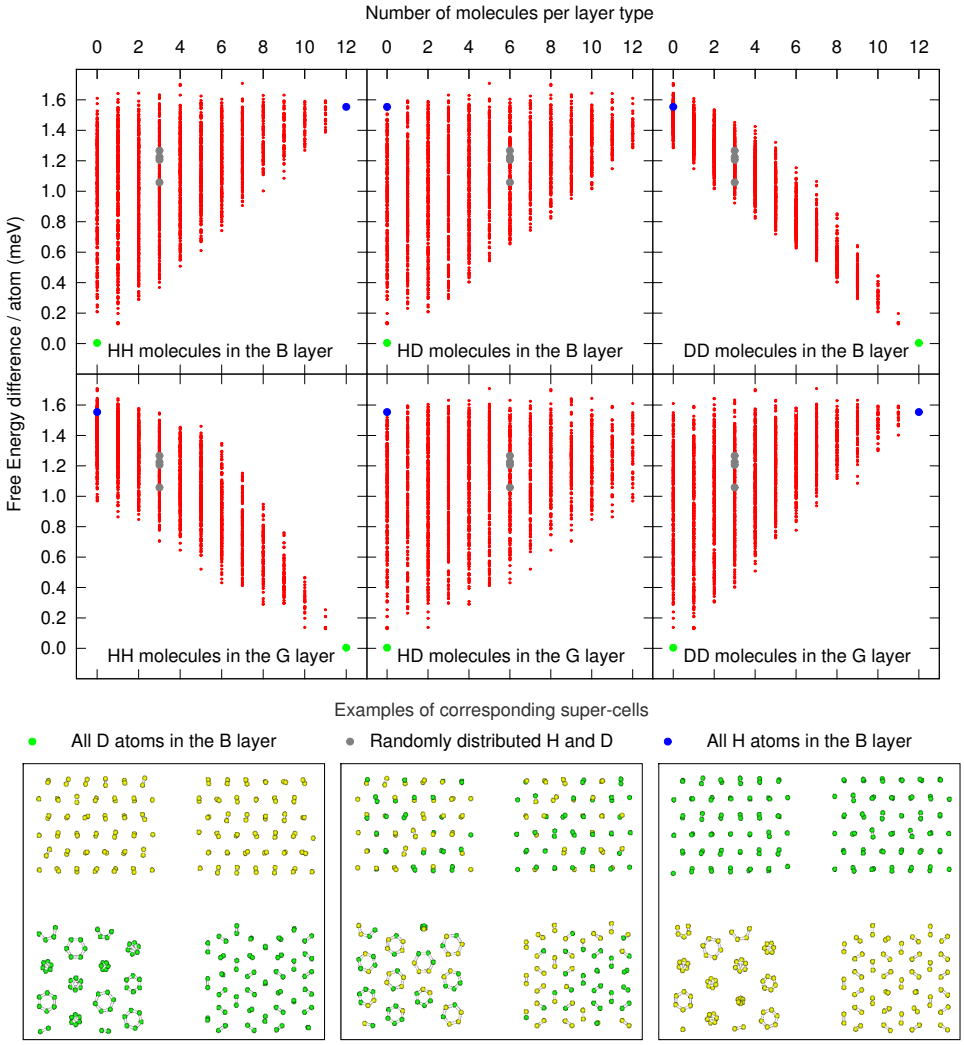

Figure 5.4 *The six panels on the top show the free energy per atom for all the 2400 samples of 50:50% hydrogen-deuterium mixtures calculated in 48 atoms Pc at 250 GPa and 300 K. The same result is plotted six times with respect to the different molecule occurrences n_{TL}. Green, gray and blue bullets label some special mixture cases with all deuteriums in the B layers, randomly distributed hydrogen and deuterium atoms and all hydrogens in the G layer, respectively. The three bottom panels illustrate examples for these mixture configurations. Note that all the free energies are given as differences with respect to the reference configuration of minimum energy.*

As I explained in the methods section 5.2.2 these calculations involved only the phonons at the gamma point, future work could repeat the simulation with a better sampling of the Brillouin zone, but this is unlikely to change the result. My result here is consistent with the experimental findings of our coworkers [2] who found strong evidence of disordered mixtures at all concentrations. Their

conclusions are based on Raman measurements which indicate six distinct vibrons originating from the three molecular species (HH, HD and DD) distributed equally in both B and G layers.

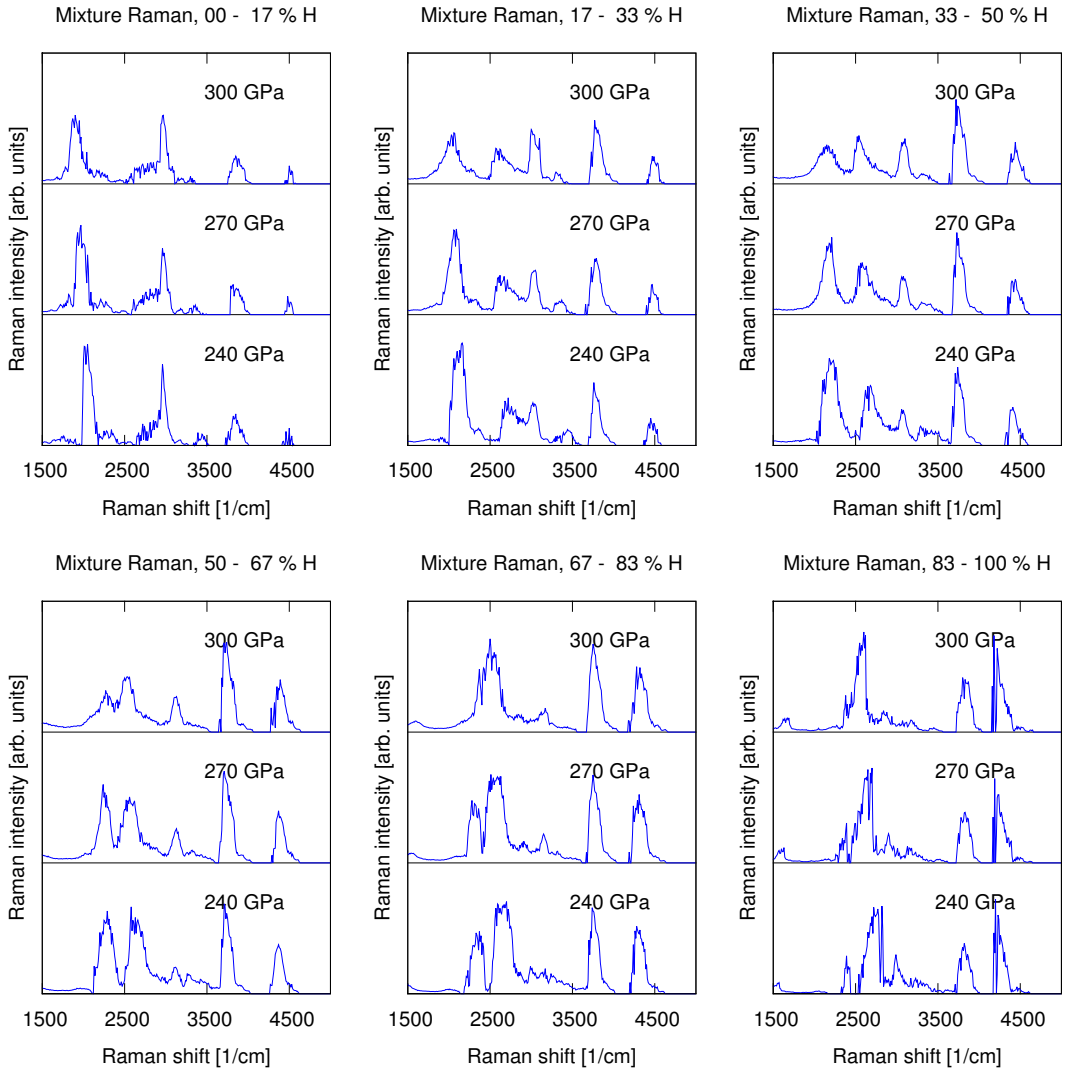

Figure 5.5 *Here I show the Raman vibrons extracted from LD in 96 atoms Pc mixtures at different pressures and concentrations. The results were grouped into intervals depending on the mixture composition, and binned into histograms.*

Here, I also calculated Raman vibrons at various concentrations and pressures as shown in figure 5.5 and I then compared these results to the experiment, later in figure 5.10. As figure 5.5 illustrates, the Raman vibrons in mixtures of phase IV are very diverse, they vary with pressure, but much stronger with hydrogen concentration. At around 30% hydrogen concentrations, one can distinguish six individual vibron peaks, but above 50% concentrations only five peaks are left. At higher and lower hydrogen (H) concentrations, respectively, the number of

vibrons decreases further. I attribute this strong dependence on concentrations to a phenomenon called phonon localization which I will discuss later in this section. Note that at 70% H two of the lower frequency peaks merge into one with increasing pressure. Many times in experiment, the disappearance, emergence and splitting of peaks is associated with structural phase transitions. As I will also discuss in the second part of this chapter 5.3 this is not always true, certainly not here. Changes to Raman and infrared spectra in mixtures should be analyzed with care, more often they are a result of mass disorder and phonon localization and do not indicate structural phase transitions.

The method of projecting selectively onto the three molecular species allowed me to identify the origin of each mode in the Raman spectra as I illustrate in figure 5.6. The signal originating from the DD molecule is clear and presents two well defined peaks that can be attributed to the two layers B and G. As expected the DDG vibron changes faster than the DDB vibron with increasing pressure, as shown by the different colors in figure 5.6. Apart from the two intense peaks, the spectra also contain smaller peaks which are the results of coupling with the HD molecule.

I found a similar behavior in the case of HD: two intense peaks that I label HDB and HDG and a few weaker peaks which originate from small coupling with the other species, mainly DD in the G layer and HH in the B layer. Finally, the HH molecule generates two vibrons at lower H concentrations, however, above 50% H concentration, the HHG peak becomes a broad band, which only depends weakly on pressures. These observations are consistent with the experimental results which show that HHG peak is broad and independent of pressure unlike the other peaks originating from the G layers.

Even though I found some evidence of coupling, it is surprising how well separated the spectra originating from the different species are. Furthermore, even within the same molecular species, not all molecules participate to the vibrations. As seen in figure 5.7 the modes seem to be localized on small groups of molecules of the same kind (e.g. all HH) separated from the rest of the crystal by molecules of different kinds (e.g. HD and DD). We call this phenomenon phonon localization because it is similar to the concept advanced by Anderson for electronic systems [140]. Phonon localization breaks the concept of a phonon band because the modes are no longer spread across the entire crystal but focused on a small island of similar molecules.

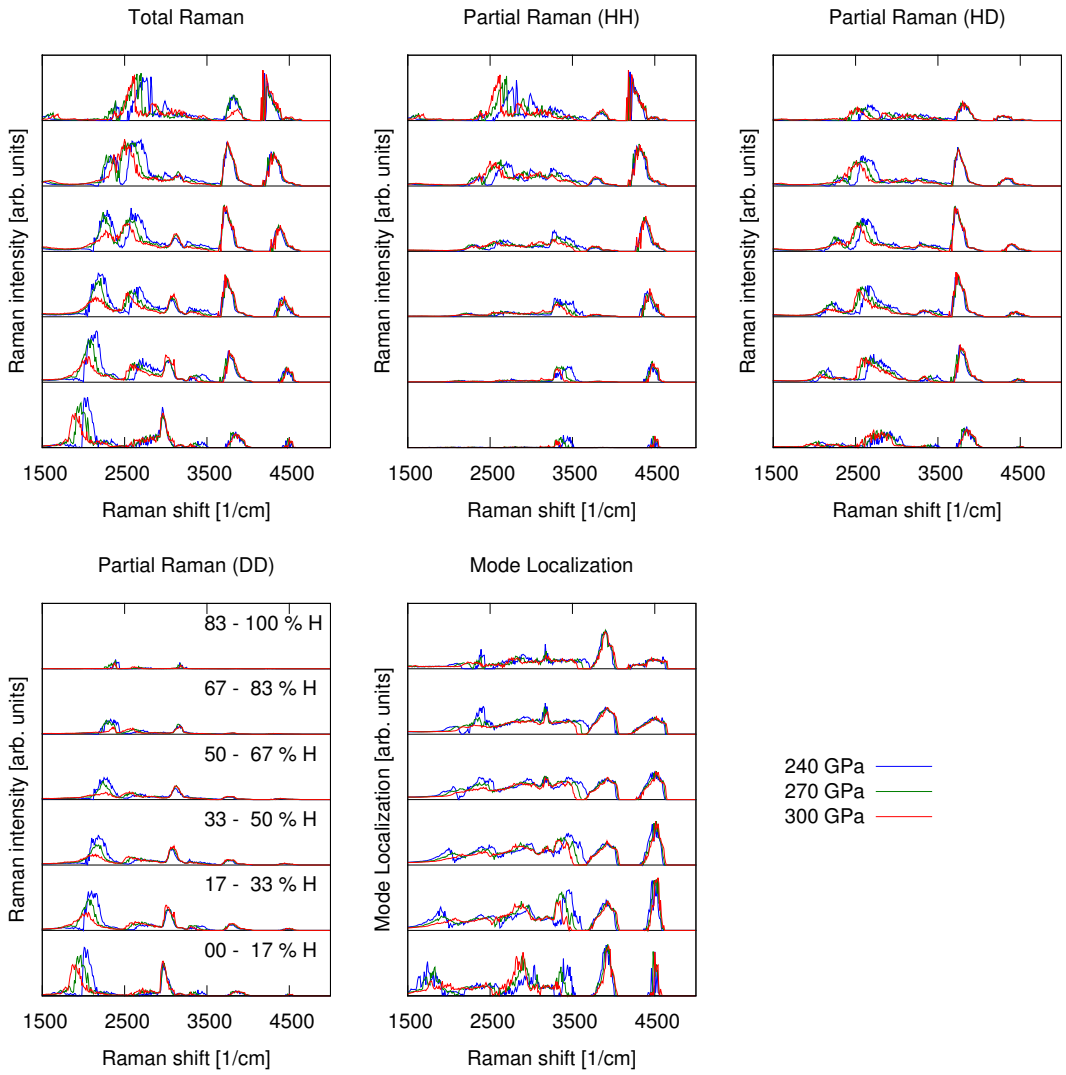

Figure 5.6 *These figures illustrate the total Raman spectra but also partial Raman contributions from different molecular species. It also shows phonon localizations. All results here were calculated from LD in 96 atoms Pc. In each figure, consecutive panels from top to bottom show results from mixtures with a decreasing concentration of hydrogen as labeled in one of the figures. The individual colors label the three different pressure points as shown in the legend. The top left figure is essentially a summary of all figures 5.5.*

Although localization itself cannot be measured experimentally, it can be gauged through its influence on the frequency of the modes. In figure 5.7 I show how the frequency of the vibrons in the B layers depends on the localization of the phonons. On one hand, more localized vibrons involve fewer molecules and generate higher frequencies. On the other hand, less localized vibrons, involve more coupling between molecules of the same kind and generate lower frequencies. The widths of the peaks are also related to the amount of localization.

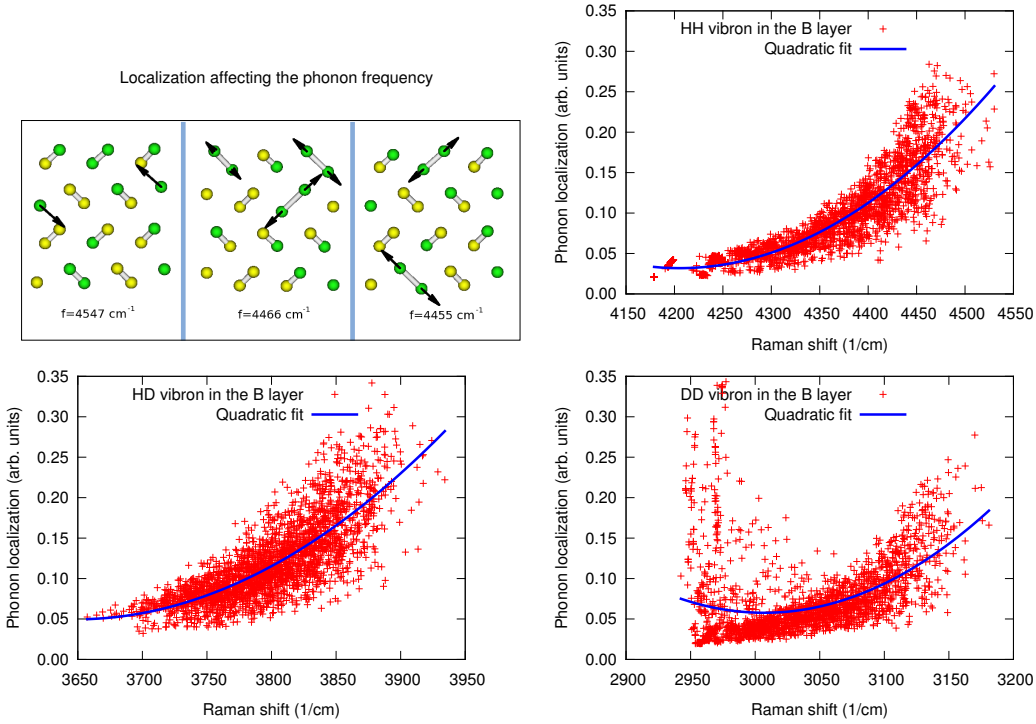

Figure 5.7 *The figure illustrates the correlation between the Raman frequency and mode localization, both calculated with LD. Here I show only results from the B layer. The top left panel shows examples of localized modes in 50:50% mixtures. The other panels show Raman versus localization for the three molecular species: HH, HD and DD, respectively. A quadratic fit appears to describe the data well.*

In figure 5.8 I show how the localization of different modes changes progressively with increasing hydrogen concentration. The vibrons originating from HHB and HDB are the highest in frequency (above 3700 cm^{-1}) and they form two bands that are independent of the rest of the modes. At low H compositions, there are few HH molecules available in the mixture and the HHB modes are highly localized, occupying a narrow band of high frequencies. As the H concentration is increased, the HHB band becomes broader and shifts to lower frequencies. The HDB band behaves in similar fashion. The band is wider and lower in frequency at high concentration of HD, while narrower and higher in frequency at lower concentration of HD. Note that the concentration of HD molecules c_{HD} is not linear in the H concentration c_H, but follows this equation in random mixtures:

$$c_{HD} = 2c_H(1 - c_H) \tag{5.21}$$

These conclusions hold true for the other vibron bands as well (e.g. DDB, HHG,

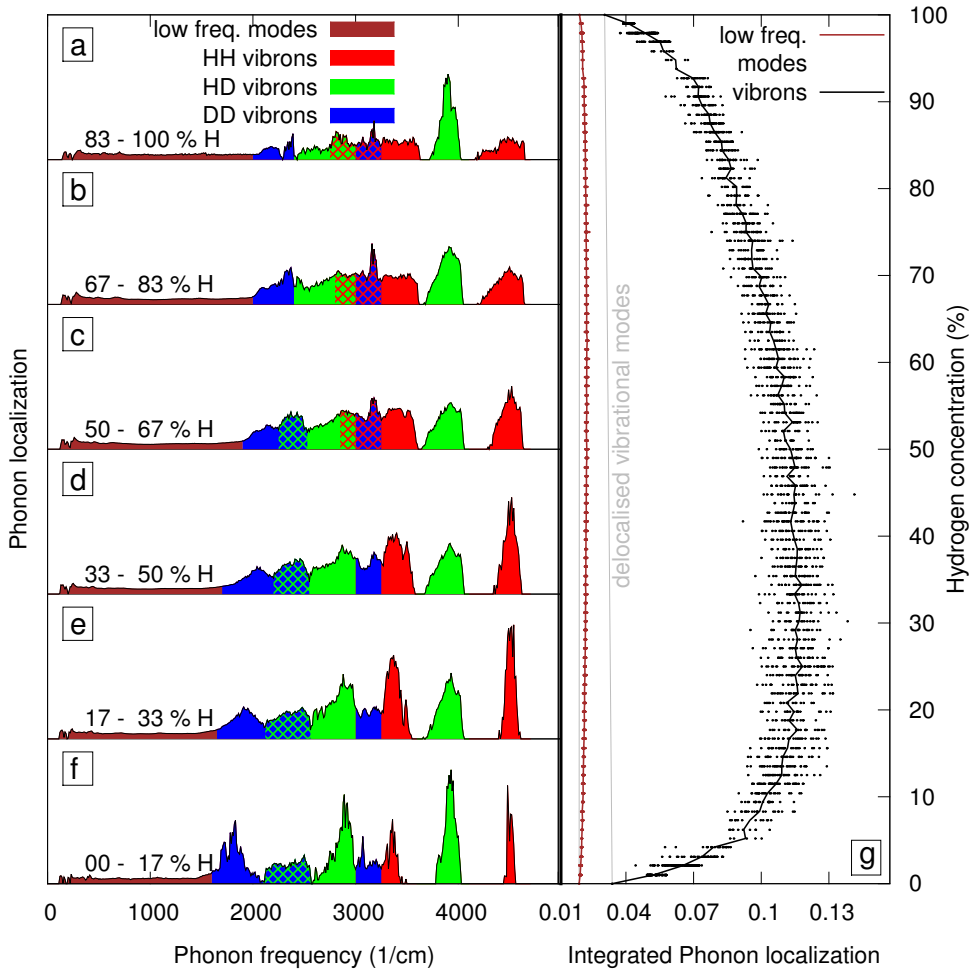

Figure 5.8 *Here I show in more detail the phonon localization calculated with LD at 270 GPa. Panels a-f report the localization per mode over intervals of decreasing hydrogen concentration. Shaded brown are the low frequency modes, which show no localization. Red, green and blue colors indicate vibrons involving mainly HH, HD and DD. I assigned the colors to peaks based on the Raman result in figure 5.10. The hashed regions illustrate overlap. Panel g shows the total localization as a function of hydrogen concentration for vibrons (black) and low frequency modes (brown), all at 270 GPa. In light gray I show the limit of entirely delocalized phonons generated by pure hydrogen and pure deuterium solids.*

HDG, DDG), however, in those cases, it is harder to analyze the peaks since all the bands merge together into one. Figure 5.8 shows all the modes not only the Raman active ones. Usually, the Raman vibrons are located on the lower frequency side within a vibron band, therefore, both softening and broadening of the band will generally result in softening the frequency of the Raman active vibrons.

Another very important observation is that low frequency modes are delocalized. As one can see in figure 5.8, the low frequency range of the spectrum comprises of homogeneous delocalized modes. This conclusion is also consistent with the experiment which finds only four Raman low frequency modes in phase IV of mixtures, the same number as in pure hydrogen/deuterium phase IV. If these modes were also localized, the experiment would find twelve of them corresponding to the three molecular species. Localization is the result of a trade-off between the strength of coupling and the mass ratio between different oscillating motives, this can easily be understood with the coupling model of oscillators. The transfer of energy from one oscillator to another is enhanced by stronger coupling and decreased by the difference in mass. This explains why vibrons are localized in space, while low frequency modes are not.

Figure 5.8 also shows integrated localization as function of H concentration at 270 GPa. Once more this shows the difference between the highly localized vibrons and delocalized low frequency modes. Vibrons of a certain molecular species are more localized at smaller concentrations of that species. However, the integral localization which is an average over all the vibrons, achieves a maximum when most of the vibrons are localized. This occurs at intermediate H concentrations. The curve is skewed with the maximum located around 20-30% H instead of 50% H. This is because HH vibrons are more localized at low hydrogen concentrations than DD vibrons are at low deuterium concentrations.

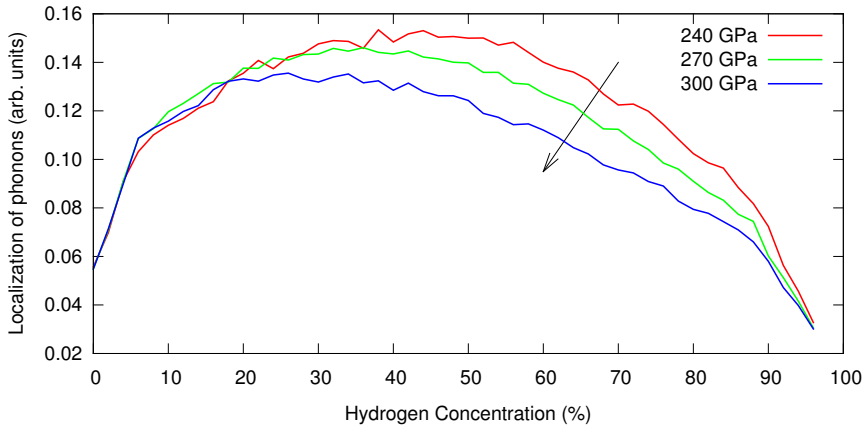

Figure 5.9 *Similar to panel g in figure 5.8 this shows the overall mode localization as function of hydrogen concentration calculated from LD. Here the result is also showed at 240 and 300 GPa for comparison. The arrow emphasizes the change in overall localization upon increasing pressure from 240 to 300 GPa.*

The shape of this overall vibron localization changes with pressure as showed in

figure 5.9. With increasing pressures, the B layers remain mostly unchanged, but in the G layers the difference between the intramolecular and intermolecular interaction strength is reduced and the DDG vibrons start coupling with other vibrons and reduce the overall localization. The HHG vibrons are affected less by changes in pressure, which is confirmed by experimental observations.

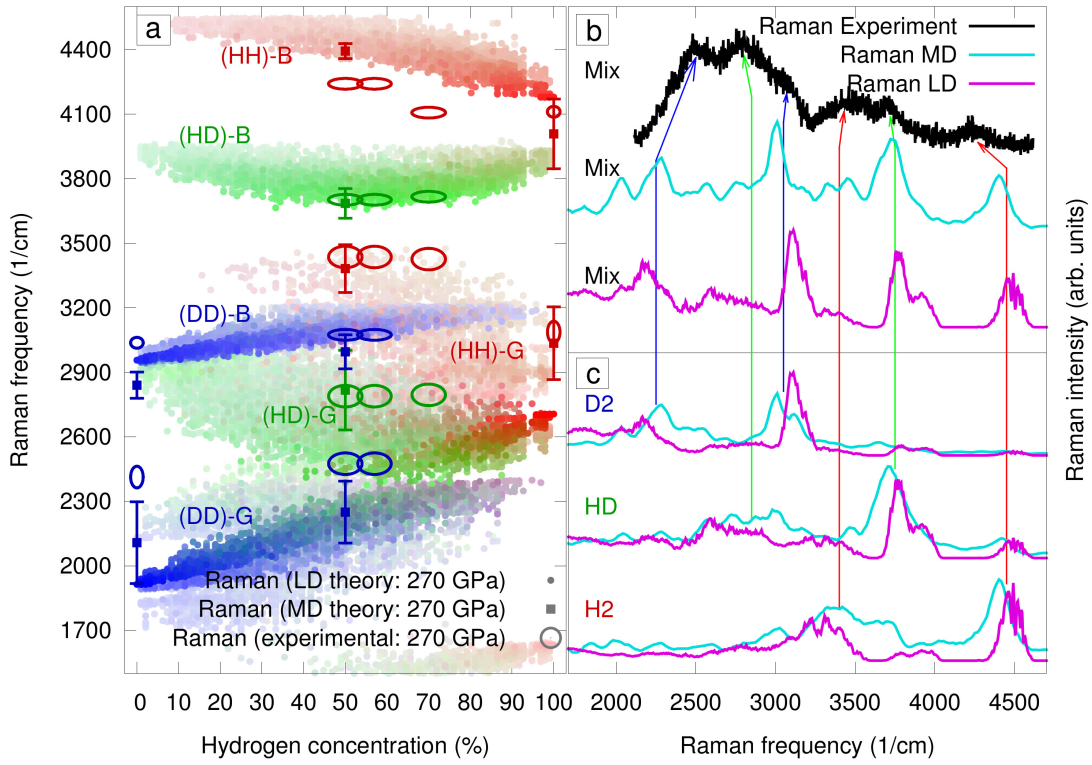

Figure 5.10 *Here I compare the Raman vibrons calculated at 270 GPa with the result from the experiment [2]. Panel a: Shows the Raman active modes from all the 2700 LD calculations (one colored solid dot per mode) at 270 GPa and various concentrations. The color of each dot is generated in RGB where the ratios of the color components are based on the partial Raman contributions from each molecular species: HH (red), HD (green) and DD (blue). The saturation of the color is proportional to the overall Raman intensity. The experimental data is plotted with ellipses, the size of which are the errors in concentration and frequency, respectively. Colors demonstrating the character of each mode in the experiment were assigned by observing the behavior of the spectra with pressure. With solid colored rectangles, I show the Raman frequencies extracted from the MD, where errors were assessed using the window method as explained in section 5.2.2. Panel b: The Raman spectra from LD, MD and experiment at 270 GPa in 50:50% hydrogen-deuterium mixtures. Panel c: Same spectra broken into individual contributions from different molecular species.*

Finally, in figure 5.10 I show a direct comparison between the theoretical and the experimental results in phase IV mixtures. All the observations above are nicely summarized in this figures. On the left, the pure vibrons originating from HH, HD and DD molecules are represented in red, green and blue, respectively, while all other coupled modes are mixed colors (e.g. brown and violet). It is clear that most vibrons originate from one type of molecules only and very little cross-coupling is observed.

The HHB and DDB bands are separated from the rest of the vibrons at all concentrations. As the H concentration increases, the proportion of HH molecules increases quadratically, the proportion of DD molecules decreases, while the proportion of HD molecules increases and then decreases with a maximum at 50% H. These changes are reflected in the frequency of the Raman vibrons through localization. When a molecular species is rare in mixtures, the vibrons from that species are more localized and their Raman frequency is higher. When the opposite is true, the vibrons are less localized and their Raman frequency is lower.

Overall, the agreement with experimental data is good. The computational results agree with the assignment of the Raman vibrons and reproduces well the changes induced by different mixture compositions. I attribute the small discrepancies in the positions of the peaks to anharmonicity [1] and inaccuracies of the PBE functional [85] as thoroughly discussed in chapter 3. Raman extracted from MD, shown on the right of figure 5.10 introduces corrections that shift the B vibrons down and the G vibrons up in frequency, which improves the agreement with experiment. The thorough analysis I gave here shows that we now understand well the nature of phase IV in both pure isotopes and mixtures. These results, further, support the BGBG layer models for phase IV and offer new insights into the changes induced by increasing pressure or varying isotopic compositions.

5.2.4 Conclusions

In this first half of the chapter, I presented my results and analysis of hydrogen-deuterium mixtures in phase IV. I carried out phonon and lattice dynamics calculations in the Pc structure, studying the effect of hydrogen-deuterium concentrations on the Raman spectra at three different pressure points: 240, 270 and 300 GPa. Additionally, I performed molecular dynamics simulations which were proven to be vital to the correct description of phase IV in pure isotopes

as I showed in chapter 3. MD calculations here improve the agreement with the experiment, showing once more that phase IV is dynamical in nature and cannot be entirely understood with static methods.

My results rule out the possibility of isotopic segregation in phase IV and show that mixtures are randomly distributed, at least at room temperatures. My conclusions are in agreement with the experimental observations. The measured Raman spectra contain six vibrons and four low frequency modes, which is consistent with a random distribution of isotopes in a Pc like crystal.

Furthermore, I found evidence of a theoretical concept called mass induced phonon localization which appears to be unique to molecular hydrogen-deuterium mixtures. While the vibron modes, which are found at high frequencies are highly localized, the low frequency modes are still extended crystal-wide. This explains why the experiment only detects four of these modes like in the case of pure isotopes, but six independent vibrons which correspond to each of the molecular species (i.e. HH, HD and DD) in each of the layer types (i.e. B and G).

Localization also explains the evolution of the Raman spectra with hydrogen composition, which can be measured directly in experiment. At low composition of a certain molecular species (e.g HH, HD or DD), that species gives rise to highly localized Raman vibrons which are higher in frequency. Conversely, at high compositions, the vibrons are more delocalized and their Raman frequency is lower. This is consistent with the experimental findings upon measuring samples with different hydrogen concentrations.

5.3 Mixtures in Phase III: Infrared Splitting

5.3.1 Introduction

In this second half of the chapter, I will focus on the new study presented in ref [114] which concerns phase III of hydrogen-deuterium mixtures at low temperatures. This experiment claims two new phase transitions in mixtures based on changes in the infrared spectra at high pressures, below room temperature. The study starts with pure hydrogen-deuteride (HD) and compresses the samples up to 350 GPa. Their findings are summarized in figure 5.11.

At low pressures they observe only vibrons originating from the HD molecule. Upon entering phase III at 150 GPa, a new vibron associated with the HH molecule is detected, yet no DD vibron is measured. Although their original figure showing the raw spectra, has been truncated at the location of the DD vibron. They attribute the new high frequency vibron to a 4% HH impurity in the original HD gas. The HD molecule has a permanent dipole moment so it is infrared active in phase II, on the other hand, the HH molecule only acquires a dipole moment in phase III.

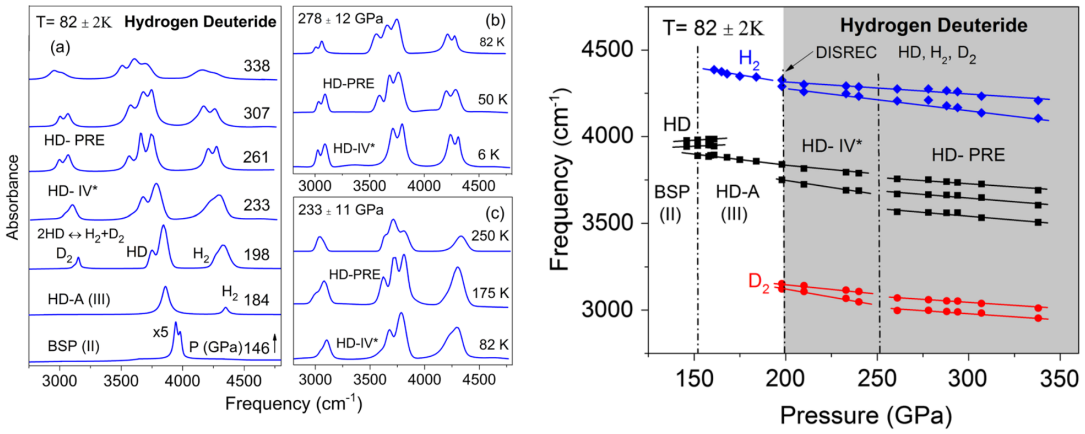

Figure 5.11 *This figure was adapted from ref [114] and shows the experimental infrared result in hydrogen-deuterium mixtures at 82 K. On the left is the fitted spectra as function of pressure and on the right are the positions of the peaks as function of pressure. The new phases were claimed on the basis of infrared splitting.*

Above 200 GPa, two new DD peaks appear and both HH and DD peaks split in two. These changes were associated with the onset of a new phase called HD-IV* in the experiment. They argue that at those conditions HD molecules also start a process of dissociation and recombination which leads to the formation of HH and DD molecules. In thermal equilibrium the HH, HD and DD molecules should be present in the ratio 1:2:1.

Around 250 GPa, yet another phase transition to HD-PRE was claimed, where the two HD vibrons split into three. The relatively small peak splittings of the order 50 cm^{-1} are close to the maximum experimental resolution which raises concerns over the accuracy of the results. It is also striking that the vibron frequencies, widths and their pressure dependence are unchanged at the claimed phase transitions. This indicates that further investigation is required to understand the behavior of the spectra.

Here I investigated hydrogen-deuterium (HD) mixtures in 50:50% concentrations

with the help of *ab initio* calculations. First I calculated the infrared spectrum in pure hydrogen using DFPT, then I used the method of isotopic substitution to extend the result to mixtures at pressures ranging from 150 to 350 GPa. The calculations were performed starting from $C2/c$ and $Cmca-12$, which are the proposed structural candidates for hydrogen phase III. The results suggest that the infrared splitting occurs naturally in $C2/c$ and this phenomenon can be explained in the absence of a structural phase transition. Additionally, the theoretical result fits well with the experiment which brings new convincing evidence that $C2/c$ is the best crystal structure of phase III, in both pure isotopes and mixtures. These results are soon to be published in a journal article [9].

5.3.2 Methods

All the calculations below use the same DFT setting as reported in the previous section 5.2.2. Here I found necessary to explain the exact procedure I used for normalization and smoothing of spectra in mixtures because I will draw conclusions based on small frequency splittings that might be sensitive to these methods. In the previous section, this was not an issue.

Infrared Calculations: Pure Hydrogen

I started with the geometry relaxation of the two structural candidates $C2/c$ and $Cmca-12$ [78] using pure hydrogen. The simulation boxes contained 24 hydrogen atoms, each, and the calculations were performed in the pressure range 150-350 GPa, every 20 GPa. For sampling the reciprocal space I used an MP k-point grid of 8x14x14 for $C2/c$ and 14x8x8 for $Cmca-12$. I also performed one additional CASTEP calculation with an increased cell of 288 atoms in $C2/c$ at 250 GPa, to study finite size effects. In this case, I used an MP grid of 4x4x6. To ensure the good convergence required by phonon calculations, I relaxed the structures using small thresholds: 10^{-9} eV/atom for energy, 10^{-6} eV/Å for force, 10^{-5} GPa for stress and 10^{-7} Å for the atomic displacement. No phase transformations were observed and all structures retained their starting symmetry.

I continued with phonon calculations using linear-response and FINE method interpolate. The calculation of Born effective charges was also enabled. All unspecified parameters were automatically set to the CASTEP default values.

The infrared spectra calculated at the gamma point ($q = 0$) for pure hydrogen structures were smoothed with non-normalised gaussians for illustration purpose:

$$I_{pure}(f_j) = \sum_{k=1}^{3N} i(f_k) \exp\left(\frac{-(f_k - f_j)^2}{2\sigma^2}\right) \tag{5.22}$$

where f_k are the $k = 3N$ frequencies calculated with CASTEP and $i(f_k)$ are the corresponding infrared intensities. I binned the data using the frequency discretization $f_j = 0 : 1 : 5000$ cm^{-1} ($5000 := N_B$). The gaussian standard deviation was set to $\sigma = 3$ cm^{-1}.

Infrared Calculations: Hydrogen-Deuterium Mixtures

I simulated the infrared spectra for $50 : 50\%$ hydrogen-deuterium mixtures using PHONONS, the CASTEP post-processing code. At every pressure point and for each of the two structural candidates $C2/c$ and $Cmca-12$, I generated $N_s = 3000$ random samples containing exactly 12 hydrogen atoms and 12 deuterium atoms at random sites. For the 288 atoms $C2/c$ calculation at 250 GPa I generated 1000 random samples. Unlike Raman (see section 5.2.2), the entire infrared spectra can easily be recalculated for every mixture sample s starting from the pure isotope dynamical matrix and using Born effective charges. For each mixture configuration, the PHONONS package was used to automatically compute new infrared spectra $i_s(f_k)$ as shown in figure 5.12. At every pressure point and for each structure, I built a histogram from the individual spectra and then normalized by the total number of samples:

$$H(f_j) = \frac{1}{N_s} \sum_{f_k \geq f_j}^{f_k < f_{j+1}} \sum_{s=1}^{N_s} i_s(f_k) \tag{5.23}$$

Finally, the data was smoothed with normalised gaussians:

$$I_{mixed}(f_{j'}) = \sum_{j=0}^{N_B} \frac{H(f_j)}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(f_j - f_{j'})^2}{2\sigma^2}\right) \tag{5.24}$$

where $f_{j'}$ represents the same discretization as f_j and $\sigma = 7$ cm^{-1}. The mixture spectra have no continuity, meaning that eigenmodes with similar frequencies

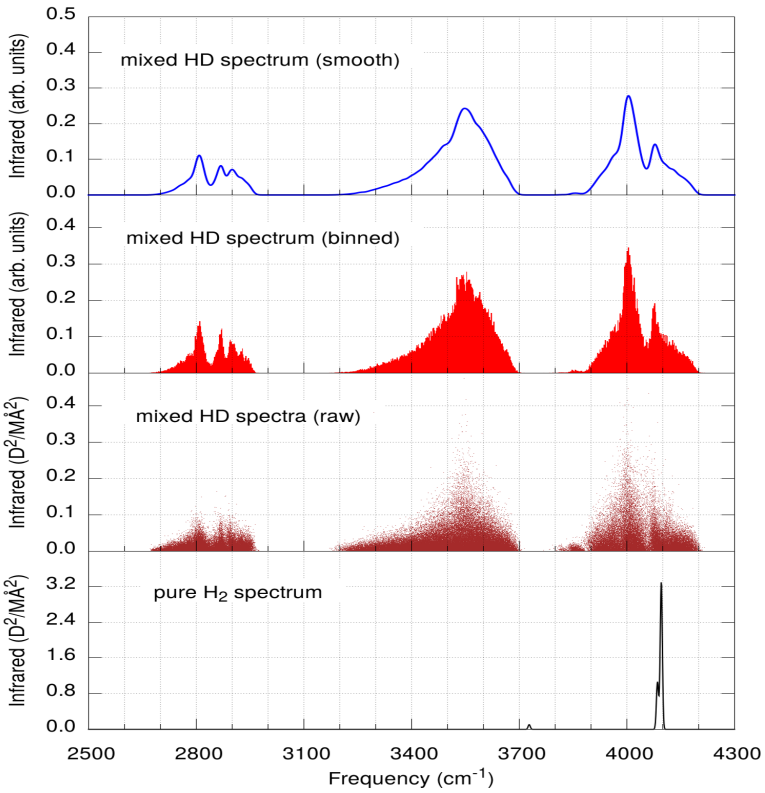

Figure 5.12 *The first panel bottom-up shows the high-frequency infrared spectrum of pure hydrogen $C2/c$ (288 atoms) at 250 GPa. The second panel shows with dots the spectra generated from 1000 random samples of hydrogen-deuterium mixtures calculated with PHONONS. Variable M in the measurement unit is the mass of the whole simulation cell expressed in atomic mass units. The third panel bottom-up shows the same data binned up in a histogram. The top most panel is the final data smoothed with gaussians.*

can have very different infrared intensities, therefore in the case of mixtures, it is not obvious what is the best way to normalize the infrared data. I found that after normalization, absolute intensities bear no meaning and only the relative intensities matter. The procedure I used to obtain the final spectra as explained above is also illustrated step by step in figure 5.12. The number of mixture samples used was sufficient to ensure the convergence of the spectra.

Localization and Raman Projections

The high-frequency Raman vibrons of each sample s can also be computed, by using the empirical rule introduced and tested in section 5.2.2 for the Pc structure. Namely, by projecting the phonon eigenmodes onto molecular stretching. The

final Raman spectrum at each pressure was binned up and smoothed, similarly to the infrared. I first collected all the $N_s = 3000$ samples into a histogram, then normalize and smoothed with gaussians:

$$R_{mixed}(f_{j'}) = \frac{1}{\sqrt{2\pi}N_s\sigma}\sum_{j=0}^{N_B}\exp\left(\frac{-(f_j - f_{j'})^2}{2\sigma^2}\right)\sum_{f_k\geq f_j}^{f_k<f_{j+1}}\sum_{s=1}^{N_s} r_s(f_k) \tag{5.25}$$

Similarly, the phonon localization $pl_s(f_k)$ for each sample was computed as in section 5.2.2. Additionally, here I also introduced a measure called hydrogen participation $hp_s(f_k)$:

$$hp_s(f_k) = \sum_{l_H=1}^{N/2}\sum_{\alpha=1}^{3}\frac{\varepsilon^{\alpha}_{l_H,k}\varepsilon^{\alpha}_{l_H,k}}{m_{l_H}} \tag{5.26}$$

The hydrogen participation sums selectively over the hydrogen atoms l_H and gives a measure of how much hydrogen participates to any given mode, compared to deuterium.

Unlike Raman and infrared data, localization and participation have continuity, so modes that are close in frequency have similar values. Therefore, when I constructed the histograms, instead of normalizing all bins with the total number of samples N_s, I divided each bin by the total number of values N_j that fall into that particular bin. N_j is different for each of the $N_B + 1$ bins and it can be smaller or larger than the number of mixture samples N_s. The final localization and hydrogen participation were obtained from:

$$PL_{mixed}(f_{j'}) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{j=0}^{N_B}\frac{1}{N_j}\exp\left(\frac{-(f_j - f_{j'})^2}{2\sigma^2}\right)\sum_{f_k\geq f_j}^{f_k<f_{j+1}}\sum_{s=1}^{N_s} pl_s(f_k) \tag{5.27}$$

$$HP_{mixed}(f_{j'}) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{j=0}^{N_B}\frac{1}{N_j}\exp\left(\frac{-(f_j - f_{j'})^2}{2\sigma^2}\right)\sum_{f_k\geq f_j}^{f_k<f_{j+1}}\sum_{s=1}^{N_s} hp_s(f_k) \tag{5.28}$$

Size Effects on Infrared Spectra

The initial cells I chose for $C2/c$ and $Cmca-12$ contained 24 atoms each, which is sufficiently large for an accurate infrared calculation in pure isotopes. In mixtures, however, a simulation box with 24 atoms only, could be too small. Finite size effects might be important for two reasons: poor sampling of randomness, on one hand, and the lack of long range interactions on the other. Every randomly mixed unit cell is in effect periodically repeated, which makes the whole crystal ordered at the large scale. The smaller the cell, the poorer the sampling and approximation of a random mass distribution.

One could address both issues simultaneously by simply performing phonon calculations on larger simulation boxes with CASTEP and then generating the mixture infrared spectra using the PHONONS code. However, this method proves to be very computationally costly. Although only the force constant matrix is needed to generalize the spectra to mixtures, CASTEP writes by default all the wave-functions to the checkfile which renders a massive file. The software PHONONS needs to load the checkfile for every random sample and this results in a huge amount of time and computing power being lost with reading and writing. This limitation restricts the number of mixture samples one can realistically calculate and leads to poor statistics. I performed, however, one expensive calculation with 288 atoms in $C2/c$ followed by 1000 randomly generated mixtures and obtained a converged spectrum at 250 GPa. Computing phonons on simulation boxes larger than 288 atoms is very costly.

I addressed the sampling issue by developing an in-house code. At the end of each phonon calculation, I prompted CASTEP to print the dynamical matrix at the gamma point and the Born effective charges in the main output file. I then expanded the simulation box and, with it, the dynamical matrix as shown on a simplified example in figure 5.13. If the interaction between atoms l and l' is within the unit cell (i.e. the distance between the atoms is less than half the cell size) the interaction is simply duplicated to all atom pairs in the image cells. If the interaction crosses a boundary that later becomes an interface, the interaction is then moved to the closest image atom in space, as best illustrated in figure 5.13.

This is a crude approximation and it does not extend the interaction cutoff, it simply deals with the sampling problem. An improvement would be to divide the interaction terms between the image atoms according to some power law,

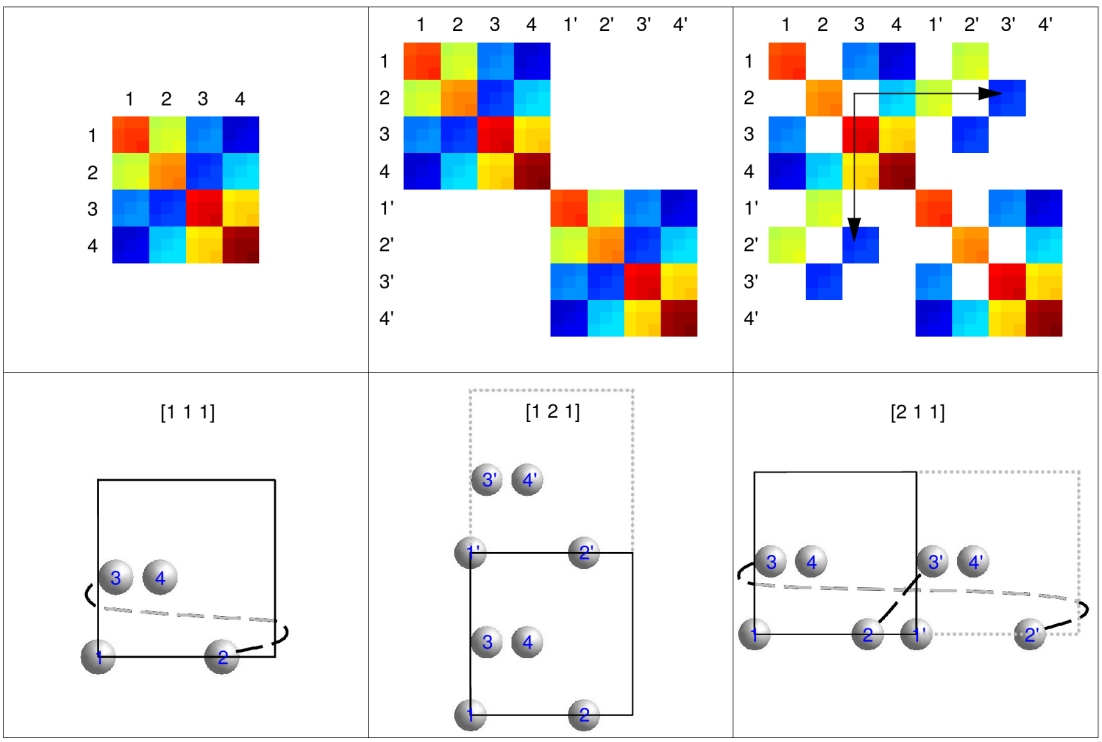

Figure 5.13 *This figure illustrates the method I used to expand the dynamical matrices while increasing the system size. On the bottom row, I show a model unit cell with four atoms. Atom 2 is connected to atom 3 via the vertical boundary. On the top row, I show the corresponding dynamical matrix. When the system is replicated along the y direction, the dynamical matrix is just repeated along the diagonal. However, when the system is reproduced along the vertical boundary, atom 2 becomes closer to the image 3' than to the original atom 3. This is then reflected in the new dynamical matrix by shifting the elements to non-diagonal blocks.*

with weaker interactions between more distant images. The distribution I used here is effectively a step function, which means that only the closest image atoms interact. The method is not perfect but it does allow testing for the size effects due to limitations in random sampling.

Following the expansion of the dynamical matrix to N_{ext} atoms, I simply generate mixed random samples of hydrogen-deuterium isotopes and re-normalize the matrix with the new atomic mases. I then diagonalize the matrix to extract the $k = 3N_{ext}$ eigenmodes. The infrared activity is then calculated by multiplying the eigenvectors with the Born effective charges $B_l^{\alpha,\alpha'}$ [145]:

$$i_s(f_k) = \sum_{\alpha'=1}^{3} \left| \sum_{\alpha=1}^{3} \sum_{l=1}^{N_{ext}} B_l^{\alpha,\alpha'} \varepsilon_{l,k}^{\alpha} \right|^2 \tag{5.29}$$

where α', like α, labels the three Cartesian components. To obtain the infrared spectra for mixtures I simply bin the data from many samples and then smoothen the normalized histogram as explained previously.

This method is much faster than the one involving the PHONONS software and it allowed me to treat larger cells with larger number of samples. I started from the dynamical matrix computed in the 24 atoms $C2/c$ phonon calculation at 250 GPa and expanded the simulation to different sizes as summarized by table 5.4.

Table 5.4 *The table shows the cell setups I used to study the size effects with the in-house code.*

FC expand	cell sides: a b c	multiply	number of atoms	number of samples
1 1 1	5.07 2.89 4.25	1	24	6000
1 2 1	5.07 5.77 4.25	2	48	5000
1 2 2	5.07 5.77 8.49	4	96	4000
1 3 2	5.07 8.66 8.49	6	144	3600
2 3 2	10.14 8.66 8.49	12	288	2000
2 4 2	10.14 11.54 8.49	16	384	1600
2 4 3	10.14 11.54 12.74	24	576	1200

Finally, I employed this method to extend the previous calculations in both $C2/c$ and $Cmca-12$ to larger simulation cells of 96 atoms, at all pressure points.

5.3.3 Results

In the case of pure hydrogen, I found two distinct infrared vibrons labeled ν_1 and ν_2 in both $C2/c$ and $Cmca-12$. The corresponding eigenmodes are illustrated in figure 5.14 and their frequency and pressure evolution are shown on the left of figure 5.15.

In $Cmca-12$, both vibrons involve eight out of the total of twelve molecules, where a group of four is out of phase with the other four. The peaks corresponding to the two modes are well resolved at all pressures and they reside more than 200 cm^{-1} apart. On the other hand, in $C2/c$ where I also found two distinct infrared peaks, the two are almost degenerate at low pressures and move further apart as the pressure is increased (figure 5.15). When inspecting the eigenmodes, it is clear that the two have very different symmetries. While ν_2 comprises of eight molecules vibrating out of phase in two groups of four, ν_1 only involves four of the molecules, two in each layer. This explains the different pressure evolution of the two peaks.

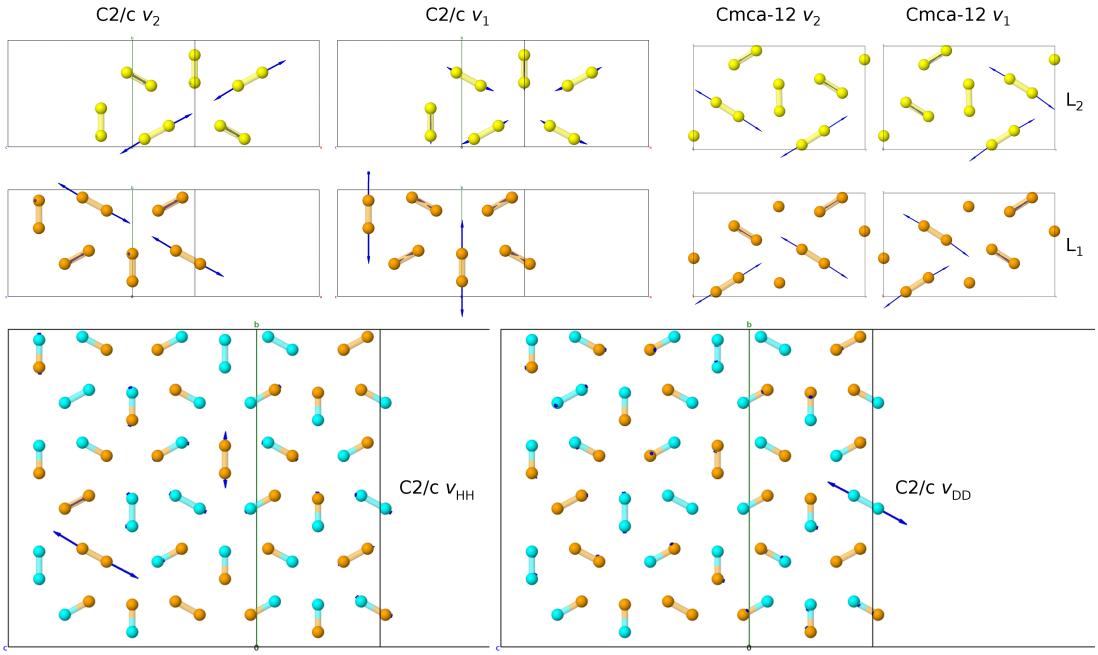

Figure 5.14 *In the top half I illustrate the most intense infrared active vibrons at 250 GPa, generated by the pure hydrogen $C2/c$ and $Cmca-12$ structures with 24 atoms. The unit cells comprise of two layers, which I show separately for clarity (L_1 - orange and L_2 - yellow). On the bottom half of the figure I show one layer from the extended $C2/c$ cell (288 atoms) of hydrogen-deuterium mixtures at 250 GPa. Hydrogen is displayed in orange, while deuterium in cyan. The two vibron modes (left and right) shown as examples are not the only infrared active ones but they are representative for the spectrum generated by mixtures. Note that the out of phase eigenvectors are less visible because they are masked by the bond.*

In the recent experimental study [114] new phase transitions were claimed based on the splitting of infrared peaks. Here I show a trivial example where infrared splitting in pure hydrogen does not require a change of symmetry. The two modes in $C2/c$ are almost degenerate in frequency at low pressures, so the experimental resolution would not be sufficient to resolve them separately. Owing to their different characteristics, the frequency of the two peaks have different pressure evolutions, and past 250 GPa they start to separate. At 350 GPa the two modes are about 30 cm^{-1} apart, while based on the raw data [114] the experiment appears to achieve a frequency resolution around 35 cm^{-1} . In this case, the splitting of infrared coincides with a pressure induced lift of accidental degeneracy and does not imply a change of cell symmetry. This is a reminder that the optical signature is only a weak indicator for crystal structure and not every change in the infrared, Raman or optical absorption indicates a phase transformation.

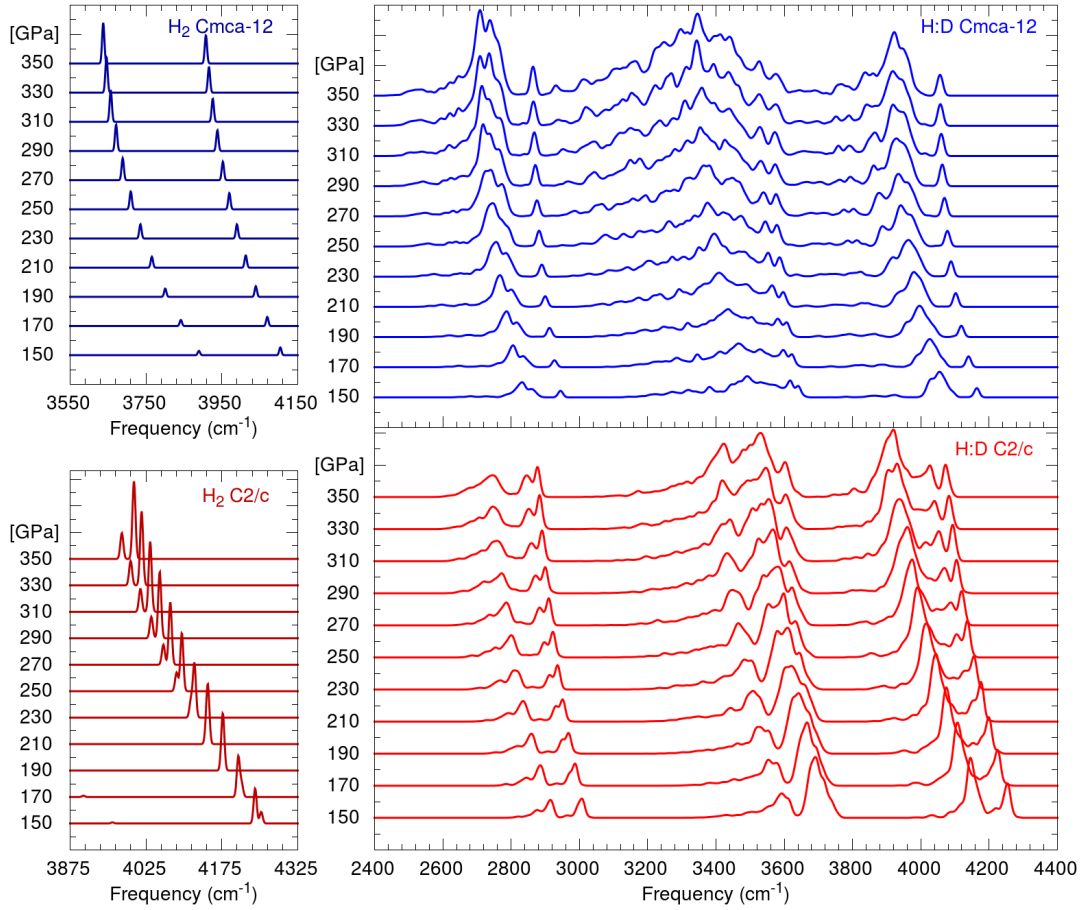

Figure 5.15 *Here I show the pressure evolution of the high-frequency infrared spectra for the two candidate structures $Cmca-12$ and $C2/c$ as calculated from the cells with 24 atoms. Pure hydrogen spectra on the left and $50:50\%$ mixture spectra, calculated with PHONONS, on the right.*

Based on the pure hydrogen infrared spectra in $C2/c$ one would expect three peaks in mixtures, corresponding to the HH, HD and DD molecules, respectively. However, here, I find six separate peaks in the high frequency regime of $C2/c$ at 200 GPa and up to nine at 350 GPa as illustrated in figure 5.15. The peak splitting at high pressures could be related to the lift of degeneracy explained above because it follows a similar pressure dependency. However, the individual peaks at low pressures are likely induced by mass disorder and provide a reliable experimental fingerprint for the $C2/c$ structure.

In $Cmca-12$ mixtures I find a large collection of infrared peaks especially in the frequency region of HD. The spectrum originating from HH and DD molecules can be separated into one higher frequency narrower peak and a broader lower

frequency collection of peaks. The two signals could be resulting from the two mode symmetries ν_1 and ν_2 in pure hydrogen $Cmca-12$.

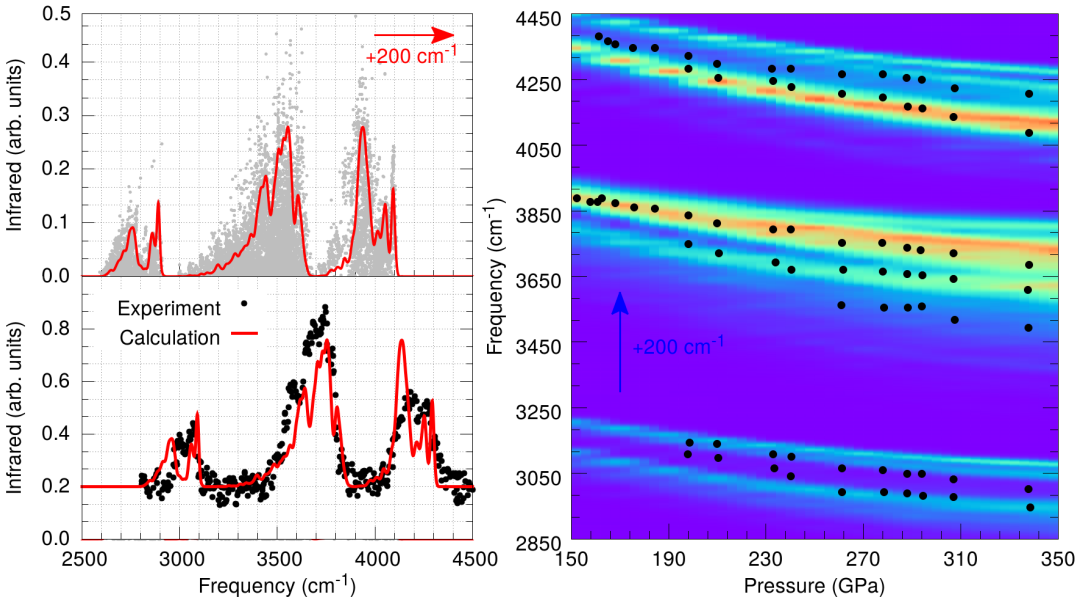

Figure 5.16 *Here I compare the computational infrared result, reported also in figure 5.15, to the recent experimental data from ref. [114]. Top left I show the mixture spectra calculated with PHONONS at 300 GPa for 24 atoms $C2/c$ (3000 randomized samples shown in gray). Bottom left is the smoothed result shifted to higher frequencies by 200 cm^{-1} and then compared to the experimental data at 307 GPa which was digitized from ref [114] using an online tool [124]. On the right I show as a color map the pressure dependence of the calculated infrared spectra shifted 200 cm^{-1} up in frequency, while in black the experimental pressure dependence of the infrared spectra from ref [114].*

Based on the result presented in figure 5.15, it is clear that $C2/c$ is the better candidate for the experimental phase III in mixtures [114], therefore in figure 5.16 I compare the experimental infrared to this theoretical result. Generated spectra appear to be 200 cm^{-1} softer than experimental spectra. This discrepancy can be explained by a combination of errors derived from the choice of exchange-correlation functional, the harmonic approximation and the experimental choice of pressure gauge [100]. Ref [85] shows that the choice of functional can change the position of the spectra by as much as 500 cm^{-1}, while my result from molecular dynamics reveals an extra 100 cm^{-1} correction in $C2/c$ at 200 GPa when anharmonicity is accounted for [1]. More important than the exact frequency match is the shape of the spectrum which is surprisingly similar to the experimental data, so I shifted the calculated infrared 200 cm^{-1} up in frequency and compared it to the experiment in figure 5.16.

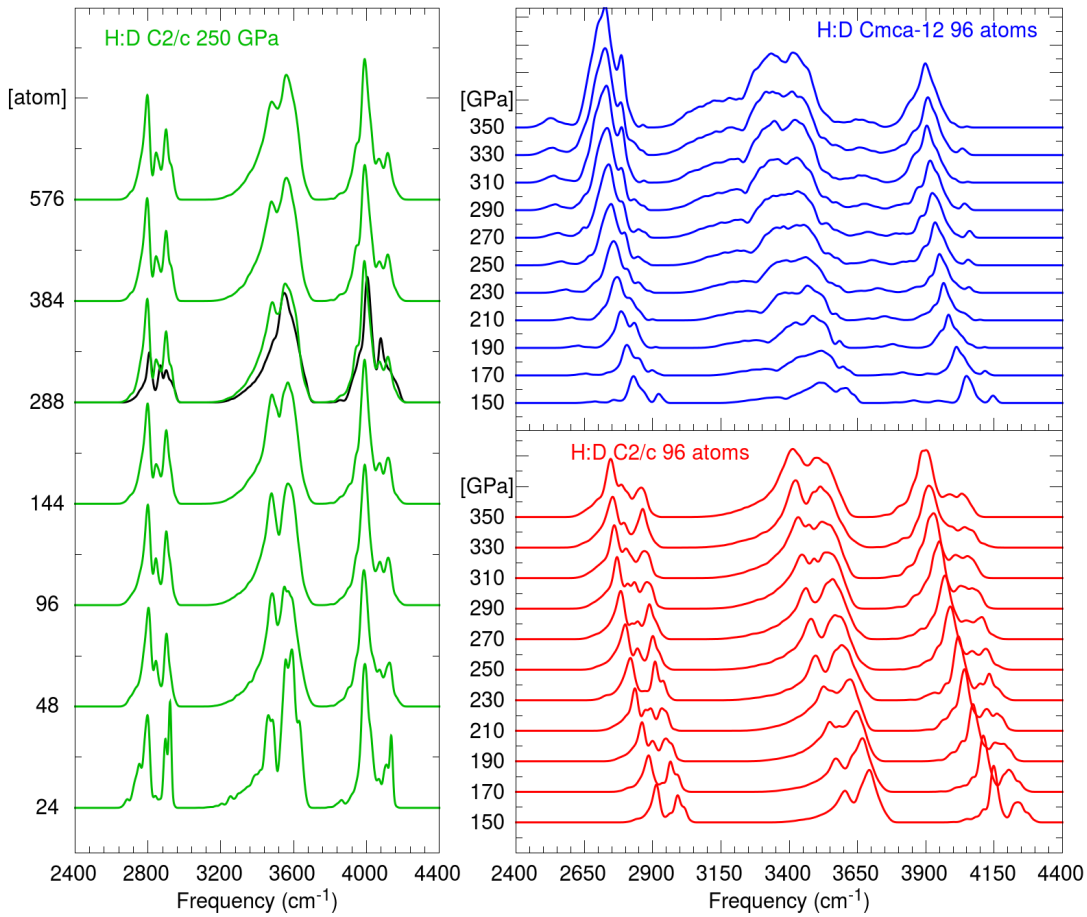

Figure 5.17 *This is the infrared spectra of mixtures calculated with my own code. The code starts from the original 24 atoms dynamical matrix calculated with CASTEP and extends it as illustrated in figure 5.13. On the left, I show the spectra in $C2/c$ for increasing cell sizes at 250 GPa. The number of atoms is shown along the y axis, while the number of random samples was reported in table 5.4. The one spectrum shown in black is the expensive calculation, where the dynamical matrix was actually calculated with CASTEP on a 288 atom cell and 1000 mixtures were generated with the PHONONS code. On the right, I report the pressure dependence of the infrared spectra as calculated with my own code on 96 atom extensions of the original cells in both $C2/c$ and $Cmca-12$.*

Notice that the calculated spectra correctly reproduces the width and shape of the peaks which result from mass disorder and phonon localization. Furthermore, the calculation agrees with the experiment on the number of peaks as shown in the left panel on figure 5.16. The pressure evolution of the experimental spectra is also well fitted by the theoretical result. The lack of DD signal at lower pressures in the experiment, most probably indicates incomplete equilibration of the isotopes. Since the experiment starts with a pure HD structure with stable molecules,

the equilibration of an HH, HD and DD mixture requires the dissociation and recombination of molecules which takes a long time to occur, especially at lower pressures. The infrared activity of HH is higher than that of DD, therefore the HH signal is measurable at relatively lower concentrations. At higher pressure, the molecule recombination is faster which facilitates the transition from pure HD to equilibrium mixtures, where DD becomes visible.

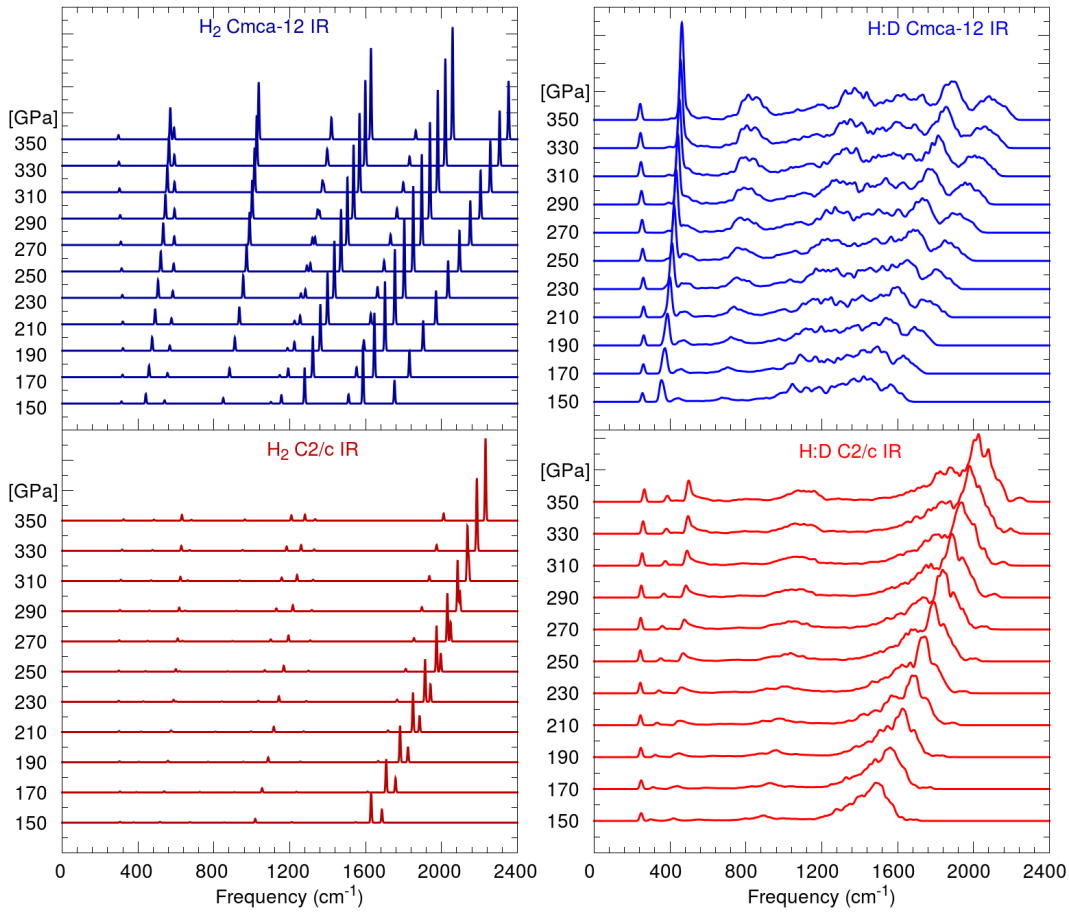

Figure 5.18 *This figure shows the same result as figure 5.15 but in the low frequency range. I report the pressure evolution of the infrared spectra for $C2/c$ and $Cmca-12$ (24 atoms). On the left, I present the spectra for pure hydrogen, and on the right, I show the data for* $50:50\%$ *mixtures calculated with PHONONS from 3000 random samples.*

The conclusions presented above hold if the theoretical result is robust. Here I investigated the possibility of finite size effects (see figure 5.17). The chosen cells for both $C2/c$ and $Cmca-12$ contain 24 atoms each, equivalent to twelve molecules. Since both crystal structures have two layers, the molecules are further divided: six in each layer. When hydrogen-deuterium mixtures are in

thermodynamical equilibrium the HH, HD and DD molecules occur with the probabilities 25%, 50% and 25%, respectively. As a result, in every given mixture sample, one is likely to find only three HH molecules in the entire cell, 1.5 in each layer. This small number of molecules raised concerns over the validity of the previous results. In figure 5.17 I investigated the infrared spectrum of $C2/c$ as function of size at 250 GPa and as a function of pressure for a 96 atoms simulation cell. The main result is unchanged: the HH, HD and DD contributions are each divided into distinct peaks. The separation between the peaks varies slightly with size but not significantly and the pressure variation is only slightly changed compared to figure 5.15. I conclude that despite the small number of molecules of similar type, due to phonon localization, only small groups of molecules contributes anyway and thus finite size effects are minimal.

In figure 5.18 I present the lower frequency infrared spectra in both $C2/c$ and $Cmca-12$, mixtures and pure hydrogen. Although presently these are experimentally inaccessible, they could become relevant for structure discrimination in the future. On one hand, the $C2/c$ pure hydrogen spectra feature two main modes that become degenerate at higher pressure. The mixtures infrared is very similar except that modes are much broader (i.e. around 300 cm^{-1}) as a result of mass disorder. On the other hand, the pure hydrogen $Cmca-12$ generates rich infrared spectra comprising of a number of peaks spread across the entire frequency range. In mixtures, the higher frequency phonons are broaden in an almost continuous band and only two of the lower frequencies persist as narrow, well defined peaks.

Hydrogen participation presented in figure 5.19, shows that in the high frequency range there is a clear distinction between HH, HD and DD vibrons, whereas in the low frequency regime the modes become gradually hydrogen dominated at progressively higher frequencies.

The localization data implies that, like in Pc [2], the high frequency modes are also localized in $C2/c$ and $Cmca-12$. For instance, the HH vibrons have a localization value of the order 0.25, which corresponds to roughly 2 molecules out of 12 participating to the mode. This is consistent with visual inspections of the modes, shown in figure 5.14, where very few molecules participate in each mode. The localization alongside mass disorder is responsible for the broadening of the modes. A delocalized mode means that atoms of the same type, from across the entire crystal, can easily couple which reduces the frequency and increases the widths of the peaks [110]. On the other hand, a localized mode is confined in

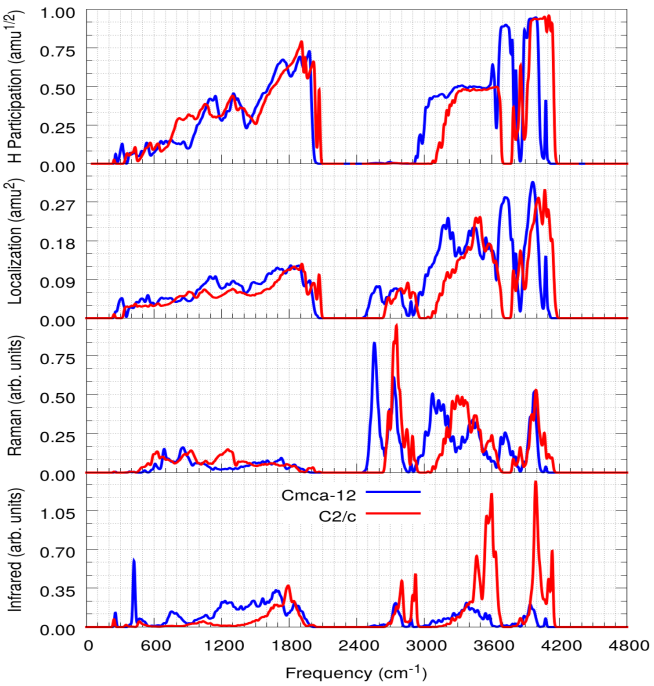

Figure 5.19 *In this figure I show for comparison the hydrogen participation, localization, Raman (molecular projection) and infrared for $C2/c$ and $Cmca-12$, beside each other. The calculation was performed with mixtures at 250 GPa, in 24 atoms cells. The Raman spectra are just approximations obtained from the molecular projection of in-phase vibrations.*

space and involves only a small group of molecules which, then, generates sharper and higher frequency peaks.

The amount of mode localization is the result of a trade-off between the isotopic mass ratio and the general strength of the intermolecular interaction as can be seen in figure 5.20. At low pressures, the mass difference wins which enhances the localization. At higher pressures, the amount of localization is reduced as the intermolecular forces increase.

Finally, in figure 5.20 I also show the Raman vibrons and their evolution with pressure, computed using the molecular projection method in both $C2/c$ and $Cmca-12$. Previously, for phase IV, I validated that molecular projection gives a good estimation of the high frequency Raman spectrum in mixtures. Here, the approximation holds in $C2/c$ but fails in $Cmca-12$. The study presented in ref [91] found only one Raman active vibron in pure hydrogen $Cmca-12$, whereas, here, I found two using the molecular projection technique. Upon careful inspection of these modes in $Cmca-12$, I discovered that only the actual Raman

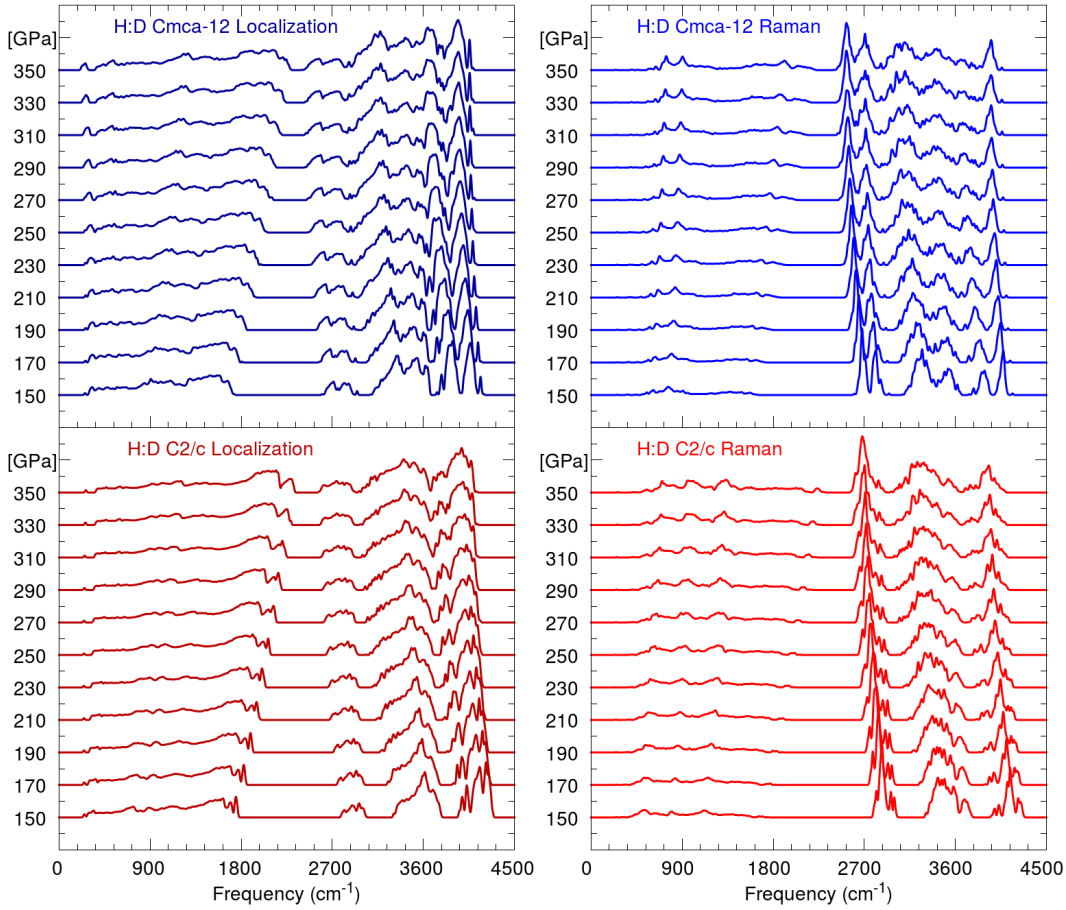

Figure 5.20 *Here I report the localization (left) and molecular projection (right) results for 24 atoms cells $Cmca-12$ (top) and $C2/c$ (bottom). The result is from mixtures which were generated with the software PHONONS. The spectra are stacked up in pressure as indicated along the y-axis.*

active mode comprises of in-phase stretching of all the molecules, the other higher frequency mode consists of two thirds of the molecules vibrating in-phase and one third out of phase. The projection method, however, wrongly identifies both modes as Raman active for pure hydrogen, which leads to six distinct modes identified as Raman active in $Cmca-12$ mixtures, three of which are artifacts. In the case of $C2/c$, the projection method correctly finds the in-phase molecular stretching as Raman active and predicts the high frequency Raman spectra in mixtures, which is yet to be measured experimentally.

5.3.4 Conclusions

In this second half of the chapter, I showed my results, mainly on the pressure evolution of the infrared spectra for hydrogen-deuterium mixtures in phase III. All calculations were performed on the structures $C2/c$ and $Cmca-12$ in the pressure range 150-350 GPa. At low pressures, I find six clear distinct infrared peaks in $C2/c$, two for each molecular species. At higher pressures (above 230 GPa), all peaks split further, but most notably the HD signal splits into three peaks. This spectrum transformations are consistent with those found experimentally [114]. Whereas the experimental study associates the infrared splitting with new structural phase transitions to phases IV* and REC, my analysis shows that these changes are a result of mass disorder and phonon localization and are not associated with structural transformations. Moreover, my results bring new, more reliable evidence that $C2/c$ is the right crystal candidate for phase III. As I showed in the case of phase IV, the entropy associated with mass disorder is small compared to k_BT, meaning that hydrogen-deuterium mixtures are likely to manifest the same crystal structures as pure isotopes. $C2/c$ is, therefore, the most likely candidate for phase III in both pure hydrogen and hydrogen-deuterium mixtures.

My hypothesis is that the experimental phases III, IV* and IV-REC are all in fact phase III, represented by the $C2/c$ crystal structure. According to this theory, the experiment should have observed six peaks from the onset of phase III at 150 GPa, which is not the case. It could be that the peaks are weak because mixtures are far from the equilibrium concentrations and only small traces of DD and HH molecules are present at the start. Another possibility is that the peaks are too close in frequency and the resolving power of the measuring device is not sufficient to distinguish them. The first idea could be tested in practice by running simulations at difference concentrations. Better quality simulations with more atoms in the cells could also give better estimates for the experimental measurements.

5.4 Summary

In this chapter, I summarized my work on hydrogen-deuterium mixtures, which consists of two studies centered around recent experimental work [2, 114].

The first study targeted hydrogen-deuterium mixtures in phase IV. Using a series of phonon and molecular dynamics calculations I show that the Raman modifications in phase IV mixtures can be understood in terms of phonon localization. This phenomenon originally proposed by Anderson [140] for electrons was realized before for sound, light and Bose-Einstein condensates in randomly disordered systems. This work reports for the first time the localization of phonons, which in this case is purely induced by mass disorder. Neglecting zero point energy, the strength of the bond and the intermolecular forces are approximately the same in all three species: HH, HD and DD. Therefore the localization is effectively the result of the large mass difference between the different isotopes.

In the second study, reported in the second half of the chapter, I performed phonon calculations in phase III mixtures, mainly to investigate the evolution of the infrared spectra with pressure. The simulation results are in good agreement with the experimental data but suggest a different interpretation. Whereas the experimental study concludes that the infrared splitting at high pressure is indicative of a series of structural phase transitions, my results show that the splitting occurs naturally in phase III as a result of mass disorder. Additionally, the spectra calculated in $C2/c$ match the experiment better than those computed in $Cmca-12$, bringing new evidence that the former crystal structure is a better candidate for phase III.

The takeaway message is that spectra in mixtures are a lot richer than those in pure isotope homologues and can help us learn a lot more about the intramolecular and intermolecular interactions and the crystal structures in general. So far mixtures have received little attention from experiment and even less from theory. My hope is that this initial studies will spark interest in this direction of research.

Chapter 6

Conclusions

6.1 Concluding Remarks

The science of high pressure is rapidly expanding and every year brings new exciting discoveries. Most elements have been exhaustively compressed to very high pressures, but mixtures of different substances offer, still, a vast territory of exploration in both experimental and theoretical condensed matter. For instance, just last year Drozdov *et al.* discovered superconductivity at record high 203 K in sulfur hydride under pressure [146].

Elemental hydrogen remains, however, one of the most studied topics in high pressure science. To date, five solid phases are known and at least two liquid phases, one of which is metallic. Solid hydrogen is yet to be metalized, and many different groups around the world are getting closer to the metallic phase every year. The phase diagram has some remarkable elements like a negative melting curve over an extended pressure, at least two entropy stabilized solids and possibly a ground state metallic liquid at very high pressures.

In this thesis, I reported my contribution to understanding the finite temperature behavior of the high pressure solid hydrogen phases. Additionally, I showed my results derived for hydrogen-deuterium mixtures which have recently sparked the interest of the experimental community. The zero temperature structure candidates for phases III and IV were previously identified in AIRSS studies [78, 91]. In my work, I focused on investigating the dynamical behavior and the connection to the experimental data through spectroscopy. In recent years

most theoretical efforts have been redirected at increasing the accuracy of energy calculations by adding layers of approximation. Most of the ongoing studies have migrated from the use of DFT to QMC on account of more accurate energetics. While this it very important for obtaining competitive crystal structure candidates, the bridge to experiment is also essential, and much less studied. While QMC is without a doubt more accurate than DFT, it is also more computationally expensive. DFT allows us to carry out long molecular dynamics simulations with large system sizes, which are essential to understanding the phases stabilized by entropy.

In chapter 3, I presented my work on solid hydrogen in phases III, IV and newly discovered phase V. The main result is that phases IV and V are similar to one another and have a dynamical behavior which cannot be understood in terms of a fixed crystal configuration. These structures comprise of different layer types, called B and G. In the B layer the molecules are almost freely rotating, similar to the quantum rotors in phase I. The B layer is quite robust and remains largely unchanged over the pressure range 250-380 GPa. Conversely, the G layer comprises of relatively weakly bonded molecules and goes through a series of transformations, gradually allowing the rotation of molecular trimers, rebonding of molecules and eventually, at high enough pressures, proton diffusion [97]. In this chapter, I also reported a new method for extracting the Raman vibron from molecular dynamics trajectories, which reduces the discrepancy between theory and experiment and provides new evidence that the nature of these phases is truly dynamical.

At higher pressure, beyond the currently known phases, the electron charge will eventually move away from the covalent bond as it has been originally predicted [19]. In that case, hydrogen can only remain a poor conductor if the electron charge localizes elsewhere (e.g. in the lattice interstices) or if the crystal organizes itself to screen the ion-ion interactions by scattering the charge density at $2k_F$ and effectively opening a pseudo-gap.

In chapter 4, I studied the behavior of hydrogen at the upper pressure limit that can be achieved currently in the experiment - 400 GPa. In molecular dynamics, many of the unstable atomic candidates, all transformed to a similar structure comprising of long polymeric chains of atoms. By comparing the XRD patterns with DoS, I showed that *Chains* could be stabilized by a charge density wave evident by the clustering of diffraction peaks near $2k_F$. Surprisingly some of the layered structures exhibit a similar behavior, which could explain how hydrogen

managed to remain a semiconductor to such high pressures. The structure *Chains*, which has been found by others in simulations before, could also be a spurious effect of sparse k-point sampling. Interestingly, this setting might be more important in MD that was previously believed and this issue needs to be addressed in future studies. At the end of the chapter, I also presented a simple thermodynamical model that summarizes the essential physics of the phase diagram and gives a broader perspective.

Finally, in chapter 5, I focused on hydrogen-deuterium mixtures where I found evidence of a textbook phenomenon - phonon localization. Owing to the large mass ratio between hydrogen and deuterium, the concept of a phonon band is broken in random mixtures, especially in the high frequency regime. Each vibron consists of a small group of vibrating molecules, which are decoupled from the rest of the system. This phenomenon is observed in phonon calculations but cannot be directly measured in the experiment. However, the dependence of the Raman spectra on the mixture composition and pressure, which was observed experimentally, is linked to phonon localization and implicitly proves it. The main message from this chapter was that hydrogen-deuterium mixtures are a much richer system to study than pure solids. For instance, by comparing the calculated infrared spectra with recent experimental results in phase III mixtures, I showed new evidence in support of the previously discovered structural candidate for this phase.

Overall, my work contributed to a better understanding of the different solid phases at high pressure, in both hydrogen and hydrogen-deuterium mixtures. Additionally, it helped us assess the quality and limitations of different theoretical methods, especially DFT-MD. All my results were published in, or will be submitted soon to journal articles and conference proceedings, while this thesis brings all the pieces together.

6.2 Future Directions of Work

Finally, let me present a few idea that I wish I had time to explore. Hopefully, this will serve as inspiration for future studies on this line of research.

One interesting problem is the shift of the Raman vibron upon heating and melting at different pressures. In their study on melting, Howie *et al.* [22]

have shown that the sign of the Raman shift changes depending on pressure. This shift cannot be measured in DFPT, since this theory is not appropriate for entropy stabilized crystals and liquids. The method I developed and presented in chapter 3 could be used, instead, to study this behavior with molecular dynamics. However, an important improvement is needed to better identify molecules before projection. This is not a trivial task in systems such as liquid hydrogen at high pressure, where the molecules probably break and reform often during the dynamics.

Good methods for computing Raman and infrared from MD are essential, because as I showed before, many of the phases, like I, IV and V, are dynamic and can only be simulated at finite temperature. Unfortunately, the method that involves projecting the velocities onto normal modes fails in hydrogen, mainly because of the free rotations and diffusion present in the systems. One idea would be to develop methods that do not require a priori calculation of the modes, but use the rich information present in the charge density during the dynamics, that is otherwise ignored. One could create a method to estimate changes in the dipole moment and polarizability directly from the dynamics of the electron density.

Another method could attempt to extract spectroscopic information from a quantum molecular dynamics simulation. The standard PIMD does not generate meaningful dynamics, so frequency cannot be extracted from a simple Fourier Transform like in MD.

While most studies focus on the high pressure, there is still plenty to learn at lower pressures. For instance, to my knowledge, there is no algorithm that can distinguish between the para and ortho states of molecular hydrogen. One could envision a quantum molecular dynamics method that encodes for the nuclear spin, and then revisit the low pressure phase diagram.

The hydrogen-deuterium mixtures offer still more possibilities for future research. For a start, I would have liked to extend my study to other crystal structure candidates and obtain some predictions that could offer guidance for future experiments. While in most pure hydrogen structures one finds a small number of Raman and infrared active modes, the mixture analogs generate an almost continuous distribution that depends on pressure and composition and can be used as a fingerprint to identify the best candidates for each solid phase. There is also the prospect for an important algorithmic improvement: while infrared can be recomputed cheaply upon isotopic substitution, the Raman spectra is difficult

to recompute. One could develop an approximate method to speed up this task. For instance, I used molecular projection to estimate the Raman activity of the vibrons, which made it possible to sample many hydrogen-deuterium distributions and converge the mixture spectra. Computing the Raman activity of the low frequency modes requires a different method.

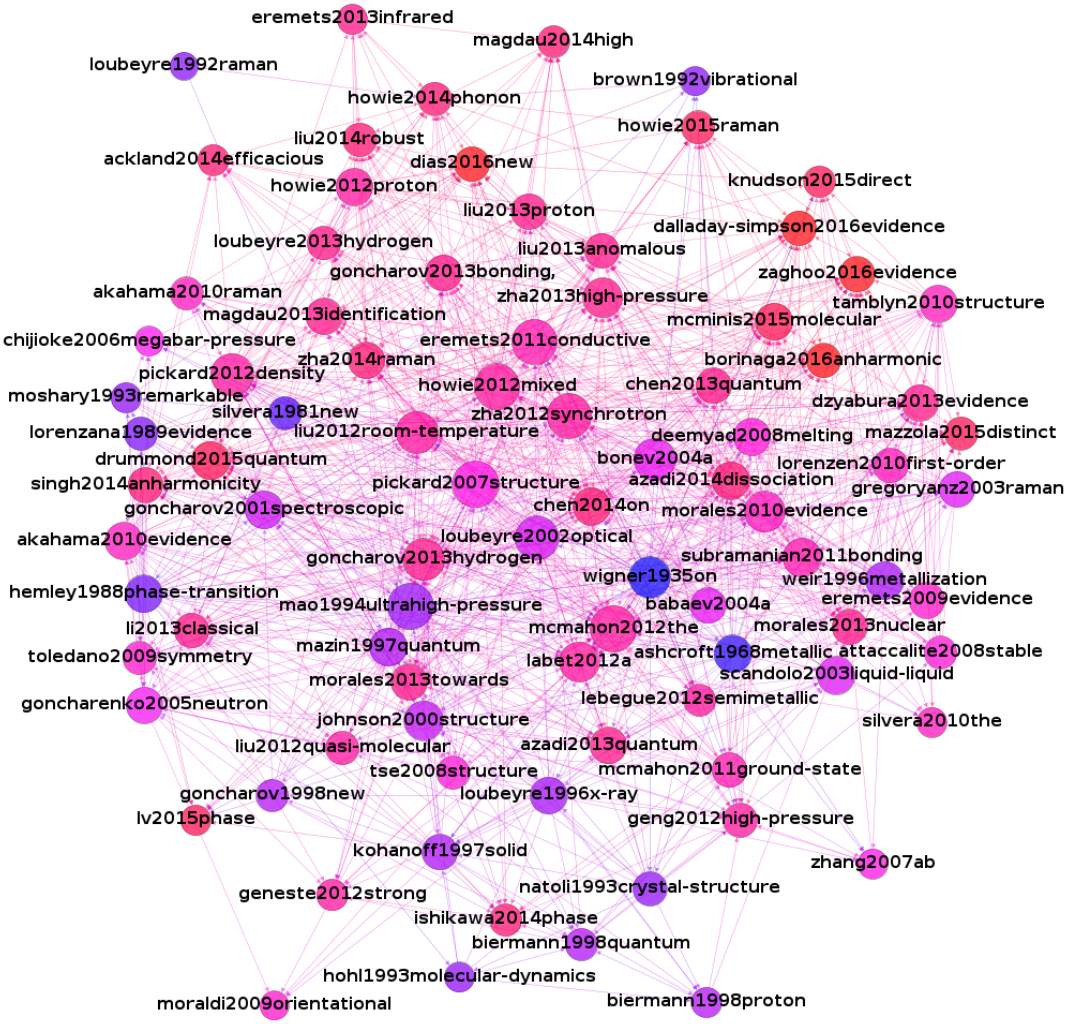

Figure 6.1 *The citation graph of some of the most relevant papers to this thesis. The color is related to the year of publication and the size to the connectivity (i.e. references and citations) in this subgraph.*

Finally, I think that the most important step that could substantially improve the quality of research in the field, is to build an open access wiki that summarizes all the essential results and which can then be kept up to date. This may seem like a tedious task but I believe it is crucial for synchronizing the collective work of all the different groups around the world. Especially in recent years, there has

been a myriad of conflicting results coming from both theory and experiment.

The page could contain an interactive phase diagram that collects all the experimental results together, it could summarize all the Raman and infrared spectra ever measured or calculated at all temperatures and all pressures. Subsequently, every new study could include their data and compare it against all previous work, at once. The page could show all crystal structure candidates with all associated, calculated properties from both DFT and QMC, such that future studies could easily benchmark their simulations. Researchers would notice at a glance where the gaps in the knowledge are, without having to browse large amounts of tedious literature.

Even a simple graph of papers and references like the one I show in figure 6.1 could show the literature interactively on the website and cluster the papers on similar topics: experiment/theory, solid/liquid etc. Newcomers in the field would take much less time to get familiar with the previous work.

Ideally, the website would provide space to upload data such as molecular dynamics trajectories or phonon calculations and means to visualize them. Additionally, a list of users could be put in place such that everyone that wants to stay updated with the field could sign up and get newsletters with the most recent related publications. I believe this wiki page could significantly increase research productivity in the field.

Bibliography

[1] I. B. Magdău and G. J. Ackland, "Identification of high-pressure phases III and IV in hydrogen: Simulating Raman spectra using molecular dynamics," *Physical Review B*, vol. 87, no. 17, p. 174110, 2013.

[2] R. T. Howie, I. B. Magdău, A. F. Goncharov, G. J. Ackland, and E. Gregoryanz, "Phonon localization by mass disorder in dense hydrogen-deuterium binary alloy," *Physical Review Letters*, vol. 113, no. 17, p. 175501, 2014.

[3] I. B. Magdău and G. J. Ackland, "High temperature Raman analysis of hydrogen phase IV from molecular dynamics," in *Journal of Physics: Conference Series*, vol. 500, p. 032012, IOP Publishing, 2014.

[4] G. J. Ackland and I. B. Magdău, "Efficacious calculation of Raman spectra in high pressure hydrogen," *High Pressure Research*, vol. 34, no. 2, pp. 198–204, 2014.

[5] G. J. Ackland and I. B. Magdău, "Appraisal of the realistic accuracy of molecular dynamics of high-pressure hydrogen," *Cogent Physics*, vol. 2, no. 1, p. 1049477, 2015.

[6] I. B. Magdău and G. J. Ackland, "Charge density wave in hydrogen at high pressure," *arXiv preprint arXiv:1511.05173*, 2015.

[7] I. B. Magdău, F. Balm, and G. J. Ackland, "Theory of high pressure hydrogen, made simple," *arXiv preprint arXiv:1512.00063*, 2015.

[8] I. B. Magdău, B. Tyson, B. Borgulya, and G. J. Ackland, "Dynamical phase transitions of solid hydrogen at high pressure," *writing*, 2016.

[9] I. B. Magdău and G. J. Ackland, "Infrared splitting in hydrogen-deuterium mixtures at high pressures," *writing*, 2016.

[10] S. Tennant, "On the nature of the diamond. by Smithson Tennant, esq. frs," *Philosophical Transactions of the Royal Society of London*, vol. 87, pp. 123–127, 1797.

[11] P. W. Bridgman, "Recent work in the field of high pressures," *Reviews of Modern Physics*, vol. 18, no. 1, p. 1, 1946.

[12] P. W. Bridgman, "Rough compressions of 177 substances to 40,000 Kg/Cm^2," in *Proceedings of the American Academy of Arts and Sciences*, vol. 76, pp. 71–87, JSTOR, 1948.

[13] P. W. Bridgman, "The resistance of 72 elements, alloys and compounds to 100,000 Kg/Cm^2," in *Proceedings of the American Academy of Arts and Sciences*, vol. 81, pp. 165–251, JSTOR, 1952.

[14] C. E. Weir, E. R. Lippincott, A. Van Valkenburg, and E. N. Bunting, "Infrared studies in the 1-to 15-micron region to 30,000 atmospheres," *J. Res. Natl. Bur. Stand. A*, vol. 63, pp. 55–62, 1959.

[15] K. Shimizu, K. Suhara, M. Ikumo, M. I. Eremets, and K. Amaya, "Superconductivity in oxygen," *Nature*, vol. 393, no. 6687, pp. 767–769, 1998.

[16] Y. Ma, M. Eremets, A. R. Oganov, Y. Xie, I. Trojan, S. Medvedev, A. O. Lyakhov, M. Valle, and V. Prakapenka, "Transparent dense sodium," *Nature*, vol. 458, no. 7235, pp. 182–185, 2009.

[17] A. Burrows, W. B. Hubbard, J. I. Lunine, and J. Liebert, "The theory of brown dwarfs and extrasolar giant planets," *Reviews of Modern Physics*, vol. 73, no. 3, p. 719, 2001.

[18] J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, "The properties of hydrogen and helium under extreme conditions," *Reviews of Modern Physics*, vol. 84, no. 4, p. 1607, 2012.

[19] E. Wigner and H. B. Huntington, "On the possibility of a metallic modification of hydrogen," *The Journal of Chemical Physics*, vol. 3, p. 764, 1935.

[20] N. W. Ashcroft, "Metallic hydrogen: A high-temperature superconductor?," *Physical Review Letters*, vol. 21, no. 26, pp. 1748–1749, 1968.

[21] S. A. Bonev, E. Schwegler, T. Ogitsu, and G. Galli, "A quantum fluid of metallic hydrogen suggested by first-principles calculations," *Nature*, vol. 431, no. 7009, pp. 669–672, 2004.

[22] R. T. Howie, P. Dalladay-Simpson, and E. Gregoryanz, "Raman spectroscopy of hot hydrogen above 200 GPa," *Nature Materials*, vol. 14, no. 5, pp. 495–499, 2015.

[23] E. Babaev, A. Sudbø, and N. W. Ashcroft, "A superconductor to superfluid phase transition in liquid metallic hydrogen," *Nature*, vol. 431, no. 7009, pp. 666–668, 2004.

[24] H. K. Mao and R. J. Hemley, "Ultrahigh-pressure transitions in solid hydrogen," *Reviews of Modern Physics*, vol. 66, no. 2, p. 671, 1994.

[25] A. Jayaraman, "Diamond anvil cell and high-pressure physical investigations," *Reviews of Modern Physics*, vol. 55, no. 1, p. 65, 1983.

[26] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Physical Review*, vol. 136, no. 3B, p. B864, 1964.

[27] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, no. 4A, p. A1133, 1965.

[28] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical Review Letters*, vol. 77, no. 18, p. 3865, 1996.

[29] C. J. Pickard and R. J. Needs, "High-pressure phases of silane," *Physical Review Letters*, vol. 97, no. 4, p. 045504, 2006.

[30] C. J. Pickard and R. J. Needs, "Ab initio random structure searching," *Journal of Physics: Condensed Matter*, vol. 23, no. 5, p. 053201, 2011.

[31] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, "First-principles simulation: ideas, illustrations and the castep code," *Journal of Physics: Condensed Matter*, vol. 14, no. 11, p. 2717, 2002.

[32] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory," *Reviews of Modern Physics*, vol. 73, no. 2, p. 515, 2001.

[33] B. J. Alder and T. E. Wainwright, "Studies in molecular dynamics. I. General method," *The Journal of Chemical Physics*, vol. 31, p. 459, 1959.

[34] R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory," *Physical Review Letters*, vol. 55, no. 22, pp. 2471–2474, 1985.

[35] J. Cao and G. A. Voth, "The formulation of quantum statistical mechanics based on the feynman path centroid density. I. Equilibrium properties," *The Journal of Chemical Physics*, vol. 100, p. 5093, 1994.

[36] J. Cao and G. A. Voth, "The formulation of quantum statistical mechanics based on the feynman path centroid density. II. Dynamical properties," *The Journal of Chemical Physics*, vol. 100, p. 5106, 1994.

[37] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *Journal of Applied Physics*, vol. 52, p. 7182, 1981.

[38] A. Laio and M. Parrinello, "Escaping free-energy minima," *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12562–12566, 2002.

[39] R. Martoňák, A. Laio, and M. Parrinello, "Predicting crystal structures: the Parrinello-Rahman method revisited," *Physical Review Letters*, vol. 90, no. 7, p. 075503, 2003.

[40] N. W. Ashcroft and N. D. Mermin, "Solid state physics," 1976.

[41] K. Refson, P. R. Tulip, and S. J. Clark, "Variational density-functional perturbation theory for dielectrics and lattice dynamics," *Physical Review B*, vol. 73, no. 15, p. 155114, 2006.

[42] F. Siebert and P. Hildebrandt, *Vibrational spectroscopy in life science*, vol. 2. Wiley-vch, 2008.

[43] D. Porezag and M. R. Pederson, "Infrared intensities and Raman-scattering activities within density-functional theory," *Physical Review B*, vol. 54, no. 11, p. 7830, 1996.

[44] M. Lazzeri and F. Mauri, "First-principles calculation of vibrational Raman spectra in large systems: Signature of small rings in crystalline SiO_2," *Physical Review Letters*, vol. 90, no. 3, p. 036401, 2003.

[45] T. A. Arias, M. C. Payne, and J. D. Joannopoulos, "Ab initio molecular dynamics: Analytically continued energy functionals and insights into iterative solutions," *Physical Review Letters*, vol. 69, no. 7, pp. 1077–1080, 1992.

[46] V. Labet, P. Gonzalez-Morelos, R. Hoffmann, and N. W. Ashcroft, "A fresh look at dense hydrogen under pressure. I. An introduction to the problem, and an index probing equalization of H–H distances," *The Journal of Chemical Physics*, vol. 136, no. 7, p. 074501, 2012.

[47] V. Labet, R. Hoffmann, and N. W. Ashcroft, "A fresh look at dense hydrogen under pressure. II. Chemical and physical models aiding our understanding of evolving H–H separations," *The Journal of Chemical Physics*, vol. 136, no. 7, p. 074502, 2012.

[48] V. Labet, R. Hoffmann, and N. W. Ashcroft, "A fresh look at dense hydrogen under pressure. III. Two competing effects and the resulting intramolecular H–H separation in solid hydrogen under pressure," *The Journal of Chemical Physics*, vol. 136, no. 7, p. 074503, 2012.

[49] V. Labet, R. Hoffmann, and N. W. Ashcroft, "A fresh look at dense hydrogen under pressure. IV. Two structural models on the road from paired to monatomic hydrogen, via a possible non-crystalline phase," *The Journal of Chemical Physics*, vol. 136, no. 7, p. 074504, 2012.

[50] G. J. Ackland, "Bearing down on hydrogen," *Science*, vol. 348, no. 6242, pp. 1429–1430, 2015.

[51] I. Silvera, "The insulator-metal transition in hydrogen," *Proceedings of the National Academy of Sciences*, vol. 107, no. 29, pp. 12743–12744, 2010.

[52] S. T. Weir, A. C. Mitchell, and W. J. Nellis, "Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar)," *Physical Review Letters*, vol. 76, no. 11, pp. 1860–1863, 1996.

[53] M. D. Knudson, M. P. Desjarlais, A. Becker, R. W. Lemke, K. R. Cochrane, M. E. Savage, D. E. Bliss, T. R. Mattsson, and R. Redmer, "Direct observation of an abrupt insulator-to-metal transition in dense liquid deuterium," *Science*, vol. 348, no. 6242, pp. 1455–1460, 2015.

[54] V. Dzyabura, M. Zaghoo, and I. F. Silvera, "Evidence of a liquid-liquid phase transition in hot dense hydrogen," *Proceedings of the National Academy of Sciences*, vol. 110, no. 20, pp. 8040–8044, 2013.

[55] M. Zaghoo, A. Salamat, and I. F. Silvera, "Evidence of a first-order phase transition to metallic hydrogen," *Physical Review B*, vol. 93, no. 15, p. 155128, 2016.

[56] G. Mazzola and S. Sorella, "Distinct metallization and atomization transitions in dense liquid hydrogen," *Physical Review Letters*, vol. 114, no. 10, p. 105701, 2015.

[57] S. Deemyad and I. F. Silvera, "Melting line of hydrogen at high pressures," *Physical Review Letters*, vol. 100, no. 15, p. 155701, 2008.

[58] M. I. Eremets and I. A. Trojan, "Evidence of maximum in the melting curve of hydrogen at megabar pressures," *JETP Letters*, vol. 89, no. 4, pp. 174–179, 2009.

[59] E. Gregoryanz, A. F. Goncharov, K. Matsuishi, H. K. Mao, and R. J. Hemley, "Raman spectroscopy of hot dense hydrogen," *Physical Review Letters*, vol. 90, no. 17, p. 175701, 2003.

[60] N. Subramanian, A. F. Goncharov, V. V. Struzhkin, M. Somayazulu, and R. J. Hemley, "Bonding changes in hot fluid hydrogen at megabar pressures," *Proceedings of the National Academy of Sciences*, vol. 108, no. 15, pp. 6014–6019, 2011.

[61] S. Scandolo, "Liquid-liquid phase transition in compressed hydrogen from first-principles simulations," *Proceedings of the National Academy of Sciences*, vol. 100, no. 6, pp. 3051–3053, 2003.

[62] C. Attaccalite and S. Sorella, "Stable liquid hydrogen at high pressure by a novel ab initio molecular-dynamics calculation," *Physical Review Letters*, vol. 100, no. 11, p. 114501, 2008.

[63] M. A. Morales, C. Pierleoni, E. Schwegler, and D. M. Ceperley, "Evidence for a first-order liquid-liquid transition in high-pressure hydrogen from ab initio simulations," *Proceedings of the National Academy of Sciences*, vol. 107, no. 29, pp. 12799–12803, 2010.

[64] M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, "Nuclear quantum effects and nonlocal exchange-correlation functionals applied to liquid hydrogen at high pressure," *Physical Review Letters*, vol. 110, no. 6, p. 065702, 2013.

[65] H. Liu, E. R. Hernandez, J. Yan, and Y. Ma, "Anomalous melting behavior of solid hydrogen at high pressures," *The Journal of Physical Chemistry C*, vol. 117, no. 22, pp. 11873–11877, 2013.

[66] J. Chen, X. Z. Li, Q. Zhang, M. J. Probert, C. J. Pickard, R. J. Needs, A. Michaelides, and E. Wang, "Quantum simulation of low-temperature metallic liquid hydrogen," *Nature Communications*, vol. 4, 2013.

[67] J. Dewar, "Sur la solidification de l'hydrogene," *Ann. Chim. Phys*, vol. 18, pp. 145–150, 1899.

[68] I. F. Silvera and R. J. Wijngaarden, "New low-temperature phase of molecular deuterium at ultrahigh pressure," *Physical Review Letters*, vol. 47, no. 1, pp. 39–42, 1981.

[69] F. Moshary, N. H. Chen, and I. F. Silvera, "Remarkable high pressure phase line of orientational order in solid hydrogen deuteride," *Physical Review Letters*, vol. 71, no. 23, pp. 3814–3817, 1993.

[70] P. Loubeyre, R. LeToullec, D. Hausermann, M. Hanfland, R. J. Hemley, H. K. Mao, and L. W. Finger, "X-ray diffraction and equation of state of hydrogen at megabar pressures," *Nature*, vol. 383, no. 6602, pp. 702–704, 1996.

[71] M. Moraldi, "Orientational structures in solid para-hydrogen in the broken symmetry phase," *Physical Review B*, vol. 80, no. 13, p. 134117, 2009.

[72] X. Z. Li, B. Walker, M. J. Probert, C. J. Pickard, R. J. Needs, and A. Michaelides, "Classical and quantum ordering of protons in cold solid hydrogen under megabar pressures," *Journal of Physics: Condensed Matter*, vol. 25, no. 8, p. 085402, 2013.

[73] C. J. Pickard and R. J. Needs, "Structures at high pressure from random searching," *Physica Status Solidi B*, vol. 246, no. 3, pp. 536–540, 2009.

[74] M. Hanfland, R. J. Hemley, and H. K. Mao, "Novel infrared vibron absorption in solid hydrogen at megabar pressures," *Physical Review Letters*, vol. 70, no. 24, p. 3760, 1993.

[75] Y. Akahama, M. Nishimura, H. Kawamura, N. Hirao, Y. Ohishi, and K. Takemura, "Evidence from x-ray diffraction of orientational ordering in phase III of solid hydrogen at pressures up to 183 GPa," *Physical Review B*, vol. 82, no. 6, p. 060101, 2010.

[76] R. J. Hemley and H. K. Mao, "Phase transition in solid molecular hydrogen at ultrahigh pressures," *Physical Review Letters*, vol. 61, no. 7, pp. 857–860, 1988.

[77] H. E. Lorenzana, I. F. Silvera, and K. A. Goettel, "Evidence for a structural phase transition in solid hydrogen at megabar pressures," *Physical Review Letters*, vol. 63, no. 19, pp. 2080–2083, 1989.

[78] C. J. Pickard and R. J. Needs, "Structure of phase III of solid hydrogen," *Nature Physics*, vol. 3, no. 7, pp. 473–476, 2007.

[79] A. F. Goncharov, E. Gregoryanz, R. J. Hemley, and H. K. Mao, "Spectroscopic studies of the vibrational and electronic properties of solid hydrogen to 285 GPa," *Proceedings of the National Academy of Sciences*, vol. 98, no. 25, pp. 14234–14237, 2001.

[80] P. Loubeyre, F. Occelli, and R. LeToullec, "Optical studies of solid hydrogen to 320 GPa and evidence for black hydrogen," *Nature*, vol. 416, no. 6881, pp. 613–617, 2002.

[81] J. S. Tse, D. D. Klug, Y. Yao, Y. Le Page, and J. R. Rodgers, "Structure and spectroscopic properties of dense solid hydrogen at 160 GPa," *Solid State Communications*, vol. 145, no. 1, pp. 5–10, 2008.

[82] Y. Akahama, H. Kawamura, N. Hirao, Y. Ohishi, and K. Takemura, "Raman scattering and x-ray diffraction experiments for phase III of solid hydrogen," in *Journal of Physics: Conference Series*, vol. 215, p. 012056, IOP Publishing, 2010.

[83] P. Tolédano, H. Katzke, A. F. Goncharov, and R. J. Hemley, "Symmetry breaking in dense solid hydrogen: Mechanisms for the transitions to phase II and phase III," *Physical Review Letters*, vol. 103, no. 10, p. 105301, 2009.

[84] C. S. Zha, Z. Liu, and R. J. Hemley, "Synchrotron infrared measurements of dense hydrogen to 360 GPa," *Physical Review Letters*, vol. 108, no. 14, p. 146402, 2012.

[85] S. Azadi and W. M. C. Foulkes, "Fate of density functional theory in the study of high-pressure solid hydrogen," *Physical Review B*, vol. 88, no. 1, p. 014115, 2013.

[86] N. D. Drummond, B. Monserrat, J. H. Lloyd-Williams, P. L. Ríos, C. J. Pickard, and R. J. Needs, "Quantum monte carlo study of the phase diagram of solid molecular hydrogen at extreme pressures," *Nature Communications*, vol. 6, 2015.

[87] M. I. Eremets and I. A. Troyan, "Conductive dense hydrogen," *Nature Materials*, vol. 10, no. 12, pp. 927–931, 2011.

[88] W. J. Nellis, A. L. Ruoff, and I. F. Silvera, "Has metallic hydrogen been made in a diamond anvil cell?," *arXiv preprint arXiv:1201.0407*, 2012.

[89] R. T. Howie, C. L. Guillaume, T. Scheler, A. F. Goncharov, and E. Gregoryanz, "Mixed molecular and atomic phase of dense hydrogen," *Physical Review Letters*, vol. 108, no. 12, p. 125501, 2012.

[90] R. T. Howie, T. Scheler, C. L. Guillaume, and E. Gregoryanz, "Proton tunneling in phase IV of hydrogen and deuterium," *Physical Review B*, vol. 86, no. 21, p. 214104, 2012.

[91] C. J. Pickard, M. Martinez-Canales, and R. J. Needs, "Density functional theory study of phase IV of solid hydrogen," *Physical Review B*, vol. 85, no. 21, p. 214114, 2012.

[92] M. I. Eremets, I. A. Troyan, P. Lerch, and A. Drozdov, "Infrared study of hydrogen up to 310 GPa at room temperature," *High Pressure Research*, no. ahead-of-print, pp. 1–4, 2013.

[93] P. Loubeyre, F. Occelli, and P. Dumas, "Hydrogen phase IV revisited via synchrotron infrared measurements in H_2 and D_2 up to 290 GPa at 296 K," *Physical Review B*, vol. 87, no. 13, p. 134101, 2013.

[94] C. S. Zha, Z. Liu, M. Ahart, R. Boehler, and R. J. Hemley, "High-pressure measurements of hydrogen phase IV using synchrotron infrared spectroscopy," *Physical Review Letters*, vol. 110, no. 21, p. 217402, 2013.

[95] C. J. Pickard, M. Martinez-Canales, and R. J. Needs, "Erratum: Density functional theory study of phase IV of solid hydrogen," *Physical Review B*, vol. 86, no. 5, p. 059902, 2012.

[96] H. Liu, L. Zhu, W. Cui, and Y. Ma, "Room-temperature structures of solid hydrogen at high pressures," *The Journal of Chemical Physics*, vol. 137, no. 7, p. 074501, 2012.

[97] H. Liu and Y. Ma, "Proton or deuteron transfer in phase IV of solid hydrogen and deuterium," *Physical Review Letters*, vol. 110, no. 2, p. 025903, 2013.

[98] A. F. Goncharov, S. T. John, H. Wang, J. Yang, V. V. Struzhkin, R. T. Howie, and E. Gregoryanz, "Bonding, structures, and band gap closure of hydrogen at high pressures," *Physical Review B*, vol. 87, no. 2, p. 024101, 2013.

[99] M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, "Towards a predictive first-principles description of solid molecular hydrogen with density functional theory," *Physical Review B*, vol. 87, no. 18, p. 184107, 2013.

[100] P. Dalladay-Simpson, R. T. Howie, and E. Gregoryanz, "Evidence for a new phase of dense hydrogen above 325 gigapascals," *Nature*, vol. 529, no. 7584, pp. 63–67, 2016.

[101] M. I. Eremets, I. A. Troyan, and A. P. Drozdov, "Low temperature phase diagram of hydrogen at pressures up to 380 GPa. A possible metallic phase at 360 GPa and 200 K," *arXiv preprint arXiv:1601.04479*, 2016.

[102] R. Dias, O. Noked, and I. F. Silvera, "New insulating low temperature phase in dense hydrogen: The phase diagram to 421 GPa," *arXiv preprint arXiv:1603.02162*, 2016.

[103] J. M. McMahon and D. M. Ceperley, "Ground-state structures of atomic metallic hydrogen," *Physical Review Letters*, vol. 106, no. 16, p. 165302, 2011.

[104] H. Y. Geng, H. X. Song, J. F. Li, and Q. Wu, "High-pressure behavior of dense hydrogen up to 3.5 TPa from density functional theory calculations," *Journal of Applied Physics*, vol. 111, no. 6, p. 063510, 2012.

[105] H. Liu, H. Wang, and Y. Ma, "Quasi-molecular and atomic phases of dense solid hydrogen," *The Journal of Physical Chemistry C*, vol. 116, no. 16, pp. 9221–9226, 2012.

[106] T. Ishikawa, H. Nagara, T. Oda, N. Suzuki, and K. Shimizu, "Phase with pressure-induced shuttlewise deformation in dense solid atomic hydrogen," *Physical Review B*, vol. 90, no. 10, p. 104102, 2014.

[107] S. Azadi, B. Monserrat, W. M. C. Foulkes, and R. J. Needs, "Dissociation of high-pressure solid molecular hydrogen: a quantum monte carlo and anharmonic vibrational study," *Physical Review Letters*, vol. 112, no. 16, p. 165501, 2014.

[108] J. McMinis, R. C. Clay III, D. Lee, and M. A. Morales, "Molecular to atomic phase transition in hydrogen under high pressure," *Physical Review Letters*, vol. 114, no. 10, p. 105305, 2015.

[109] A. F. Goncharov, R. J. Hemley, and H. K. Mao, "Vibron frequencies of solid H_2 and D_2 to 200 GPa and implications for the P–T phase diagram," *The Journal of Chemical Physics*, vol. 134, no. 17, p. 174501, 2011.

[110] D. M. Brown and W. B. Daniels, "Vibrational Raman spectra of hydrogen and deuterium mixtures at high pressures," *Physical Review A*, vol. 45, no. 9, p. 6429, 1992.

[111] J. Kranendonk, *Solid Hydrogen: Theory of the Properties of Solid H_2, HD, and D_2*. Springer Science & Business Media, 2012.

[112] P. Loubeyre, R. LeToullec, and J. P. Pinceaux, "Raman measurements of the vibrational properties of H_2 as a guest molecule in dense helium, neon, argon, and deuterium systems up to 40 GPa," *Physical Review B*, vol. 45, no. 22, p. 12844, 1992.

[113] A. Chijioke and I. F. Silvera, "Megabar-pressure infrared study of hydrogen deuteride," *Physical Review Letters*, vol. 97, no. 25, p. 255701, 2006.

[114] R. P. Dias, O. Noked, and I. F. Silvera, "New phases and dissociation-recombination of hydrogen deuteride to 3.4 Mbar," *Physical Review Letters*, vol. 116, no. 14, p. 145501, 2016.

[115] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne, "First principles methods using castep," *Zeitschrift für Kristallographie-Crystalline Materials*, vol. 220, no. 5/6, pp. 567–570, 2005.

[116] C. J. Pickard and G. J. Ackland. personal communication.

[117] U. Pinsook and G. J. Ackland, "Calculation of anomalous phonons and the hcp-bcc phase transition in zirconium," *Physical Review B*, vol. 59, no. 21, p. 13642, 1999.

[118] H. Liu, J. Tse, and Y. Ma, "Robust diffusive proton motions in phase IV of solid hydrogen," *The Journal of Physical Chemistry C*, vol. 118, no. 22, pp. 11902–11905, 2014.

[119] C. S. Zha, R. E. Cohen, H. K. Mao, and R. J. Hemley, "Raman measurements of phase transitions in dense solid hydrogen and deuterium to 325 GPa," *Proceedings of the National Academy of Sciences*, vol. 111, no. 13, pp. 4792–4797, 2014.

[120] W. Humphrey, A. Dalke, and K. Schulten, "VMD: visual molecular dynamics," *Journal of molecular graphics*, vol. 14, no. 1, pp. 33–38, 1996.

[121] Jmol, "an open-source java viewer for chemical structures in 3d," *Jmol web page: http://www. jmol. org/, last accessed*, vol. 15, 2013.

[122] B. Monserrat, N. D. Drummond, and R. J. Needs, "Anharmonic vibrational properties in periodic systems: energy, electron-phonon coupling, and stress," *Physical Review B*, vol. 87, no. 14, p. 144302, 2013.

[123] R. Singh, S. Azadi, and T. D. Kühne, "Anharmonicity and finite-temperature effects on the structure, stability, and vibrational spectrum of phase III of solid molecular hydrogen," *Physical Review B*, vol. 90, no. 1, p. 014110, 2014.

[124] A. Rohatgi, "Webplotdigitizer," *URL http:// arohatgi.info/ WebPlotDigitizer/ app*, 2011.

[125] S. Lebègue, C. M. Araujo, D. Y. Kim, M. Ramzan, H. K. Mao, and R. Ahuja, "Semimetallic dense hydrogen above 260 GPa," *Proceedings of the National Academy of Sciences*, vol. 109, no. 25, pp. 9766–9769, 2012.

[126] M. Marqués, G. J. Ackland, L. F. Lundegaard, G. Stinton, R. J. Nelmes, M. I. McMahon, and J. Contreras-Garcia, "Potassium under pressure: a pseudobinary ionic compound," *Physical Review Letters*, vol. 103, no. 11, p. 115501, 2009.

[127] M. Gatti, I. V. Tokatly, and A. Rubio, "Sodium: a charge-transfer insulator at high pressures," *Physical Review Letters*, vol. 104, no. 21, p. 216404, 2010.

[128] M. Marqués, M. I. McMahon, E. Gregoryanz, M. Hanfland, C. L. Guillaume, C. J. Pickard, G. J. Ackland, and R. J. Nelmes, "Crystal structures of dense lithium: a metal-semiconductor-metal transition," *Physical Review Letters*, vol. 106, no. 9, p. 095502, 2011.

[129] D. Hohl, V. Natoli, D. M. Ceperley, and R. M. Martin, "Molecular dynamics in dense hydrogen," *Physical Review Letters*, vol. 71, no. 4, p. 541, 1993.

[130] S. Biermann, D. Hohl, and D. Marx, "Proton quantum effects in high pressure hydrogen," *Journal of Low Temperature Physics*, vol. 110, no. 1-2, pp. 97–102, 1998.

[131] S. Biermann, D. Hohl, and D. Marx, "Quantum effects in solid hydrogen at ultra-high pressure," *Solid State Communications*, vol. 108, no. 6, pp. 337–341, 1998.

[132] S. A. Bonev and G. J. Ackland. personal communication.

[133] S. Fleming and A. Rohl, "Gdis: a visualization program for molecular and periodic systems," *Zeitschrift Fur Kristallographie*, vol. 220, no. 5-6, pp. 580–584, 2005.

[134] G. Kresse and J. Hafner, "Ab initio molecular dynamics for liquid metals," *Physical Review B*, vol. 47, no. 1, p. 558, 1993.

[135] A. Savin, O. Jepsen, J. Flad, O. K. Andersen, H. Preuss, and H. G. von Schnering, "Electron localization in solid-state structures of the elements: the diamond structure," *Angewandte Chemie International Edition in English*, vol. 31, no. 2, pp. 187–188, 1992.

[136] G. J. Ackland, I. Magdău, and F. Balm, "Theory of high pressure hydrogen, made simple/airapt 2015 data," 2016.

[137] A. N. Jackson and G. J. Ackland, "Lattice-switch monte carlo simulation for binary hard-sphere crystals," *Physical Review E*, vol. 76, no. 6, p. 066703, 2007.

[138] J. W. Eaton, D. Bateman, and S. Hauberg, *GNU Octave version 3.0.1 manual: a high-level interactive language for numerical computations*. CreateSpace Independent Publishing Platform, 2009. ISBN 1441413006.

[139] V. V. Kechin, "Melting curve equations at high pressure," *Physical Review B*, vol. 65, no. 5, p. 052102, 2001.

[140] P. W. Anderson, "Absence of diffusion in certain random lattices," *Physical Review*, vol. 109, no. 5, p. 1492, 1958.

[141] H. Hu, A. Strybulevych, J. H. Page, S. E. Skipetrov, and B. A. van Tiggelen, "Localization of ultrasound in a three-dimensional elastic network," *Nature Physics*, vol. 4, no. 12, pp. 945–948, 2008.

[142] G. Roati, C. D'Errico, L. Fallani, M. Fattori, C. Fort, M. Zaccanti, G. Modugno, M. Modugno, and M. Inguscio, "Anderson localization of a non-interacting Bose–Einstein condensate," *Nature*, vol. 453, no. 7197, pp. 895–898, 2008.

[143] M. Segev, Y. Silberberg, and D. N. Christodoulides, "Anderson localization of light," *Nature Photonics*, vol. 7, no. 3, pp. 197–204, 2013.

[144] C. Monthus and T. Garel, "Anderson localization of phonons in dimension d= 1, 2, 3: Finite-size properties of the inverse participation ratios of eigenstates," *Physical Review B*, vol. 81, no. 22, p. 224208, 2010.

[145] P. R. Tulip, *Dielectric and lattice dynamical properties of molecular Crystals via density functional perturbation theory: implementation within a first principles code*. PhD thesis, Durham University, 2004.

[146] A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov, and S. I. Shylin, "Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system," *Nature*, vol. 525, no. 7567, pp. 73–76, 2015.