Title
Combining Crystal Structure Prediction and Simulated Spectroscopy to Investigate Challenging High Pressure Phases

Author
Sontising, Watit

Publication Date
2020

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Combining Crystal Structure Prediction and Simulated Spectroscopy to Investigate Challenging High Pressure Phases

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Watit Sontising

June 2020

Dissertation Committee:

Dr. Gregory J.O. Beran, Chairperson
Dr. Chia-en A. Chang
Dr. Christopher J. Bardeen

The Dissertation of Watit Sontising is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisor, without whose help, I would not have been here.

To my parents, friends, and mentors for all the support.

ABSTRACT OF THE DISSERTATION

Combining Crystal Structure Prediction and Simulated Spectroscopy to Investigate Challenging High Pressure Phases

by

Watit Sontising

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, June 2020
Dr. Gregory J.O. Beran, Chairperson

Solid carbon dioxide and nitrogen exhibit rich phase diagrams at high pressure. The large number of viable packing motifs stems from their small size and weak, non-polar intermolecular interactions, which make many packing arrangements and orientations energetically competitive. Experimental observation and characterization of high-pressure polymorphs have proved challenging, not only because of flat energy landscape, but also their kinetic path-dependence and hysteresis in the phase transitions. As a result, high-quality experimental data are difficult to obtain, leaving many high-pressure crystal structures of nitrogen to remain unknown over decades, or creating ambiguities in the nature of some carbon dioxide phases.

This thesis employs a combination of high-level fragment-based electronic structure method and Raman simulation to study high-pressure polymorphs in these systems. First, we investigate the nature of carbon dioxide phases III and VII. We provide evidence that the long-accepted structure of phase III is problematic from comparison of large-basis-set quasi-harmonic second-order Møller-Plesset and experimental data. The experimental phase

III and VII structures both relax to the same phase VII structure. Furthermore, Raman spectra predicted for phase VII are in good agreement with those observed experimentally for both phase III and VII, while those for the purported phase III structure contradict experimental observations. Crystal structure prediction is employed to search for other potential structures which might account for phase III, but none are found. Together, these results suggest that phases III and VII are likely identical.

Second, we revisit nitrogen phase λ, one of the high-pressure solid nitrogen forms that was discovered by combining experimental monoclinic lattice parameters with atomic positions from an earlier, computationally predicted structure that had similar unit cell dimensions. Crystal structure prediction is performed to demonstrate that the reported $P2_1/c$ structure is indeed the likeliest candidate for the λ phase. Furthermore, we provide further evidence for the structural assignment by demonstrating good agreement between its predicted and experimental structural parameters and Raman spectra. Finally, the thermodynamic stability of the λ phase relative to other phases has been uncertain, but the calculations do suggest that it may be the thermodynamically most stable phase for at least part of the pressure range over which it has been observed.

Lastly, we perform crystal structure prediction using *ab initio* random structure searching and density functional theory to identify candidate structures for nitrogen phase ζ, the phase whose structure remains unknown decades after it was first observed spectroscopically, despite numerous experimental and theoretical investigations. The candidates are then analyzed for consistency with experiment in terms of their simulated x-ray diffraction patterns and Raman spectra. While none of the structures generated is a clear match

for the phase ζ experimental data, several of the candidates do exhibit features in common with the experiments and could provide an interesting starting point for future studies. The techniques here also rule out several candidate ζ nitrogen structures that have been identified previously. Finally, one of the structures might be considered a candidate for phase κ, whose structure is also unknown.

Contents

List of Figures **xi**

List of Tables **xv**

1 Introduction **1**
- 1.1 Crystal Structure Prediction 3
  - 1.1.1 Crystal Structure Generation 4
  - 1.1.2 Energy Ranking 6
- 1.2 Raman Spectroscopy 7
- 1.3 Outline of This Dissertation 9

2 Methods **11**
- 2.1 Outline 11
- 2.2 Fragment-based hybrid many-body interaction (HMBI) 12
- 2.3 Quasi-Harmonic Approximation 14
- 2.4 Structure Refinement 15
- 2.5 Raman Simulation 17

3 Theoretical Predictions Suggest Carbon Dioxide Phases III and VII are Identical **19**
- 3.1 Introduction 19
- 3.2 Methods 23
  - 3.2.1 Quasi-Harmonic structure optimizations 23
  - 3.2.2 Electronic Structure Calculations 24
  - 3.2.3 Raman Spectra Predictions 24
  - 3.2.4 Crystal Structure Prediction 25
  - 3.2.5 Data Analysis 26
- 3.3 Results and Discussion 26
  - 3.3.1 Comparison of predicted and experimental lattice parameters 26
  - 3.3.2 Equations of state 29
  - 3.3.3 Phase III optimization with fixed experimental cell 31
  - 3.3.4 Impact of basis set on the predicted Raman spectra 32

3.3.5 Pressure dependence of the Raman spectra 34
3.3.6 Crystal structure prediction results 38
3.3.7 Comparing carbon dioxide phase III and VII 40
3.4 Conclusions 48

4 Theoretical Assessment of the Structure and Stability of the λ Phase of Nitrogen 50
4.1 Introduction 50
4.2 Methods 54
4.2.1 Density functional theory calculations 54
4.2.2 Fragment-based hybrid many-body interaction (HMBI) calculations 55
4.2.3 Structure refinement and quasi-harmonic approximation 56
4.2.4 Simulated powder X-ray diffraction (PXRD) and Raman spectra . . 58
4.3 Results and Discussion 59
4.3.1 DFT Pseudopotential and Convergence Tests 59
4.3.2 Crystal Structure Prediction and Structure Ranking 67
4.3.3 Crystal energy landscape 74
4.3.4 Spectroscopic comparisons 79
4.3.5 Thermodynamic stability 83
4.4 Conclusions 85

5 Combining crystal structure prediction and simulated spectroscopy in pursuit of the unknown nitrogen phase ζ crystal structure 87
5.1 Introduction 87
5.2 Methods 91
5.2.1 AIRSS structure generation: 91
5.2.2 Density functional theory structure optimization and enthalpies: . . 92
5.2.3 Simulated Raman spectra: 94
5.3 Results and Discussion 96
5.3.1 Lattice constants for known crystalline phases of nitrogen 96
5.3.2 Crystal structure prediction and structure ranking 96
5.3.3 Comparison of x-ray diffraction patterns and Raman spectra 101
5.3.4 Equation of state for phase ϵ 104
5.3.5 Symmetry breaking of structure #19 105
5.3.6 Structures #12 and #19 as candidates for phase κ 106
5.3.7 Crystal energy landscape 109
5.3.8 Simulated x-ray diffraction and Raman spectra 118
5.3.9 Predicted enthalpies 122
5.4 Conclusions 126

6 Conclusions 128

List of Figures

1.1 A sample crystal energy landscape. Ideally, the most stable structure will correspond to the experimental form 3
1.2 Low frequency Raman spectrum for solid carbon dioxide phase I (left) and II (right). 8

2.1 Fragment of many-body contribution terms from the cluster of 1-body (left), 2-body (middle), and 3-body (right). 12
2.2 The hybrid many-body interaction model for crystal. 13
2.3 Fitting Murnaghan equation of state with predicted enthalpy-volume curve from series geometry at external pressure. 16

3.1 Phase diagram of carbon dioxide up to 40 GPa, and structure overlay of the experimental crystal structures for phases III (blue) and VII (gray). Root-mean-square deviation[1] = 0.24 Å. 20
3.2 Comparison between the 11.8 GPa experimental phase III structure[2] (gray) and the MP2/aug-cc-pVDZ optimized one (red) when the optimization is performed with the experimental lattice parameters held fixed. 32
3.3 Comparison of the MP2/aug-cc-pVDZ and aug-cc-pVTZ predicted Raman spectra for phase I for the quasiharmonic MP2/CBS structure obtained at 14.5 GPa GPa and room temperature. 33
3.4 Comparison of predicted MP2/aug-cc-pVDZ Raman spectra when the cell is fully optimized with MP2/aug-cc-pVDZ (and no quasiharmonic approximation) instead of using the quasiharmonic MP2/CBS lattice parameters for three different examples. 34
3.5 Pressure dependence of the librational Raman modes for phase I at room temperature. Comparison of MP2/aug-cc-pVDZ predictions based on quasiharmonic MP2/CBS structures from this work, results entirely at the MP2/aug-cc-pVDZ level with no quasiharmonic approximation from Li et al,[3] and PBE DFT results from Bonev et al.[4] Points indicate experimental data,[5–8] while lines correspond to the predicted frequencies. 35

3.6 Pressure dependence of the librational Raman modes for phase II at room temperature. Comparison of MP2/aug-cc-pVDZ predictions based on quasi-harmonic MP2/CBS structures from this work and PBE DFT from Bonev et al[4] (lines) against experimental data (points).[9–11] 37
3.7 Pressure dependence of the librational Raman modes for phase III at room temperature. Comparison of MP2/aug-cc-pVDZ predictions based on quasi-harmonic MP2/CBS structures from this work and PBE DFT results from Bonev et al[4] (lines) against experimental data from Olijnyk and Jephcoat (points).[7] 38
3.8 Simulated powder X-ray diffraction patterns for the 25 predicted crystal structures which lie within 10 kJ/mol of the stable phase II one. All were optimized at the PBE-D2 level and 11.8 GPa, including those identified as Phase I, II, or VII. These patterns are compared against the actual experimental phase III powder X-ray diffraction pattern (purple dotted line) and the simulated phase III spectrum based on the claimed experimental structure.[2] 39
3.9 Overlays and root-mean-square deviations (RMSD) between MP2-predicted (red) and experimental crystal structures for (a) phase I, (b) phase II, (c) phase III, and (c) phase VII. Note the unit cell discrepancy for phase III. Predicted structures were optimized at the same temperature and pressure as the experimental structure. 40
3.10 Comparison of predicted MP2/CBS (lines) and experimental (points) room temperature isotherms for (a) phase I,[2,8,12,13] (b) phase II,[10,11] and (c) phase III[14] carbon dioxide. Vinet equation of state[15] parameters fitted to the MP2 volumes are listed. Tables 3.1–3.4 show good agreement between the predicted and literature values for specific structures and the equation of state data. (d) Errors in the predicted lattice constants versus experiment. The shaded band indicates ±1% error. 41
3.11 (a) Comparison of MP2-predicted and experimental lattice phonon Raman spectra for phases I,[7] II,[11], IV,[16] and VII[17] carbon dioxide. (b) Comparison of Raman spectra for phases III and VII,[7,17] including predicted spectra using either MP2/CBS or experimentally determined unit cell parameters. (c) Pressure dependence of the experimental phase III[7] and predicted phase VII Raman spectra. All spectra are at room temperature unless otherwise indicated. *The Phase IV spectrum employs an MP2/aug-cc-pVDZ optimized cell instead of a quasi-harmonic MP2/CBS one. 43
3.12 (a) PBE-D2 crystal energy landscape for potential carbon dioxide crystal structures at 11.8 GPa with $Z = 2$ or 4 molecules in the unit cell. (b) Comparison for simulated powder X-ray diffraction patterns for the lowest-lying PBE-D2 predicted structures against the actual[2] (purple) and simulated experimental phase III (using the purported structure) ones. Aside from the purported phase III structure, only the phase VII structure plausibly corresponds to the experimental data. Simulated diffraction patterns for all 25 low-energy structures are provided in Section 3.3.6. 46

4.1 Phase diagram of nitrogen. The λ phase has been observed over the conditions highlighted in red, though its thermodynamic stability relative to the other phases remains unclear. 51
4.2 Convergence of the total energy per molecule with the planewave cutoff for several known N_2 crystal structures. 62
4.3 N_2 bond length vs. planewave basis set cutoff at two different pressures and for four different experimentally known crystal structures. 63
4.4 Overlay of experimental and predicted B86bPBE-XDM crystal structures for (left) phase α at ambient pressure, rmsd 0.026 Å, and (right) phase γ at 0.04 GPa, rmsd 0.047 Å. The calculations employed PAW potentials, an 80 Ry cutoff, and a 6×6×6 k-point grid. 64
4.5 The 15 lowest-enthalpy structures (S1–S15) at the B86bPBE-XDM level, as generated by the AIRSS search. 72
4.6 Crystal energy landscape for the low-energy crystal structures at 34 GPa with the B86bPBE-XDM (red Z=4/orange Z=2), MP2/CBS + pHF (dark blue Z=4/light blue Z=2), and MP2/CBS + AMOEBA (dark green Z=4/light green Z=2) levels of theory. Open symbols correspond to further CCSD(T) refinements of the structures. The experimentally inferred molar volume is indicated in purple. Enthalpies at each level of theory are plotted relative to the lowest-energy structure. 75
4.7 Structure overlay (top) and simulated powder x-ray diffraction spectrum (bottom) comparing the experimental,[18] CCSD(T)/CBS + periodic HF, and CCSD(T)/CBS + AMOEBA crystal structures. 77
4.8 Comparison of the predicted and experimentally observed[18] equations of state for λ N_2. 79
4.9 Comparison of the predicted and experimentally observed[18] Raman spectra for λ nitrogen in the librational region. 80
4.10 Crystal structures of N_2 (top left) phase α[19] (top right) phase γ[20] (bottom left) phase ϵ[21] (bottom right) phase ι[22] 82
4.11 Comparison of CCSD(T) enthalpies at 0 K for the α, γ, ϵ, λ, and ι phases as a function of pressure using either periodic HF (solid lines) or AMOEBA (dotted lines) for the many-body treatment. 83

5.1 Overlay of structure A2 (red) and #3 (blue). They two are virtually identical in half the cell, but differ in the symmetry operations relating the other half of the cell. 101
5.2 Comparison between predicted and experimental (a) powder X-ray diffraction pattern (λ=0.3683 Å) and (b) Raman spectrum for nitrogen phase ζ. 102
5.3 Comparison between predicted and experimental (a) powder X-ray diffraction pattern (λ=0.3683 Å) and (b) Raman for nitrogen phase ζ. 103
5.4 Comparison of the B86bPBE-XDM predicted equation of state against experimental data for nitrogen phase ϵ. 104
5.5 At 30 GPa, structure #19 breaks symmetry slightly, relaxing from $P2_12_12$ to P_2 symmetry and lowering the enthalpy by ∼0.3 kJ/mol). 105

5.6 Structure #12 optimized within the experimental κ phase lattice parameters at 130 GPa. 107
5.7 Comparison of the simulated powder X-ray diffraction patterns of structures #12 and #19 before and after constraining the lattice constants to match the experimentally reported values. All spectra employ a wavelength of $\lambda =$ 0.3683 Å. 108
5.8 Predicted crystal energy landscape for orthorhombic molecular N_2 phases at 80 GPa. Structures that have been reported in earlier structure prediction studies are shown in orange, while candidates in blue are new structures discussed in detail below. The enthalpy and volume for the optimized ϵ phase is shown for comparison, along with the reported molar volume of phase ζ. 110
5.9 Selected candidate crystal structures discussed in the text, as optimized with DFT at 80 GPa. 111
5.10 Comparison of predicted equations of state for all predicted structures against experimental data for nitrogen phase ζ. Dashed lines correspond to the B86bPBE-XDM predicted equations of state for all candidate structures from Figure 5.8, with the lines for structures B8, #12 and #19 highlighted in green, blue, and red. Solid lines show the new equation of state for the selected structures after fragment-based MP2/CBS + pHF refinement. . . . 113
5.11 Comparison between predicted and experimental lattice constants for several known phases of nitrogen and two ζ-phase candidates. The calculation were performed using B86bPBE-XDM with 6×6×6 k-point grid at the experimental pressure. Where experimental data was available at multiple pressures, errors are shown as computed for each pressure. See Table 5.1 for details. . 114
5.12 (a) Simulated powder x-ray diffraction patterns at 80 GPa (λ = 0.3683 Å) and (b) predicted Raman spectra at 30 GPa for selected candidate structures, compared against the experimental data for ζ N_2.[23,24] See Section 5.3.3 for the complete set of simulated spectra. 117
5.13 (a) Reducing the pressure used to compute the predicted spectrum to 20 GPa improves the agreement between theory and experiment considerably for the known ϵ phase. (b) Comparison of the 20 GPa spectra predicted for structures B8 and #19 against the unknown ζ phase spectrum at 30 GPa. Experimental spectra at 293 K for ϵ and 32 K for ζ taken from Ref 24. . . . 119
5.14 Comparison of the simulated powder X-ray diffraction patterns of structures #12 and #19 before and after constraining the lattice constants to match the experimentally reported values. All spectra employ a wavelength of $\lambda =$ 0.3683 Å. 121
5.15 DFT enthalpies versus pressure for the predicted candidate structures and several experimentally known phases. Colored lines correspond to the key structures; dashed gray lines correspond to other predicted structures from the CSP landscape which are not discussed in detail. 123

List of Tables

3.1 Comparison of predicted and experimental lattice parameters for phase I carbon dioxide ($Pa\bar{3}$ space group) at ambient temperature and selected pressures. 27
3.2 Comparison of predicted and experimental lattice parameters for phase II carbon dioxide ($P4_2/mnm$ space group). 28
3.3 Comparison of predicted and experimental lattice parameters for phase III & VII carbon dioxide ($Cmca$ space group). 28
3.4 Summary of predicted and experimental room-temperature bulk modulus data for phases I–III. Literature data typically employed the Birch-Murnaghan EOS, while the fits in this work employed the Vinet EOS. 30

4.1 Comparison of nitrogen bond length and unit cell parameters obtained using US-PP and PAW core treatments with B86bPBE-XDM, an 80 Ry cutoff, a 6×6×6 k-point grid, and 34 GPa external pressure. 64
4.2 Comparison of unit cell parameters for nitrogen phase α and γ predicted from different computational methods. DFT calculations were performed with B86BPBE-XDM functional, 80 Ry cutoff, 6x6x6 k-point grid. The obtained DFT structures were refined further with wave function correlation method at CCSD(T). The complete basis set was used by extrapolating consistent-correlation triple and quadruple zeta valence of Dunning basis sets, aug-cc-pVXZ. All calculations include quasi-harmonic thermal expansion. 65
4.3 Comparison of unit cell parameters for nitrogen phase ϵ and λ predicted from different computational methods. DFT calculations were performed with B86BPBE-XDM functional, 80 Ry cutoff, 6x6x6 k-point grid. The obtained DFT structures were refined further with wave function correlation method at CCSD(T). The complete basis set was used by extrapolating consistent-correlation triple and quadruple zeta valence of Dunning basis sets, aug-cc-pVXZ. All calculations include quasi-harmonic thermal expansion. 66
4.4 The unit-cell parameters and atomic positions of the twelve lowest DFT-energy structures, S1-S12, generated by the random structure search. Structures and relative enthalpies per molecule are reported at the B86bPBE-XDM level of theory under 34 GPa of external pressure. Atomic positions are reported in fractional coordinate. 69

4.5 The unit-cell parameters and atomic positions of the remaining structures S13-S22 generated by the random structure search. Structures and relative enthalpies per molecule are reported at the B86bPBE-XDM level of theory under 34 GPa of external pressure. Atomic positions are reported in fractional coordinate. 70
4.6 Relative enthalpies for each predicted structure at 34 GPa with B86bPBE-XDM and fragment-based MP2 methods. The structures are ordered based on the enthalpies at the DFT level; numbers in parentheses indicate the new ranking after energy refinement. Despite considerable reordering of the other structures, the structure ascribed to phase λ remains the most stable with all three models. 71
4.7 Comparison of experimental and predicted lattice parameters for λ N_2 at 34 GPa. 73

5.1 Comparison between predicted and experimental lattice constants for several known phases of nitrogen and two ζ-phase candidates. 97
5.2 The unit-cell parameters and atomic positions of AIRSS structures #1–#9, optimized at 80 GPa with B86bPBE-XDM. Atomic positions are reported in fractional coordinates. The remaining atoms can be generated via the appropriate space group symmetry operations. 98
5.3 The unit-cell parameters and atomic positions of AIRSS structures #10–#15, optimized at 80 GPa with B86bPBE-XDM. Atomic positions are reported in fractional coordinates. The remaining atoms can be generated via the appropriate space group symmetry operations. 99
5.4 The unit-cell parameters and atomic positions of AIRSS structures #16–#22, optimized at 80 GPa with B86bPBE-XDM. Atomic positions are reported in fractional coordinates. The remaining atoms can be generated via the appropriate space group symmetry operations 100
5.5 Comparison of the B86bPBE-XDM lattice constants for several predicted structures compared to the experimentally reported values for ζ nitrogen from Ref 23. Only structures #12 and #19 exhibit lattice parameters that are similar to the experimental ones. 115

Chapter 1

Introduction

Pressure is a physical force that can tune the interatomic distances in matter. Adjusting this distance allows changes in chemical bonding or crystal packing. The most well-known example of a pressure-induced bonding change is turning graphite into precious diamond under high pressure and temperature. More advance applications are also under study such as superconductors[25–28], hydrogen storage[29–32], insulatosr[33], etc. Recently, ground breaking, metallic hydrogen has been synthesized under extremely high pressure (more than 100 GPa).[34]

Molecular crystal polymorphism is also accessible via different pressure conditions. Relatively speaking, it is easier to change molecular crystal packing compared to making diamond or metallic hydrogen. Phase transitions for these molecular polymorphs usually occur under milder condition or instrumentally accessible region due to the weaker interactions such as hydrogen bonding, dispersion, electrostatics, or $\pi - \pi$ stacking interactions. Since it requires less effort to change from one polymorph to another, one of the challenge lies in ensuring one obtains the desired crystal form.

It is important to understand how to access the desirable polymorphs. Incomplete knowledge of the phase diagram could greatly affect life or finance, especially for pharmaceutical company. For example, Ritonavir, the anti-HIV drug, was previously mass-produced by being ground into tablet. This grounding process exerted high pressure on Ritonavir crystal turning it into insoluble form.[35,36] The accidental production forced the drugs to be removed and costed Abbott Laboratories's company estimably $250 million leaving patients out of their medications.[37]

High-pressure crystallizations are sometimes employed during solid-form screening, as in the pharmaceuticals Dalcetrapib[38] and Galunisertib[39]. High pressure conditions can complicate the study of polymorphism, however. Pressure condenses the crystal to be tightly packed, making it harder to solve the structure with current X-ray diffraction techniques. Besides the issues associate with small size, sample graininess and low X-ray scattering intensities may also hinder observations.[18] Complementary techniques such as Infrared or Raman spectrometry could be used for characterizing high-pressure samples.

Crystal structure prediction is another complementary tool used to study crystal structures computationally. It is a powerful method that has led to many important discoveries in areas such as drug design, material discovery, high-pressure chemistry, and mineralogy of the Earth's and planetary interiors.[40] The work presented in this thesis focuses on combining simulated Raman spectroscopy with crystal structure prediction to provide additional information to distinguish polymorphs.

1.1 Crystal Structure Prediction

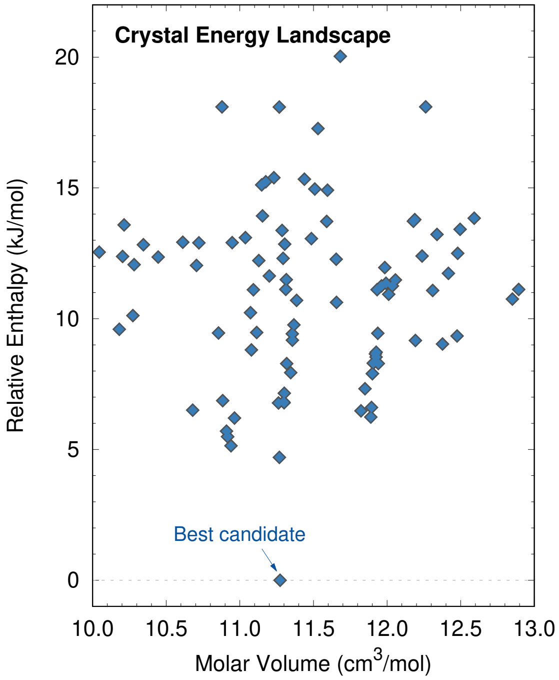


Figure 1.1: A sample crystal energy landscape. Ideally, the most stable structure will correspond to the experimental form

Crystal structure prediction is the art of searching for most energetically stable crystal packing. Typically, crystal energy landscape is mapped out and the best candidate is determined from the most energetically stable structure as shown in Figure 1.1. It is a powerful tool to predict possible crystal structures starting from only the molecular structures. Many examples can be found in the literature where crystal structures were predicted first and then later found experimental tally, such as the stable N_8[41] and nitrogen phase λ.[18,42] Moreover, computational predictions of crystal structures have been used in aiding the characterization of polymorphs from powder X-ray diffraction.[43] The crystal structure prediction at finite temperature and pressure can be used to construct phase diagrams that provides insight into the conditions under which a desired crystal packing will be thermodynamically preferred. Thus, it has potential to be a valuable tool in advance materials, pharmaceutical science, agriculture science, and mineralogy.

Unfortunately, despite decades of developments, there is still no perfect crystal structure prediction approach. The prediction suffers from several limitations. For example, it is feasible to generate all possible crystal packing but increasing numbers of atoms, chirality, and conformational flexibility lead to unimaginable numbers of degrees of freedom. Evaluating all of them exhaustively would be tremendously demanding computationally, so the searching method must be efficient enough and developing criteria for when to stop the searching can also be important. Predicted crystal structures are typically ranked by energy. However, this requires high accuracy since lattice energy differences between polymorphs can be as low as 1 kJ/mol.[44] An energy ranking method that predicts accurate rankings at ambient conditions may be inaccurate at high pressure.[45] Even if the energy ranking is accurate, there is possibility that the experimentally observed crystals are kinetically accessible metastable crystal forms, rather than the thermodynamically stable ones. In order to understand crystal structure prediction, the process is divided roughly into two categories. 1) crystal structure generation and 2) the energy ranking.

1.1.1 Crystal Structure Generation

The idea of generating crystal structures from a known molecular structure is conceptually simple. To create a crystal structure, simply pick the unit cell dimensions and insert molecules into the cell. The challenge lies in the massive number of potential crystal structures, which grows exponentially with the number of degrees of freedom in the system. The unit cell could have a variety of shapes and sizes, and it could hold differing numbers of molecules. The orientations and conformations of those molecules can differ as well. Two common approach to generate crystal structures are random generation and genetic and evolutionary algorithm.

Ab Initio Random Structure Searching (AIRSS) is one of the random generation method. It is a thorough method that potentially captures every possible combination of crystal packing. Truly random searching can be efficient in small systems, but the search space quickly becomes too large as the complexity of the system increases. The method can be tweaked by constraining with some biases to make it less exhaustive such as using known lattice parameters, space group, crystal cell type, number of molecules, etc. Many crystal structures have been predicted successfully using the AIRSS approach for example, ice,[46] methanol,[47] and nitrogen crystals.[18,42]

Genetic and evolutionary algorithm is a combination of crystal structure generation and energy ranking process. The initial structures are randomly generated. After energy ranking, the low energy structures are selected to be parents for the next generation. Parents are recombined with one another partially or mutated to create new offspring. These offspring become the next generation of initial structures and go through the same process until the most stable structure are obtained. Hence, the genetic and evolutionary algorithm offer more efficiency way to search for candidate. The only common problem is that sometimes the search space can be too narrow and may get trapped in local minima instead of finding the global minima.

1.1.2 Energy Ranking

Lattice energies are typically used to rank crystal structure candidates. The crystal cells are fully relaxed at some chosen level of theory. The computational methods used for energy evaluation include molecular mechanics, semi-empirical, periodic density functional theory, or *ab Initio* with fragment-based methods. The former provide typically less computing demanding but also less accuracy. A common strategy employs less accurate, lower cost methods to pre-screen candidate crystal structures and then refines the most promising structures with more accurate method.

Density functional theory (DFT) is the most widely used electronic structure method for crystal structure modeling. Conventional semi-local density functional lack long-range correlation. Without a dispersion correction, such functionals describe intermolecular interactions poorly. This is a significant problem for crystal, where the energy is sensitive to intermolecular interaction. Many van-der Waals dispersion corrections have been developed to overcome this deficiency by adding a correction term such as Grimmes's dispersion correction,[48–51] exchange-hole dipole moment model (XDM),[52] Tkatchenko-Scheffler (TS),[53,54] or many-body dispersion (MBD).[55–57] When paired with a high-quality dispersion correction, DFT can often predict crystal energies accurately with reasonable computational cost.[58] The accuracy could be further improved with better functionals coming from higher Jacob's ladder.[59–61] However, using expensive functional such as meta-GGA or hybrid and generalized random-phase approximation, demands more computing resource. Some studies also shown that these expensive functionals failed to locate potential surface energy for conformational structures of ROY.[62,63]

With the inconsistent performance of DFT, one would like to use all-electron approaches such as second-order Møller-Plesset perturbation theory (MP2) or even coupled cluster models (CCSD(T)). On the other hand, CCSD(T) in particular is limited to systems with only a few tens of atoms due to its steep $O(N^7)$ computational cost scaling.

Fortunately, fragment methods such as the hybrid many-body interaction (HMBI) model make it feasible to achieve the high accuracy promised by these wave function methods without the extreme computational cost. HMBI partitions a crystal into small fragments. The properties such as electronic energy can be combined from calculations on individual fragments and small groups of fragments according to many-body expansion theory. Since each fragment is relatively small compared to the whole crystal, the computational cost for computing these fragments will be cheaper.

There are number of studies demonstrated that HMBI could predict lattice energy of molecular crystal as accurate with sub kJ/mol.[64–67] This accuracy is important to discriminate polymorph apart. HMBI also allows accurate predictions of chemical and spectroscopic properties. It is an excellent tool to study crystal structure.

1.2 Raman Spectroscopy

Spectroscopic instruments are essential to observe microscopic properties. Combining additional spectroscopic observables such as infrared, Raman, or nuclear magnetic resonance spectra, could remarkably increase confidence in the structural assignments.[44,68–71] For example, Hirata and co-workers helped resolve several controversies surrounding the interpretation of experiments on two difference ice phases through *ab initio* simulation of structures and vibrational spectra.[71–73]

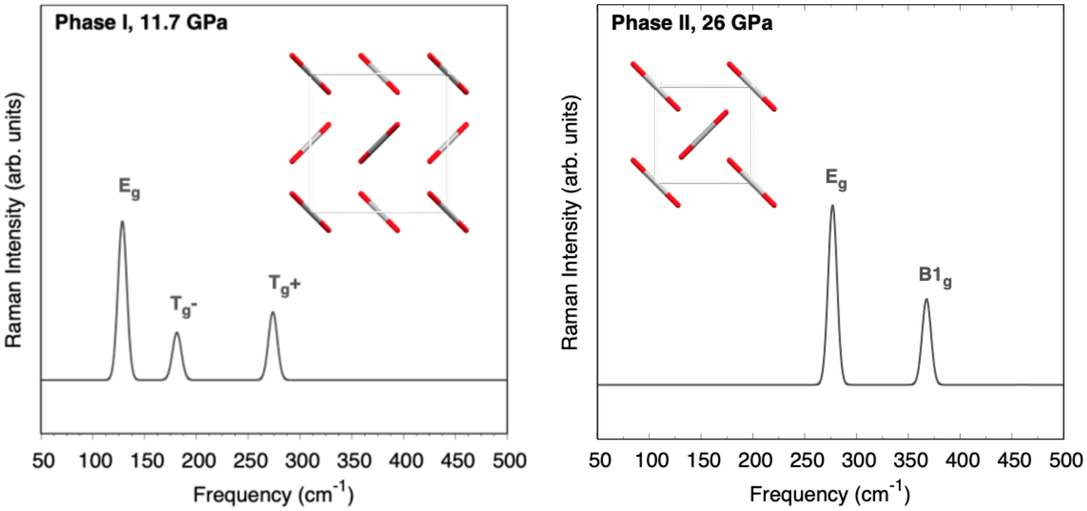


Figure 1.2: Low frequency Raman spectrum for solid carbon dioxide phase I (left) and II (right).

Raman spectra for the same compound can be different depend on crystal packing sd shown in Figure 1.2. Because low frequency phonons are sensitive to molecular packing and Raman spectrum provide molecular fingerprints for these region, Raman spectroscopy is essential complementary tool to observe crystal structure. Incorporating Raman simulation with crystal structure study could prove to be beneficial. For crystal structure prediction, it is common to find many polymorphs with very similar lattice energies. Structure determination based on electronic energy alone is often an insufficient descriptor. For example, crystal structure prediction was employed to determine structures for form I and II phenytoin, an anti-epileptic drug.[74] The study found crystal structures that matched grazing incidence X-ray diffraction pattern for both form. However, the actual crystal structure for form II were the higher energy structure and was noticed only by comparing the experimental and

DFT Raman spectra.[75]

1.3 Outline of This Dissertation

Solid carbon dioxide and nitrogen are governed by weak and non-polar intermolecular interactions. Many packing arrangements and orientations are inevitable raising a large number of viable packing motifs at high pressure. Experimentally, observing and characterizing of high-pressure polymorphs have proved challenging. Flat energy landscape is often observed. Some of high-pressure polymorphs are kinetic path-dependence and hysteresis in the phase transitions. As a results, high-quality experimental data are difficult to be obtained leaving many high-pressure crystal structures of nitrogen to remain unknown over decades, or creating controversies such as possibility that carbon dioxide phases III and VII are identical.

In this thesis, high-level fragment-based electronic structure method and Raman simulation are implemented to accurately provide more robust tools for studying high-pressure crystal structures. The next chapter will describe the conceptual and methodology to study electronic structure of crystal using fragment-based calculation, the approach to obtain accurate crystal structure at low-cost, and Raman simulation.

The results of this dissertation will be divided into three subsequential chapters. First, we will investigate controversy surrounding the identity of carbon dioxide phases III and VII. Large-basis-set quasi-harmonic second-order Møller-Plesset are performed to examine these crystal structures. Raman spectra are predicted to provide molecular fingerprints. Furthermore, crystal structure prediction is employed to search for other potential structures which might account for phase III or VII. All evidences to identify the nature of phases III and VII will be discussed.

Second, we will revisit high-pressure solid nitrogen. Phase λ is a recently discovered high-pressure solid nitrogen form whose structure was solved by combining experimental monoclinic lattice parameters with atomic positions from an earlier, computationally predicted structure that had similar unit cell dimensions. The AIRSS is performed to demonstrate the capability of our method to find candidate structures for the λ phase. Raman spectra is incorporated to help narrow down the candidate structures. Furthermore, the thermodynamic stability of the λ phase relative to other phases will be examined. Lastly, AIRSS is combined with Raman simulation to identify candidate structures for nitrogen phase ζ, the phase which remains unknown decades after it was first observed spectroscopically, despite numerous experimental and theoretical investigations. The crystal structure prediction results will be discussed along with some potential starting candidate for phase κ, whose structure is also unknown. The conclusions of this dissertation will be summarized in the last chapter.

Chapter 2

Methods

2.1 Outline

In order to model crystal structures at high pressure, one need to consider how to study them appropriately. High-level $ab-initio$ methods can describe such systems accurately in principle, though the computational costs can be steep. Fortunately, fragment-based methods, such as HMBI, reduce the computational cost by taking advantage of the many-body expansion. It help studying crystal system more affordable while remain high accuracy. Here, we describe the fragment-based calculation of HMBI, which is used for both electronic structure energy calculations an the calculation of vibrational frequencies. We also try to describe the crystal at finite temperature, employing the quasi-harmonic approximation, to account for thermal expansion. Next, we will illustrate structure refinement scheme to predict crystal structure at low cost. Lastly, the implement of Raman spectra simulation will be described.

2.2 Fragment-based hybrid many-body interaction (HMBI)

Crystal structure modeling requires accounting for both the central unit cell and its periodic images in order to describe the crystaline environment correctly. Thus, the computational cost can be demanding. To make the study more affordable, fragment-based methods introduce efficient way to model crystalline systems by exploiting the many-body expansion.[76–78]

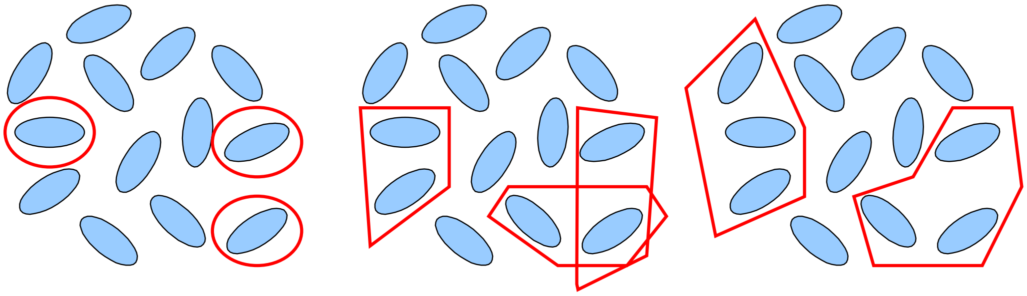


Figure 2.1: Fragment of many-body contribution terms from the cluster of 1-body (left), 2-body (middle), and 3-body (right).

Figure 2.1 shows example of interactions involving particles in the cluster. The total energy of the system can be expanded in to fragment contributions as following

$$E_{total} \quad = \quad E_{1-body} + E_{2-body} + E_{3-body} + ... \tag{2.1}$$

The n-body contributions can be expressed as following

$$E_{1-body} = \sum_{i} E_i \tag{2.2}$$

$$E_{2-body} = \sum_{i,j} (E_{i,j} - E_i - E_j) \tag{2.3}$$

$$E_{3-body} = \sum_{i,j,k} (E_{i,j,k} - E_{i,j} - E_{i,k} - E_{j,k} - E_i - E_j - E_k) \tag{2.4}$$

where E_i, E_j, and E_k are energies of monomer i, j, and k (Figure 2.1(left)), $E_{i,j}$, $E_{i,k}$, and $E_{j,k}$ are energies of dimer i,j, i,k, and j,k (Figure 2.1(middle)), and $E_{i,j,k}$ is energy of trimer i,j,k (Figure 2.1(right))

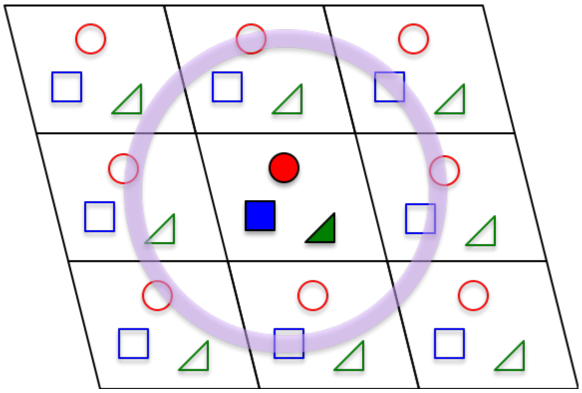


Figure 2.2: The hybrid many-body interaction model for crystal.

Adopting many-body expansion, HMBI decomposes the total energy of the crystal into a summation of fragment contributions as follows:

$$E_{crystal}^{HMBI} = E_{1-body} + E_{SR,2-body} + E_{LR,2-body} + E_{many-body} \tag{2.5}$$

where 1-body terms correspond to the intramolecular energies of the individual molecules in the unit cell, two-body terms to the pairwise interaction energies between pairs of molecules (both within the central unit cell and involving periodic image molecules), and many-body terms to the non-additive three-body and higher contributions. The 2-body also can be partitioned into short-range and long-range interactions based on the intermolecular distance.

In Figure 2.2, the 1-body and short-range 2-body terms constitute the dominant contribution to the crystal lattice energy. These two contributions need to be evaluated accurately. The typical level of theory used to study this contributions are MP2 or CCSD(T). The long range 2-body and many-body contributions are weaker and can often be estimated at a lower level of theory, either with quantum mechanics such as periodic-HF or with molecular mechanics such as the Amoeba force field.[79]

2.3 Quasi-Harmonic Approximation

Thermal expansion is a phenomena where the crystal volume changes with temperature. It can be important to include this effect when dealing with crystals at low pressure.[80] To include thermal expansion effects, crystal structure optimizations can be performed by minimizing the free energy as a function of temperature and pressure, $G(T, P)$:[81]

$$G(T, P) = U_{el} + PV + F_{vib}(T) \tag{2.6}$$

where U_{el} is the electronic energy (Equation 2.5), PV is the pressure-volume term, and F_{vib} is the Helmholtz vibrational free energy as a function of temperature in the standard harmonic approximation.

$$F_{vib}(T) = N_a \sum_i \left(\frac{\hbar\omega_{k,i}}{2} + k_b T \ln \left[1 - \exp\left(-\frac{\hbar\omega_{k,i}}{k_b T} \right) \right] \right), \tag{2.7}$$

where N_a is number of atoms, $\hbar$ is the Dirac constant, k_b is the Boltzmann constant, and $\omega_{k,i}$ is frequency term.

According to the quasi-harmonic approximation, unit-cell volume changes systematically with vibrational frequencies. Mode-specific Grüneisen parameters $\gamma_{k,i}$ for each mode i were approximated at each reciprocal lattice vector k via finite difference of the frequencies using modestly compressed and expanded cells. The $\omega_{k,i}$ at an arbitrary volume were then computed from the reference frequencies, reference volume, and Grüneisen parameters,

$$\omega_{k,i} = \omega_{k,i}^{ref} \left(\frac{V}{V^{ref}} \right)^{-\gamma_{k,i}} \tag{2.8}$$

where

$$\gamma_{k,i} = -\left(\frac{\partial \ln \omega_{k,i}}{\partial \ln V} \right) \approx -\frac{\ln(\omega_{k,i,2}) - \ln(\omega_{k,i,1})}{\ln(V_2) - \ln(V_1)} \tag{2.9}$$

2.4 Structure Refinement

HMBI-based MP2 or CCSD(T) calculations can potentially provide higher-accuracy structures compared to inexpensive DFT models. Fragment-based methods are far less expensive compared to periodic-MP2 or CCSD(T). However, full crystal relaxations with HMBI-based MP2 or CCSD(T) are still somewhat more computationally expensive than those with DFT, especially with the larger basis sets needed. To make the prediction further less expensive, one would like to refine the structures at the MP2/CCSD(T) levels of theory from the cheap structure prediction, such as DFT calculation.[82]

The structure refinement follows the simplified quasi-1-D optimization approach. First, the crystal structures are optimized with DFT at a series of external pressures. With standard generalized gradient approximation functionals, DFT usually gives accurate structures despite less reliable energy prediction. Second, single-point HMBI electronic energies are computed with MP2 or CCSD(T) and different many-body treatments at each DFT geometry. This step is to extract energy term as a function of volume $E(V)$. A PV pressure-volume contribution was added to the resulting, $E(V)$ to obtain enthalpies versus volume: $H = E(V) + PV$. The enthalpy-volume curves are fitted to the Murnaghan

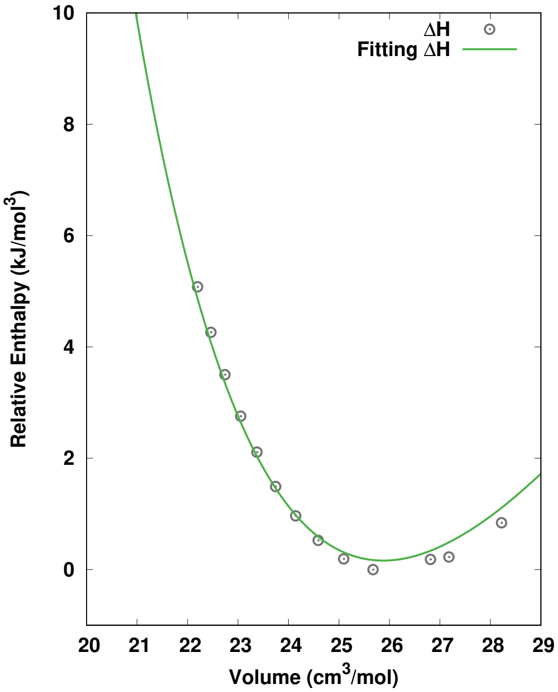


Figure 2.3: Fitting Murnaghan equation of state with predicted enthalpy-volume curve from series geometry at external pressure.

equation of state as shown in Figure 2.3. The Murnaghan equation of state is written as follows:

$$H(V) \quad = \quad H_0 + \frac{B_0 V}{B_0'}\left[\frac{(V_0/V)^{B_0'}}{B_0' - 1} + 1\right] - \frac{B_0 V_0}{B_0' - 1} \tag{2.10}$$

where the enthalpy (H_0), volume (V_0), bulk modulus (B_0), and its first pressure derivative (B_0') at zero pressure are the fitting parameters. The optimal volume V_0 and corresponding enthalpy H_0 are extracted from the minimum of the fit. Atomic coordinates can be obtained via interpolation of the DFT fractional coordinates to the optimal volume extracted from the equation of state fit. As a results, one can predict high-accuracy structure from the refinement approach.

This approach is similar to how the quasi-harmonic approximation (QHA) is sometimes performed from Section 2.3. Here, zero-point and thermal vibrational contributions are neglected. This neglect is reasonable at higher pressures where thermal expansion effects are expected to be small. Otherwise, the quasi-harmonic approximation should be included to improve accuracy.

2.5 Raman Simulation

Raman spectroscopy is one of the essential tools to characterize crystal structures experimentally. Simulating Raman spectrum for the crystal systems required even more computing resources due to the astronomic cost of computing third derivatives. Fortunately, the cost can be reduced using a fragment-based scheme such as HMBI. Zone-center ($\mathbf{k} = 0$) phonons can be computed from decomposition of the mass-weighted Hessian as following

$$H_{crystal}^{HMBI} = H_{1-body} + H_{SR,2-body} + H_{LR,2-body} + H_{many-body} \tag{2.11}$$

In the harmonic oscillator approximation, the potential or mass-weighted Hessian can be written as

$$H_{crystal}^{HMBI} = \frac{1}{2} \sum_{N} f_i q_{i,j}^2 \tag{2.12}$$

The diagonalized product of Equation 2.12 is mass-weighted force constant, f at N degrees of freedom. It is used to determined vibrational frequencies as following

$$\omega_i = \frac{1}{2\pi} \sqrt{f_i} \tag{2.13}$$

Raman intensities, I_{Raman}, are then approximated at the 1- and 2-body level from polarizability derivatives.[72]

$$\tilde{\alpha} \approx \tilde{\alpha}_{1-body} + \tilde{\alpha}_{SR,2-body} \tag{2.14}$$

$$I_{Raman} = 45\tilde{\alpha}^2 + 7\tilde{\gamma}^2 \tag{2.15}$$

$$\tilde{\alpha} = \frac{1}{3}\sum_{pp} \tilde{\alpha_{pp}} \tag{2.16}$$

$$\tilde{\gamma}^2 = \frac{1}{2}\sum_{pq} (\tilde{\alpha_{pp}} - \tilde{\alpha_{qq}})^2 + \frac{3}{4}\sum_{pq} (\tilde{\alpha_{pq}} + \tilde{\alpha_{qp}})^2 \tag{2.17}$$

where $\tilde{\alpha}$ is polarizability tensor along axis p and q. For simplicity, many-body effects are neglected in Equation 2.14. This approximation is reasonably given the non-polar nature of all crystals studied in this work.

Chapter 3

Theoretical Predictions Suggest Carbon Dioxide Phases III and VII are Identical

3.1 Introduction

Carbon dioxide is one of the most fundamental chemical species on earth, yet its solid-phase behavior at high pressures continues to confound. Starting with the 1994 powder X-ray diffraction structure of phase III,[2] researchers have mapped out a rich phase diagram with 8-10 crystalline phases that range from molecular crystals at lower pressures to extended covalent and ionic phases at high pressures.[83] Experimental characterization of these crystal structures and their solid-solid phase boundaries has often proved challenging, with considerable kinetic path-dependence and hysteresis in the phase transitions, difficulty

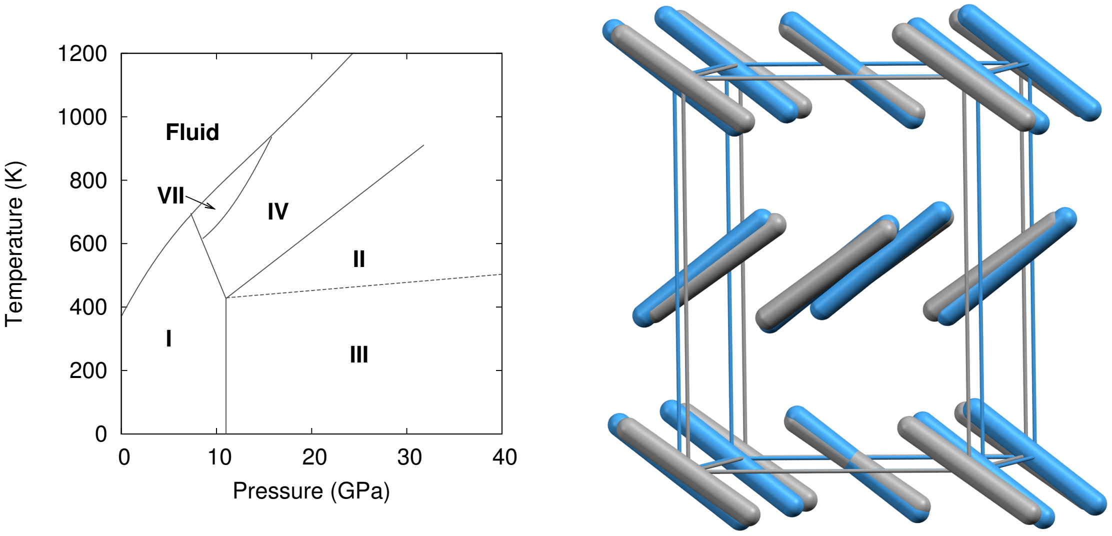


Figure 3.1: Phase diagram of carbon dioxide up to 40 GPa, and structure overlay of the experimental crystal structures for phases III (blue) and VII (gray). Root-mean-square deviation[1] = 0.24 Å.

in obtaining high-quality diffraction data, and sharp pressure gradients within samples that complicate spectroscopic measurements. As a result, the literature on high-pressure carbon dioxide contains numerous contradictory experimental interpretations. Fortunately, substantial advances in computational modeling of molecular materials[44] mean that *ab initio* calculations can help resolve such controversies and play an integral role in molecular crystallography. Using high-level electronic structure calculations, the present study investigates several molecular crystal phases of carbon dioxide and demonstrates quantitative agreement between predicted and experimentally observed structural, mechanical, and spectroscopic properties for several of them. However, the same theoretical evidence indicates that the long-accepted structure of phase III carbon dioxide is inconsistent with spectroscopic data, and that phases III and VII are likely identical.

Controversy has long surrounded the high-pressure phase diagram of carbon dioxide (Figure 3.1). Substantial research developed a framework in which lower pressure molecular crystal phases I, III, and VII transition to "intermediate bonding" phases II and IV at moderate pressures (∼20–40 GPa), before eventually forming extended covalent or ionic phases at higher pressures.[83] The intermediate bonding phases purportedly exhibit elongated and/or bent carbon dioxide molecules and abnormally large bulk moduli.[9,10,84] However, subsequent experiments[11,16] and density functional theory (DFT) calculations[4,85] have challenged this interpretation, suggesting instead more traditional molecular crystal structures at intermediate pressures, with linear carbon dioxide molecules and typical bulk moduli.

Phases III and VII represent another conundrum. X-ray diffraction studies suggest that both phases adopt similar $Cmca$ space group structures (Figure 3.1).[2,17] The primary differences lie in effectively swapping the a and b lattice constants and slightly altering the angle the molecule forms relative to the c crystallographic axis. Phase III can be formed at room temperature by compressing phase I (dry ice) to pressures above ∼12 GPa, though the precise phase boundary remains unclear due to the sluggish martensitic phase change.[83] Obtaining high-quality diffraction data for phase III has proved challenging, and the currently accepted structure was extracted from powder X-ray diffraction on a sample believed to contain a mixture of phases I and III.[2] Despite routinely being included in the low-temperature region of the phase diagram, phase III is actually believed to be metastable and monotropic relative to phase II. It converts to phase II upon annealing to ∼500 K at 12 GPa or above.[9,11]

Phase VII occurs in a narrow pressure and temperature region around 15–17 GPa and 750 K,[83] and producing it experimentally can also be challenging.[17,86] Nevertheless, its structure was determined via X-ray diffraction on single crystals grown from the melt. Given the difficulty of obtaining quality diffraction data for phase III and the correspondingly poor constraints on its structure,[7,14,17] the possibility that phases III and VII were actually the same phase was raised immediately.[17] However, the non-contiguous existence domains for III and VII in the phase diagram and subtle differences in the Raman spectra were cited in favor of there being two distinct phases.[17]

Resolving these issues experimentally has proved challenging. *Ab initio* crystallography plays an increasingly important role in molecular crystals, materials, and even biological systems. Computational refinement of experimental crystal structures has long been integral in many studies, and advances in crystal structure prediction[87,88] have made *ab initio* structure determination even more viable. Unfortunately, energy alone is often an insufficient descriptor—one commonly predicts multiple potential structures whose energies are sufficiently close so as to prevent clear discrimination. By predicting and comparing additional spectroscopic observables such as infrared, Raman, or nuclear magnetic resonance spectra, however, one can markedly increase confidence in the structural assignments.[44,68–71]

Here, several molecular crystalline phases of carbon dioxide are revisited with large basis, quasi-harmonic second-order Møller-Plesset perturbation theory (MP2) electronic structure theory calculations.[80,81] These high-level calculations quantitatively reproduce structures, mechanical properties, and Raman spectra across most of the phases considered. However, these models do not predict a distinct phase III structure whatsoever. Moreover, even if the experimental structure were correct, the predictions here indicate that its Raman spectrum would differ from the experimentally observed spectra. In contrast, the predictions for phase VII are highly consistent with those observed experimentally for phase III.

These high-level *ab initio* calculations are made feasible for crystalline carbon dioxide using the fragment-based hybrid many-body interaction (HMBI) model.[77,89] The HMBI model partitions the crystal into molecular fragments. Unit cell monomers and short-range dimers are treated quantum mechanically, while longer-range dimers and many-body effects are approximated with a classical AMOEBA[90,91] polarizable force field. A quasi-harmonic phonon treatment was employed to incorporate zero-point vibrational energy and thermal vibrational effects computed at the same MP2 level of theory into the model. Previous quasi-harmonic MP2 and coupled cluster HMBI calculations on phase I carbon dioxide predicted the thermal volume expansion within 2%, the sublimation enthalpy within 1.5 kJ/mol, and the sublimation entropy within 2 J/mol K between 0–200 K.[81] The sublimation point was predicted within 3 degrees Celsius.[80]

3.2 Methods

3.2.1 Quasi-Harmonic structure optimizations

Crystal structure optimizations were performed by minimizing the free energy as a function of temperature and pressure, $G(T, P)$:

$$G(T, P) = U_{el} + PV + F_{vib}(T) \tag{3.1}$$

where U_{el} is the electronic energy, PV is the pressure-volume term, and F_{vib} is the Helmholtz vibrational free energy in the standard harmonic approximation. The phonon frequencies were computed via the quasi-harmonic approximation: Frequencies for the electronic-energy minimized structure were computed via lattice dynamics on a $3 \times 3 \times 3$ Monkhorst-Pack grid and a $3\times3\times3$ supercell. Mode-specific Grüneisen parameters $\gamma_{k,i}$ for each mode i were approximated at each reciprocal lattice vector k via finite difference of the frequencies using modestly compressed and expanded cells. The frequencies $\omega_{k,i}$ at an arbitrary volume were then computed from the reference frequencies, reference volume, and Grüneisen parameters,

$$\omega_{k,i} = \omega_{k,i}^{ref} \left(\frac{V}{V^{ref}} \right)^{-\gamma_{k,i}} \tag{3.2}$$

See Refs 81 and 80 for details.

3.2.2 Electronic Structure Calculations

The electronic energy calculations were performed using the fragment-based HMBI model.[77?] One-body and short-range two-body contributions were computed at the density-fitted MP2/CBS level using Molpro 2012,[92] while the long-range and many-body contributions were evaluated using the AMOEBA force field, as implemented in Tinker 6.3.[93] CBS extrapolation of the energies and forces was performed via standard two-point extrapolation of the aug-cc-pVTZ and aug-cc-pVQZ[94] results. Carbon dioxide force field parameters were generated using Poltype 1.1.3.[91] Space group symmetry is exploited throughout to reduce the number of monomer and dimer fragments that need to be computed.[95]

3.2.3 Raman Spectra Predictions

To compute Raman spectra, the quasi-harmonic HMBI MP2/CBS structures (or the experimental structures in selected cases) were electronic energy-minimized at the HMBI

MP2/aug-cc-pVDZ level with lattice parameters held fixed. Zone-center ($\mathbf{k} = 0$) phonons were computed at the same level of theory. Raman intensities were then approximated at the 1- and 2-body level.[72] Many-body effects were neglected in computing the polarizability derivatives that are used to evaluate the Raman intensities, but these effects are generally small in carbon dioxide. Space group symmetry was again employed. To minimize numerical noise associated with finite differencing the polarizability derivatives, individual fragment MP2 frequency calculations were performed via analytic second derivatives as implemented in Gaussian 09.[96] The polarizability derivatives were evaluated via finite difference of the polarizabilities. Simulated spectra were plotted as a sum of Gaussian functions with a full width at half maximum of 10 cm^{-1}.

3.2.4 Crystal Structure Prediction

Potential carbon dioxide crystal structures were generated via the evolutionary algorithms implemented in USPEX.[97] Each of six runs was seeded with ten random structures from randomly chosen space groups, containing either two or four molecules in the unit cell (the unit cell sizes for phases I, II, and III/VII). Structures were relaxed at ambient pressure and energies computed using Tinker and the OPLS-AA force field.[98] New structures were constructed for 15–20 generations via heredity, coordinate/rotational mutations, or lattice mutation.[97] This process generated 660 structures with $Z = 2$ and 1083 structures with $Z = 4$ were generated, though many of these were redundant or clearly energetically unfavorable.

After removal of obvious duplicates, the most stable 91 structures were then refined under 11.8 GPa of pressure (the pressure for the experimental powder X-ray diffraction of phase III) in Quantum Espresso[99] using periodic PBE-D2,[100,101] an 80 Ry planewave cutoff, a 7×7×7 Monkhorst-Pack *k*-point grid, and ultrasoft pseudopotentials C.pbe-rrkjus.UPF and O.pbe-rrkjus.UPF from http://www.quantum-espresso.org. Structures were analyzed in terms of energy, unit cell volume, root-mean-square deviations in atomic positions, and by comparing simulated powder X-ray diffraction patterns (wavelength 0.6888 Å) generated by Mercury.[102] These were then compared against the experimental phase III diffraction pattern and the simulated pattern for the experimentally reported phase III crystal structure.[2]

3.2.5 Data Analysis

Experimental crystal structures and Raman data was taken from the literature. When tabulated data was unavailable, Raman spectra and equation of state data were digitized from published figures. Root-mean-square deviations between predicted and optimized structures employ 15-molecule clusters[1], as implemented in Mercury.[102]) Bulk moduli were computed by fitting to the Vinet equation of state,[15] which proves more numerically robust than the Birch-Murnaghan one here.[81]

3.3 Results and Discussion

3.3.1 Comparison of predicted and experimental lattice parameters

Tables 3.1–3.3 compare predicted lattice parameters against reported experimental structures. Selected predicted structures from the literature are included as well. Matlab/GNU Octave scripts are provided separately which contain the quasi-harmonic room temperature HMBI MP2/CBS + AMOEBA structures for the pressures calculated explicitly and which can interpolate the structures at arbitrary pressures 1–20 GPa for phase I, and 0–60 GPa for phases II and III/VII. Because of the smooth variation in lattice parameters and fractional coordinates, interpolation reproduces the structures well.

Table 3.1: Comparison of predicted and experimental lattice parameters for phase I carbon dioxide ($Pa\bar{3}$ space group) at ambient temperature and selected pressures.

Method	Pressure	Temperature	a (Å)	Volume (cm^3/mol)	Source
Expt. (1998)	1 GPa	293 K	5.4942(2)	24.969(6)	Ref 103
MP2/CBS	1 GPa	296 K	5.504	25.11	this work
PBE-D3(BJ)[a]	1 GPa	n/a	5.515	27.59	Ref 85
Expt. (1994)	7.46 GPa	295 K	5.056(1)	19.46(2)	Ref 2
MP2/CBS	7.46 GPa	296 K	5.063	19.54	this work
Expt. (1994)	11.8 GPa	295 K	4.939(10)	18.14(11)	Ref 2
MP2/CBS	11.8 GPa	296 K	4.941	18.16	this work
MP2/aDZ[b]	11.8 GPa	n/a	4.91	17.82	Ref 3

[a] Thermal expansion was not included in the modeling.
[b] Using binary interaction model, without Counterpoise correction or thermal expansion.

Table 3.2: Comparison of predicted and experimental lattice parameters for phase II carbon dioxide ($P4_2/mnm$ space group).

Method	Pressure	Temperature	a (Å)	c (Å)	Volume (cm^3/mol)	$r_{C=O}$ (Å)	Source
Expt. (2002)	28 GPa	680 K	3.5345	4.1401	15.57	1.331(3)	Ref 10
MP2/CBS	28 GPa	680 K	3.504	4.125	15.25	1.155	this work
PBE-D3(BJ)[a]	28 GPa	n/a	3.51	4.06	15.09	1.157	Ref 85
Expt. (2014)	25.8 GPa	295 K	3.516(2)	4.104(2)	15.28	1.14	Ref 11
MP2/CBS	25.8 GPa	298 K	3.515	4.124	15.34	1.155	this work

[a] Thermal expansion was not included in the modeling.

Table 3.3: Comparison of predicted and experimental lattice parameters for phase III & VII carbon dioxide ($Cmca$ space group).

Method	Pressure	Temperature	a (Å)	b (Å)	c (Å)	Volume (cm^3/mol)	Source
Expt. III	11.8 GPa	295 K	4.330(15)	4.657(5)	0.5963(9)	18.11(8)	Ref 2
MP2/CBS	11.8 GPa	298 K	4.635	4.285	5.953	17.80	this work
MP2/aDZ[a]	11.8 GPa	n/a	4.63	4.33	5.80	17.82	Ref 3
PBE-D3(BJ)[b]	11.8 GPa	n/a	–	–	–	18.10	Ref 85
Expt. VII	12.1 GPa	726 K	4.746(1)	4.313(1)	5.948(1)	18.33	Ref 17
MP2/CBS	12.1 GPa	726 K	4.680	4.316	5.989	18.22	this work

[a] Using binary interaction model, without Counterpoise correction or thermal expansion.
[b] Thermal expansion was not included in the modeling; lattice parameters were not reported.

3.3.2 Equations of state

The P-V isotherms generated here were fitted to the Vinet equation of state (EOS),[15]

$$P = 3B_0 \frac{(1-\tilde{V})}{\tilde{V}^2} \exp\left[\frac{3}{2}(B_0' - 1)(1-\tilde{V})\right] \tag{3.3}$$

via non-linear least squares fitting to extract V_0, B_0, and B_0'. Here, $\tilde{V} = (V/V_0)^{1/3}$. As discussed in our earlier work,[81] fits using the Birch-Murnaghan EOS proved ill-constrained, with very different sets of V_0, B_0, and B_0' giving similar quality fits. The Vinet EOS behaves much better numerically in our experience, so it was used here instead.

Table 3.4 summarizes the equations of state obtained here and compares them against other theoretical and experimental values found in the literature. The agreement of quasiharmonic MP2 predictions for the bulk modulus of phase I has been discussed extensively previously.[81] The phase I bulk modulus decreases several fold between 0 K and room temperature. The B_0 obtained from quasi-harmonic MP2/CBS as a function of temperature is reasonable, as exemplified by the good agreement between it and the experimental data at room temperature shown here. Note that the bulk modulus reported by Yoo et al[10] appears to represent an unphysical fit, with a zero-pressure volume at room temperature that is smaller than even experimentally known volumes near 0 K.[81]

Other MP2 and dispersion-corrected density functional theory predictions that neglect thermal expansion typically obtain bulk modulus values that are several-fold larger than the room temperature value.[85,104] The PBE result of Bonev and co-workers[4] is fortuitously good, because the neglect of dispersion artificially makes the intermolecular interactions overly repulsive and expands the unit cell, which partially compensates for the neglected thermal expansion.

Table 3.4: Summary of predicted and experimental room-temperature bulk modulus data for phases I–III. Literature data typically employed the Birch-Murnaghan EOS, while the fits in this work employed the Vinet EOS.

Method	Source	V_0 (cm^3/mol)	B_0 (GPa)	B'_0
Phase I				
PBE[a]	Bonev et al[4]	31.9	3.2	8.1
PBE-D3(BJD)[a]	Gohr et al[85]	27.6	9.9	5.4
MP2/aDZ[a,b]	Li et al[104]	23.7	16.1	6.9
MP2/aTZ[a,b]	Li et al[104]	24.5	12.1	7.7
MP2/CBS[a]	Heit et al[81]	24.6	10.6	7.8
MP2/CBS @ 296 K	Heit et al[81]	29.4	3.3	9.0
Experiment	Giordano et al[8]	30.1±1	3±1	8.4±0.8
Experiment	Giordano et al[8c]	30.7±1	2±1	9±1
Experiment	Giordano et al[8d]	30.1±1	3±5	9±3
Experiment	Liu[13]	31.4	2.9	7.8
Experiment	Yoo et al[10]	25.1	6.2	6.2
Phase II				
PBE[a]	Bonev et al[4]	30.8	4.4	6.7
PBE-D3(BJD)[a]	Gohr et al[85]	27.6	7.5	6.3
MP2/CBS @ 298 K	this work	28.1	4.1	8.4
Experiment	Yoo et al[10]	18.0	131.5	2.1
Experiment	Datchi et al[11]	25±1	16±5	5.2±0.4
Experiment	Datchi et al[11]	26.7±0.2	8.5±0.3	6.3[e]
Phase III				
PBE[a]	Bonev et al[4]	33.0	3.53	7.12
PBE-D3(BJD)[a]	Gohr et al[85]	28.4	6.1	6.9
MP2/CBS @ 298 K	this work	28.0	4.6	5.5
Experiment	Yoo et al[10]	19.8	87	3.3

[a] No quasiharmonic approximation/finite temperature.

[b] No Counterpoise correction.

[c] Re-fit of Olinger data.[12]

[d] Re-fit of Liu data.[13]

[e] B'_0 was fixed to Gohr PBE-D3(BJD) value.

For phase II, the room-temperature MP2/CBS predictions are in reasonable agreement with the data PBE-D3 data of Gohr et al,[85] though the MP2 values are again slightly smaller (presumably due to neglect of thermal expansion in the PBE-D3 work). Again, PBE without dispersion produces a somewhat softer crystal that fortuitously agrees reasonably well with the MP2/CBS results. The MP2 predictions are in moderate agreement with the experimental results of Datchi and co-workers. As noted in that work, the lack of experimental low-pressure data for phase II introduces considerable uncertainty into their EOS fits. In contrast, the predicted B_0 of 131.5 from Yoo et al[10] is a clear outlier. The 18.0 cm^3/mol molar volume at zero pressure is only slightly larger than the 16.6 cm^3/mol volume at 20.6 GPa, in marked contrast to the predictions and the fit to Datchi.

For phase III, the three different sets of predict values are in reasonable agreement, with the PBE-D3 calculations that neglect thermal expansion again predicting a bulk modulus that is moderately larger, and the DFT calculations without dispersion predicting a cell that is too large and soft. Experimentally, only one value has been reported to our knowledge.[10] As with the phase I and II EOS data from the same study, the experimental phase III EOS data disagrees substantially from the calculations and suggests a material that is much harder (and similar to the bulk moduli of many metals). Given that all three phases I–III exhibit molecular crystalline structures, the much smaller bulk moduli predicted in three separate studies seem more likely to be correct.

3.3.3 Phase III optimization with fixed experimental cell

To investigate the claimed phase III structure, MP2/aug-cc-pVDZ optimization of the experimental phase III structure was performed with fixed lattice constants. As shown in Figure 3.2, the predicted structure agrees very well with the claimed experimental one, with an rmsd15 of 0.03 Å. For comparison, the rmsd15 between the experimental structures of phase III and VII is an order of magnitude larger at 0.24 Å. This phase III structure lies 4 kJ/mol above the fully relaxed phase VII-like structure, and it predicts a Raman spectrum that is appreciably different from the observed one (will be discussed in Section 3.3.7).

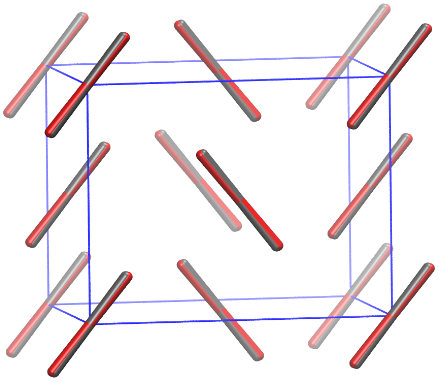


Figure 3.2: Comparison between the 11.8 GPa experimental phase III structure[2] (gray) and the MP2/aug-cc-pVDZ optimized one (red) when the optimization is performed with the experimental lattice parameters held fixed.

3.3.4 Impact of basis set on the predicted Raman spectra

3.3.4.1 MP2/aug-cc-pVDZ optimization for phase I at experimental cell

In this study, Raman spectra were calculated at the MP2/aug-cc-pVDZ, albeit using frozen lattice parameters either optimized with quasiharmonic MP2/CBS or taken from experimentally reported crystal structures. MP2/aug-cc-pVDZ was used to relax the atomic positions within that fixed cell to a stationary point at the same level of theory and compute the phonon modes and Raman intensities. As shown in Figure 3.3, using the larger aug-cc-pVTZ basis to relax the atomic positions and compute the Raman spectra has only a modest impact on the predicted peak positions and intensities in the low-frequency region. For the phase I at 14.5 GPa and room temperature, it shifts the Raman-active lattice phonon modes down by 2% (2–6 cm^{-1}). These changes slightly improve agreement with experiment the T_{g-} and E_g modes, while making it slightly worse for the T_{g+} mode (see Figure 3.5 below).

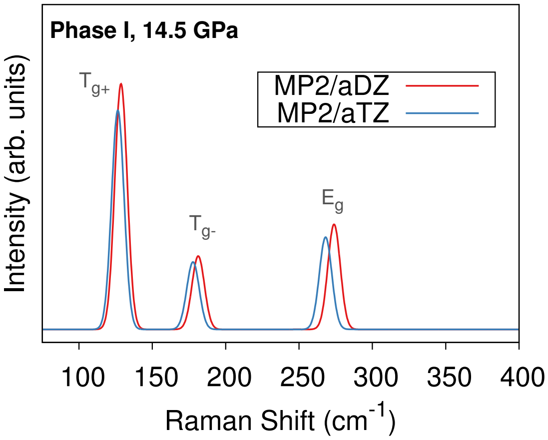


Figure 3.3: Comparison of the MP2/aug-cc-pVDZ and aug-cc-pVTZ predicted Raman spectra for phase I for the quasiharmonic MP2/CBS structure obtained at 14.5 GPa GPa and room temperature.

3.3.4.2 MP2/aug-cc-pVDZ fully optimization for phase IV

Raman spectra for phase IV were fully optimized (both lattice parameters and atomic positions) at the MP2/aug-cc-pVDZ level (Figure 3.11a, will be discussed in Section). While this approach is not expected to be as accurate as the MP2/CBS quasiharmonic calculations for determining the lattice parameters, it actually performs better than one might expect due to fortuitous error cancellation between the underbinding of the crystal at the MP2/aug-cc-pVDZ level and the neglect of thermal expansion. The former overestimates the cell volume, while the latter underestimates it. Detailed analysis of basis set and thermal expansion for phase I has been reported previously.[81] As shown in Figure 3.4, this inexpensive approximation performs reasonably. Other examples of the performance of this approximation in carbon dioxide can be found in Refs 3 and 104

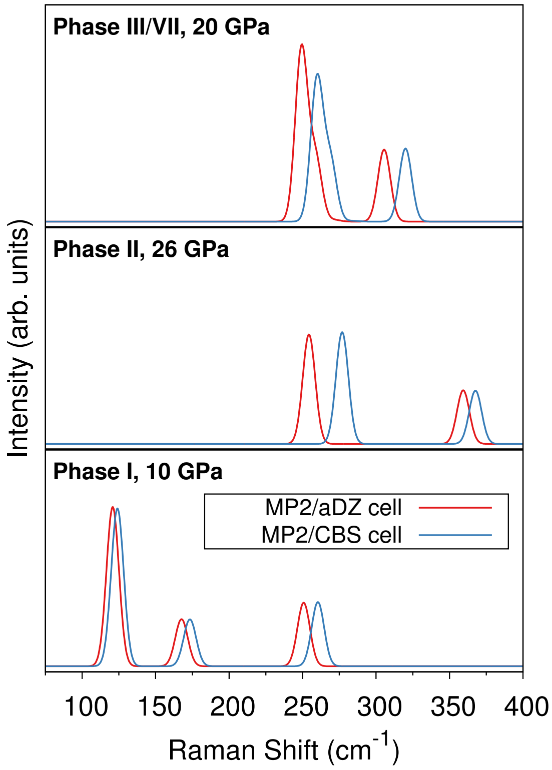


Figure 3.4: Comparison of predicted MP2/aug-cc-pVDZ Raman spectra when the cell is fully optimized with MP2/aug-cc-pVDZ (and no quasiharmonic approximation) instead of using the quasiharmonic MP2/CBS lattice parameters for three different examples.

3.3.5 Pressure dependence of the Raman spectra

3.3.5.1 Phase I

Figure 3.5a compares our predicted MP2/aug-cc-pVDZ Raman shift frequencies from the quasiharmonic approximation (QHA) MP2/CBS unit cells against experiment[8] at room temperature as a function of pressure for phase I. Figure 3.5b reports MP2/aug-cc-pVDZ results from Li et al[3] which employ a slightly different fragment approach. The calculations of Li et al optimized the unit cell at the MP2/aug-cc-pVDZ level instead of MP2/CBS, and no quasi-harmonic approximation was employed. Comparison of Figures 3.5a and 3.5b indicate that the simpler model used by Li et al works well at higher pressure, but it performs worse at lower pressures where thermal expansion is more significant. Figure 3.5c plots the PBE density functional theory (DFT) results from Bonev et al.[4] Those results are also in generally good agreement with experiment. However, those 2003 results did not employ a dispersion correction, which leads to artificial expansion of the unit cell. This is observed in the systematic underestimation of the prediction phonon frequencies.

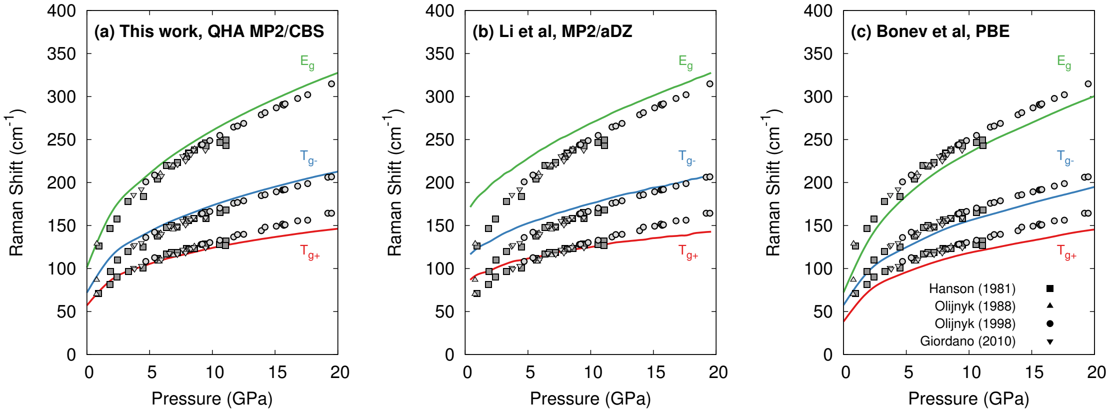


Figure 3.5: Pressure dependence of the librational Raman modes for phase I at room temperature. Comparison of MP2/aug-cc-pVDZ predictions based on quasiharmonic MP2/CBS structures from this work, results entirely at the MP2/aug-cc-pVDZ level with no quasi-harmonic approximation from Li et al,[3] and PBE DFT results from Bonev et al.[4] Points indicate experimental data,[5–8] while lines correspond to the predicted frequencies.

3.3.5.2 Phase II

Experimental pressure-dependent Raman data for phase II largely comes from Iota et al (2001).[9] At the time, the structure of phase II was unclear, and the low-frequency Raman spectrum was fitted to three distinct modes, with highest-frequency mode ("mode C" in Ref 9) fitted to the shoulder of the experimentally observed peak that is now attributed to the B_{1g} mode. Subsequent data from Yoo et al (2002)[10] and Datchi et al (2014)[11] fitted to only two distinct modes (B_{1g} and E_g) as expected from the space group and confirmed by theoretical calculations here and from Bonev et al.[4]

Ignoring the spurious "mode C", the agreement between theory and experiment is fairly good across the pressure range for both the MP2 calculations reported here and the earlier DFT ones from Bonev et al.[4]

3.3.5.3 Phase III/VII

Figure 3.7 compares pressure dependent experimental Raman study of Olijnyk and Jephcoat[7] against MP2 predictions from this work and PBE DFT calculations by Bonev et al.[4] Note that Li et al[3] also predicted phase III phonons as a function of pressure, though they are not included here because the symmetry character of the mode assignments in Figure 6 of Li et al is not clear. Nevertheless, the Raman plot in Figure 5 of Li et al shows three larger peaks and one very weak one around 300 cm^{-1} at 18 GPa, which is generally consistent with our predictions.

Although Bonev et al did not report Raman intensities, the mode symmetries follow the same pattern in both the MP2 and PBE results. At low pressure, the modes are,

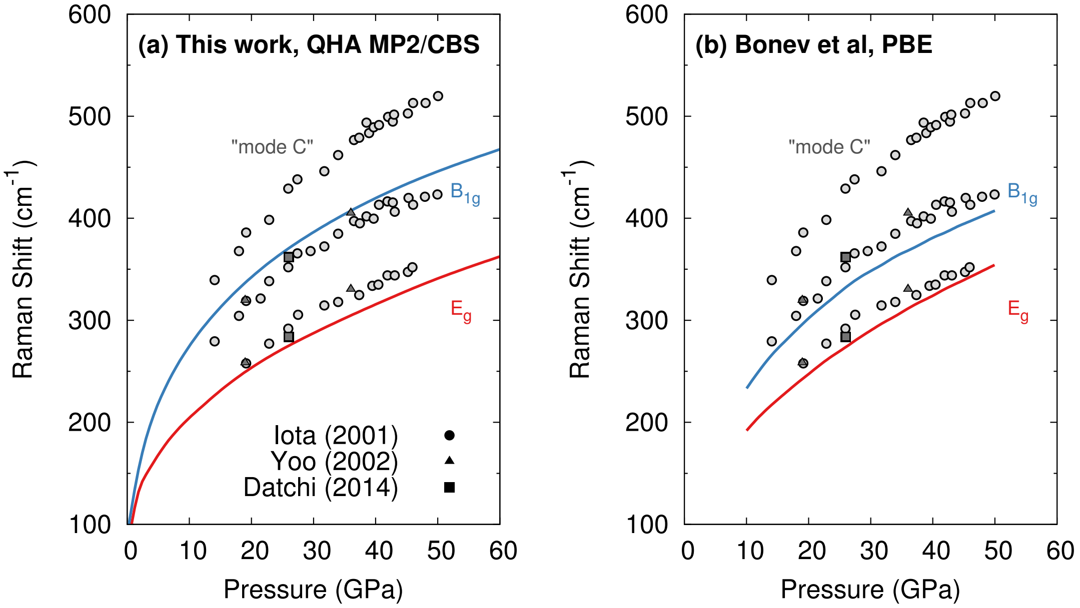


Figure 3.6: Pressure dependence of the librational Raman modes for phase II at room temperature. Comparison of MP2/aug-cc-pVDZ predictions based on quasiharmonic MP2/CBS structures from this work and PBE DFT from Bonev et al[4] (lines) against experimental data (points).[9–11]

in order of increasing frequency: A_g, B_{1g}, B_{3g}, and B_{2g}. Around 25–30 GPa, the A_g and B_{1g} modes cross. Experimentally, the lowest-frequency mode is difficult to observe after 20 GPa.[7] The two middle experimental modes (b and c in Figrue 3.11b) cross around 25 GPa, just as for the predicted A_g and B_{1g} modes. Finally, the highest-frequency experimental mode (d in Figure b) agrees well with the B_{2g} mode predicted here. We hypothesize that the lowest-frequency experimental mode is artifactual, and the fourth expected librational mode is actually the low-intensity B_{3g} mode. Because of the lack of dispersion in the Bonev calculations the phonon modes are artificially shifted to lower frequencies, giving a nominal appearance of agreement with the experimental assignment. However, we believe the symmetry character and crossing behavior of the modes is more consistent with the MP2 predictions in the current study. Repeating the DFT calculations with a dispersion-correction theory would likely shift the modes to higher frequencies, in better agreement with our MP2 predictions.

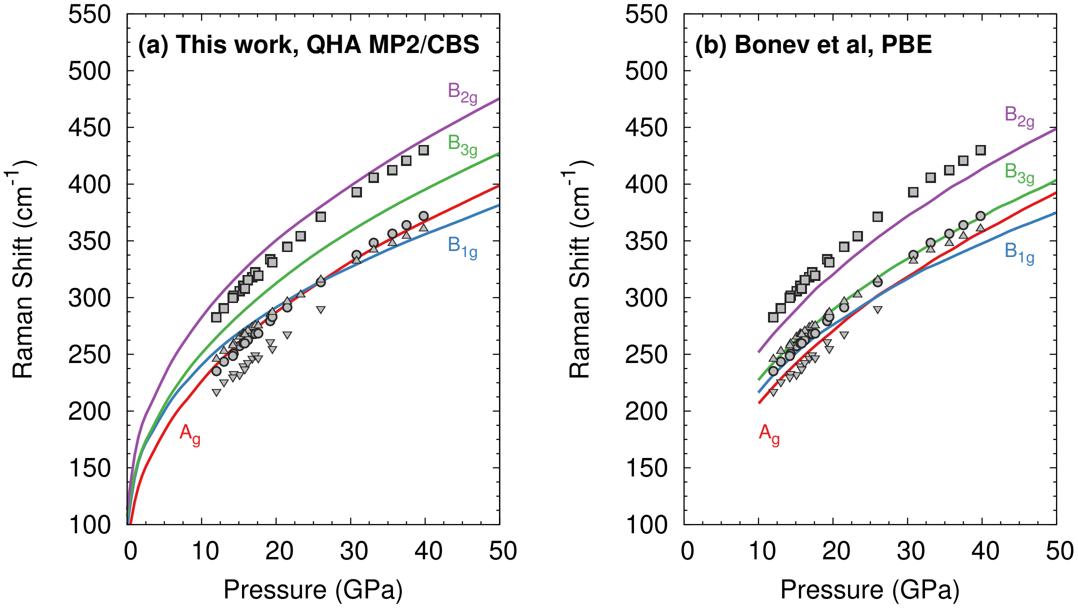


Figure 3.7: Pressure dependence of the librational Raman modes for phase III at room temperature. Comparison of MP2/aug-cc-pVDZ predictions based on quasiharmonic MP2/CBS structures from this work and PBE DFT results from Bonev et al[4] (lines) against experimental data from Olijnyk and Jephcoat (points).[7]

3.3.6 Crystal structure prediction results

Crystal structure prediction was performed using USPEX to generate potential candidate structures which might better account for phase III. Evolutionary algorithms were employed to generate more than 1700 structures, though large numbers of these were duplicates or energetically unfavorable. The 91 lowest structures were relaxed with PBE-D2 under 11.8 GPa of pressure, and many additional structures either coalesced into a single structure or proved energetically unfavorable. In the end, 25 structures had enthalpies within 10 kJ/mol of the most stable one (phase II) at this pressure, as shown in Figure 3.12. Figure 3.8 plots simulated powder X-ray diffraction (PXRD) patterns for all 25 structures. The crystal structure prediction generated phase I, II, and VII structures. It did not generate phase IV, which has a larger number of molecules in the unit cell.

The PXRD experiment was performed on a sample believed to consist of a mixture of phase I and III.[2] Examination of all the patterns reveals that none of the structures which are not experimentally known provides a good match to the experimental PXRD. The best matches come from either the claimed phase III structure or the phase VII one (with some mixture of phase I). CIF Structure files for all 25 structures are provided separately.

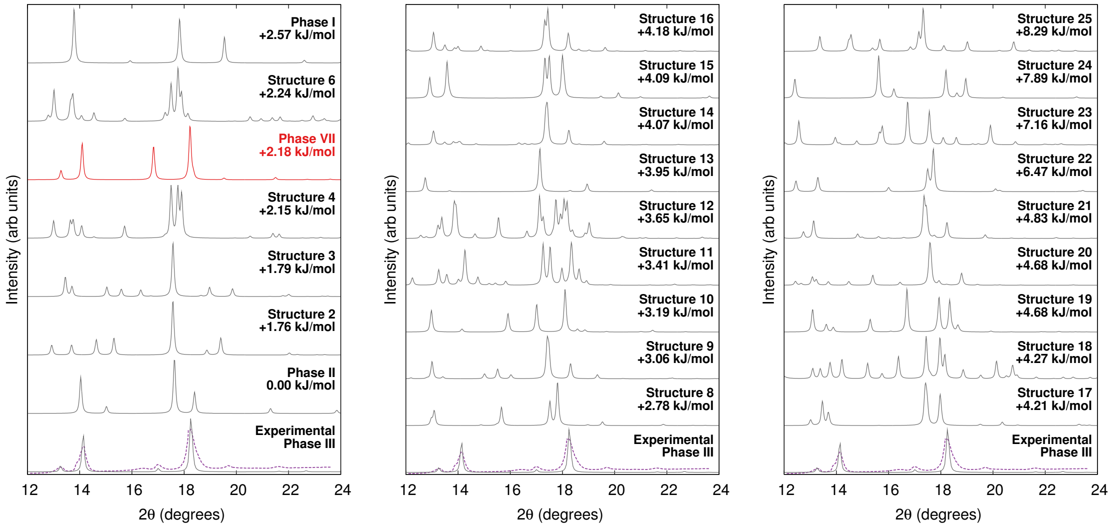


Figure 3.8: Simulated powder X-ray diffraction patterns for the 25 predicted crystal structures which lie within 10 kJ/mol of the stable phase II one. All were optimized at the PBE-D2 level and 11.8 GPa, including those identified as Phase I, II, or VII. These patterns are compared against the actual experimental phase III powder X-ray diffraction pattern (purple dotted line) and the simulated phase III spectrum based on the claimed experimental structure.[2]

3.3.7 Comparing carbon dioxide phase III and VII

To begin, compare the predicted and experimental structures for several different molecular crystalline phases. As can be seen in Figures 3.9 and 3.10, quasi-harmonic complete-basis-set (CBS) MP2 predicts the phase I, II, and VII lattice parameters and unit cell volumes in excellent agreement with experiment, with root-mean-square deviations of only 0.01–0.05 Å. The errors in the predicted room-temperature lattice constants never exceed 1% across a broad pressure range. The phase II predictions indicate a normal 1.155 Å C=O bond length at 25.8 GPa, in excellent agreement with the 2014 crystal structure[11] and DFT calculations,[4,85] and contrary to earlier suggestions of an "intermediate bonding" structure with extended 1.33 Å bonds.[10] Only one experimental crystal structure has been reported for phase VII (at 726 K and 12.1 GPa). Despite the high temperature which is more challenging for a quasi-harmonic approximation, only the 1.4% error in the *a* lattice constant exceeds this 1% error threshold.

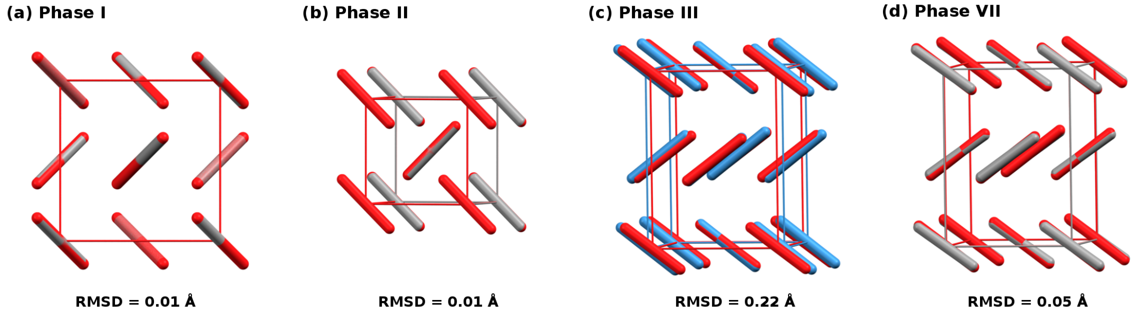


Figure 3.9: Overlays and root-mean-square deviations (RMSD) between MP2-predicted (red) and experimental crystal structures for (a) phase I, (b) phase II, (c) phase III, and (c) phase VII. Note the unit cell discrepancy for phase III. Predicted structures were optimized at the same temperature and pressure as the experimental structure.

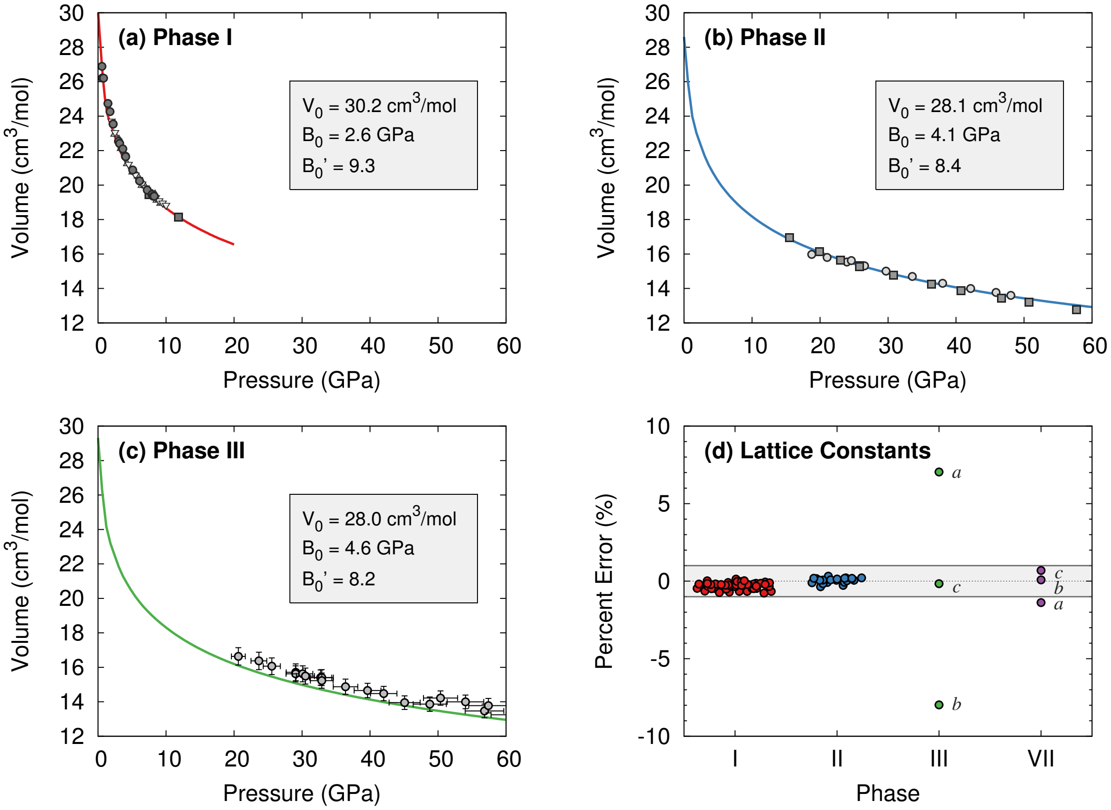


Figure 3.10: Comparison of predicted MP2/CBS (lines) and experimental (points) room temperature isotherms for (a) phase I,[2,8,12,13] (b) phase II,[10,11] and (c) phase III[14] carbon dioxide. Vinet equation of state[15] parameters fitted to the MP2 volumes are listed. Tables 3.1–3.4 show good agreement between the predicted and literature values for specific structures and the equation of state data. (d) Errors in the predicted lattice constants versus experiment. The shaded band indicates ±1% error.

The situation for phase III is completely different. As shown in Figures 3.10c–d, the predicted volumes are modestly smaller than the experimentally reported values, and the errors in the *a* and *b* lattice parameters relative to the experimental crystal structure are 5–10 fold larger than those for the other phases. The 0.22 Å RMSD between the experimental and predicted phase III structure is twenty-times larger than that for phases I and II.

Both conventional electronic and quasi-harmonic MP2 free energy relaxation of the experimental phase III structure always converge to the phase VII structure, regardless of temperature or pressure. In fact, no reported electronic structure calculation on phase III predicts a distinct phase III structure.[3,4,85,105,106] Even if phase III is only metastable relative to phase II as inferred experimentally, it should exist as a local minimum on the free energy surface that is distinct from phase VII.

Raman spectroscopy provides further evidence for problems with the phase III structure. Consider the librational phonons, which are the sensitive to crystal packing and do not suffer from the anharmonic modeling complexities[107] of the Fermi resonance that occurs at higher frequencies. Figure 3.11 compares room-temperature experimental Raman spectra against those computed at the MP2/aug-cc-pVDZ level using unit cell parameters from the quasi-harmonic MP2/CBS-limit calculations. Note that switching to the larger aug-cc-pVTZ basis has only a small effect on the predicted spectrum as shown in Figure 3.3.

Figure 3.11a shows that the predicted Raman peak positions and intensities generally agree very well with experiment for several different molecular phases of carbon dioxide. For phases I and II, the predicted peaks lie within $\sim$10 cm^{-1} of experiment. Similarly good results are obtained at other pressures as well (Figures 3.5–3.6). The phase II calculations confirm the assignments[4,11] of the broad band near 300 cm^{-1} to the doubly-degenerate E_g mode, and the band near 365 cm^{-1} to the B$_{1g}$ mode, contrary to earlier reports.[10] Note that the experimental broadening of the E_g mode is sample dependent, and it has been attributed to microscopic strains which may lift the two-fold degeneracy via orthorhombic distortion[4,11]—factors which are not present in the modeling.

For phase VII, agreement between the predicted and experimental structures is

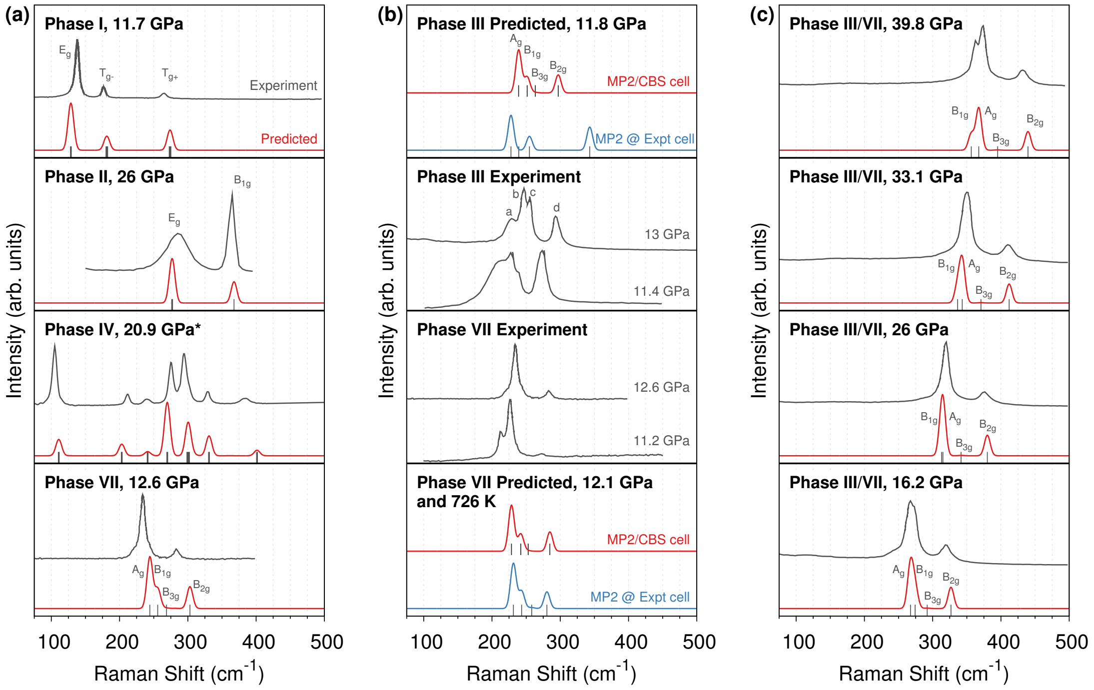


Figure 3.11: (a) Comparison of MP2-predicted and experimental lattice phonon Raman spectra for phases I,[7] II,[11], IV,[16] and VII[17] carbon dioxide. (b) Comparison of Raman spectra for phases III and VII,[7,17] including predicted spectra using either MP2/CBS or experimentally determined unit cell parameters. (c) Pressure dependence of the experimental phase III[7] and predicted phase VII Raman spectra. All spectra are at room temperature unless otherwise indicated. *The Phase IV spectrum employs an MP2/aug-cc-pVDZ optimized cell instead of a quasi-harmonic MP2/CBS one.

also reasonably good, with the predicted frequencies ∼10–15 cm^{-1} too high. Although this study does not perform complete-basis MP2 quasi-harmonic calculations for the larger phase IV unit cell, even the Raman spectrum predicted for the MP2/aug-cc-pVDZ optimized phase IV cell is in good agreement with the experiment spectrum (due to error cancellation between the small basis and the neglect of thermal expansion; Figure 3.4).

Consider next the comparison of phase III and VII spectra in Figure 3.11b. The *Cmca* structures should exhibit four Raman-active lattice modes. Experimentally, these modes have been assigned to the four peaks which are labeled *a*–*d* in the phase III spectrum at 13 GPa. The MP2 predictions concur with the three higher-frequency modes, *b*–*d*. However, they suggest the fourth Raman-active mode is not *a*, but rather a low-intensity B_{3g} mode in between the *c* and *d* modes.

Focus first on the three major peaks *b*–*d* which are clearly present in both the experimental and predicted spectra. The positions and intensities of these peaks in the MP2/CBS phase VII structure are consistent with both the experimental phase III and VII spectra. Of course, that consistency alone is insufficient to rule out the phase III structure. However, theory allows one to predict what the Raman spectrum would look like if the phase III structure were correct. Although the phase III structure is not a stationary point on the free energy surface, one can start with the purported experimental structure of phase III, freeze the lattice constants, relax the atomic positions, and predict the Raman frequencies and intensities. This differs from the other Raman calculations here only in obtaining the lattice constants from the experimental structure instead of from quasi-harmonic MP2/CBS calculations. The phase III structure optimized with fixed experimental lattice constants reproduces the claimed experimental structure very well (RMSD 0.03 Å, Figure 3.2). Strikingly, adopting the purported experimental structure shifts the highest-frequency band ~50 cm^{-1} to almost 350 cm^{-1} at 11-12 GPa (Figure 3.11b, top panel), versus below 300 cm^{-1} experimentally. That ~50 cm^{-1} disagreement between theory and experiment is several-fold larger than the disagreements observed between theory and experiment for any of the other phases examined here.

For comparison, performing the same procedure on the 726 K experimental phase VII structure[17] results in a predicted spectrum that is in excellent agreement with both the MP2/CBS cell Raman spectrum at the same temperature and the experimental room-temperature Raman at ∼11-13 GPa. In other words, while the experimentally reported phase VII structure is consistent with the predicted one in terms of both structure and Raman activity, the putative phase III structure is neither a stationary point on the free energy surface, nor is its predicted Raman spectrum compatible with the experimentally observed one.

If the phase III structure is incorrect, might some other unknown structure account for the experimental data? To investigate this possibility, evolutionary algorithm-driven crystal structure prediction was performed with the OPLS-AA force field[98] to generate potential carbon dioxide crystal structures with either two or four molecules in the unit cell at ambient pressure, followed by subsequent structural relaxation at 11.8 GPa with planewave PBE-D2. The crystal structure prediction generated Phase I, II, and VII (phase IV has more than four molecules in the unit cell), along with 22 other potential structures within 10 kJ/mol of the most stable one. However, none of the other structures has a simulated powder X-ray diffraction pattern that is plausibly consistent with the experimental phase III one (Section 3.3.6). Of course, the potential for structures with a different number of molecules in the cell or otherwise missed by this crystal structure prediction cannot be ruled out.

Nevertheless, in the absence of other viable phase III structures, the most obvious alternative is that phase III and VII are in fact the same, as was first raised (and subsequently discounted) by Giordano and Datchi.[17] Raman spectroscopic evidence supports this hypothesis. The phase VII predictions quantitatively reproduce the pressure dependence of

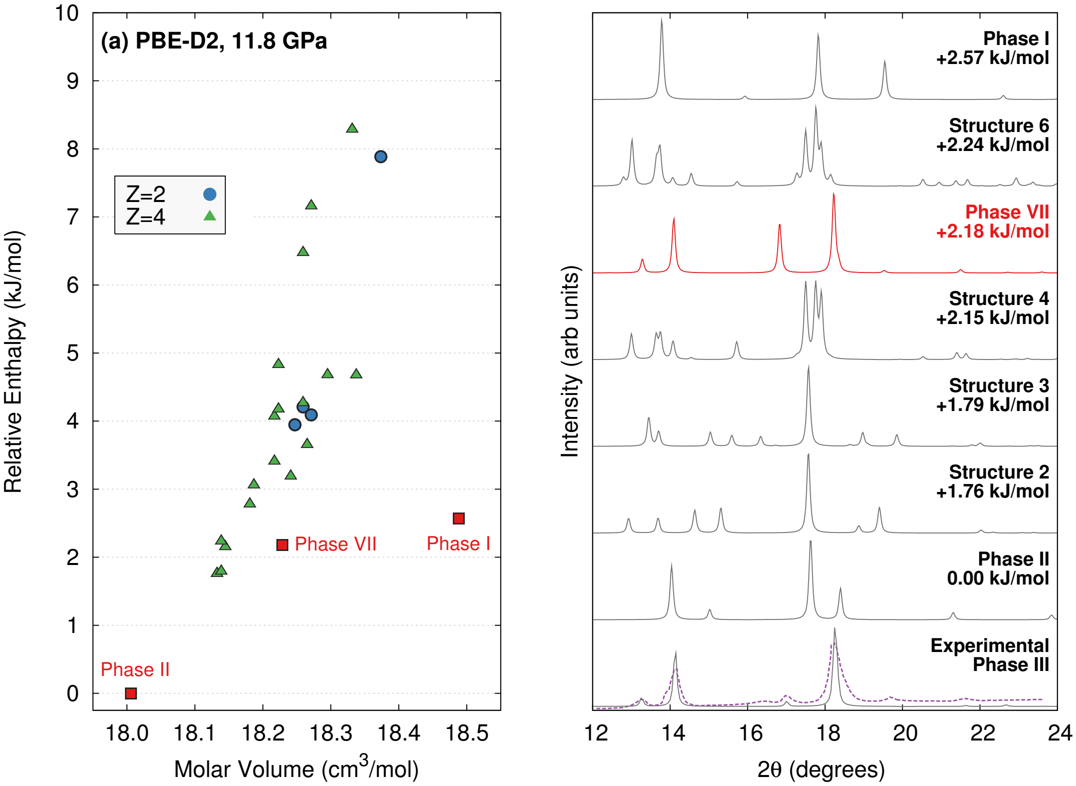


Figure 3.12: (a) PBE-D2 crystal energy landscape for potential carbon dioxide crystal structures at 11.8 GPa with $Z = 2$ or 4 molecules in the unit cell. (b) Comparison for simulated powder X-ray diffraction patterns for the lowest-lying PBE-D2 predicted structures against the actual[2] (purple) and simulated experimental phase III (using the purported structure) ones. Aside from the purported phase III structure, only the phase VII structure plausibly corresponds to the experimental data. Simulated diffraction patterns for all 25 low-energy structures are provided in Section 3.3.6.

the phase III Raman spectra over tens of GPa (Figure 3.11c). The predicted phase VII A_g and B_{1g} modes cross at 25 GPa, just like experimentally observed b and c modes for phase III (Figure 3.7).

The most significant disagreement between theory and experiment for phases III and VII stems from mode a, which appears as a lower-frequency shoulder on mode b in some experiments. Neither our calculations nor previous MP2 calculations[3] reproduce this shoulder. Instead, they predict the low-intensity fourth B_{3g} librational mode between experimental modes *c* and *d*. Although the DFT frequencies predicted by Bonev et al[4] provide nominal agreement with the experimental frequencies for modes *a*–*d*, closer inspection of their symmetry character and the errors expected from the neglect of van der Waals dispersion in those calculations suggests they are actually more consistent with the MP2 results here (Section 3.3.5).

Several experimental details also support the possibility that the peak ascribed to mode *a* has erroneously been attributed to phase III. First, Raman spectra in the phase I–III transition region cannot always be described as a simple combination of the modes for the two phases[7] The transition between phases I and III is notoriously sluggish over a broad pressure range, and other difficult-to-assign peaks are common during the transition. The Raman spectra is sensitive to strain and behaves very differently upon pressure loading and unloading. For example, the intensity of mode *a* is considerably stronger upon loading to 16.8 GPa than it is upon unloading to 16.2 GPa.[7] See also the significant differences in the experimental intensity of the mode *a* peak for phase VII at 11.2 and 12.6 GPa (Figure 3.11b).[17]

Second, the intensity of mode *a* decreases rapidly as pressure is increased beyond the phase transition region, and it is difficult to observe above ∼20 GPa.[7] Assigning it to one of the four Raman-active librational modes would require that its intensity exhibit substantially larger pressure dependence than is exhibited by any of the other modes experimentally or theoretically. Third, experimental Raman spectrum of the analogous *Cmca* phases of CS_2[108] and C_2H_2[109] support the theoretical assignment for CO_2. In both species, only three of the expected four librational modes exhibit appreciable intensity, and the lowest two frequency modes are the ones that cross as a function of pressure (instead of the middle two modes b and c according to the experimental assignments for CO_2). MP2/aug-cc-pVDZ calculations on CS_2 suggest the fourth unobserved mode is indeed a weak-intensity one, just as predicted for CO_2. Taken together, this evidence bolsters the case that mode a is an artifact of the phase transition rather than a librational mode of phase III, and that the experimental phase III Raman spectra are well-reproduced by the phase VII structure instead of the phase III one.

3.4 Conclusions

To summarize, large-basis, quasi-harmonic MP2 electronic structure calculations accurately reproduce experimentally observed structural, mechanical, and spectroscopic properties for several different molecular crystal phases of carbon dioxide across broad pressure range. However, theoretical optimization of the purported structure for phase III relaxes directly to phase VII. Even if phase III is only metastable, the existence of two distinct phases should translate to two separate free energy basins, but only one is found computationally. Furthermore, the claimed experimental phase III structure would produce a Raman spectrum in the librational region that disagrees with the experimentally observed ones. In contrast, Raman spectra predicted for phase VII agree well with the experimentally observed phase III and VII ones over a broad pressure range.

Based on the above results and the failure to identify a plausible alternative structure, we propose that phases III and VII are in fact the same. Although this hypothesis accounts for the data discussed above, questions remain. While the subtle differences in the experimental phase III and VII Raman spectra might be attributed to the variations arising from microstrain or other experimental complications, the apparent disconnect between the phase III and VII regions in the phase diagram is more difficult to rationalize. On the one hand, there is no contradiction in having phase III/VII be kinetically accessible in the phase III region and thermodynamically stable in the phase VII region. Instead of phase III being monotropically related to phases II and/or IV, phase III/VII would be enantiotropically related to them.

On the other hand, if the two phases are the same, why is phase VII seemingly difficult to form from phase I (it is typically formed from the melt instead),[86] while phase III forms readily? Perhaps given the sluggish nature of the phase I→III/VII transition and the narrow region of phase VII stability, the transformation from I→VII upon isothermal compression near 725 K is incomplete before one enters the region of phase IV stability. Similarly, why can one not form phase III kinetically via isothermal compression of phase I, then heat it to the phase VII region of thermodynamic stability without it transforming to phase II? If the kinetic barrier to transforming phase III→II is relatively small, maybe heating metastable phase III/VII from ambient temperatures provides sufficient thermal energy to convert to the more stable phase II before one reaches the regime of phase III/VII thermodynamic stability. Interestingly, one can quench phase VII down to room temperature, suggesting that the rate of heating may be significant. New experiments that investigate the crystal structure of phase III and its relationship to phase VII are clearly needed.

Chapter 4

Theoretical Assessment of the Structure and Stability of the λ Phase of Nitrogen

4.1 Introduction

Solid nitrogen exhibits fascinating phase behavior at high pressures, with at least 13 molecular and polymeric phases reported.[18] Some of these phases exhibit well-defined regions of thermodynamic stability. Others are kinetically accessible but thermodynamically metastable polymorphs, as evidenced by the often overlapping temperature and pressure conditions reported for different phases. These factors make mapping out the phase diagram (Figure 4.1) challenging. Furthermore, solving the crystal structure for high-pressure phases can also be difficult. The structures for the ζ,[23,110–112] θ,[113] and κ phases[23] are currently

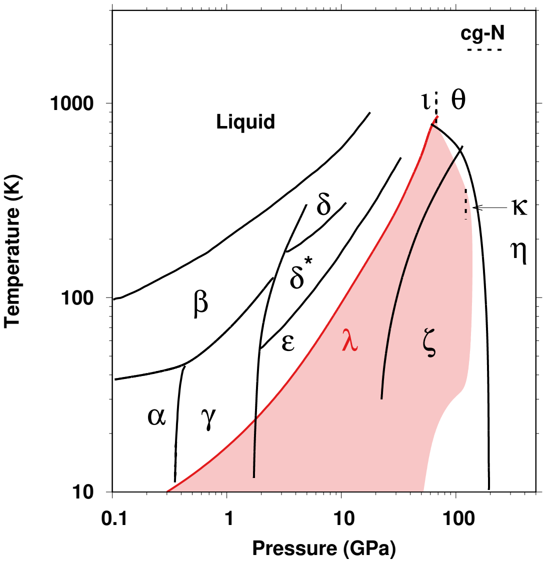


Figure 4.1: Phase diagram of nitrogen. The λ phase has been observed over the conditions highlighted in red, though its thermodynamic stability relative to the other phases remains unclear.

unknown, for example. The structure of the ι phase was reported only in 2018.[22]

The monoclinic λ phase was discovered in 2016.[18] This molecular nitrogen phase can be synthesized by compression of high-purity liquid nitrogen at low temperatures. It has been stabilized between 0.3-110 GPa at 77 K and between 32-140 GPa at 300 K. Like the ζ-phase, λ nitrogen transforms into the η phase at pressures above 100 GPa. This exceptionally wide range of pressure stability means that the λ phase can coexist with nine other phases: the γ, ϵ, ζ, ι, θ, κ, amorphous η, polymeric *cg*, and polymeric layered phases. It remains unclear whether the λ phase is thermodynamically preferred or only metastable relative to these other phases in this region of the phase diagram.

Solving the crystal structure of λ nitrogen experimentally proved challenging. Full Rietveld refinement of the structure from powder X-ray diffraction data was not possible due to the small sample sizes, sample graininess, and low X-ray scattering intensities for nitrogen.[18] Instead, the crystal structure determination relied heavily on an earlier density functional theory (DFT) crystal structure prediction study.[42] That study identified several low-enthalpy monoclinic and orthorhombic crystal structures which had not previously been observed experimentally. Frost and co-workers[18] found that fitting their experimental diffraction data to a monoclinic cell produced lattice parameters in fairly good agreement with the previously predicted $P2_1/c$ structure. Their final reported crystal structure was then constructed by combining the experimental lattice parameters with the previously predicted fractional coordinates of the atoms.

While this structural determination approach is reasonable, further evidence for the determined structure would be beneficial. For example, the long-accepted structure of phase III carbon dioxide, which was solved from relatively challenging powder x-ray diffraction data,[2] has recently been challenged. High-quality modeling and comparison against a variety of experimental data found in the literature suggest that carbon dioxide phases III and VII are actually identical, with phase VII being the true structure.[114] Density functional theory predictions of the structures and Raman spectrum were used to help confirm the ι structure as well.[22]

Here, we re-visit the λ phase of nitrogen computationally to provide further structural and spectroscopic evidence for the reported crystal structure and to assess its overall thermodynamic stability. Modeling molecular nitrogen in the solid state can be difficult. The weak, non-specific non-covalent interactions between molecules produce a relatively flat crystal energy landscape, with many possible crystal structures exhibiting similar energetic stabilities. For example, the 2009 DFT crystal structure prediction study[42] whose results were used to help solve the λ N_2 structure found several low-enthalpy structures within $\sim$0.5 kJ/mol of each other at 40 GPa.

When discriminating between such closely ranked crystal structures, it is important to model the interactions carefully. The neglect of van der Waals dispersion interactions in the 2009 DFT study could impact the relative energies, for example. Further insights and energy refinement can be gained by employing higher-level correlated wave function techniques such as second-order Møller-Plesset perturbation theory (MP2) or even coupled cluster models. Periodic local MP2 calculations have provided insight into the phase transitions between the α, γ, ϵ, and polymeric *cg* phases,[115,116] but such calculations are relatively expensive, and nuclear gradients that would allow geometry optimizations and other response properties to be predicted readily have not been implemented.

Alternatively, fragment-based methods[44,117–119] such as the hybrid many-body interaction (HMBI) model[76–78] provide an computationally practical strategy for applying high-level correlated wave function methods to periodic systems. In HMBI, individual molecules and their short-range pairwise intermolecular interactions are modeled with MP2 or other high-level electronic structure methods, while longer-range pairwise interactions and non-additive many-body intermolecular interactions arising from the infinite lattice are approximated with a polarizable force field. When coupled with a quasi-harmonic treatment of thermal expansion, the HMBI model predicts structural, mechanical, and spectroscopic properties of several phases of carbon dioxide in excellent agreement with experiment.[80–82,114] HMBI-predicted structural and spectroscopic data was used to support the aforementioned argument that carbon dioxide phases III and VII are the same phase.[114] The HMBI fragment approach also predicts the polymorphic phase diagram of methanol with ∼0.5 kJ/mol accuracy,[66,67] and it has been applied to larger polymorphic organic crystals such as aspirin[65] and oxalyl dihydrazide.[64] The related binary interaction fragment model[71,120] has proved similarly effective for modeling molecular crystal structures and properties in solid hydrogen fluoride,[121,122] carbon dioxide,[3,104,107,123] and ices.[72,73,124]

Here, we investigate the λ phase of nitrogen using a mixture of periodic planewave DFT theory and fragment-based MP2 calculations. We confirm the previously reported structure of the λ phase through a combination of DFT-based crystal structure prediction (employing *ab initio* random structure searching, or AIRSS) and higher-level refinement with MP2. Further support for the λ phase structure is provided by comparison between the predicted and experimentally observed Raman spectra. Finally, to investigate whether the λ phase is a thermodynamically stable phase on the phase diagram (rather than a kinetically accessible metastable one), the thermodynamic stability of the λ phase relative to several other experimentally known phases which can exist under the same thermodynamic conditions is investigated.

4.2 Methods

4.2.1 Density functional theory calculations

DFT calculations were employed to optimize structures and provide an initial stability ranking. The calculations were performed using the B86bPBE density functional[100,125] with the exchange-hole dipole moment (XDM) dispersion correction,[126] an

80 Ry planewave cutoff, and a 6×6×6 Monkhorst-Pack k-point grid, as implemented in Quantum Espresso version 6.2.1.[99,127] Projector-augmented wave (PAW) potentials for nitrogen atoms were produced using A. Dal Corso's Atomic code v6.1. External pressure was applied to the variable cell optimizations to mimic the experimental conditions. See Section 4.3.1 for validation and convergence testing of the DFT models.

4.2.2 Fragment-based hybrid many-body interaction (HMBI) calculations

The structures and stability rankings of the DFT structures were subsequently refined with correlated wave function methods via the HMBI model.[76–78] HMBI decomposes the total energy of the crystal according to a many-body expansion,

$$\begin{aligned} E^{HMBI}_{crystal} &= E^{QM}_{1\text{-}body} + E^{QM}_{SR\ 2\text{-}body} + E^{MM}_{LR\ 2\text{-}body} \\ &\quad + E^{MM}_{many\text{-}body} \end{aligned} \tag{4.1}$$

where 1-body terms correspond to the energies of individual nitrogen molecules in the unit cell, two-body terms to the interaction energies between pairs of molecules (both within the central unit cell and involving periodic image molecules), and many-body terms to the non-additive three-body and higher contributions. The 1-body and short-ranged two-body terms (i.e. dimers separated by no more than 6 Å) were computed with MP2 or CCSD(T), while the long-range two-body and many-body terms are approximated using the periodic Hartree-Fock (pHF) or AMOEBA polarizable force field[79]c calculations under periodic boundary conditions. The number of monomer and dimer fragments that need to be computed is reduced by exploiting space group symmetry.[95]

The density-fitted MP2 calculations were performed using Molpro 2012,[92] CCSD(T) calculations (with core electrons frozen) were employed using PSI4 v1.0,[128] the pHF calculations were carried out using CRYSTAL17,[129] and the polarizable force field calculations were conducted using the Tinker version 6.3.[93] Existing AMOEBA force field parameters were used for the N_2 molecule.[79] Single-point MP2 and CCSD(T) fragment energies were computed with counterpoise correction at the extrapolated complete basis set (CBS) limit. The CBS results were estimated via the combination of HF/aug-cc-pVQZ energies plus two-point extrapolation of the correlation energy contributions[130] obtained from the aug-cc-pVTZ and aug-cc-pVQZ basis sets. The impact of basis set on predicted energies in small molecule crystals has been studied extensively elsewhere.[44,80,82] Single-point pHF fragment energies were computed using the pob-TZVP basis set,[131] which is version of the popular def2-TZVP basis set[132] that has been adapted for periodic calculations. Empirical testing in small nitrogen clusters suggests that this basis set, with no counterpoise correction, provides intermolecular many-body HF interaction energies that agree fairly well with those obtained from much larger basis sets.

4.2.3 Structure refinement and quasi-harmonic approximation

Because the predicted molar volume is sensitive to the level of theory and because the HMBI-based MP2 or CCSD(T) calculations can potentially provide higher-accuracy structures than those from the DFT model, one would like to refine the structures at the MP2/CCSD(T) levels of theory. Full crystal relaxations with MP2 or CCSD(T) are somewhat more computationally expensive than those with DFT (especially with the larger basis sets needed for the correlated wave function models). To make the initial structure refinements less expensive, the following simplified quasi-1-D optimization approach was employed: First, the geometries were optimized with DFT at a series of external pressures. Second, single-point HMBI electronic energies were computed with MP2 or CCSD(T) and different many-body treatments at each DFT geometry. A PV pressure-volume contribution was added to the resulting energies to obtain enthalpies versus volume: $H = E(V) + PV$. The enthalpy-volume curves were fitted to the Murnaghan equation of state,

$$H(V) = H_0 + \frac{B_0 V}{B_0'}\left[\frac{(V_0/V)^{B_0'}}{B_0' - 1} + 1\right] - \frac{B_0 V_0}{B_0' - 1} \tag{4.2}$$

where the enthalpy (H_0), volume (V_0), bulk modulus (B_0), and its first pressure derivative (B_0') at zero pressure are the fitting parameters. The optimal volume V_0 and corresponding enthalpy H_0 were extracted from the minimum of the fit. Atomic coordinates were obtained via interpolation of the DFT fractional coordinates to the optimal volume extracted from the equation of state fit. This approach is similar to how the quasi-harmonic approximation (QHA) is sometimes performed (see Ref 82, for example), except the zero-point and thermal vibrational contributions are neglected here. This neglect is reasonable at higher pressures where thermal expansion effects are expected to be small.

For the most promising λ phase structure candidate, full QHA calculations including phonon contributions were performed. DFT geometry optimizations were carried out over a wide pressure range (at 28 pressures ranging 0–150 GPa). Again, MP2 and CCSD(T) single-point energy refinements were computed using HMBI, as described above. Harmonic Γ-point phonons were computed at the DFT level, and the volume-dependence of the phonon frequencies was approximated using mode-specific Grüneisen parameters that were computed via finite-difference, as described previously.[82] Optimal volumes and energies were then obtained by minimizing the Gibbs free energies,

$$G(T, P) = E(V) + PV + F_{vib}(T) \tag{4.3}$$

where F_{vib} is the standard harmonic Helmholtz vibrational free energy. This combination of DFT geometries and phonons plus higher-level energies in quasi-harmonic calculations effectively reproduced thermal expansion and thermochemical properties in an earlier study on several small-molecule crystals.[82]

4.2.4 Simulated powder X-ray diffraction (PXRD) and Raman spectra

PXRD and Raman spectra were modeled at the HMBI MP2 level for comparison with experiment. To do so, the crystal structure was relaxed at the HMBI MP2/aug-cc-pVDZ + AMOEBA level using fixed unit cell parameters determined from the HMBI MP2/CBS QHA calculations described above. Previous work on carbon dioxide[114] demonstrated that when the unit cell is constrained with the lattice parameters obtained from a high level of theory, relaxing the atomic positions and predicting spectroscopic properties with MP2 in a smaller basis set does not introduce substantial errors compared to larger basis set results, since the unit cell dimensions constrain the packing density.

Simulated PXRD spectra were generated using Mercury[102] and the same 0.42418 Å wavelength as the experiments. Simulated Raman spectrum are based on Γ-point MP2/aug-cc-pVDZ + AMOEBA harmonic phonon frequencies, using the analytical Hessian algorithms implemented in Gaussian 09.[96] The use of analytical Hessians for each fragment contribution helps reduce numerical artifacts associated with summing contributions from many fragment Hessians in the HMBI model. Raman intensities were calculated from finite difference of polarizability derivatives, which were approximated via the QM 1- and 2-body contributions only (no AMOEBA or periodic HF many-body contribution). Because intermolecular many-body contributions are relatively small in nitrogen, the effect on the Raman intensities from neglecting the many-body contributions to the polarizability derivatives should be small. This approach has been described previously.[72,114] Peaks in the simulated Raman spectra are plotted with an arbitrary full width at half maximum of 10 cm^{-1}.

4.3 Results and Discussion

4.3.1 DFT Pseudopotential and Convergence Tests

The short, relatively unique triple bonds found in the nitrogen molecule make it potentially challenging to model accurately with pseudopotential DFT. Here, some tests are performed to examine the impact of the pseudopotential choice and to demonstrate the convergence of the predictions with respect to the planewave basis cutoff.

First, the effect of pseudopotential type was studied. Table 4.1 compares the DFT-optimized N_2 bond lengths and unit cell lattice parameters as computed from two different treatments of the core: the N.pbe-n-rrkjus_psl.0.1.UPF ultrasoft pseudopotential (US-PP) downloaded from http://www.quantum-espresso.org and the projector augmented wave (PAW) potentials. The B86BPBE-XDM density functional, XDM dispersion correction, an 80 Ry planewave cutoff, and a 6×6×6 Monkhorst-Pack k-point grid were used. For consistency with the high pressures that are of interest for the study, the calculations were performed under 34 GPa of external hydrostatic pressure. All calculations were performed with Quantum Espresso version 6.2.1.

As is evident from Table 4.1, both core electron approximations predict virtually identical N_2 triple bond lengths. The predicted lattice parameters from both sets of calculations are also quite similar, though the volumes obtained from the US-PP are slightly larger (by less than 2%). Both US-PP and PAW potential approaches underestimate the experimental λ unit cell volume by 8–10%.

We also compare the predicted bond lengths for gas-phase N_2 with both the PAW approach and all-electron methods. The all-electron calculation was performed with B86bPBE in PSI4 v1.1[128] with the def2-QZVP Gaussian basis set. For the planewave-PAW approach, the single N_2 molecule was placed in a large $20\times19\times21$ Å unit cell to ensure significant vacuum spacing on all sides and optimized using the parameters described above. No XDM dispersion correction was included, since dispersion contributes little to the intramolecular structure of N_2, and XDM is not available in PSI4. The PAW and all-electron methods predict bond lengths of 1.106 Å and 1.065 Å, respectively, compared against the experimental value of 1.098 Å.[133] In other words, the PAW potential results appear to be in good agreement with both all-electron results for the gas-phase and US-PP results for the solid state.

Next, the convergence of the energies and structures was investigated with respect to planewave cutoff, focusing on the PAW potentials and other DFT parameters. For all four crystal structures, excellent convergence in the total energy per molecule is obtained by the 80 Ry cutoff used here. For all four forms, the 80 Ry results are within 0.15 kJ/mol of the 100 Ry results for the 0 GPa structures, and within 0.10 kJ/mol for the 34 GPa structures. The N-N triple bond lengths converge even more rapidly. By 40 Ry, the bond lengths are converged to within less than 0.001 Å of the 100 Ry basis set result (Figure 4.3). These predicted bond lengths matched with those of the experiment[20,21] within ∼0.01 Å at low pressure.

Finally, comparison of predicted and experimental crystal structures is performed. Figure 4.4 compares the 80 Ry cutoff results for the α and γ phases at low pressures. Both predicted structures match the experimental ones to within rmsd 0.05 Å or better. Table 4.2– 4.3 compares unit-cell parameters with different computational method including quasi-harmonic thermal expansion. At low temperatures & pressures, QHA DFT actually agrees more closely with experiment than the QHA CCSD(T) results. At higher temperatures and pressures, however, the higher level methods perform better.

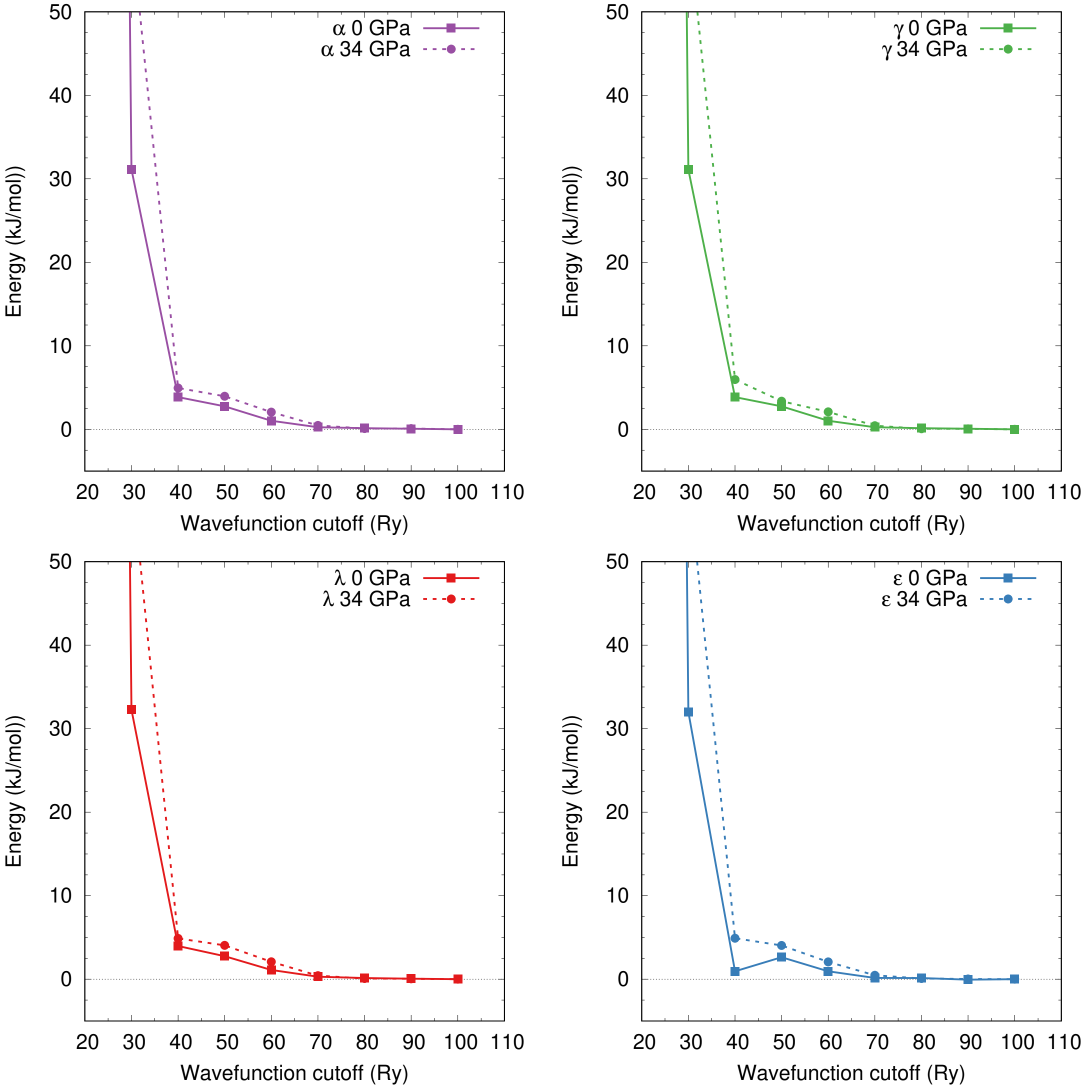


Figure 4.2: Convergence of the total energy per molecule with the planewave cutoff for several known N_2 crystal structures.

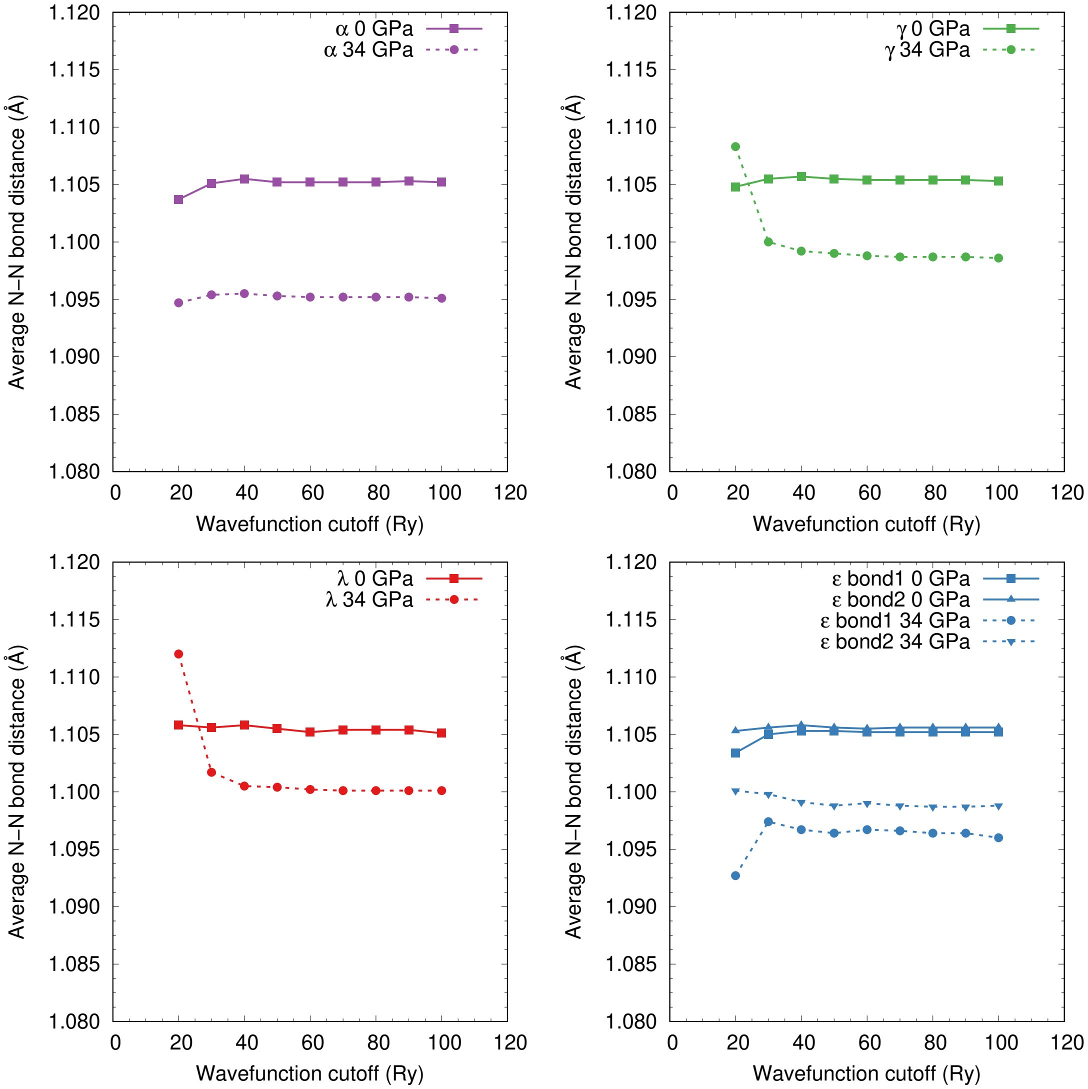


Figure 4.3: N_2 bond length vs. planewave basis set cutoff at two different pressures and for four different experimentally known crystal structures.

Table 4.1: Comparison of nitrogen bond length and unit cell parameters obtained using US-PP and PAW core treatments with B86bPBE-XDM, an 80 Ry cutoff, a 6×6×6 k-point grid, and 34 GPa external pressure.

Phase	Pseudo Potential	bond length (Å)	a (Å)	b (Å)	c (Å)	*beta* (°)
α	US-PP	1.0954	4.258	4.258	4.258	90.0
	PAW	1.0952	4.235	4.235	4.235	90.0
γ	US-PP	1.0988	2.997	2.997	4.132	90.0
	PAW	1.0987	2.978	2.978	4.114	90.0
ϵ	US-PP	1.0966, 1.0989	7.094	7.094	10.133	90.0
	PAW	1.0964, 1.0987	7.052	7.052	10.081	90.0
λ	US-PP	1.1001	2.966	2.940	5.650	132.4
	PAW	1.1001	2.949	2.919	5.621	132.4

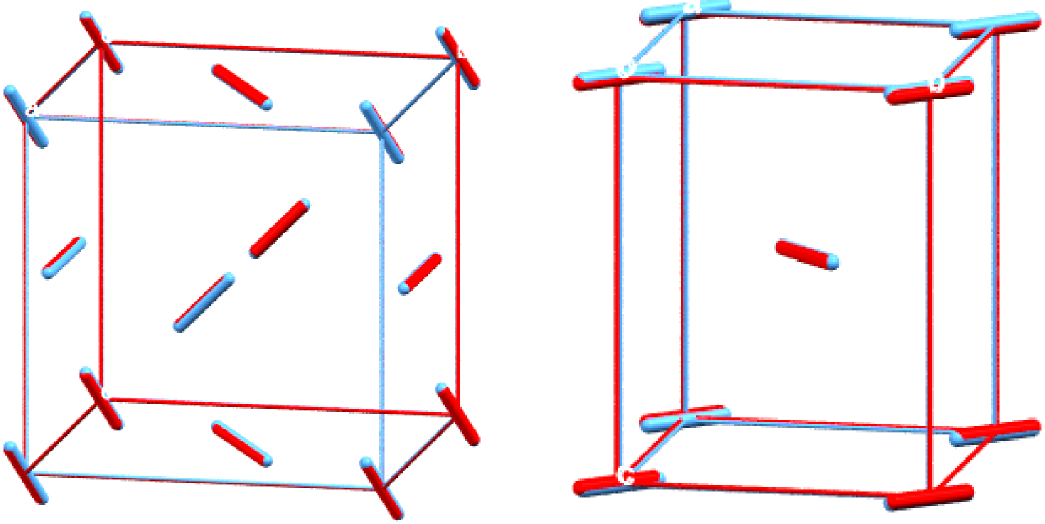


Figure 4.4: Overlay of experimental and predicted B86bPBE-XDM crystal structures for (left) phase α at ambient pressure, rmsd 0.026 Å, and (right) phase γ at 0.04 GPa, rmsd 0.047 Å. The calculations employed PAW potentials, an 80 Ry cutoff, and a 6×6×6 k-point grid.

Table 4.2: Comparison of unit cell parameters for nitrogen phase α and γ predicted from different computational methods. DFT calculations were performed with B86BPBE-XDM functional, 80 Ry cutoff, 6x6x6 k-point grid. The obtained DFT structures were refined further with wave function correlation method at CCSD(T). The complete basis set was used by extrapolating consistent-correlation triple and quadruple zeta valence of Dunning basis sets, aug-cc-pVXZ. All calculations include quasi-harmonic thermal expansion.

Phase	Temperature (K)	Pressure (GPa)	Methods	*a* (Å)	*b* (Å)	*c* (Å)	*beta* (°)
α	4.2	ambient	QHA-DFT	5.582	5.582	5.582	90.0
			QHA-CCSD(T)/CBS + AMOEBA	5.495	5.495	5.495	90.0
			QHA-CCSD(T)/CBS + pHF	5.873	5.873	5.873	90.0
			Experiment	5.644	5.644	5.644	90.0
	20	ambient	QHA-DFT	5.583	5.583	5.583	90.0
			QHA-CCSD(T)/CBS + AMOEBA	5.495	5.495	5.495	90.0
			QHA-CCSD(T)/CBS + pHF	5.876	5.876	5.876	90.0
			Experiment	5.649	5.649	5.649	90.0
	20	0.269	QHA-DFT	5.500	5.500	5.500	90.0
			QHA-CCSD(T)/CBS + AMOEBA	5.422	5.422	5.422	90.0
			QHA-CCSD(T)/CBS + pHF	5.734	5.734	5.734	90.0
			Experiment	5.495	5.495	5.495	90.0
γ	20.4	0.4	QHA-DFT	3.962	3.962	5.222	90.0
			QHA-CCSD(T)/CBS + AMOEBA	3.911	3.911	5.178	90.0
			QHA-CCSD(T)/CBS + pHF	4.131	4.131	5.425	90.0
			Experiment	3.957	3.957	5.109	90.0

Table 4.3: Comparison of unit cell parameters for nitrogen phase ϵ and λ predicted from different computational methods. DFT calculations were performed with B86BPBE-XDM functional, 80 Ry cutoff, 6x6x6 k-point grid. The obtained DFT structures were refined further with wave function correlation method at CCSD(T). The complete basis set was used by extrapolating consistent-correlation triple and quadruple zeta valence of Dunning basis sets, aug-cc-pVXZ. All calculations include quasi-harmonic thermal expansion.

Phase	Temperature (K)	Pressure (GPa)	Methods	a (Å)	b (Å)	c (Å)	$beta$ (°)
ϵ	110	7.8	QHA-DFT	9.086	9.086	12.553	90.0
			QHA-CCSD(T)/CBS + AMOEBA	8.071	8.071	11.308	90.0
			QHA-CCSD(T)/CBS + pHF	8.155	8.155	11.404	90.0
			Experiment	8.020	8.020	11.104	90.0
	ambient	16.3	QHA-DFT	7.695	7.695	11.964	90.0
			QHA-CCSD(T)/CBS + AMOEBA	8.598	8.598	10.872	90.0
			QHA-CCSD(T)/CBS + pHF	7.673	7.673	10.847	90.0
			Experiment	7.605	7.605	10.622	90.0
λ	300	34	QHA-DFT	2.967	2.935	5.656	132.3
			QHA-CCSD(T)/CBS + AMOEBA	3.046	3.023	5.768	132.0
			QHA-CCSD(T)/CBS + pHF	3.005	2.979	5.701	132.2
			Experiment	3.051	3.066	5.705	131.6

4.3.2 Crystal Structure Prediction and Structure Ranking

The AIRSS search was performed over all 13 monoclinic space groups with Z=2. Within these monoclinic space group, the $C2$, Cm, Cc, and $C2/c$ space groups were searched with Z=4 because generated cell with Z=2 was not amenable. Structures were generated by placing a single nitrogen molecule at one possible Wyckoff position with random orientation and then applying the space group symmetry operators to populate the other molecules in the unit cell. Lattice parameters were randomized within the constraint that the unit cell volume remained within ±40% of the experimentally reported value of 39.9 $Å^3$ (or 12.0 cm^3/mol) at 34 GPa. At least 100 random structures were generated for each space group, for a total of over 1,300 structures. Each structure was fully relaxed with B86BPBE-XDM at 34 GPa of external pressure using the same computational method from previous section at 80 Ry cutoff. Some of the random structures adopted covalent/polymeric forms upon relaxation and were discarded from this work. In the end, the AIRSS procedure generated 636 molecular structures with Z=2 and 246 with Z=4. After clustering to remove duplicate structures, 22 unique structures remained. These unique structures are summarized in Table 4.4 and Table 4.5. The structures of the 15 structures with the lowest enthalpy at the B86bPBE-XDM level are pictured in Figure 4.5.

Discriminating closely ranked crystal structures can be challenging. Here, fragment-based refinements were carried out at the MP2/CBS level, using either an AMOEBA or periodic HF many-body treatment, as reported in Figure 4.6. Table 4.6 provides the relative enthalpies at each level of theory, highlighting the significant re-ranking that occurs with the different refinement schemes. Fortunately, in all cases, structure S1 corresponding to λ

nitrogen remains the most stable by ∼5 kJ/mol or more across all three methods.

Table 4.4: The unit-cell parameters and atomic positions of the twelve lowest DFT-energy structures, S1-S12, generated by the random structure search. Structures and relative enthalpies per molecule are reported at the B86bPBE-XDM level of theory under 34 GPa of external pressure. Atomic positions are reported in fractional coordinate.

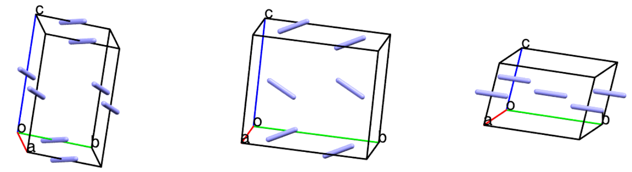
Structure	Space group	a (Å)	b (Å)	c (Å)	β (°)	Atoms	Multiplicity	Wyckoff position	x	y	z	Relative Enthalpy (kJ/mol)
Phase λ (S1)	$P2_1/c$	2.951	2.916	4.671	116.8	N1	4	e	0.15967	0.62161	0.04666	0.00
S2	$C2/c$	3.579	4.835	4.198	90.04	N1	8	f	0.75004	0.35080	0.43924	5.09
S3	$C2/m$	4.472	3.728	2.763	128.1	N1	4	h	0.00000	0.85240	0.50000	5.43
S4	$P2_1/m$	2.910	4.345	2.912	100.2	N1	2	e	0.65994	0.25000	0.04554	5.64
						N2	2	e	-0.04550	0.25000	0.3403	
S5	$C2/c$	4.862	2.925	5.688	115.8	N1	8	f	0.74734	0.12827	0.42538	6.15
S6	$C2/m$	3.763	4.381	2.889	130.6	N1	4	i	0.35357	0.00000	-0.00012	6.81
S7	$P2_1$	2.680	5.255	2.929	116.9	N1	2	a	0.25965	0.02787	0.02403	8.75
						N2	2	a	0.41221	0.88002	0.32115	
S8	$C2/m$	4.173	3.689	2.706	120.1	N1	4	i	0.64701	0.00000	0.16635	9.40
S9	$P\bar{1}$	2.928	4.780	5.471	103.6	N1	2	i	0.39699	0.81053	0.32737	10.18
						N2	2	i	0.10234	0.85871	0.17939	
						N3	2	i	0.39880	0.64229	0.82172	
						N4	2	i	0.10302	0.68860	0.67334	
S10	Pc	2.977	4.656	5.453	102.8	N1	2	a	0.20714	0.08389	0.47441	11.05
						N2	2	a	0.89683	0.08378	0.32125	
						N3	2	a	0.70722	0.58381	0.47454	
						N4	2	a	0.39691	0.58386	0.32138	
S11	$C2/m$	5.150	2.654	2.819	106.0	N1	4	i	-0.09203	0.00000	0.34500	13.87
S12	$C2/c$	4.470	5.456	3.037	90.06	N1	8	f	0.24990	0.32349	0.62374	15.06

Table 4.5: The unit-cell parameters and atomic positions of the remaining structures S13-S22 generated by the random structure search. Structures and relative enthalpies per molecule are reported at the B86bPBE-XDM level of theory under 34 GPa of external pressure. Atomic positions are reported in fractional coordinate.

Structure	Space group	a (Å)	b (Å)	c (Å)	β (°)	Atoms	Multiplicity	Wyckoff position	x	y	z	Relative Enthalpy (kJ/mol)
S13	$P2_1/c$	2.618	5.337	3.086	120.6	N1	4	e	0.06493	0.42753	0.64483	15.18
S14	$C2/m$	3.815	5.12	2.69	134.9	N1	4	h	0.00000	0.89292	0.50000	15.34
S15	$C2/m$	3.060	4.434	2.832	103.1	N1	4	i	0.14651	0.00000	0.65352	18.04
S16	$P2$	2.970	2.887	8.630	90.02	N1	2	e	0.11653	0.01742	0.87121	19.70
						N2	2	e	0.38381	0.75311	0.87095	
						N3	2	e	0.11609	0.51626	0.37099	
						N4	2	e	0.38340	0.25198	0.37114	
S17	$P2_1/m$	2.867	9.343	2.870	100.15	N1	4	f	0.44531	0.38352	0.85348	30.14
						N2	4	f	0.14697	0.38351	0.55428	
S18	$P2/c$	2.962	2.822	5.369	123.3	N1	4	g	0.13643	0.86661	0.50040	32.33
S19	Pm	2.644	10.814	2.987	115.6	N1	2	c	0.02968	0.66285	0.53000	36.64
						N2	2	c	0.85611	0.60401	0.19824	
						N3	2	c	0.69256	0.16224	0.85723	
						N4	2	c	0.51994	0.10392	0.52331	
S20	Pm	2.829	9.521	2.918	101.2	N1	2	c	0.15990	0.86902	0.48832	37.74
						N2	2	c	0.85843	0.86884	0.18776	
						N3	2	c	0.15891	0.36905	-0.01200	
						N4	2	c	0.85745	0.36883	0.68743	
S21	Pm	2.734	10.437	2.997	115.36	N1	2	c	0.54158	0.15966	0.88867	39.62
						N2	2	c	0.35234	0.10694	0.53773	
						N3	2	c	0.04098	0.65963	0.88720	
						N4	2	c	0.85183	0.60694	0.53621	
S22	Pm	2.607	10.977	2.977	114.2	N1	2	c	0.86719	0.60277	0.82414	39.99
						N2	2	c	0.69781	0.66082	0.49516	
						N3	2	c	0.36738	0.10276	0.32405	
						N4	2	c	0.19790	0.16080	-0.00496	

Table 4.6: Relative enthalpies for each predicted structure at 34 GPa with B86bPBE-XDM and fragment-based MP2 methods. The structures are ordered based on the enthalpies at the DFT level; numbers in parentheses indicate the new ranking after energy refinement. Despite considerable reordering of the other structures, the structure ascribed to phase λ remains the most stable with all three models.

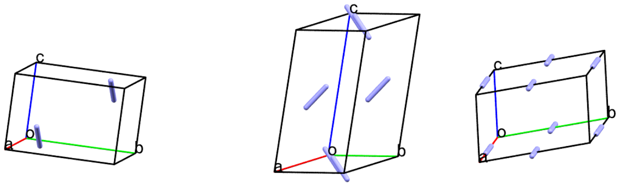
Structure	B86BPBE-XDM	MP2/CBS + AMOEBA	MP2/CBS + pHF
Phase λ (S1)	0.0	0.0 (#1)	0.0 (#1)
S2	5.09	6.61 (#4)	4.70 (#2)
S3	5.43	7.90 (#6)	6.82 (#5)
S4	5.64	7.32 (#5)	6.78 (#3)
S5	6.15	6.48 (#3)	8.28 (#7)
S6	6.81	11.11 (#12)	11.49 (#10)
S7	8.75	6.24 (#2)	7.94 (#6)
S8	9.40	8.30 (#8)	6.79 (#4)
S9	10.18	8.65 (#9)	9.42 (#8)
S10	11.05	8.72 (#10)	10.70 (#9)
S11	13.87	9.44 (#11)	13.06 (#11)
S12	15.06	11.96 (#15)	17.27 (#15)
S13	15.18	8.29 (#7)	14.96 (#13)
S14	15.34	11.48 (#14)	14.91 (#12)
S15	18.04	11.26 (#13)	15.33 (#14)



S1: P2_1/c (Z=2)

S2: C2/c (Z=4)

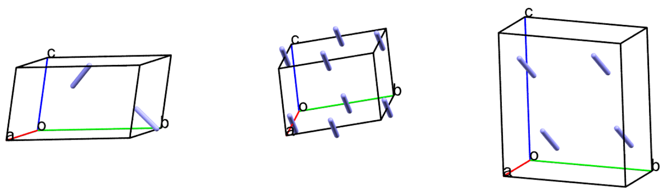
S3: C2/m (Z=2)



S4: P2_1/m (Z=2)

S5: C2/c (Z=4)

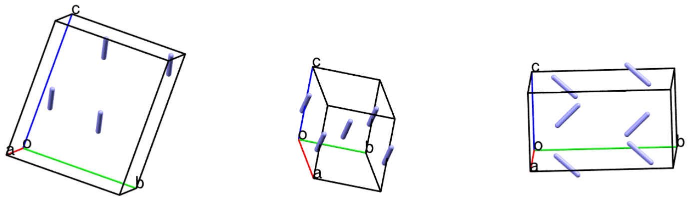
S6: C2/m (Z=2)



S7: P2_1 (Z=2)

S8: C2/m (Z=2)

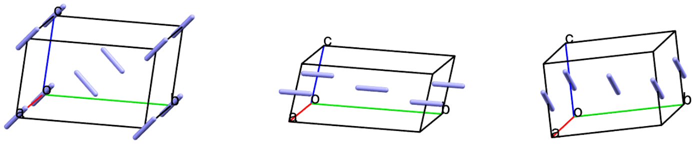
S9: P-1 (Z=4)



S10: Pc (Z=4)

S11: C2/m (Z=2)

S12: C2/c (Z=4)



S13: P2_1/c (Z=2)

S14: C2/m (Z=2)

S15: C2/m (Z=4)

Figure 4.5: The 15 lowest-enthalpy structures (S1–S15) at the B86bPBE-XDM level, as generated by the AIRSS search.

Table 4.7: Comparison of experimental and predicted lattice parameters for λ N_2 at 34 GPa.

	a (Å)	b (Å)	c (Å)	β (°)	Volume (cm^3/mol)
B86bPBE-XDM	2.951	2.916	5.638	132.32	10.80
MP2/CBS + pHF	2.975	2.945	5.663	132.71	11.05
QHA MP2/CBS + pHF (300 K)	2.985	2.957	5.674	132.26	11.16
QHA CCSD(T)/CBS + pHF (300 K)	3.005	2.979	5.701	132.20	11.38
QHA MP2/CBS + AMOEBA (300 K)	3.035	3.010	5.750	132.09	11.74
QHA CCSD(T)/CBS + AMOEBA (300 K)	3.046	3.023	5.768	132.05	11.87
Experiment (300 K)[18]	3.051(7)	3.066(5)	5.705(13)	131.65(5)	12.01

4.3.3 Crystal energy landscape

To explore the landscape of potential crystal structures, crystal structure prediction was performed via *ab initio* random structure searching (AIRSS).[134] Employing experimental constraints can facilitate AIRSS by narrowing the search space, as demonstrated for the unusually complex structure of ι N_2.[22] Experimental reports indicate that the λ N_2 phase adopts monoclinic space group symmetry with two molecules in the unit cell (Z=2).[18] Accordingly, the AIRSS search was performed over all 13 monoclinic space groups with Z=2. Structures were generated by placing a single nitrogen molecule at one possible Wyckoff position with random orientation and then applying the space group symmetry operators to populate the other molecules in the unit cell. Lattice parameters were randomized within the constraint that the unit cell volume remained within ±40% of the experimentally reported value of 39.9 Å^3 (or 12.0 cm^3/mol) at 34 GPa. Because the $C2$, Cm, Cc, and $C2/c$ space groups are not amenable to Z=2, searches in those groups were run with Z=4. The Z=4 searches produced a mixture of new structures and structures which were supercell representations of previously predicted Z=2 structures.

In the end, at least 100 random structures were generated for each space group, for a total of over 1,300 structures. Each structure was fully relaxed with B86bPBE-XDM at 34 GPa of external pressure. Some of the random structures adopted covalent/polymeric forms upon relaxation. At 34 GPa, the DFT enthalpies of the polymeric structures are significantly higher than those of molecular forms, so they were discarded. In the end, the AIRSS procedure generated 636 molecular structures with Z=2 and 246 with Z=4. After

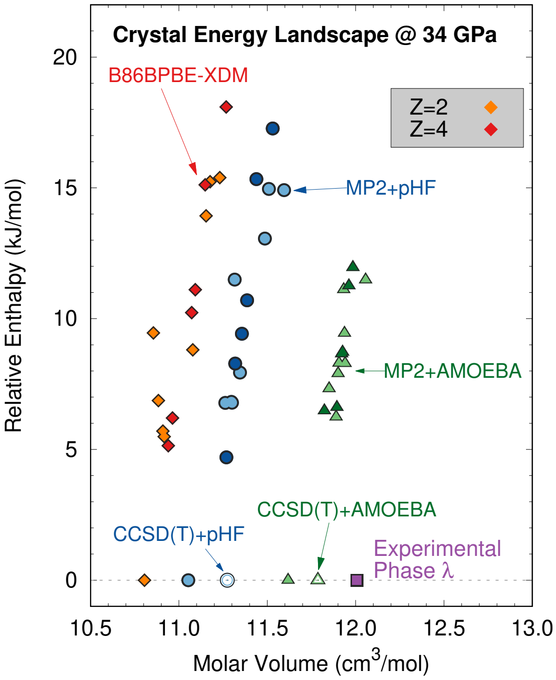


Figure 4.6: Crystal energy landscape for the low-energy crystal structures at 34 GPa with the B86bPBE-XDM (red Z=4/orange Z=2), MP2/CBS + pHF (dark blue Z=4/light blue Z=2), and MP2/CBS + AMOEBA (dark green Z=4/light green Z=2) levels of theory. Open symbols correspond to further CCSD(T) refinements of the structures. The experimentally inferred molar volume is indicated in purple. Enthalpies at each level of theory are plotted relative to the lowest-energy structure.

clustering to remove duplicate structures, 22 unique structures remained.

Figure 4.6 compares the crystal energy structure for the 15 structures whose B86bPBE-XDM energies lie within 20 kJ/mol of the most stable structure. Higher-energy structures are unlikely to occur experimentally, since the typical energy window for polymorphism is ∼10 kJ/mol.[135,136] Pictures, lattice parameters, atomic coordinates, and energetics for these predicted structures are provided earlier in the Section 4.3.2. All of the DFT-predicted structures exhibit molar volumes ranging 10.8–11.3 cm^3/mol, which is 6–10% smaller than the 12.0 cm^3/mol inferred from the diffraction experiments.[18] The neglect of expansion arising from zero-point energy and thermal vibrational contributions will likely cause some volume underestimation, but those effects should be small at 34 GPa. Large volume errors are potentially problematic when comparing different nitrogen phases, since many nitrogen phase changes at high-pressure occur with volume changes of as little as 2%.

Refinement of the structures with correlated fragment-based MP2/CBS with either pHF or AMOEBA many-body treatments alters the relative enthalpies of the different candidate structures, but all models predict the same densely packed $P2_1/c$ structure to be the most stable one. This structure is identical to the $P2_1/c$ structure predicted by Pickard and Needs[42] and which has been ascribed to the λ phase.[18] Though the relative enthalpies differ depending on the model, all models predict that this structure is more stable than the second-lowest structure by ∼5 kJ/mol or more.

MP2 + pHF and MP2 + AMOEBA refinement also increase the molar volumes by ∼2% and ∼9%, on average, to 11.1–11.6 and 11.6–12.1 cm^3/mol, respectively (Figure 4.6). With both many-body treatments, the volume of the lowest-energy structure shifts closer toward the experimental value of 12.0 cm^3/mol. Switching from MP2 to CCSD(T) increases the molar volumes further, by about 0.2 cm^3/mol. At the CCSD(T)/CBS + AMOEBA level, the molar volume of 11.8 cm^3/mol is 2% smaller than the experimentally reported one.

These results demonstrate some sensitivity of the predictions to the many-body treatment. With the monomer and dimer contributions in the fragment model treated with large-basis MP2 or CCSD(T), the largest remaining source of error probably lies in the more approximate many-body treatment. One would typically expect the quantum me-

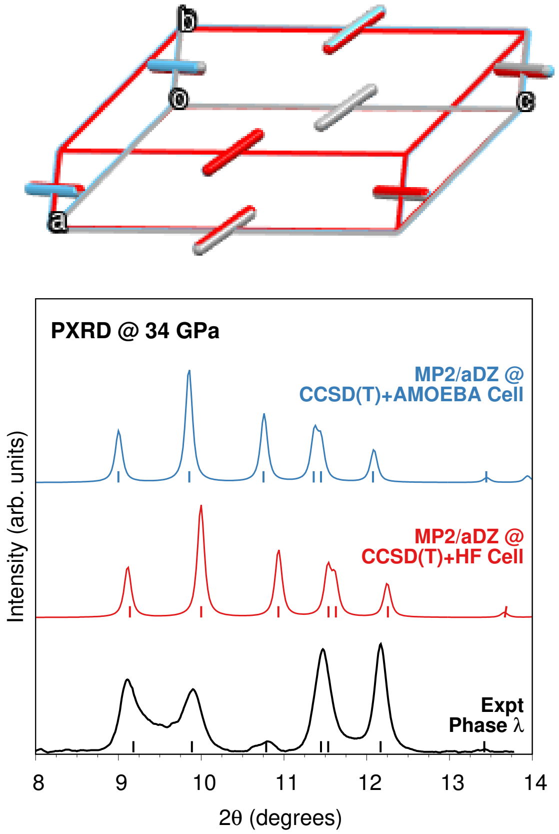


Figure 4.7: Structure overlay (top) and simulated powder x-ray diffraction spectrum (bottom) comparing the experimental,[18] CCSD(T)/CBS + periodic HF, and CCSD(T)/CBS + AMOEBA crystal structures.

chanical pHF treatment to perform better than the AMOEBA one, as has been observed in other small-molecule crystals.[67,82] At high pressures, the HF many-body exchange-overlap description should be superior to that of the simplified short-range polarization damping model employed AMOEBA force field (and which was parameterized at ambient conditions). For this nitrogen phase, however, the AMOEBA many-body treatment predicts volumes closer to experiment at 34 GPa. As will be discussed below, however, the periodic HF many-body treatment does predict molar volume in closer agreement with experiment at higher pressures. Interestingly, the range of relative enthalpies for the higher-lying structures is considerably smaller with the MP2 + AMOEBA model than for the MP2 + pHF or DFT models. In the end, results from both many-body treatments are presented throughout the remainder of the paper to help quantify the uncertainties in the predictions with respect to the treatment of the many-body interactions.

As noted above, the slight underestimation of the molar volumes here is consistent with the neglect of zero-point and thermal vibrational contributions. Quasi-harmonic calculations that estimate these contributions were performed on this most-stable structure. Table 4.7 compares the lattice parameters predicted from several different models against experiment. As expected given the high pressure, thermal expansion effects are small, increasing individual lattice constants by ∼0.01 Å and the total molar volume by only ∼0.1 cm^3/mol. Earlier work on phase I carbon dioxide found that expansion effects due to vibrational contributions were most significant below ∼10–20 GPa.[81] To enable the prediction of additional properties, the atomic positions here obtained from the initial DFT optimizations were relaxed at the MP2/aug-cc-pVDZ + AMOEBA level of theory, holding the lattice parameters fixed at their predicted QHA CCSD(T)/CBS values.

With either many-body treatment, the final QHA CCSD(T) predictions at 300 K and 34 GPa reproduce the experimental lattice constants within 0.06 Å and 0.4°, and in the case of the QHA CCSD(T) + AMOEBA, reproduce the volume to to within ∼0.1 cm^3/mol. The rmsd15,[1] which measures of the root-mean-square deviations for a 15-molecule cluster taken from the crystal, is an excellent 0.037 Å between this structure and the experimental one. Overlays of the predicted and experimental structures are shown in Figure 4.7. Similarly, the simulated PXRD patterns predicted from these CCSD(T) structures reproduce

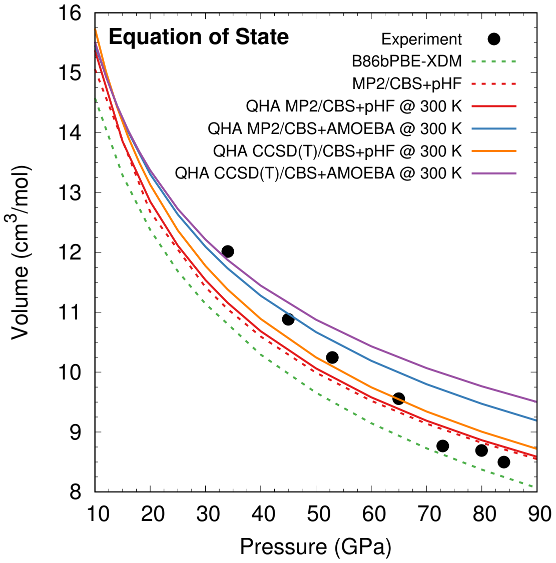


Figure 4.8: Comparison of the predicted and experimentally observed[18] equations of state for λ N_2.

the experimental peak positions fairly well, especially for the CCSD(T)/CBS + AMOEBA cell whose lattice constants are in best agreement with those reported experimentally. The combination of good agreement between the predicted and experimentally reported structure at 34 GPa and the enthalpic stability of the $P2_1/c$ structure relative to other candidate structures generated by AIRSS crystal structure prediction support the assignment of this structure to the λ phase.

4.3.4 Spectroscopic comparisons

Further insight into this phase is gained by looking at the how the structure and Raman spectrum changes with pressure. Figure 4.8 plots the equation of state predictions for several different models against experimentally reported volumes. None of the models quite reproduce the experimental volume data. B86bPBE-XDM underestimates

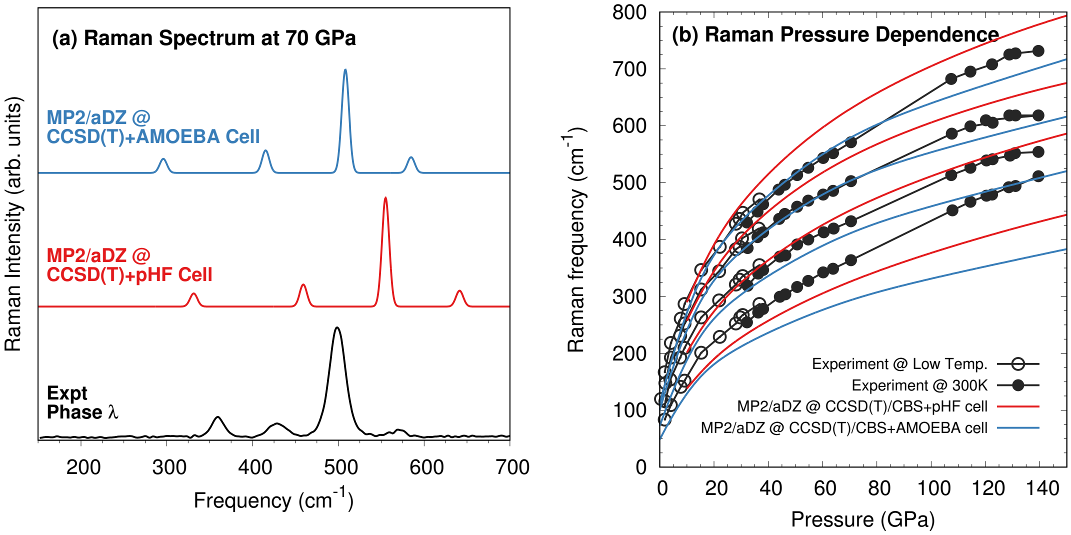


Figure 4.9: Comparison of the predicted and experimentally observed[18] Raman spectra for λ nitrogen in the librational region.

the molar volume throughout the pressure range. The QHA MP2 and CCSD(T) results with AMOEBA many-body terms predict consistently larger molar volumes, which end up being closer to the experimental values near ∼30–40 GPa. The molar volumes from the models with periodic HF many-body treatments are consistently smaller and agree better with experiment at higher pressures (as one might expect from the better HF treatment of many-body exchange-overlap at high pressures). For a given many-body treatment, CCSD(T) predicts a slightly larger volume than MP2. The additional expansion obtained by including the QHA treatment is small throughout the pressure range. The impact of the QHA approximation would be more noticeable at even lower pressures, of course.

Using the 300 K QHA CCSD(T) structures, MP2/aug-cc-pVDZ + AMOEBA harmonic phonons and Raman intensities were computed at various pressures. As described in the Methods section, the atomic positions were relaxed within the fixed QHA CCSD(T) unit cells at this same level of theory to ensure stationarity of the energy with respect to atomic position, as required by the harmonic phonon approximation. Figure 4.9a plots representative spectra at 70 GPa, while Figure 4.9b predicts how the Raman-active librational mode frequencies vary with pressure. The librational modes provide a useful fingerprint for crystal packing. The number of Raman active modes and their relative intensities in the predicted spectra agree well with the experimental ones. The agreement between the predicted and experimental frequencies is rather good at low pressures, but the errors reach up to ∼50–100 cm^{-1} at high pressures.

While the overall agreement between theory and experiment in the pressure-volume and Raman data is fairly good, it is unclear why the disagreement between the models and experiment is as large as it is. Earlier work on high-pressure phases of carbon dioxide in the ∼10–60 GPa range found that a very similar fragment-based modeling approach generally reproduced the molar volumes to within 1%, while the positions of the Raman-active librational modes were reproduced to within 10–15 cm^{-1}.[114] The treatment of one- and two-body interactions at the CCSD(T)/CBS level is likely well converged (it differs only modestly from the MP2/CBS results). The treatment of the many-body effects appears to be a larger problem, and perhaps computing those contributions with a larger basis set and/or higher level of theory would be helpful. It's possible, for example, that terms such as the generally repulsive Axilrod-Teller-Muto three-body dispersion term or other, higher-order exchange overlap contributions that are missing from HF become important at these high pressures. Alternatively, given the small size and low-orientational specificity of the intermolecular interactions in the crystals, anharmonic and/or dynamical contributions may be more important in high-pressure nitrogen phases than in carbon dioxide.

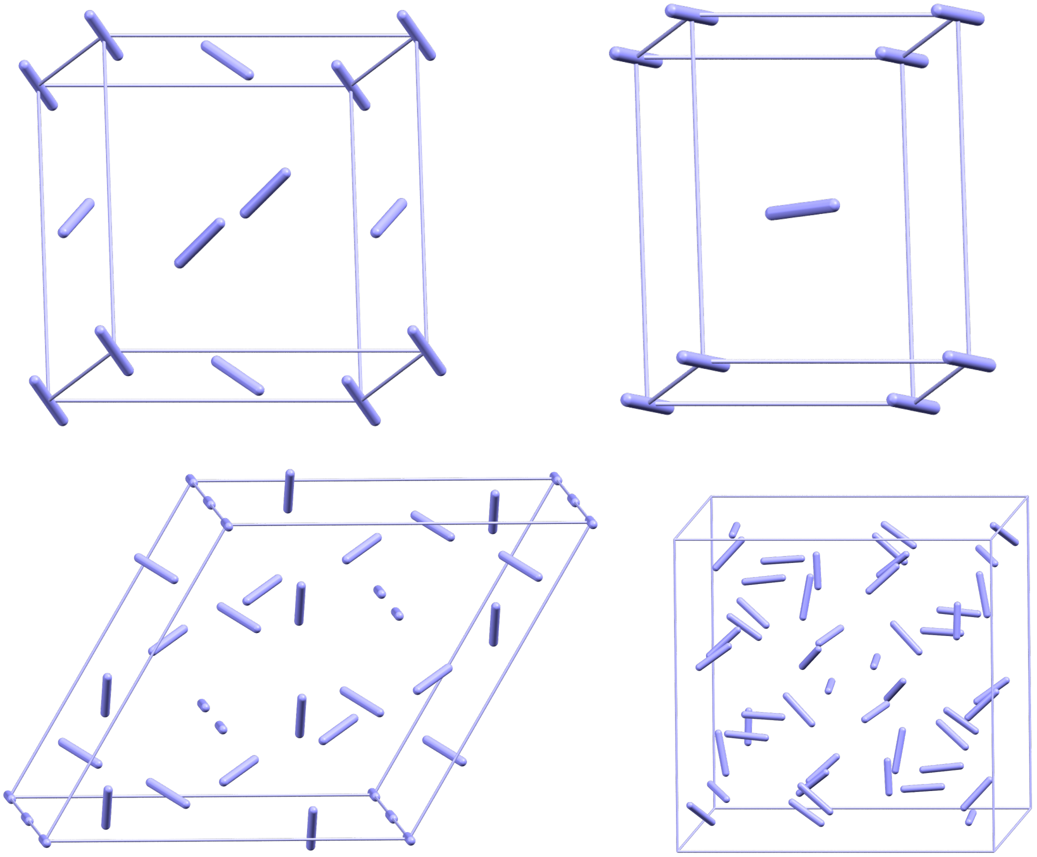


Figure 4.10: Crystal structures of N_2 (top left) phase α[19] (top right) phase γ[20] (bottom left) phase ϵ[21] (bottom right) phase ι[22]

Beyond possible errors in the modeling, problems with the experimental measurements and their interpretation cannot be ruled out. Pressure gradients, inhomogeneous samples (e.g. due to partial phase transformations), and other factors can impact high-pressure studies such as those used to characterize this phase. Despite the moderate disagreements between theory and experiment, the overall collection of crystal structure and property predictions here support assignment of this $P2_1/c$ structure to the λ phase of nitrogen.

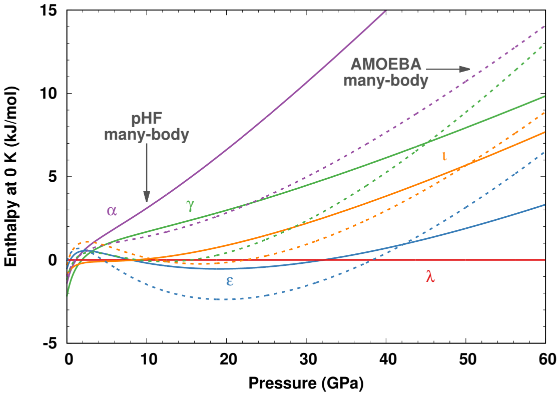


Figure 4.11: Comparison of CCSD(T) enthalpies at 0 K for the α, γ, ϵ, λ, and ι phases as a function of pressure using either periodic HF (solid lines) or AMOEBA (dotted lines) for the many-body treatment.

4.3.5 Thermodynamic stability

Finally, we turn to the question of the thermodynamic stability of the λ phase compared to other potential phases it overlaps with on the phase diagram. The original experimental study indicated that the λ phase potentially coexists with nine other phases on the phase diagram: γ, ϵ, ζ, ι, θ, κ, η, cg, and the layered polymeric phase.[18] The fragment-based approach used here is not well-suited to model the network covalent phases, so the cg and layered polymeric phases are not considered further. Those network covalent nitrogen phases are not likely viable until higher pressures anyways (at least 50 GPa, if not above 100 GPa[23,137,138]). Similarly, the κ and η phases only occur at much higher pressures (above 60 GPa). The ζ and θ phases can be quenched to around ~30 GPa where λ occurs, but modeling of their stabilities is not possible because their structures are currently unknown. Therefore, we focus on the molecular α, γ, ϵ, and ι phases, which are present in the relevant pressure regime (Figure 4.1) and have known structures (Figure 4.10).

These other crystal structures were modeled using the same techniques as those used above: the structures were optimized at under various external pressures using B86bPBE-XDM. Single-point energy refinements were performed with fragment-based CCSD(T) calculations, and the optimal 0 K enthalpies determined at each pressure (including Γ-point harmonic vibrational contributions). Figure 4.11 plots the relative stabilities of the four phases with both periodic HF and AMOEBA many-body contributions.

Experimentally, the low-pressure α phase is most stable until 0.36 GPa.[139] As pressure increases, experiments indicate that γ nitrogen becomes the preferred phase until $\sim$2 GPa,[140,141] at which point it transforms to the ϵ phase. Earlier fully periodic MP2 calculations reproduced this behavior nicely, predicting these transitions to occur at 0.42 and 2.25 GPa, respectively.[115] Here, the CCSD(T)/CBS + AMOEBA fragment-based enthalpies at 0 K predict the $\alpha \longrightarrow \gamma$ and $\gamma \longrightarrow \epsilon$ transitions to occur at 0.45 and 3.0 GPa, respectively, which is also in rather good agreement with the experimentally reported transitions. Switching to the pHF many-body treatment stabilizes the γ phase relative to α and ϵ somewhat. The γ-ϵ transition shifts only slightly, to 3.2 GPa, but the α phase (incorrectly) becomes less stable than γ all the way down to zero pressure (by 0.8 kJ/mol). The AMOEBA many-body treatment is nominally performing better than periodic HF for these transitions, though given the subtle differences in energetics between models, this could also represent fortuitous error cancellation. The calculations here also predict the ι phase to be less stable than ϵ in the $\sim$10–60 GPa range, especially with the periodic HF many-body treatment. This result contrasts earlier PBE DFT calculations,[22] which found ι to be more stable than ϵ. In both cases, however, the energy differences between the two phases are only a few kJ/mol or less, and the calculations are performed at 0 K, rather than the elevated temperatures where ι nitrogen has been observed experimentally.

Regarding the λ phase, both sets of CCSD(T) calculations in Figure 4.11 suggest that λ may be more stable than the α, γ, ϵ, and ι phases under certain pressure conditions, though the specific pressure windows varies with the many-body treatment. While the uncertainties in the models and the omission of several other phases with unknown or network covalent structures prevent definitive statements, it appears plausible that λ nitrogen may represent a thermodynamically stable phase on the phase diagram, at least at low temperatures.

4.4 Conclusions

To summarize, the experimentally reported crystal structure for λ nitrogen reflected a mixture of partial solution powder x-ray diffraction data together with a previously predicted structure which had similar lattice parameters. Here, a detailed study of λ nitrogen has been carried out to confirm this structure, using both dispersion-corrected density functional theory and higher-level fragment-based MP2 and CCSD(T) calculations. AIRSS crystal structure prediction over monoclinic space groups predicted this λ structure as the most stable monoclinic one by $\sim$5 kJ/mol or more at 34 GPa. While the lattice constants predicted by B86bPBE-XDM were already in decent agreement with experiment, refining the structure further with MP2 and CCSD(T) brought them even closer to the experimental one. Further support for assigning this structure to the λ phase was provided by comparison between the predicted and experimentally reported equation of state and Raman spectra over a broad pressure range. Taken together, the body of evidence provides support for the λ nitrogen structure first proposed by Frost et al.[18]

Because it has been unclear whether λ nitrogen is the thermodynamically stable or only a kinetically accessible phase under the conditions where it has been observed, the stability of this λ phase relative to that of several other experimental phases which are known to co-exist in the same temperature and pressure conditions was compared. Though the sensitivity/uncertainties associated with energetics are relatively large and not all possible phases have been considered, the results do suggest that the λ phase may well be the thermodynamically most stable phase at low temperatures and moderate pressures.

Finally, on the methodological side, the fragment-based MP2 and CCSD(T) models do generally out-perform the B86bPBE-XDM density functional for these nitrogen phases. However, the predicted structures and properties at high pressure are somewhat sensitive to the treatment of many-body interactions. For pressures near ~100 GPa and above in particular (e.g. where the largest errors in the predicted Raman data occur), it may be important in the future to employ even better treatments of the many-body interactions. Periodic local MP2 corrected with coupled cluster calculations have been demonstrated, for example.[142] Pragmatic approaches that account for many-body dispersion effects might also prove useful.[143,144] More careful assessment of these issues should be carried out in the future.

Chapter 5

Combining crystal structure prediction and simulated spectroscopy in pursuit of the unknown nitrogen phase ζ crystal structure

5.1 Introduction

Nitrogen exhibits fascinating behavior in the solid state, with at least 15 experimentally reported phases: α, β, γ, δ, δ_{loc}, ϵ, ζ, ζ', η, θ, ι, κ, λ, *cg*, *lp*/*bp*, and *hlp*)[18,20–24,113,139,141,145–154] Most of these phases are molecular and ordered, but some are disordered (β,[20] δ[151]), non-molecular (cg,[150] η[148,149], lp/bp[152,154], hlp[153]) and/or amorphous (η). The large number of viable molecular packing motifs stems from the small size of molecular N_2 and its weak, non-polar intermolecular interactions which make many packing arrangements and orientations energetically competitive.

Mapping out the high-pressure phase diagram of nitrogen has proved challenging due to the overlapping temperature and pressure regions in which individual phases have been observed. In some cases, the observed phases are the thermodynamically stable ones, while in others they are kinetically accessible along certain temperature and pressure pathways through the phase diagram. For example, despite being discovered only very recently, the λ phase coexists with nine other known nitrogen phases,[18] and it is believed to be the thermodynamically most stable phase over a significant portion of its temperature and pressure range.[18,155]

Solving the crystal structures of several high-pressure phases has also proved difficult experimentally. The λ[18] and ι[22] phases were solved only recently via combination of experiment and crystal structure prediction. The structures of the ζ, κ, and θ phases remain unknown, however. The structure of molecular phase ζ has attracted particular attention, as it marks the "frontier" phase in the transition from molecular to non-molecular phases at high-pressure.

Room-temperature compression of phase ϵ induces the transition to phase ζ around 60 GPa.[156] At low temperatures, this transition occurs at even lower pressures, around 25 GPa.[24,110,141] This transition was first observed optically via Raman spectroscopy,[24,141,156] and later confirmed via x-ray diffraction.[23,111,112] It remains stable until around 115 GPa, after which it converts to the monoclinic κ phase at ambient temperature[23] or to non-molecular phases at high temperatures.[150]

Initial proposals suggested that phase ζ had $R3c$ symmetry,[24,141] though this subsequently proved inconsistent with low-temperature Raman and infrared data.[110] In 2004, Eremets et al attempted to solve the crystal structure of phase ζ using powder X-ray diffraction.[111] Their proposed orthorhombic crystal structure contained 4 molecules in unit cell and adopted one of the $P222_1$, $P2_12_12$, or $P2_12_12_1$ space groups, with $P222_1$ being considered most likely. However, a follow-up study argued against the $P222_1$ space group based on further analysis of Raman and x-ray diffraction data.[112] In 2007, Gregoryanz et al re-examined the crystal structure of phase ζ by using single crystal X-ray diffraction.[23] They suggested that the orthorhombic unit cell contains 8 molecules instead of 4 and adopts a $Pmma$ space group. They remained unable to solve for the atomic positions, however.

Computationally, Hooper et al employed *ab initio* crystal structure prediction to search for the ζ-phase structure.[157] Their genetic algorithm search relaxed and ranked structures using planewave density functional theory (DFT) and the PBE functional. They produced two groups of structures based on the space groups reported in the experimental studies above. However, the study proved inconclusive. The lowest-energy structures did not adopt the appropriate space groups or cell types. In the end, they narrowed their list of candidates to four most promising structures: A1 ($Pbcn$), A2 ($P2_12_12_1$), B1 ($Immm$), and B8 ($Pnma$).

Around the same time, Pickard and Needs employed *ab initio* random structure searching (AIRSS)[134] to study high pressure crystal structures of nitrogen.[42] They predicted four low-enthalpy structures at high pressure using PBE calculations.[42]. One of these structures, monoclinic $P2_1/c$, was later found to correspond to the experimentally observed λ phase.[18] More recently, DFT-driven AIRSS contributed to solving the structure of ι nitrogen, which contains 48 N_2 molecules per cell.[22]

Despite these successes, crystal structure prediction alone is not always sufficient. Inaccuracies in the theoretical models and/or the ambiguity surrounding whether an experimentally observed phase is thermodynamically stable or metastable can mean that the most stable candidate structure(s) will not necessarily correspond to the experimentally observed structure. It becomes valuable, therefore, to consider other "orthogonal" observables that can be used to independently assess the agreement between a candidate structure and experimental observations. At ambient pressure, for example, NMR crystallography seeks crystal structure candidates whose predicted chemical shifts match the experimentally observed spectra.[158] In high pressure systems like nitrogen, simulated x-ray diffraction data and Raman spectra (particularly for the lower-frequency lattice modes which are sensitive to crystal packing) can provide insightful when analyzing candidate structures.

For example, Hirata and co-workers helped resolve several controversies surrounding the interpretation of experiments on two difference ice phases through *ab initio* simulation of structures and vibrational spectra.[71–73] We employed structure prediction and Raman techniques in arguing that carbon dioxide phases III and VII are actually the same phase,[114] and that argument has received subsequent support from *ab initio* modeling of the phase diagram.[159] In nitrogen, simulated Raman spectra played a role in confirming the ι phase structure, while we used a combination of AIRSS plus simulated powder x-ray diffraction and Raman spectroscopy to help confirm the structure of λ nitrogen.[155]

The present computational study attempts once again to solve the structure of ζ N_2, this time combining structure prediction and simulated spectroscopy. Structures consistent with experimental constraints were generated randomly via the AIRSS protocol. These structures were then refined with dispersion-corrected DFT. Further comparison of the predicted lattice constants, powder x-ray diffraction patterns, and Raman spectra narrows the list of predicted candidates down to just a handful of plausible structures. Moreover, several ζ nitrogen candidates identified in previous studies can be ruled out based on poor spectroscopic agreement. In the end, none of the candidate structures studied here provides a clear match for the experimental data, but a couple are potentially promising enough to merit further study.

5.2 Methods

5.2.1 AIRSS structure generation:

With sufficient searching, AIRSS ensures broad, unbiased structure prediction. Unfortunately, the complexity of the search space grows exponentially with the number of degrees of freedom. It is therefore common to constrain the search space using whatever experimental information is available. Discovering the 48-molecule unit cell of ι nitrogen would have been virtually impossible without experimental constraints on the space group and lattice parameters to help focus the search, for example.[22]

The AIRSS search here was focused by restraining the randomly generated structures based on experimentally inferred information about the unit cell type and volume.

Specifically, all recent experimental evidence suggests that ζ nitrogen adopts an orthorhombic unit cell with four[111,112] or, more likely, eight[23] molecules per unit cell. Therefore, the AIRSS search was performed over all 59 orthorhombic space groups and cells containing eight N_2 molecules. After selecting the orthorhombic space group at random, the lattice parameters were randomized subject to the constraint that the unit cell volume lie within ±40% of the experimentally reported volume of 114.8 $Å^3$ (8.6 cm^3/mol) at 80 GPa.[23] These constraints considerably narrow the random search space and ideally facilitate the structure prediction. On the other hand, the search will likely fail to find the true structure if the experimentally inferred constraints are incorrect.

Once the cell dimensions and space group were selected, a nitrogen molecule was placed at a randomly selected Wyckoff position associated with that space group. The cell was populated with additional symmetry-equivalent molecules by employing the space group symmetry operators. This procedure of random molecule placement and symmetry-based replication of the molecules was repeated until the cell contained eight molecules. Crystal packings which resulted in intermolecular N_2 distances of less than 1.65 Å were discarded.[157]

5.2.2 Density functional theory structure optimization and enthalpies:

The structures generated by the AIRSS approach were then geometry optimized under 80 GPa of external pressure via periodic planewave DFT using the B86bPBE density functional[100,125] and the exchange-hole dipole moment (XDM) dispersion correction,[126] as implemented in Quantum Espresso version 6.2.1.[99,127] Note that while the XDM dispersion correction is included, the impacts of the correction are small at the high pressures considered here. The DFT calculations employed an 80 Ry plane-wave cutoff and projector-augmented wave (PAW) potentials for nitrogen atoms that were produced using A. Dal Corso's Atomic code v6.1. Assessment and convergence testing of the planewave cutoff and PAW potentials for solid state nitrogen phases has been performed previously.[155]

Structures were initially optimized with a 3×3×3 Monkhorst-Pack *k*-point grid. After clustering to remove duplicate structures, the structures were refined further with a 6×6×6 *k*-point grid. Larger, anisotropic *k*-point meshes were tested for selected structures with small individual lattice constants, but the structures and enthalpies did not change appreciably. The use of adaptive grids that target consistent *k*-point density would be more efficient computationally. Space groups for the optimized structures were determined using FINDSYM version 7.[160,161] Simulated powder x-ray diffraction (PXRD) spectra were generated for each DFT-optimized structure using Mercury[102] and the same 0.3683 Å wavelength as the experiments.[23] The rmsd15 metric[1] overlaying clusters of 15 molecules from the crystal was employed for selected structure comparisons.

The enthalpies of the candidate structures were estimated by combining the DFT electronic energy with the pressure-volume term, $H = E_{elec} + PV$. Vibrational contributions to the molar volume and enthalpy are neglected here. After identifying candidate structures at 80 GPa, the structures were also optimized at 0, 2, 4, 6, 8, 10, 15, 20, 40, 60, 100, 120, and 150 GPa. The equation of state was then interpolated between these data points via cubic splines.

5.2.3 Simulated Raman spectra:

Previous work has found that fragment-based correlated wavefunction calculations employing second-order Møller-Plesset perturbation theory (MP2) can predict Raman spectra in good agreement with experiment[114,155] These calculations are made feasible via the fragment-based hybrid many-body interaction (HMBI) model. HMBI partitions the total energy of the nitrogen crystal into contributions arising from individual molecules (1-body), their pairwise intermolecular interactions (2-body terms), and the remaining non-pairwise-additive many-body contributions. The 1-body and shorter-range 2-body terms (out to 6 Å) are computed with MP2, while the longer-range pairwise and many-body terms are approximated at a lower level of theory,

$$\begin{aligned} E_{crystal}^{HMBI} &= E_{1\text{-}body}^{MP2} + E_{SR\ 2\text{-}body}^{MP2} + E_{LR\ 2\text{-}body}^{Low} \\ &\quad + E_{many\text{-}body}^{Low} \end{aligned} \tag{5.1}$$

Only symmetry-unique monomer and dimer fragments need to be computed in evaluating Eq 5.1.[95] The many-body contributions here are largely modeled with the AMOEBA polarizable force field[79] under periodic boundary conditions. Polarizable force field calculations were conducted using Tinker version 6.3.[93] with existing AMOEBA force field parameters for the N_2 molecule.[79]

To simulate Raman spectra, the atomic positions were first relaxed via the MP2/aug-cc-pVDZ + AMOEBA HMBI approximation within fixed DFT unit cell parameters. Harmonic phonon frequencies were then computed at the zone center ($\mathbf{k} = 0$) by diagonalizing the Hessian matrix computed from the second derivative of Eq 5.1 with respect to atomic positions. The 1-body and 2-body fragment contributions to the Hessian were computed analytically using Gaussian 09.[96] Raman intensities were approximated via finite difference of the polarizability derivatives using only the 1- and 2-body MP2 terms from Eq 5.1.[72] Neglecting the intermolecular many-body contributions is not expected to have a large impact on the Raman intensities for non-polar N_2. This protocol is very similar to ones that proved effective in previous simulated Raman studies of high-pressure nitrogen and carbon dioxide phases.[114,155]

Finally, fragment-based calculations were performed to refine the unit cell volumes of selected crystal structures. These calculations combined complete-basis-set MP2 monomer and dimer fragments with a periodic Hartree-Fock/pob-TZVP basis treatment of the long-range and many-body terms using a protocol described previously.[155] However, the periodic Hartree-Fock calculations with the Gaussian basis set proved difficult to converge above ∼30 GPa of pressure. Accordingly, data computed at 0, 2, 4, 6, 8, 10, 15, 20, 25, and 30 GPa was fitted to a Murnaghan equation of state,

$$H(V) = H_0 + \frac{B_0 V}{B_0'}\left[\frac{(V_0/V)^{B_0'}}{B_0' - 1} + 1\right] - \frac{B_0 V_0}{B_0' - 1} \tag{5.2}$$

where the enthalpy (H_0), volume (V_0), bulk modulus (B_0), and its first pressure derivative (B_0') at zero pressure are fitting parameters. The limited pressure range potentially limits the reliability of the equation of state at higher pressures.

5.3 Results and Discussion

5.3.1 Lattice constants for known crystalline phases of nitrogen

To assess the level of agreement one should expect for high pressure nitrogen phases with the DFT model used here, structure optimizations were carried out under experimental pressure conditions. When lattice constants were available at multiple pressures, the calculations were performed at each. Table 5.1 summarizes the results of these calculations which were used to generate Figure 5.11.

5.3.2 Crystal structure prediction and structure ranking

The AIRSS search generated 32 unique molecular structures. Tables 5.2-5.4 summarize the crystal structures for the 22 lowest-enthalpy ones, which lie within 40 kJ/mol of the most stable structure. They were optimized with B86bPBE-XDM at 80 GPa.

Table 5.1: Comparison between predicted and experimental lattice constants for several known phases of nitrogen and two ζ-phase candidates.

Phase	Experimental Reference	Pressure (GPa)	Experiment a (Å)	Experiment b (Å)	Experiment c (Å)	Predicted a (Å)	Predicted b (Å)	Predicted c (Å)
α	Ref 145	0	5.644	5.644	5.644	5.626	5.626	5.626
	Ref 20	0.269	5.495	5.495	5.495	5.467	5.467	5.467
γ	Ref 146	0.4	3.957	3.957	5.109	3.899	3.899	5.166
ϵ	Ref 21	5.4	8.208	8.208	11.362	8.179	8.179	11.448
	Ref 21	7.8	8.020	8.020	11.104	7.974	7.974	11.199
	Ref 21	9.8	7.859	7.859	11.029	7.840	7.840	11.035
	Ref 147	16.3	7.605	7.605	10.622	7.529	7.529	10.661
	Ref 147	27.6	7.241	7.241	10.261	7.192	7.192	10.252
λ	Ref 18	34	3.051	3.066	5.705	2.951	2.916	5.638
ι	Ref 22	56	9.899	8.863	8.726	9.794	8.622	8.919
ζ	Ref 23 (vs #12)	90	6.533	2.574	6.844	6.331	2.740	6.802
	Ref 23 (vs #19)					6.407	2.768	6.713

Table 5.2: The unit-cell parameters and atomic positions of AIRSS structures #1–#9, optimized at 80 GPa with B86bPBE-XDM. Atomic positions are reported in fractional coordinates. The remaining atoms can be generated via the appropriate space group symmetry operations.

Structure	Space group	a (Å)	b (Å)	c (Å)	Atoms	Multiplicity	Wyckoff position	x	y	z	Relative Enthalpy (kJ/mol)
#1 (PN1)	$P2_12_12_1$	2.691	7.645	2.690	N1	4	a	0.71652	0.85079	0.69430	0.000
					N2	4	a	0.44416	0.89922	-0.03356	
#2 (A1)	$Pbcn$	2.671	4.240	9.793	N1	8	d	0.86705	0.77493	0.29285	0.840
					N2	8	d	0.36717	-0.00143	0.45712	
#3	$P2_12_12_1$	3.939	10.489	2.704	N1	4	a	0.39411	-0.09163	-0.00113	6.927
					N2	4	a	0.52930	-0.03296	0.72548	
					N3	4	a	0.31950	0.71407	0.78261	
					N4	4	a	0.45401	0.65851	0.49886	
#4	$Cmcm$	3.769	3.867	7.703	N1	8	f	0.00000	0.86637	0.02453	8.699
					N2	8	g	0.14553	0.41735	0.25000	
#5 (B1)	$Immm$	2.612	3.064	3.474	N1	4	i	0.00000	0.00000	0.84132	9.131
#6	$Cmca$	3.133	4.439	8.049	N1	8	f	0.00000	0.07366	0.16211	10.692
					N2	8	f	0.00000	0.28296	0.08878	
#7	$Cmca$	3.164	4.549	3.913	N1	8	f	0.00000	0.10709	0.56510	11.559
#8	$Pbcm$	2.659	10.561	3.989	N1	4	d	0.69672	0.38740	0.25000	11.849
					N2	4	d	-0.03286	0.46631	0.25000	
					N3	4	d	0.45280	0.13643	0.25000	
					N4	4	d	0.18445	0.21571	0.25000	
#9 (B8)	$Pnma$	5.294	7.952	2.664	N1	4	c	0.27949	0.25000	0.14449	12.430
					N2	4	c	0.12220	0.25000	0.87491	
					N3	8	d	0.42125	0.50077	0.13418	

Table 5.3: The unit-cell parameters and atomic positions of AIRSS structures #10–#15, optimized at 80 GPa with B86bPBE-XDM. Atomic positions are reported in fractional coordinates. The remaining atoms can be generated via the appropriate space group symmetry operations.

Structure	Space group	a (Å)	b (Å)	c (Å)	Atoms	Multiplicity	Wyckoff position	x	y	z	Relative Enthalpy (kJ/mol)
#10	$P2_12_12$	5.109	8.048	2.730	N1	4	c	0.50974	0.21684	0.27142	14.338
					N2	4	c	0.41983	0.29346	0.55979	
					N3	4	c	0.70542	0.44798	-0.05029	
					N4	4	c	0.79992	0.52278	0.23569	
#11	$Pbca$	5.043	5.272	4.229	N1	8	c	0.44239	0.21205	0.35139	14.434
					N2	8	c	0.21099	-0.06439	0.84756	
#12	$P2_12_12_1$	6.820	2.698	6.1760	N1	4	a	0.41702	0.54361	0.40282	14.744
					N2	4	a	0.53885	0.79864	0.43667	
					N3	4	a	0.21364	0.87923	0.17781	
					N4	4	a	0.14210	0.13636	0.28932	
#13	$Abm2$	7.859	3.607	3.981	N1	8	d	0.87479	0.59751	0.24224	15.006
					N2	4	c	0.65128	0.25000	0.18565	
					N3	4	c	0.59954	0.25000	0.44201	
#14	$Pnma$	9.135	4.693	2.624	N1	4	c	0.28847	0.25000	0.76621	16.317
					N2	4	c	0.19244	0.25000	0.51267	
					N3	8	d	0.04200	-0.03174	0.86125	
#15	$Pbca$	4.965	4.514	5.074	N1	8	c	0.64484	-0.03657	0.66468	17.507
					N2	8	c	0.57989	0.86590	0.85244	

Table 5.4: The unit-cell parameters and atomic positions of AIRSS structures #16–#22, optimized at 80 GPa with B86bPBE-XDM. Atomic positions are reported in fractional coordinates. The remaining atoms can be generated via the appropriate space group symmetry operations

Structure	Space group	a (Å)	b (Å)	c (Å)	Atoms	Multiplicity	Wyckoff position	x	y	z	Relative Enthalpy (kJ/mol)
#16	$Pbcn$	5.286	5.088	4.222	N1	8	d	0.62070	0.83839	0.04342	18.335
					N2	8	d	0.80678	-0.08434	0.11183	
#17	$P2_12_12_1$	4.331	4.396	5.977	N1	4	a	0.40427	0.82407	0.03469	18.924
					N2	4	a	0.42697	0.82940	0.21731	
					N3	4	a	-0.04166	0.13154	0.10095	
					N4	4	a	0.11821	0.31556	0.14516	
#18	$Pna2_1$	4.287	5.985	4.431	N1	4	a	0.47232	0.10179	0.53304	19.038
					N2	4	a	0.64689	0.15183	0.70054	
					N3	4	a	0.56392	0.54126	0.71839	
					N4	4	a	0.55523	0.71260	0.63009	
#19	$P2_12_12$	6.624	6.298	2.720	N1	4	c	0.46958	0.67013	0.72292	21.374
					N2	4	c	0.58365	0.72498	0.45923	
					N3	4	c	0.76236	0.47044	0.19221	
					N4	4	c	0.85074	0.55757	-0.08383	
#20	$Cmmm$	5.001	6.501	3.525	N1	8	p	0.28481	0.32980	0.00000	25.659
					N2	4	j	0.00000	-0.08510	0.50000	
					N3	4	l	0.00000	0.50000	0.34502	
#21	$Pnna$	4.799	9.206	2.604	N1	8	e	0.55339	0.54119	0.38385	27.828
					N2	8	e	0.71397	0.29793	0.87568	
#22	$Fddd$	2.388	6.175	7.885	N1	16	f	0.12500	0.71365	0.12500	39.550

As mentioned in the text, structure A2 was not found in the $Z = 8$ search, but the closely related structure #3 was found. Figure 5.1 compares the two structures. Structure A2 overlays almost perfectly with the right half of the structure #3 cell in Figure 5.1. However, the left half of structure #3 is related to the right half via two-fold screw rotation, rather than the simple translation it would be in a $Z = 8$ supercell of A2. Because structure #3 is about 5 kJ/mol more stable than A2, it seems likely that this structure was found preferentially in the AIRSS search.

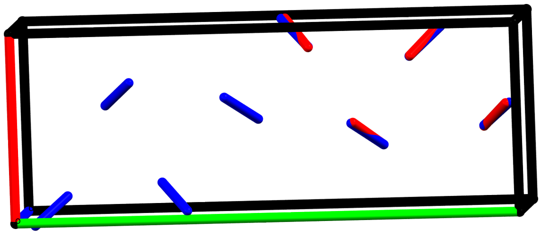


Figure 5.1: Overlay of structure A2 (red) and #3 (blue). They two are virtually identical in half the cell, but differ in the symmetry operations relating the other half of the cell.

5.3.3 Comparison of x-ray diffraction patterns and Raman spectra

Simulated powder x-ray diffraction (PXRD) and Raman spectra were generated for each the 22 lowest-enthalpy candidate. Figures 5.2 and 5.3 present the complete set of simulated spectra.

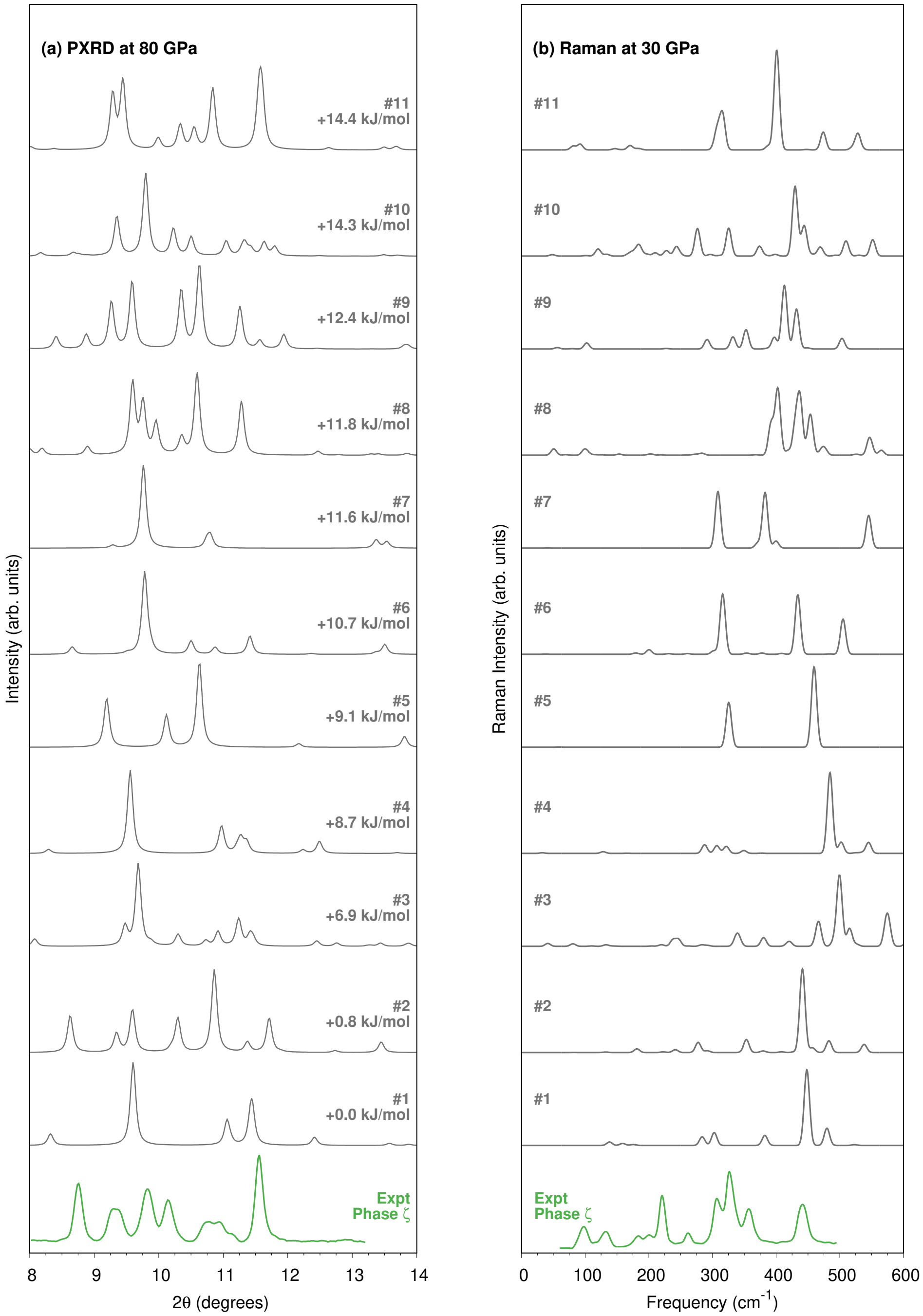


Figure 5.2: Comparison between predicted and experimental (a) powder X-ray diffraction pattern (λ=0.3683 Å) and (b) Raman spectrum for nitrogen phase ζ.

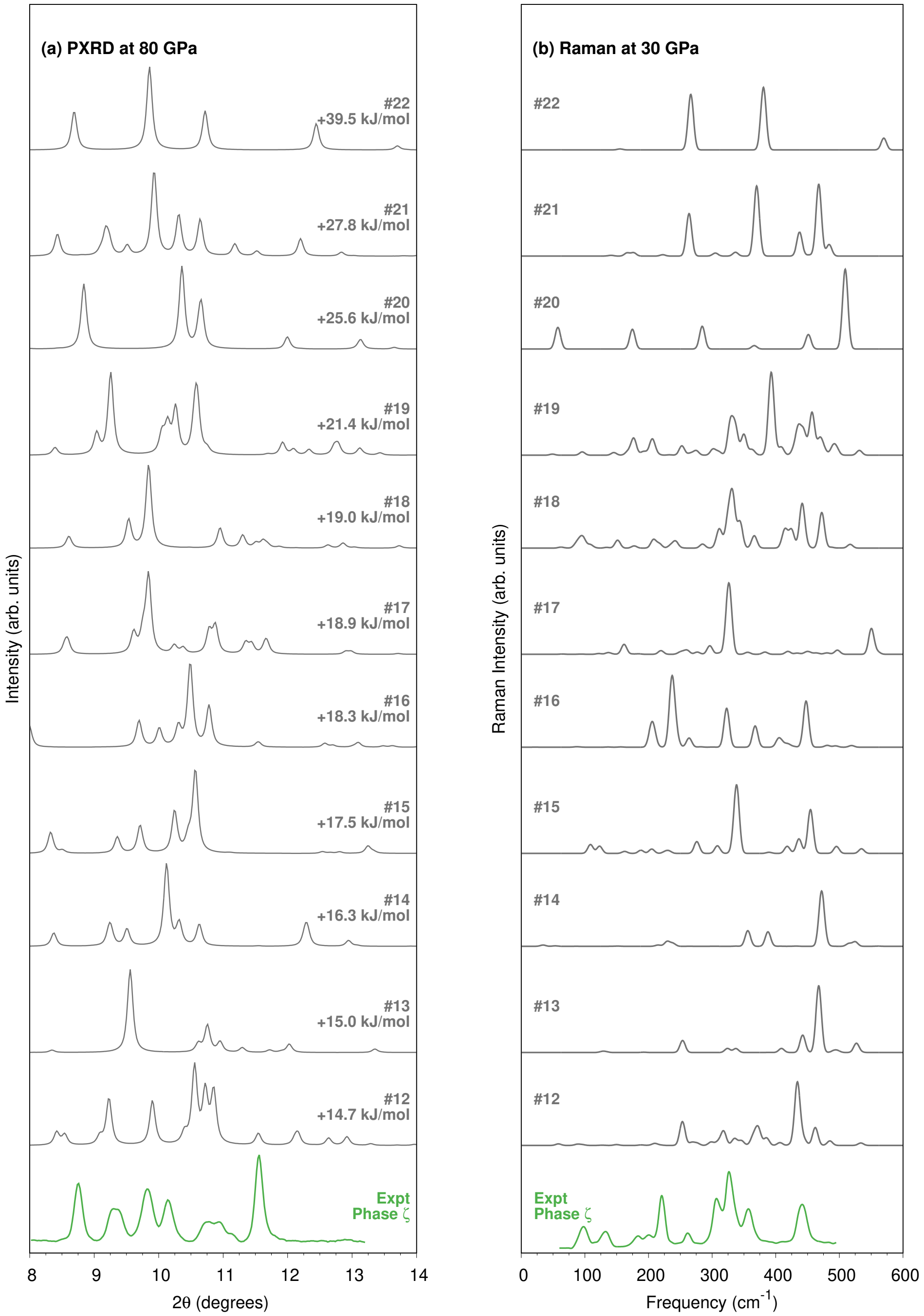


Figure 5.3: Comparison between predicted and experimental (a) powder X-ray diffraction pattern (λ=0.3683 Å) and (b) Raman for nitrogen phase ζ.

5.3.4 Equation of state for phase ϵ

For comparison with phase ζ data, structure optimizations were also performed for phase ϵ. The periodic planewave DFT were carried on at the same B86bPBE-XDM level of theory at 0, 2, 4, 6, 8, 10, 15, 20, 40, 60, 80, 100, 120, and 150 GPa. Figure 5.4 compares the predicted and experimental equations of state. Similar to what was found for the ζ phase, the DFT calculations underestimate the molar volume, especially at higher pressures.

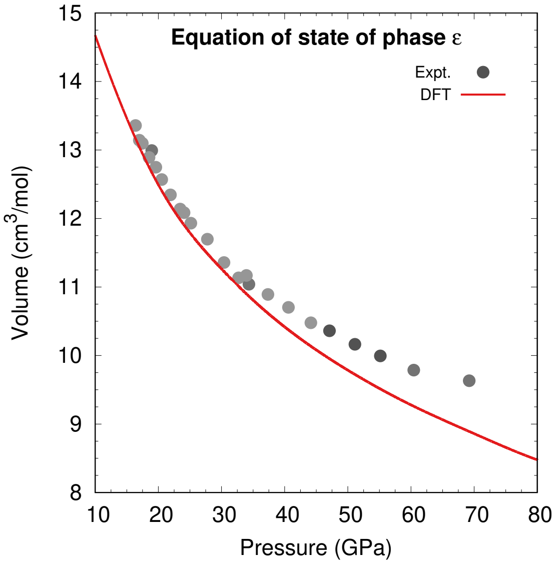


Figure 5.4: Comparison of the B86bPBE-XDM predicted equation of state against experimental data for nitrogen phase ϵ.

5.3.5 Symmetry breaking of structure #19

To eliminate a small imaginary vibrational frequency for structure #19 found with the MP2-based Raman frequencies in the DFT unit cell, a 1-D scan was performed along the imaginary mode in order to find the local minima. This process discovered a new pair of minima that lie ~0.3 kJ/mol below the original structure. These structures result from a symmetry breaking that lowers the symmetry from $P2_12_12$ to $P2$. Figure 5.5 plots the energy along this symmetry-breaking coordinate.

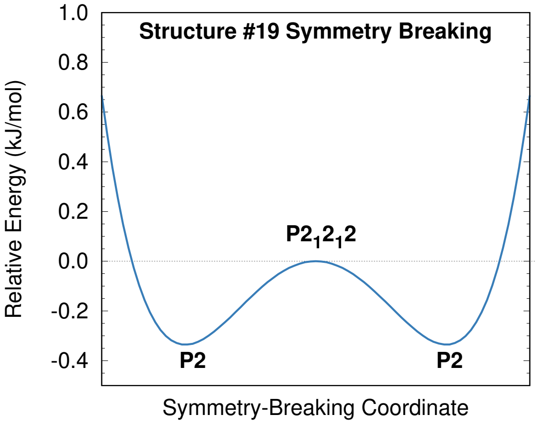


Figure 5.5: At 30 GPa, structure #19 breaks symmetry slightly, relaxing from $P2_12_12$ to P_2 symmetry and lowering the enthalpy by ~0.3 kJ/mol).

5.3.6 Structures #12 and #19 as candidates for phase κ

Given the similarity of the experimentally reported lattice constants between phase ζ and κ, atomic coordinate of structure #12 and #19 were fitted into the experimentally reported cells[23] for phase ζ at 80 GPa and κ at 130 GPa. For orthorhombic ζ, the atomic coordinates were fitted into the cell with lattice parameters a=6.85 Å, b=6.51 Å, c=2.58 Å. For monoclinic κ, the atomic coordinates were fitted into the cell with lattice parameters a=6.92 Å, b=6.20 Å, c=2.29 Å, and $\beta = 91.8°$. The structures were then relaxed with DFT subject to fixed lattice constants.

Structure #12 changed only slightly upon relaxation in the κ cell, with an rmsd15[1] of only 0.2 Å relative to the original 80 GPa structure for the 15-molecule cluster. Structure #19 changed considerably more, and the only 8 out of 15 molecules overlapped with the original coordinates when using distance and angle tolerances of 20% and 20° via Mercury's Packing Similarity feature[102]. The structure of #12 within the κ lattice parameters is shown in Figure 5.6. CIF files for both structures #12 and #19 in the κ unit cell are provided separately.

Comparison of the simulated and experimental x-ray diffraction patterns are shown in Figure 5.7. While neither structure #12 nor #19 is a perfect match for the experimental κ phase data, there are a number of similarities between #12 and experiment in particular. These results make #12 an interesting candidate for κ nitrogen, though more experimental data would be needed before drawing further conclusions.

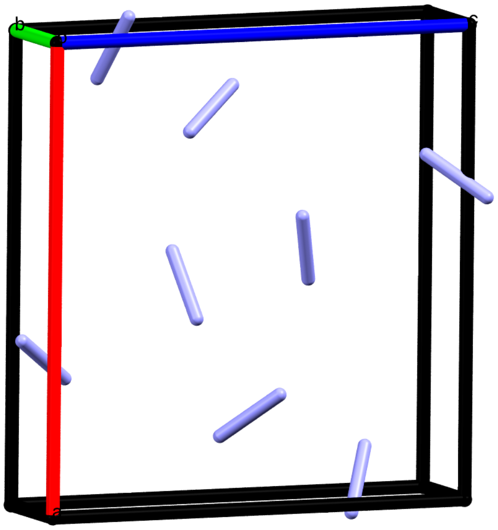


Figure 5.6: Structure #12 optimized within the experimental κ phase lattice parameters at 130 GPa.

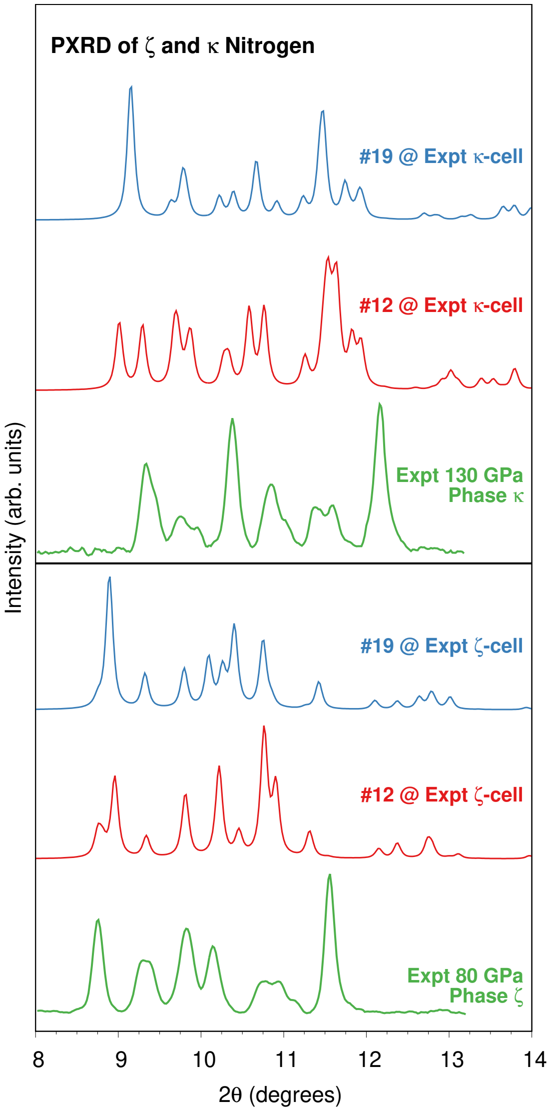


Figure 5.7: Comparison of the simulated powder X-ray diffraction patterns of structures #12 and #19 before and after constraining the lattice constants to match the experimentally reported values. All spectra employ a wavelength of $\lambda = 0.3683$ Å.

5.3.7 Crystal energy landscape

Nearly 1,300 structures with orthorhombic space groups were generated randomly for the AIRSS search. After initial B86bPBE-XDM relaxation at 80 GPa of external pressure with the modest 3×3×3 *k*-point grid, about half of the random structures optimized to network covalent/polymeric forms. Those structures were discarded from this study due to (1) their significantly higher enthalpies compared to molecular forms at 80 GPa and (2) because all experimental evidence suggests ζ nitrogen is a molecular phase.[23,24,111,112,141] Space group symmetry was not enforced during the relaxations. Of the 636 remaining molecular structures, 553 relaxed orthorhombic space groups. The other 83 relaxed to monoclinic, triclinic, or tetragonal structures and were discarded. After clustering based on crystal packing similarity, 34 unique molecular crystal structures remained. Each of these 34 structures was generated multiple times during the AIRSS search, and several of them were generated tens of times.

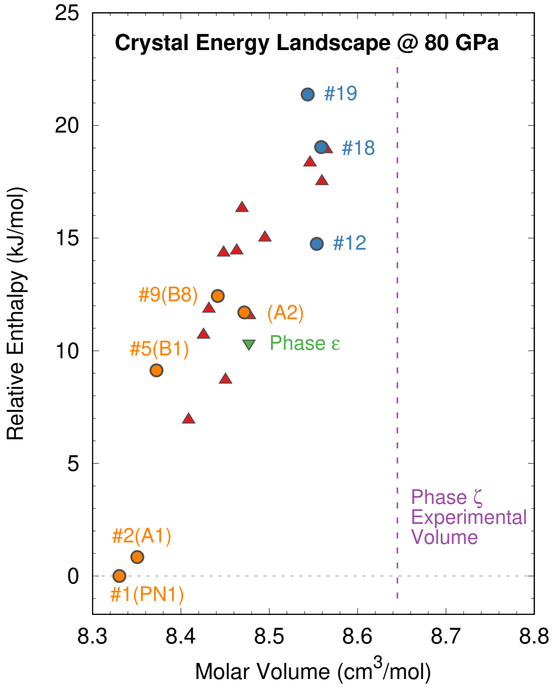
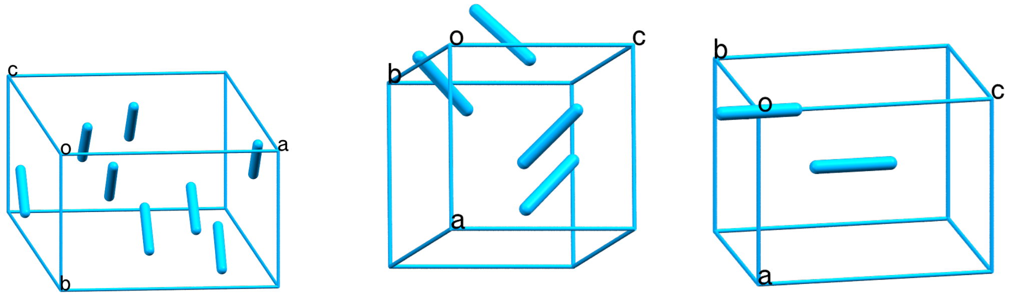
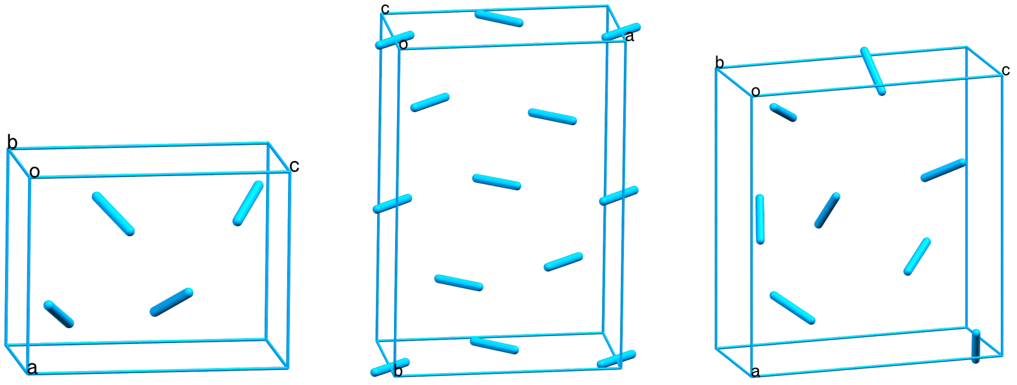


Figure 5.8: Predicted crystal energy landscape for orthorhombic molecular N_2 phases at 80 GPa. Structures that have been reported in earlier structure prediction studies are shown in orange, while candidates in blue are new structures discussed in detail below. The enthalpy and volume for the optimized ϵ phase is shown for comparison, along with the reported molar volume of phase ζ.

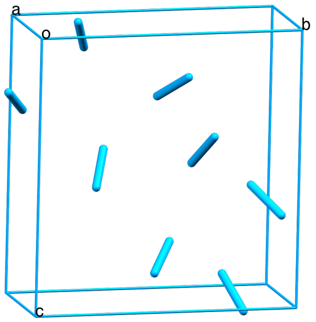
The crystal structures of these 34 candidates were subsequently refined with the denser 6×6×6 k-point grid. This additional relaxation caused two more pairs of structures to coalesce, leaving 32 unique structures. For convenience, these structures will be referred to by number in order of increasing enthalpy, with Structure #1 being the most stable. 19 of these 32 predicted structures lie within 25 kJ/mol of Structure #1 at the B86bPBE-XDM level of theory, as shown in the crystal energy landscape plotted in Figure 5.8. The remaining structures have even higher relative enthalpies and seem very unlikely to occur at this pressure experimentally. Figure 5.9 shows selected structures that will be discussed in detail below.



(a) PN1 (b) A1 (c) B1



(d) A2 (e) B8 (f) #12



(g) #19

Figure 5.9: Selected candidate crystal structures discussed in the text, as optimized with DFT at 80 GPa.

From Figure 5.8, it is clear that the higher-density structures are generally more stable, as one expects under high external pressures. It is also notable that the DFT-predicted structures underestimate the experimental volume by up to 4% at 80 GPa. This volume underestimation extends across a broad pressure range, as seen from the equations of state plotted in Figure 5.10.

However, the volume errors here do not indicate that the predicted structures are inconsistent with experiment. Similar B86bPBE-XDM volume underestimation is seen for the ϵ (Section 4 5.3.4) and λ N_2 phases[155] as well. That study showed how refining the geometries with correlated wave function methods improved the agreement of the predicted structures with the experimental molar volume, and similar results are obtained here. Figure 5.10 plots Murnaghan equations of state for three candidate structures (B8, #12, #19) that were fitted to predictions from the same fragment-based MP2/CBS + pHF approach used in Ref 155. Subject to the caveats noted in the Methods section about their accuracy at high pressure, the MP2 results suggest that the candidate structures here do exhibit molar volumes that are consistent with the experimental data for phase ζ.

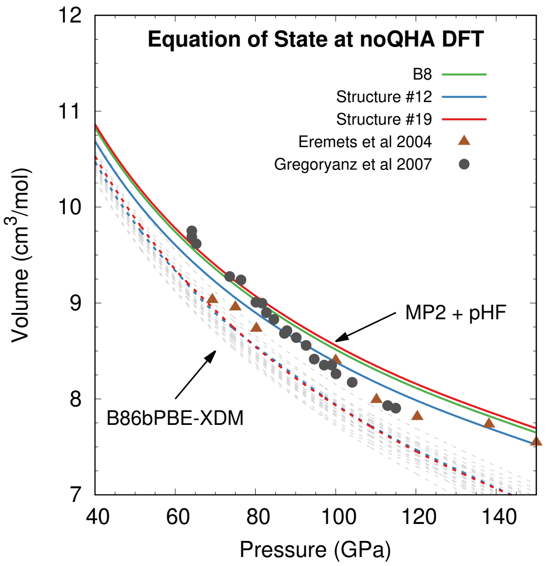


Figure 5.10: Comparison of predicted equations of state for all predicted structures against experimental data for nitrogen phase ζ. Dashed lines correspond to the B86bPBE-XDM predicted equations of state for all candidate structures from Figure 5.8, with the lines for structures B8, #12 and #19 highlighted in green, blue, and red. Solid lines show the new equation of state for the selected structures after fragment-based MP2/CBS + pHF refinement.

The 32 unique structures generated here include several structures that were identified in previous crystal structure prediction studies of nitrogen. Structure #1 corresponds to the 4-molecule $P4_12_12$ structure found by Pickard and Needs.[42] They also found it to be the most stable structure in this pressure regime. It will be referred to as structure PN1 for the remainder of the present work. Structure #2 matches the four-molecule $Pbcn$ structure of Pickard and Needs[42] and what Hooper et al[157] called structure A1. The search here also found structures corresponding to Hooper et al's two-molecule $Immm$ B1 (#5) and eight-molecule $Pnma$ B8 (#9) structures.

The $Z = 8$ search did not discover a supercell corresponding to the A2 structure ($P2_12_12_1$, $Z = 4$) from Hooper et al. This does not reflect any fundamental problem of the AIRSS search algorithm: a $Z = 4$ search in the $P2_12_12_1$ space group found A2 readily. Rather, the larger $Z = 8$ cell used here allows the closely related but more stable structure #3 to form instead. When compared against a $Z = 8$ supercell of A2, half of the structure #3 cell is virtually identical to A2, while the other half is related via a two-fold screw axis instead of simple translation (see Section 5.3.3). Because Structure #3 lies 5 kJ/mol lower in energy than A2, it was presumably found preferentially in the $Z = 8$ search. Both structures are included in Figure 5.8.

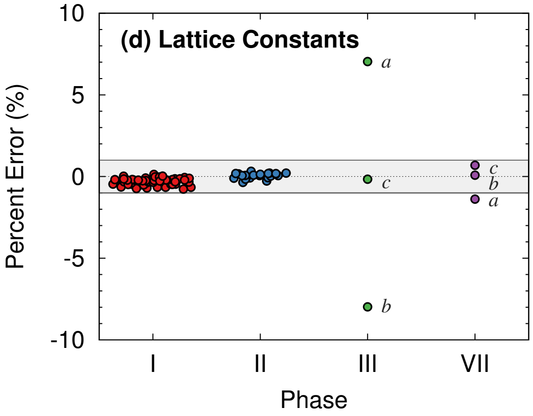


Figure 5.11: Comparison between predicted and experimental lattice constants for several known phases of nitrogen and two ζ-phase candidates. The calculation were performed using B86bPBE-XDM with $6\times6\times6$ k-point grid at the experimental pressure. Where experimental data was available at multiple pressures, errors are shown as computed for each pressure. See Table 5.1 for details.

Table 5.5: Comparison of the B86bPBE-XDM lattice constants for several predicted structures compared to the experimentally reported values for ζ nitrogen from Ref 23. Only structures #12 and #19 exhibit lattice parameters that are similar to the experimental ones.

Lattice constants	PN1 $P4_12_12$	A1 $Pbcn$	B1 $Immm$	B8 $Pnma$	#12 $P2_12_12_1$	#18 $Pna2_1$	#19 $P2_12_12$	Expt (Ref 23)
				80 GPa				
a (Å)	2.690	4.241	10.450	5.294	6.176	4.287	6.298	6.507
b (Å)	5.382	2.671	3.474	7.952	2.698	5.985	2.721	2.578
c (Å)	7.645	9.793	3.064	2.664	6.820	4.431	6.624	6.846
				90 GPa				
a (Å)	2.658	4.185	10.473	5.241	6.042	4.237	6.217	6.533
b (Å)	5.318	2.639	3.440	7.848	2.658	5.929	2.685	2.574
c (Å)	7.556	9.696	2.972	2.630	6.808	4.370	6.558	6.844

From Figure 5.8, PN1 and A1 are considerably more stable enthalpically than any of the other predicted structures. However, as shown in Table 5.5, none of the aforementioned structures exhibit lattice parameters that are similar to the experimentally reported ones.[23] Looking through the complete set of structures, we find two structures, #12 ($P2_12_12_1$) and #19 ($P2_12_12$), which have lattice constants in good agreement with the experimentally reported lattice constants. The lattice parameters for structure #19 match experiment to within 0.2 Å, while those for #12 match within 0.3 Å. Figure 5.11 compares the errors between the predicted and experimental lattice constants for several other known N_2 phases computed with the same DFT method. This figure shows that the lattice constant agreement for Structures #12 and #19 relative to experiment is consistent with what is found for other high-pressure phases like ϵ, λ, and ι. The space groups of these two structures also match space groups that were originally suggested for ζ nitrogen,[111] even if subsequent work identified an alternative space group.[23] None of the other candidate structures generated in the search here appear viable based on lattice constants (See Section 2).

While the lattice constant agreement for structures #12 and #19 is encouraging, their enthalpies lie 15 and 21 kJ/mol higher than structure 1, respectively. For crystallization from solution at ambient conditions, the typical energy window for polymorphism is ~10 kJ/mol,[136] and structures lying more than 10 kJ/mol above the most stable structures are deemed unlikely to crystallize experimentally.[162] The kinetics of high-pressure solid-state phase transitions differ considerably from those of solution-phase crystallizations and can depend on the sample history, but crystal forms lying far outside the 10 kJ/mol window would still seem unlikely to occur experimentally. The stabilities of various candidate structures will be discussed in more detail below.

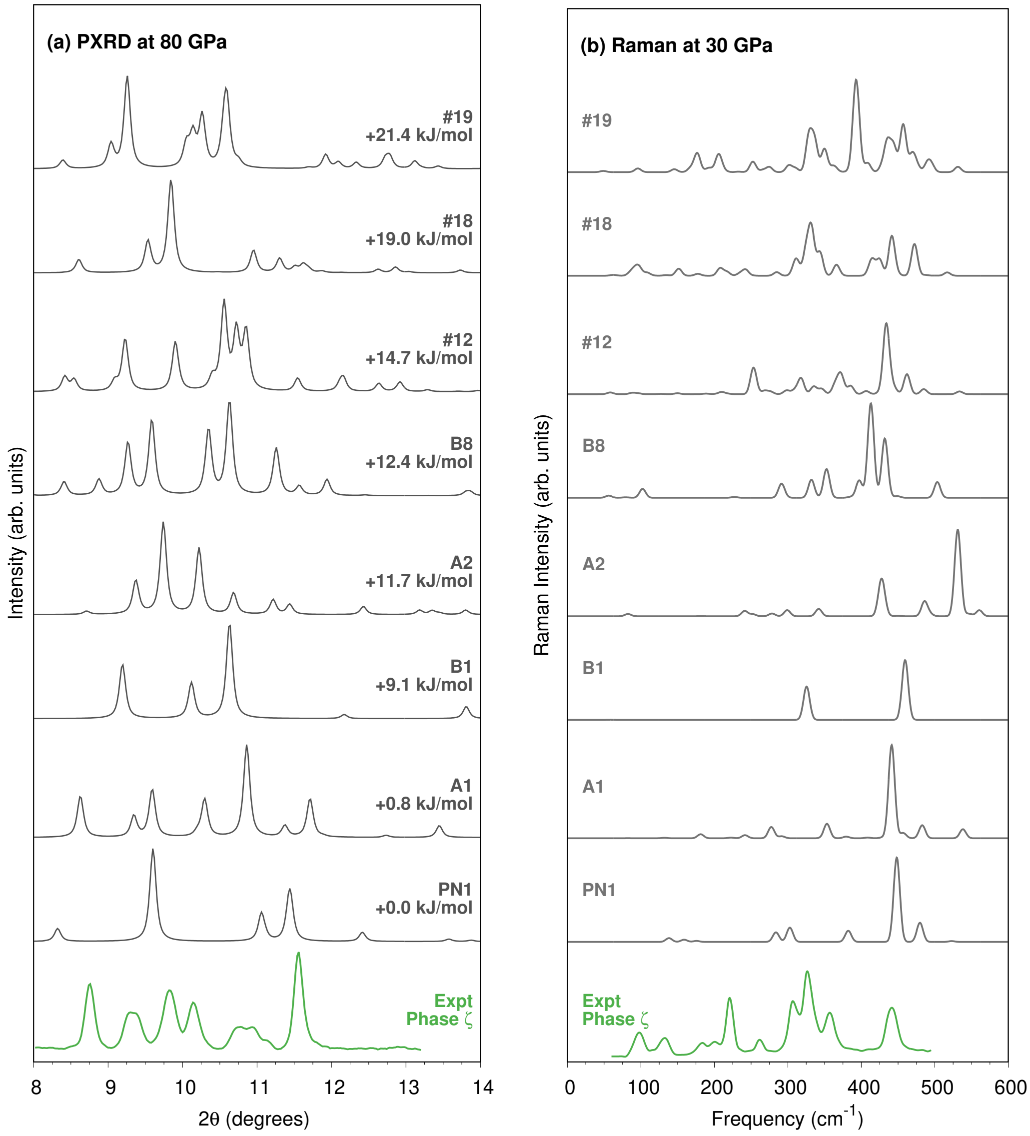


Figure 5.12: (a) Simulated powder x-ray diffraction patterns at 80 GPa ($\lambda = 0.3683$ Å) and (b) predicted Raman spectra at 30 GPa for selected candidate structures, compared against the experimental data for ζ N_2.[23,24] See Section 5.3.3 for the complete set of simulated spectra.

5.3.8 Simulated x-ray diffraction and Raman spectra

Unit cell representations are not unique, and simple comparison of lattice parameters may not be sufficient to determine crystal structure similarity. Accordingly, the simulated and experimental powder X-ray diffraction (PXRD) patterns are compared for all candidate structures shown in the crystal energy landscape (Figure 5.8). Selected examples are shown in Figure 5.12a; all others can be found in Section 5.3.3. The simulated diffraction patterns for PN1 and B1 clearly differ significantly from experiment, as do many of the other ones shown in SI. Structures A1, B8, #12, #18, and #19 exhibit some PXRD features in common with experiment with regard to peak position and, to a lesser extent, peak intensity, though none is a clear match. The PXRD comparison is complicated by the fact that experimentally observed peak intensities can be problematic and sometimes depend on sample history in high-pressure molecular systems like ζ N_2.

Further insight is gained by simulating Raman spectra for the candidate structures and comparing them against the low-temperature (32 K) experimental spectrum at 30 GPa.[24] Selected species are shown in Figure 5.12b; all others can be found in the Section 5.3.3. The DFT underestimation of the molar volumes inhibits intermolecular vibrations and tends to shift the lattice phonon modes toward higher frequencies. For comparison, consider phase ϵ. Similar to the ζ-phase candidate structures here, the molar volume of phase ϵ is underestimated across the pressure range (Section 5.3.4), and this leads to considerable blue shift in the predicted Raman spectrum. Testing in smaller unit cells suggests that the small MP2/aug-cc-pVDZ basis set used in computing the frequencies also contributes to the frequencies being overestimated. Using the structure predicted at 20

GPa instead of 30 GPa increases the molar volume and shifts the Raman spectrum toward the red, improving the agreement between the predicted peaks and experimental Raman spectra considerably (Figure 5.13a).

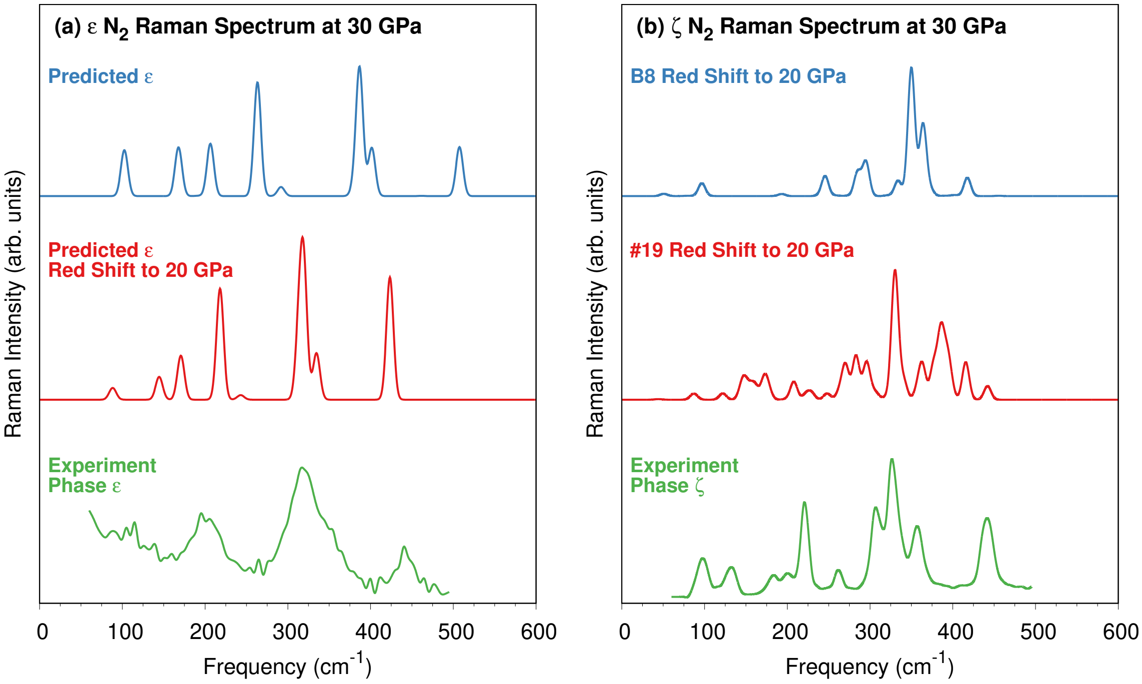


Figure 5.13: (a) Reducing the pressure used to compute the predicted spectrum to 20 GPa improves the agreement between theory and experiment considerably for the known ϵ phase. (b) Comparison of the 20 GPa spectra predicted for structures B8 and #19 against the unknown ζ phase spectrum at 30 GPa. Experimental spectra at 293 K for ϵ and 32 K for ζ taken from Ref 24.

For the ζ phase candidates, the vast majority of the predicted structures exhibit Raman spectra that are clearly inconsistent with the experimental spectrum, even if one similarly corrects for the artificial blue shift. Structures such as PN1, A1, B1, and A2 exhibit far fewer Raman-active modes than the experimental spectrum, making them poor candidates for phase ζ. Structure #12 has more Raman-active modes than those others, though the peaks are shifted much too far to the blue and would lack any significant Raman activity below 200 cm^{-1} even after applying a red shift.

Only three of the structures considered here show plausible resemblance to the experimental Raman spectrum: B8, #18, and #19. Structure #18 has several peaks in excellent agreement with experiment, such as the trio of peaks just above 300 cm^{-1} and the high-intensity modes around 450 cm^{-1} (Figure 5.12b). On the other hand, it lacks the higher-intensity mode(s) around 225 cm^{-1}. Red-shifting the structure #18 spectrum using the 20 GPa structure would only reduce agreement further.

The Raman spectra of both B8 and #19 appear shifted too far to the blue, but using the spectrum predicted at 20 GPa instead of 30 GPa improves the agreement with experiment in both cases (Figure 5.13b). The shifted B8 spectrum exhibits excellent agreement with the higher-frequency end of the experimental spectrum, particularly above 300 cm^{-1}. On the other hand, it exhibits little appreciable Raman activity in the ~100–200 cm^{-1} region, in clear contrast to the experimental spectrum.

The shifted spectrum for Structure #19 exhibits some agreement with experiment (Figure 5.13b), including the pair peaks around 100 cm^{-1} in the experiment that appear near 150 cm^{-1} in the predicted spectrum, plus a high density of peaks in the ~300–400 cm^{-1} region. However, structure #19 exhibits clear disagreements in other portions of the spectrum, even after the red shift. Interestingly, the $P2_12_12$ Structure #19 initially had one imaginary phonon frequency after fragment-based MP2 geometry optimization. Further relaxing the structure along this imaginary coordinate lowers the symmetry to $P2$ and stabilizes the structure by 0.3 kJ/mol (See Section 5.3.5). This shallow double-well potential could introduce appreciable anharmonicity into the phonon modes that is not accounted for here. While the monoclinic $P2$ symmetry is inconsistent with the experimentally inferred space groups, the barrier is so low that the two structures could readily interconvert (or ground state might even lie above the barrier between the two P_2 states), meaning that system could adopt the $P2_12_12$ structure on average.

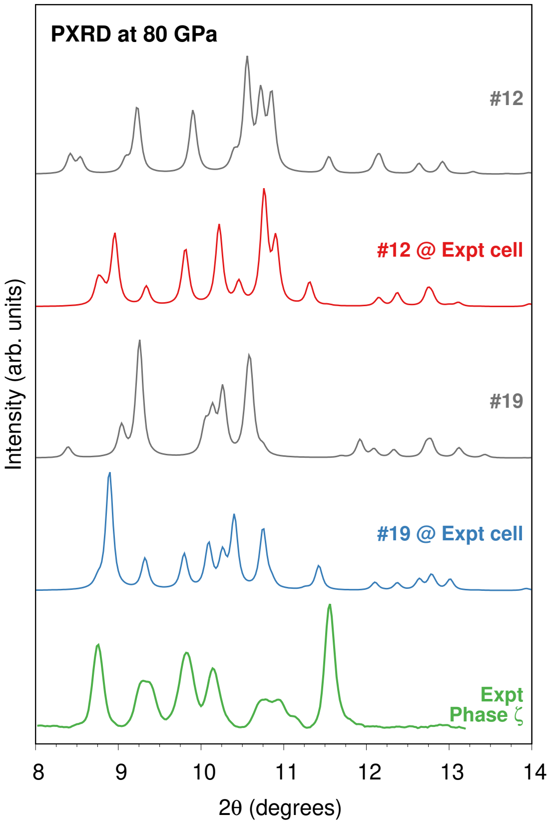


Figure 5.14: Comparison of the simulated powder X-ray diffraction patterns of structures #12 and #19 before and after constraining the lattice constants to match the experimentally reported values. All spectra employ a wavelength of $\lambda = 0.3683$ Å.

Given the apparent promise of structures #12 (based on lattice parameters) and #19 (lattice constants, space group, PXRD, and Raman), we examine the impact of constraining the lattice constants of those structures to equal the experimental lattice constants reported at 80 GPa and then relaxing the atomic positions. Figure 5.14 shows that constraining the lattice constants in this manner improves the agreement with the experimental peak positions considerably, as one would expect, though the intensity patterns still differ somewhat (especially for the experimental peak near 11.5°). Of the two structures, Structure #19 gives the closer match for the experimental PXRD spectrum.

As an aside, we note that the structure of κ nitrogen is also unknown, but preliminary fitting to the experimental diffraction data at 130 GPa suggested a monoclinic cell with lattice parameters a=6.92 Å, b=6.20 Å, c=2.29 Å, and $\beta = 91.8°$.[23] Because these values are also quite similar to the constants predicted for structures #12 and #19, analogous constrained optimizations were performed on these two structures with the κ lattice parameters at 130 GPa. As described in Section 5.3.6 placing structure #12 in the experimentally reported κ lattice parameters and relaxing it altered the structure only slightly and produced a simulated x-ray diffraction pattern with some similarities to the experimental data. More experimental data would be helpful to assess the structure further, but structure #12 could be worth pursuing for phase κ as well.

5.3.9 Predicted enthalpies

Finally, the pressure-dependent stabilities of these candidate phases relative to several experimentally known phases are considered. Experimentally, the ζ phase is formed by compressing ϵ N_2 above ~60 GPa at room temperature, or at 25 GPa at low temperatures.[23,24,111,112,156] Around 110–130 GPa and 2000 K, ζ nitrogen transforms to the polymeric cubic gauche (cg) phase.[150? ?] The λ phase[18] has proved more difficult to form kinetically, but it overlaps with ϵ, ζ and several other phases and is believed to be more stable than many of those throughout much of the lower-pressure regions of the phase diagram.[18,155]

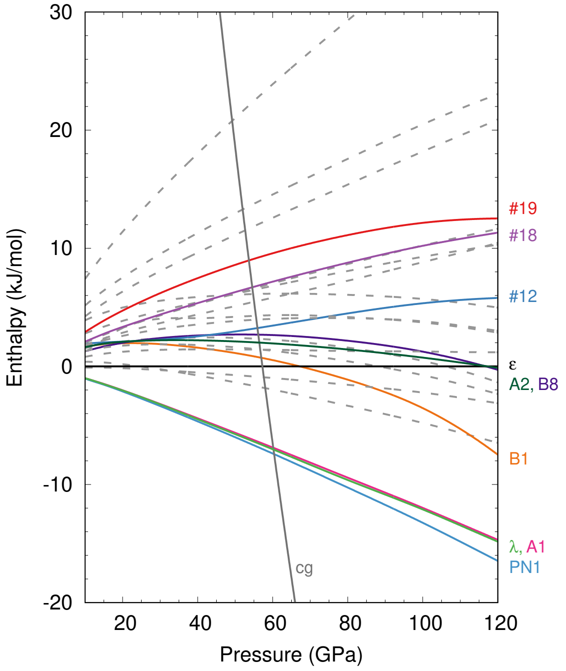


Figure 5.15: DFT enthalpies versus pressure for the predicted candidate structures and several experimentally known phases. Colored lines correspond to the key structures; dashed gray lines correspond to other predicted structures from the CSP landscape which are not discussed in detail.

Figure 5.15 compares the B86bPBE-XDM enthalpies of the candidate structures predicted here and several experimental phases relative to phase ϵ. These enthalpies lack the vibrational free energy contributions that impact phase stability, but they can still provide helpful insights into the stability regimes of the different phases. Typical vibrational free energy contributions to polymorph energy differences rarely exceed 2 kJ/mol for organic crystals at ambient conditions, for example.[136]

Several features are notable in Figure 5.15. First, λ is indeed found to be more stable than ϵ throughout for all pressures considered here. Moreover, structures A1 and PN1 are also more stable than the ϵ phase over the 10–120 GPa range, and they become increasingly stabilized at higher pressures. So while A1 and PN1 are not experimentally known and are spectroscopically inconsistent with ζ nitrogen, their excellent stability raises the prospect that they might be realized experimentally in the future.

Second, the polymeric *cg* phase becomes more stable than ϵ around 60 GPa. This predicted transition pressure is consistent with earlier theoretical studies,[42,116,137? ? ,138] but it is inconsistent with the experimental transition happening around ∼110–130 GPa.[150? ?] This discrepancy likely arises from a mixture of factors: (1) The entropic effects neglected here stabilize the molecular phases considerably (the entropic difference between molecular and non-molecular phases can be much larger than between two different molecular phases).[116] Their inclusion would likely shift the predicted thermodynamic transition point to somewhat higher pressures. (2) The experimentally observed transition temperature and pressure likely reflect the conditions necessary to overcome the kinetic activation barrier, as evidenced by the fact that the transition pressure varies with the heating method.[150? ?]

Third, Structures #12, #18, and #19 become increasingly less stable relative to ϵ nitrogen as the pressure is increased to 120 GPa (though the slope of the enthalpy curve suggests that Structure #19 might begin stabilizing relative to ϵ shortly beyond 120 GPa). If one of those structures did correspond to the ζ phase, there would not be any obvious thermodynamic preference to drive the $\epsilon \rightarrow \zeta$ transition near the 25–60 GPa experimental phase transition pressures (depending on temperature). It seems unlikely that the neglected vibrational free energy effects would dramatically reduce the stability difference.

Fourth, B1, A2, and B8 are the only candidates discussed above which become more stable than ϵ at high pressures. For B1, this occurs around 70 GPa, while for A2 and B8 it occurs just before 120 GPa. These latter two transitions occur well after the regime where *cg* becomes thermodynamically preferred, though again, that may not be a problem if the experimental transition to the *cg* phase is kinetically controlled. On the other hand, they also occur well after the experimental $\epsilon \rightarrow \zeta$ transition. However, stabilizing those forms by ∼2–3 kJ/mol relative to ϵ would be sufficient to bring the predicted $\epsilon \rightarrow \zeta$ transition back to the appropriate pressure regime. That amount is plausibly within the errors one might expect from the combination of GGA-type DFT functionals and neglecting entropic contributions.

Given all these results, what is the structure of ζ nitrogen? Several of the structures considered here exhibit features that might make them viable candidates. At the same time, each displays some inconsistencies versus experiment that argue against that candidate being the correct structure. PN1 and A1 are clearly the most stable structures identified here (and in previous work), and they have not been ascribed to any experimental phase. However, their spectroscopic properties are clearly inconsistent with experimental observations for the ζ phase. Structures B1 and A2 similarly fail to reproduce the spectroscopic observations. Structure #12 agrees well with the experimental lattice constants, though its Raman spectrum in particular appears to be a poor match for experiment. It is also considerably less stable than ϵ nitrogen. Structure #18 has some Raman features that are consistent with experiment, but the powder X-ray diffraction spectrum and poor stability argue against it. None of these structures appears to be an appropriate candidate.

Structures B8 and #19 are the best two candidates considered here. Both exhibit plausible simulated X-ray diffraction data and share several features in common with the experimental Raman spectrum (though they each exhibit discrepancies with the Raman spectrum as well). The lattice constants for Structure #19 are in good agreement with experiment, and its $P2_12_12$ symmetry matches one of space groups suggested in earlier work on the ζ phase.[111] On the other hand, the enthalpy of #19 is far higher than one expects for a viable structure. Even if entropy stabilizes structure #19 relative to the ϵ phase, it would likely be too small of a contribution to achieve a phase transition in the relevant pressure regime. B8 is considerably more stable than #19 and even eventually becomes more stable than ϵ at high pressure, though the predicted transition pressure appears to be unrealistically high. Accepting the B8 structure would also require an alternative explanation for the low-frequency Raman bands in the ~100–200 cm^{-1} region that are absent in its simulated spectrum.

5.4 Conclusions

The AIRSS crystal structure prediction approach was employed to search for crystal structures that are consistent with the experimentally reported structural and spectroscopic data for the ζ phase of nitrogen. The structure search did not definitively identify the structure of the ζ phase, but candidates such as PN1, A1, A2, and B1 that have been suggested previously and a number of new structures predicted here can be ruled out based on their spectroscopic properties. The most promising candidates examined here are B8 and #19, though neither gives a perfect match for the experimentally observed properties. On the other hand, perhaps these two candidates would provide promising starting points for refining the experimental x-ray diffraction data. That strategy proved helpful in solving the structure of λ nitrogen, for example.[18] The possible connection between structure #12 and the κ phase of nitrogen was also raised, though too little experimental data is available to draw firm conclusions.

If none of these candidates can account for the entire body of experimental evidence, then what is the true ζ structure? After multiple independent structure prediction searches in the literature, it seems unlikely that there exists another low-energy orthorhombic structure with $Z \leq 8$ that has been missed. The possibility of an incorrect interpretation of the experimental diffraction data cannot be ruled out, though that data too has already been examined in a number of studies.[23,111,112] Despite reaching no definitive structural determination for phase ζ, the ability to rule out a number of proposed candidates here demonstrates once again how using the combination of crystal structure prediction and simulated spectroscopy is far more powerful than pure crystal structure prediction alone.

Chapter 6

Conclusions

In this dissertation, we presented the computational approach to study high-pressure polymorphs by combining crystal structure prediction with spectroscopy. We demonstrated its usefulness to solve chemical problems for solid carbon dioxide and nitrogen. HMBI provided high-accuracy for energy and able to accurately reproduce experimentally observed structural, mechanical, and spectroscopic properties for number of known-phase polymorphs . With distinct spectroscopic characterization, we were able to suggestion that carbon dioxide phase III and VII are identical. HMBI can also be used for energy ranking as it could search and confirmed crystal structure of nitrogen phase λ. For phase ζ, we were able to eliminate some candidate structures and propose new ones, which could help lead to the eventual discovery of the true structure.

On the methodological side, HMBI is a valuable tool to study crystal structures, but there is room to improve its robustness. First, we have many-body treatment issues for structure refinement approach. The structure refinement accelerates structural prediction

However, the many-body treatment wasn't accurate enough to describe across pressure range resulting in inaccuracy of prediction at high pressure regime. The AMOEBA many-body treatment was typically parameterized at ambient pressure and wasn't able to give correct interaction description at high pressure. On the other hand, all *ab initio* method could be independent to external pressure. DFT with dispersion correction appeared to overbind lattice energy leading to under predict crystal volume at high pressure. Periodic-HF gave the best prediction but we could not converge the self-consistent field equations in large basis sets at high pressure. Thus, the periodic HF many-body treatment in this work was performed by extrapolation of the enthalpy-volume curves toward high pressures, instead of interpolation. The identification of large basis sets with better convergence at high pressures is needed to improve accuracy.

Second, we need faster Raman prediction. With faster Raman prediction, it may be practical to perform Raman simulation to assist crystal structure prediction for pharmaceutical molecules. Raman intensities are obtained from derivatives of the polarizability tensor. The polarizability derivative calculation is performed via analytical approach. Computing time could be reduced by implementing machine learning scheme to acquire this derivative. Lastly, anharmonicity can be implemented to HMBI to improve spectroscopy analysis. We found a symmetry breaking issue with the key structures during crystal structure prediction for phase ζ nitrogen. The inability to access anharmonicity precisely may cloud the judgement regard to spectroscopy analysis. With this improvement, we could revisit the crystal structure search for nitrogen phase ζ.

For the future work, we should search for crystal structure of solid nitrogen phase κ

and θ. These structures remain unresolved experimentally, and theory could help. However, more experimental spectroscopic data would facilitate this process. We should also continue explore small molecular crystals including H_2, He, H_2O, CH_4, and such. These fundamental crystals are important in astronomy, biology, and environmental science. Their phase diagrams are rich and complex. With nature of the compound, they are hard to study in their solid form experimentally.[163] Majority of work that been done computationally were employed using DFT, which may lack the most important long-range interactions for these molecular systems. For example, hydrogen is the most abundant molecule in the universe. Four experimentally well-established phases including phase I, II, III, and IV. Despite DFT crystal structure search, crystal structures of phases II and III are unknown. Next, we could study flexible molecules by performing crystal structure prediction aiding by Raman spectra analysis. Molecules as flexible as Dalcetrapib, with 10 torsional degree of freedom, are most likely to have competitive crystal energy landscape.[38] Ideally, fast Raman prediction is needed to make the narrow down candidate more efficient. It could help targeting crystal structure for pharmaceutical molecule where the energy ranking are ambiguous.

Bibliography

[1] James Alexander Chisholm and W D S Motherwell. COMPACK: A program for identifying crystal structure similarity using distances. *J. Appl. Crystall.*, 38(1):228–231, jan 2005. ISSN 0021-8898. doi: 10.1107/S0021889804027074. URL `http://scripts.iucr.org/cgi-bin/paper?S0021889804027074`.

[2] K Aoki, H Yamawaki, M Sakashita, Y Gotoh, and K Takemura. Crystal Structure of the High-Pressure Phase of Solid CO_2. *Science*, 263(5145):356–358, 1994. URL `http://www.jstor.org/stable/2882774`.

[3] Jinjin Li, Olaseni Sode, Gregory A Voth, and So Hirata. A solid-solid phase transition in carbon dioxide at high pressures and intermediate temperatures. *Nature Commun.*, 4(May):2647, jan 2013. ISSN 2041-1723. doi: 10.1038/ncomms3647. URL `http://www.ncbi.nlm.nih.gov/pubmed/24145659`.

[4] S A Bonev, F Gygi, T Ogitsu, and G Galli. High-Pressure Molecular Phases of Solid Carbon Dioxide. *Phys. Rev. Lett.*, 91(6):065501, aug 2003. ISSN 0031-9007. doi: 10.1103/PhysRevLett.91.065501. URL `http://link.aps.org/doi/10.1103/PhysRevLett.91.065501`.

[5] R C Hanson and L H Jones. Infrared and Raman studies of pressure effects on the vibrational modes of solid CO_2. *J. Chem. Phys.*, 75(3):1102–1112, aug 1981. ISSN 0021-9606. doi: 10.1063/1.442183. URL `http://aip.scitation.org/doi/10.1063/1.442183`.

[6] H. Olijnyk, H. Däufer, H.-J. Jodl, and H. D. Hochheimer. Effect of pressure and temperature on the Raman spectra of solid CO_2. *J. Chem. Phys.*, 88(7):4204–4212, apr 1988. ISSN 0021-9606. doi: 10.1063/1.453828. URL `http://scitation.aip.org/content/aip/journal/jcp/88/7/10.1063/1.453828http://aip.scitation.org/doi/10.1063/1.453828`.

[7] H. Olijnyk and A. P. Jephcoat. Vibrational studies on CO_2 up to 40 GPa by Raman spectroscopy at room temperature. *Phys. Rev. B*, 57(2):879–888, jan 1998. ISSN 0163-1829. doi: 10.1103/PhysRevB.57.879. URL `https://link.aps.org/doi/10.1103/PhysRevB.57.879`.

[8] Valentina M. Giordano, Frédéric Datchi, Federico A. Gorelli, and Roberto Bini. Equation of state and anharmonicity of carbon dioxide phase I up to 12 GPa and 800 K. *J. Chem. Phys.*, 133(14):144501, 2010. ISSN 00219606. doi: 10.1063/1.3495951.

[9] V. Iota and C. S. Yoo. Phase diagram of carbon dioxide: Evidence for a new associated phase. *Phys. Rev. Lett.*, 86(26 I):5922–5925, 2001. ISSN 00319007. doi: 10.1103/PhysRevLett.86.5922.

[10] C. Yoo, H. Kohlmann, H. Cynn, M. Nicol, V. Iota, and T. LeBihan. Crystal structure of pseudo-six-fold carbon dioxide phase II at high pressures and temperatures. *Phys. Rev. B*, 65(10):1–6, 2002. ISSN 0163-1829. doi: 10.1103/PhysRevB.65.104103.

[11] Frédéric Datchi, Bidyut Mallick, Ashkan Salamat, Gwenaëlle Rousse, Sandra Ninet, Gaston Garbarino, Pierre Bouvier, and Mohamed Mezouar. Structure and compressibility of the high-pressure molecular phase II of carbon dioxide. *Phys. Rev. B*, 89 (14):144101, 2014. ISSN 1550235X. doi: 10.1103/PhysRevB.89.144101.

[12] Bart Olinger. The compression of solid CO_2 at 296 K and 10 GPa. *J. Chem. Phys.*, 77(12):6255–6258, 1982. ISSN 00219606. doi: 10.1063/1.443828.

[13] L. Liu. Compression and phase behavior of solid CO_2 to half a megabar. *Earth Planet. Sci. Lett.*, 71(1):104–110, nov 1984. ISSN 0012821X. doi: 10.1016/0012-821X(84)90056-6. URL `http://linkinghub.elsevier.com/retrieve/pii/0012821X84900566`.

[14] C S Yoo, H Cynn, F Gygi, G Galli, V Iota, M Nicol, S Carlson, D. Häusermann, and C Mailhiot. Crystal Structure of Carbon Dioxide at High Pressure: “Superhard” Polymeric Carbon Dioxide. *Phys. Rev. Lett.*, 83(26):5527–5530, dec 1999. ISSN 0031-9007. doi: 10.1103/PhysRevLett.83.5527. URL `http://link.aps.org/doi/10.1103/PhysRevLett.83.5527`.

[15] P Vinet, J R Smith, J Ferrante, and J H Rose. Temperature effects on the universal equation of state. *Phys. Rev. B*, 35(4):1945–1953, 1987.

[16] Frédéric Datchi, Valentina M. Giordano, Pascal Munsch, and A. Marco Saitta. Structure of Carbon Dioxide Phase IV: Breakdown of the Intermediate Bonding State Scenario. *Phys. Rev. Lett.*, 103(18):185701, 2009. ISSN 00319007. doi: 10.1103/PhysRevLett.103.185701.

[17] V. M Giordano and F Datchi. Molecular carbon dioxide at high pressure and high temperature. *Europhys. Lett.*, 77(4):46002, feb 2007. ISSN 0295-5075. doi: 10.1209/0295-5075/77/46002. URL `http://stacks.iop.org/0295-5075/77/i=4/a=46002?key=crossref.988987a595b1849e39f78f4c59b9f040`.

[18] Mungo Frost, Ross T Howie, Philip Dalladay-Simpson, Alexander F Goncharov, and Eugene Gregoryanz. Novel high-pressure nitrogen phase formed by compres-

sion at low temperature. *Phys. Rev. B*, 93(2):024113, jan 2016. ISSN 2469-9950. doi: 10.1103/PhysRevB.93.024113. URL `http://link.aps.org/doi/10.1103/PhysRevB.93.024113`.

[19] R W G Wyckoff. *Crystal Structures - Volume 1.* Interscience Publishers, New York, 1963.

[20] A. F. Schuch and R. L. Mills. Crystal Structures of the Three Modifications of Nitrogen 14 and Nitrogen 15 at High Pressure. *J. Chem. Phys.*, 52(12):6000–6008, jun 1970. ISSN 0021-9606. doi: 10.1063/1.1672899. URL `http://aip.scitation.org/doi/10.1063/1.1672899`.

[21] R. L. Mills, Bart Olinger, and D. T. Cromer. Structures and phase diagrams of N $_2$ and CO to 13 GPa by x-ray diffraction. *J. Chem. Phys.*, 84(5):2837–2845, mar 1986. ISSN 0021-9606. doi: 10.1063/1.450310. URL `http://aip.scitation.org/doi/10.1063/1.450310`.

[22] Robin Turnbull, Michael Hanfland, Jack Binns, Miguel Martinez-Canales, Mungo Frost, Miriam Marqués, Ross T. Howie, and Eugene Gregoryanz. Unusually complex phase of dense nitrogen at extreme conditions. *Nature Commun.*, 9(1):4717, dec 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-07074-4. URL `http://www.nature.com/articles/s41467-018-07074-4`.

[23] Eugene Gregoryanz, Alexander F. Goncharov, Chrystele Sanloup, Maddury Somayazulu, Ho-kwang Mao, and Russell J. Hemley. High P-T transformations of nitrogen to 170GPa. *J. Chem. Phys.*, 126(18):184505, may 2007. ISSN 0021-9606. doi: 10.1063/1.2723069. URL `http://aip.scitation.org/doi/10.1063/1.2723069`.

[24] Roberto Bini, Lorenzo Ulivi, Jörg Kreutz, and Hans J. Jodl. High-pressure phases of solid nitrogen by Raman and infrared spectroscopy. *J. Chem. Phys.*, 112(19):8522–8529, may 2000. ISSN 0021-9606. doi: 10.1063/1.481455. URL `http://scitation.aip.org/content/aip/journal/jcp/112/19/10.1063/1.481455http://aip.scitation.org/doi/10.1063/1.481455`.

[25] Yoshikazu Mizuguchi, Fumiaki Tomioka, Shunsuke Tsuda, Takahide Yamaguchi, and Yoshihiko Takano. Superconductivity at 27K in tetragonal FeSe under high pressure. *Applied Physics Letters*, 93(15):152505, 2008. doi: 10.1063/1.3000616. URL `https://doi.org/10.1063/1.3000616`.

[26] 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*, 525(7567):73–76, 2015. ISSN 1476-4687. doi: 10.1038/nature14964. URL `https://doi.org/10.1038/nature14964`.

[27] A P Drozdov, P P Kong, V S Minkov, S P Besedin, M A Kuzovnikov, S Mozaffari,

L Balicas, F F Balakirev, D E Graf, V B Prakapenka, E Greenberg, D A Knyazev, M Tkacz, and M I Eremets. Superconductivity at 250 K in lanthanum hydride under high pressures. *Nature*, 569(7757):528–531, 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1201-8. URL `https://doi.org/10.1038/s41586-019-1201-8`.

[28] José A Flores-Livas, Lilia Boeri, Antonio Sanna, Gianni Profeta, Ryotaro Arita, and Mikhail Eremets. A perspective on conventional high-temperature superconductors at high pressure: Methods and materials. *Physics Reports*, 856:1–78, 2020. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2020.02.003. URL `http://www.sciencedirect.com/science/article/pii/S0370157320300363`.

[29] Nobuhiko Takeichi, Hiroshi Senoh, Tomoyuki Yokota, Hidekazu Tsuruta, Kenjiro Hamada, Hiroyuki T Takeshita, Hideaki Tanaka, Tetsu Kiyobayashi, Toshio Takano, and Nobuhiro Kuriyama. “Hybrid hydrogen storage vessel”, a novel high-pressure hydrogen storage vessel combined with hydrogen storage material. *International Journal of Hydrogen Energy*, 28(10):1121–1129, 2003. ISSN 0360-3199. doi: https://doi.org/10.1016/S0360-3199(02)00216-1. URL `http://www.sciencedirect.com/science/article/pii/S0360319902002161`.

[30] Hirotomo Nishihara, Peng-Xiang Hou, Li-Xiang Li, Masashi Ito, Makoto Uchiyama, Tomohiro Kaburagi, Ami Ikura, Junji Katamura, Takayuki Kawarada, Kazuhiko Mizuuchi, and Takashi Kyotani. High-Pressure Hydrogen Storage in Zeolite-Templated Carbon. *The Journal of Physical Chemistry C*, 113(8):3189–3196, feb 2009. ISSN 1932-7447. doi: 10.1021/jp808890x. URL `https://doi.org/10.1021/jp808890x`.

[31] Yury Gogotsi, Cristelle Portet, Sebastian Osswald, Jason M Simmons, Taner Yildirim, Giovanna Laudisio, and John E Fischer. Importance of pore size in high-pressure hydrogen storage by porous carbons. *International Journal of Hydrogen Energy*, 34(15):6314–6319, 2009. ISSN 0360-3199. doi: https://doi.org/10.1016/j.ijhydene.2009.05.073. URL `http://www.sciencedirect.com/science/article/pii/S0360319909008015`.

[32] Jinyang Zheng, Xianxin Liu, Ping Xu, Pengfei Liu, Yongzhi Zhao, and Jian Yang. Development of high pressure gaseous hydrogen storage technologies. *International Journal of Hydrogen Energy*, 37(1):1048–1057, 2012. ISSN 0360-3199. doi: https://doi.org/10.1016/j.ijhydene.2011.02.125. URL `http://www.sciencedirect.com/science/article/pii/S0360319911005283`.

[33] F J Manjón, R Vilaplana, O Gomis, E Pérez-González, D Santamaría-Pérez, V Marín-Borrás, A Segura, J González, P Rodríguez-Hernández, A Muñoz, C Drasar, V Kucek, and V Muñoz-Sanjosé. High-pressure studies of topological insulators Bi2Se3, Bi2Te3, and Sb2Te3. *physica status solidi (b)*, 250(4):669–676, apr 2013. ISSN 0370-1972. doi: 10.1002/pssb.201200672. URL `https://doi.org/10.1002/pssb.201200672`.

[34] Ranga P Dias and Isaac F Silvera. Observation of the Wigner-Huntington transition to metallic hydrogen. *Science*, 355(6326):715 LP – 718, feb 2017. doi: 10.1126/science.aal1579. URL `http://science.sciencemag.org/content/355/6326/715.abstract`.

[35] Sanjay R Chemburkar, John Bauer, Kris Deming, Harry Spiwek, Ketan Patel, John Morris, Rodger Henry, Stephen Spanton, Walter Dziki, William Porter, John Quick, Phil Bauer, John Donaubauer, B A Narayanan, Mauro Soldani, Dave Riley, and Kathyrn McFarland. Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development. *Organic Process Research & Development*, 4(5):413–417, sep 2000. ISSN 1083-6160. doi: 10.1021/op000023y. URL `https://doi.org/10.1021/op000023y`.

[36] John Bauer, Stephen Spanton, Rodger Henry, John Quick, Walter Dziki, William Porter, and John Morris. Ritonavir: An Extraordinary Example of Conformational Polymorphism. *Pharmaceutical Research*, 18(6):859–866, 2001. ISSN 1573-904X. doi: 10.1023/A:1011052932607. URL `https://doi.org/10.1023/A:1011052932607`.

[37] Sherry L Morissette, Stephen Soukasene, Douglas Levinson, Michael J Cima, and Örn Almarsson. Elucidation of crystal form diversity of the HIV protease inhibitor ritonavir by high-throughput crystallization. *Proceedings of the National Academy of Sciences*, 100(5):2180 LP – 2184, mar 2003. doi: 10.1073/pnas.0437744100. URL `http://www.pnas.org/content/100/5/2180.abstract`.

[38] M A Neumann, J van de Streek, F P A Fabbiani, P Hidber, and O Grassmann. Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. *Nature Communications*, 6(1):7793, 2015. ISSN 2041-1723. doi: 10.1038/ncomms8793. URL `https://doi.org/10.1038/ncomms8793`.

[39] Rajni M Bhardwaj, Jennifer A McMahon, Jonas Nyman, Louise S Price, Sumit Konar, Iain D H Oswald, Colin R Pulham, Sarah L Price, and Susan M Reutzel-Edens. A Prolific Solvate Former, Galunisertib, under the Pressure of Crystal Structure Prediction, Produces Ten Diverse Polymorphs. *Journal of the American Chemical Society*, 141(35):13887–13897, sep 2019. ISSN 0002-7863. doi: 10.1021/jacs.9b06634. URL `https://doi.org/10.1021/jacs.9b06634`.

[40] Artem R Oganov. Crystal structure prediction: reflections on present status and challenges. *Faraday Discussions*, 211(0):643–660, 2018. ISSN 1359-6640. doi: 10.1039/C8FD90033G. URL `http://dx.doi.org/10.1039/C8FD90033G`.

[41] Barak Hirshberg, R Benny Gerber, and Anna I Krylov. Calculations predict a stable molecular crystal of N8. *Nature Chemistry*, 6(1):52–56, 2014. ISSN 1755-4349. doi: 10.1038/nchem.1818. URL `https://doi.org/10.1038/nchem.1818`.

[42] Chris J. Pickard and R. J. Needs. High-Pressure Phases of Nitrogen. *Phys. Rev. Lett.*, 102(12):125702, 2009. ISSN 0031-9007. doi: 10.1103/PhysRevLett.102.125702. URL `http://link.aps.org/doi/10.1103/PhysRevLett.102.125702`.

[43] Sarah L Price. The computational prediction of pharmaceutical crystal structures and polymorphism. *Advanced Drug Delivery Reviews*, 56(3):301–319, 2004. ISSN 0169-409X. doi: https://doi.org/10.1016/j.addr.2003.10.006. URL `http://www.sciencedirect.com/science/article/pii/S0169409X03002199`.

[44] G.J.O. Beran. Modeling Polymorphic Molecular Crystals with Electronic Structure Theory. *Chem. Rev.*, 116(9), 2016. ISSN 15206890. doi: 10.1021/acs.chemrev.5b00648.

[45] Kamil F. Dziubek, Martin Ende, Demetrio Scelta, Roberto Bini, Mohamed Mezouar, Gaston Garbarino, and Ronald Miletich. Crystalline polymeric carbon dioxide stable at megabar pressures. *Nature Communications*, 9(1):5–10, 2018. ISSN 20411723. doi: 10.1038/s41467-018-05593-8. URL `http://dx.doi.org/10.1038/s41467-018-05593-8`.

[46] Jeffrey M McMahon. Ground-state structures of ice at high pressures from ab initio random structure searching. *Physical Review B*, 84(22):220104, dec 2011. doi: 10.1103/PhysRevB.84.220104. URL `https://link.aps.org/doi/10.1103/PhysRevB.84.220104`.

[47] Tzu-Jen Lin, Cheng-Rong Hsing, Ching-Ming Wei, and Jer-Lai Kuo. Structure prediction of the solid forms of methanol: an ab initio random structure searching approach. *Physical Chemistry Chemical Physics*, 18(4):2736–2746, 2016. ISSN 1463-9076. doi: 10.1039/C5CP06583F. URL `http://dx.doi.org/10.1039/C5CP06583F`.

[48] Vincenzo Barone, Maurizio Casarin, Daniel Forrer, Michele Pavone, Mauro Sambi, and Andrea Vittadini. Role and effective treatment of dispersive forces in materials: Polyethylene and graphite crystals as test cases. *Journal of Computational Chemistry*, 30(6):934–939, apr 2009. ISSN 0192-8651. doi: 10.1002/jcc.21112. URL `https://doi.org/10.1002/jcc.21112`.

[49] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics*, 132(15):154104, apr 2010. ISSN 0021-9606. doi: 10.1063/1.3382344. URL `https://doi.org/10.1063/1.3382344`.

[50] Stefan Grimme, Andreas Hansen, Jan Gerit Brandenburg, and Christoph Bannwarth. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chemical Reviews*, 116(9):5105–5154, may 2016. ISSN 0009-2665. doi: 10.1021/acs.chemrev.5b00533. URL `https://doi.org/10.1021/acs.chemrev.5b00533`.

[51] Eike Caldeweyher, Christoph Bannwarth, and Stefan Grimme. Extension of the D3 dispersion coefficient model. *The Journal of Chemical Physics*, 147(3):34112, jul 2017. ISSN 0021-9606. doi: 10.1063/1.4993215. URL `https://doi.org/10.1063/1.4993215`.

[52] A Otero-de-la Roza and Erin R Johnson. Many-body dispersion interactions from the exchange-hole dipole moment model. *The Journal of Chemical Physics*, 138(5):54103, feb 2013. ISSN 0021-9606. doi: 10.1063/1.4789421. URL `https://doi.org/10.1063/1.4789421`.

[53] Alexandre Tkatchenko and Matthias Scheffler. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Physical Review Letters*, 102(7):73005, feb 2009. doi: 10.1103/PhysRevLett.102.073005. URL `https://link.aps.org/doi/10.1103/PhysRevLett.102.073005`.

[54] Anthony M Reilly and Alexandre Tkatchenko. van der Waals dispersion interactions in molecular materials: beyond pairwise additivity. *Chemical Science*, 6(6):3289–3301, 2015. ISSN 2041-6520. doi: 10.1039/C5SC00410A. URL `http://dx.doi.org/10.1039/C5SC00410A`.

[55] Robert A DiStasio, O Anatole von Lilienfeld, and Alexandre Tkatchenko. Collective many-body van der Waals interactions in molecular systems. *Proceedings of the National Academy of Sciences*, 109(37):14791 LP – 14795, sep 2012. doi: 10.1073/pnas.1208121109. URL `http://www.pnas.org/content/109/37/14791.abstract`.

[56] Alexandre Tkatchenko, Robert A DiStasio, Roberto Car, and Matthias Scheffler. Accurate and Efficient Method for Many-Body van der Waals Interactions. *Physical Review Letters*, 108(23):236402, jun 2012. doi: 10.1103/PhysRevLett.108.236402. URL `https://link.aps.org/doi/10.1103/PhysRevLett.108.236402`.

[57] Alberto Ambrosetti, Anthony M Reilly, Robert A DiStasio, and Alexandre Tkatchenko. Long-range correlation energy calculated from coupled atomic response functions. *The Journal of Chemical Physics*, 140(18):18A508, feb 2014. ISSN 0021-9606. doi: 10.1063/1.4865104. URL `https://doi.org/10.1063/1.4865104`.

[58] W Kohn and L J Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 140(4A):A1133–A1138, nov 1965. doi: 10.1103/PhysRev.140.A1133. URL `https://link.aps.org/doi/10.1103/PhysRev.140.A1133`.

[59] Jiří Klimeš, David R Bowler, and Angelos Michaelides. Van der Waals density functionals applied to solids. *Physical Review B*, 83(19):195131, may 2011. doi: 10.1103/PhysRevB.83.195131. URL `https://link.aps.org/doi/10.1103/PhysRevB.83.195131`.

[60] Lars Goerigk and Stefan Grimme. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Physical Chemistry Chemical Physics*, 13(14):6670–6688, 2011. ISSN 1463-9076. doi: 10.1039/C0CP02984J. URL `http://dx.doi.org/10.1039/C0CP02984J`.

[61] Yan Zhao and Donald G Truhlar. Benchmark Databases for Nonbonded Interactions and Their Use To Test Density Functional Theory. *Journal of Chemical Theory and Computation*, 1(3):415–432, may 2005. ISSN 1549-9618. doi: 10.1021/ct049851d. URL `https://doi.org/10.1021/ct049851d`.

[62] Sajesh P Thomas and Mark A Spackman. The Polymorphs of ROY: A Computational Study of Lattice Energies and Conformational Energy Differences*. *Australian Journal of Chemistry*, 71(4):279–284, 2018. URL `https://doi.org/10.1071/CH17620`.

[63] Jonas Nyman, Lian Yu, and Susan M Reutzel-Edens. Accuracy and reproducibility in crystal structure prediction: the curious case of ROY. *CrystEngComm*, 21(13): 2080–2088, 2019. doi: 10.1039/C8CE01902A. URL `http://dx.doi.org/10.1039/C8CE01902A`.

[64] S. Wen and G.J.O. Beran. Crystal polymorphism in oxalyl dihydrazide: Is empirical DFT-D accurate enough? *J. Chem. Theory Comput.*, 8(8):2698–2705, 2012. ISSN 15499618 15499626. doi: 10.1021/ct300484h.

[65] S. Wen and G.J.O. Beran. Accidental degeneracy in crystalline aspirin: New insights from high-level ab initio calculations. *Cryst. Growth Des.*, 12(5):2169–2172, 2012. ISSN 15287483 15287505. doi: 10.1021/cg300358n.

[66] Ctirad Červinka and Gregory J O Beran. Ab Initio Thermodynamic Properties and Their Uncertainties for Crystalline α-Methanol. *Phys. Chem. Chem. Phys.*, pages 7–9, 2017. ISSN 1463-9076. doi: 10.1039/C7CP06605H. URL `http://pubs.rsc.org/en/Content/ArticleLanding/2017/CP/C7CP06605H`.

[67] Ctirad Červinka and Gregory J. O. Beran. Ab initio prediction of the polymorph phase diagram for crystalline methanol. *Chem. Sci.*, pages 4622–4629, 2018. ISSN 2041-6520. doi: 10.1039/C8SC01237G. URL `http://xlink.rsc.org/?DOI=C8SC01237G`.

[68] Elodie Salager, Graeme M Day, Robin S Stein, Chris J Pickard, Bénédicte Elena, and Lyndon Emsley. Powder Crystallography by Combined Crystal Structure Prediction and High-Resolution 1 H Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.*, 132 (8):2564–2566, mar 2010. ISSN 0002-7863. doi: 10.1021/ja909449k. URL `http://pubs.acs.org/doi/abs/10.1021/ja909449k`.

[69] Charlotte Martineau, Jürgen Senker, and Francis Taulelle. Chapter One - NMR Crystallography. volume 82, pages 1–57. Academic Press, 2014. ISBN 0066-

4103. doi: https://doi.org/10.1016/B978-0-12-800184-4.00001-1. URL `http://www.sciencedirect.com/science/article/pii/B9780128001844000011`.

[70] Mark D Eddleston, Katarzyna E Hejczyk, Erica G Bithell, Graeme M Day, and William Jones. Polymorph identification and crystal structure determination by a combined crystal structure prediction and transmission electron microscopy approach. *Chem. Eur. J.*, 19(24):7874–82, jun 2013. ISSN 1521-3765. doi: 10.1002/chem.201204368. URL `http://www.ncbi.nlm.nih.gov/pubmed/23592444`.

[71] So Hirata, Kandis Gilliard, Xiao He, Jinjin Li, and Olaseni Sode. Ab initio molecular crystal structures, spectra, and phase diagrams. *Acc. Chem. Res.*, 47(9):2721–30, sep 2014. ISSN 1520-4898. doi: 10.1021/ar500041m. URL `http://www.ncbi.nlm.nih.gov/pubmed/24754304`.

[72] Xiao He, Olaseni Sode, Sotiris S Xantheas, and So Hirata. Second-order many-body perturbation study of ice Ih. *J. Chem. Phys.*, 137(20):204505, nov 2012. ISSN 1089-7690. doi: 10.1063/1.4767898. URL `http://www.ncbi.nlm.nih.gov/pubmed/23206017`.

[73] Kandis Gilliard, Olaseni Sode, and So Hirata. Second-order many-body perturbation and coupled-cluster singles and doubles study of ice VIII. *J. Chem. Phys.*, 140(17):174507, may 2014. ISSN 1089-7690. doi: 10.1063/1.4873919. URL `http://www.ncbi.nlm.nih.gov/pubmed/24811646`.

[74] Doris E. Braun, Arianna Rivalta, Andrea Giunchi, Natalia Bedoya-Martinez, Benedikt Schrode, Elisabetta Venuti, Raffaele Guido Della Valle, and Oliver Werzer. Surface Induced Phenytoin Polymorph. 1. Full Structure Solution by Combining Grazing Incidence X-ray Diffraction and Crystal Structure Prediction. *Crystal Growth and Design*, 19(11):6058–6066, nov 2019. ISSN 15287505. doi: 10.1021/acs.cgd.9b00857.

[75] Natalia Bedoya-Martínez, Benedikt Schrode, Andrew O.F. Jones, Tommaso Salzillo, Christian Ruzié, Nicola Demitri, Yves H. Geerts, Elisabetta Venuti, Raffaele Guido Della Valle, Egbert Zojer, and Roland Resel. DFT-Assisted Polymorph Identification from Lattice Raman Fingerprinting. *Journal of Physical Chemistry Letters*, 8(15), 2017. ISSN 19487185. doi: 10.1021/acs.jpclett.7b01634.

[76] Gregory J O Beran. Approximating quantum many-body intermolecular interactions in molecular clusters using classical polarizable force elds. *J. Chem. Phys.*, 130:164115, 2009. doi: 10.1063/1.3121323.

[77] Gregory J. O. Beran and Kaushik Nanda. Predicting Organic Crystal Lattice Energies with Chemical Accuracy. *J. Phys. Chem. C*, 1:3480–3487, dec 2010. ISSN 1948-7185. doi: 10.1021/jz101383z. URL `http://dx.doi.org/10.1021/jz101383z`.

[78] Gregory J O Beran, Joshua D. Hartman, and Yonaton N. Heit. Predicting Molecular

Crystal Properties from First Principles: Finite-Temperature Thermochemistry to NMR Crystallography. *Accounts of Chemical Research*, 49(11):2501–2508, 2016. ISSN 15204898. doi: 10.1021/acs.accounts.6b00404.

[79] Pengyu Ren, Chuanjie Wu, and Jay W Ponder. Polarizable Atomic Multipole-based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.*, 7(10):3143–3161, oct 2011. ISSN 1549-9626. doi: 10.1021/ct200304d. URL `http://www.ncbi.nlm.nih.gov/pubmed/22022236`.

[80] Yonaton N. Heit and Gregory J.O. Beran. How important is thermal expansion for predicting molecular crystal structures and thermochemistry at finite temperatures? *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 72(4):514–529, 2016. ISSN 20525206. doi: 10.1107/S2052520616005382.

[81] Yonaton N. Heit, Kaushik D. Nanda, and Gregory J. O. Beran. Predicting finite-temperature properties of crystalline carbon dioxide from first principles with quantitative accuracy. *Chem. Sci.*, 7:246–255, 2016. ISSN 2041-6520. doi: 10.1039/C5SC03014E. URL `http://xlink.rsc.org/?DOI=C5SC03014E`.

[82] Jessica L Mckinley and Gregory J O Beran. Faraday Discussions Identifying pragmatic quasi-harmonic electronic struc- ture approaches for modeling molecular crystal thermal expansion. 2018. doi: 10.1039/C8FD00048D. URL `http://pubs.rsc.org/en/content/articlepdf/2018/fd/c8fd00048d`.

[83] C. S. Yoo. Physical and chemical transformations of highly compressed carbon dioxide at bond energies. *Phys. Chem. Chem. Phys.*, 15(21):7949–66, 2013. ISSN 1463-9084. doi: 10.1039/c3cp50761k. URL `http://www.ncbi.nlm.nih.gov/pubmed/23615853`.

[84] J.-H. Park, C. S. Yoo, V. Iota, H. Cynn, M. F. Nicol, and T. Le Bihan. Crystal structure of bent carbon dioxide phase IV. *Phys. Rev. B*, 68(1):014107, jul 2003. ISSN 0163-1829. doi: 10.1103/PhysRevB.68.014107. URL `http://link.aps.org/doi/10.1103/PhysRevB.68.014107`.

[85] Sebastian Gohr, Stefan Grimme, Tilo Söhnel, Beate Paulus, and Peter Schwerdtfeger. Pressure dependent stability and structure of carbon dioxide-A density functional study including long-range corrections. *J. Chem. Phys.*, 139(17):174501, 2013. ISSN 1089-7690. doi: 10.1063/1.4826929. URL `http://www.ncbi.nlm.nih.gov/pubmed/24206310`.

[86] Choong-Shik Yoo, Amartya Sengupta, and Minseob Kim. Phase diagram of carbon dioxide: update and challenges. *High Press. Res.*, 31(1):68–74, 2011. ISSN 0895-7959. doi: 10.1080/08957959.2010.523000.

[87] Anthony M. Reilly, Richard I. Cooper, Claire S. Adjiman, Saswata Bhattacharya, A. Daniel Boese, Jan Gerit Brandenburg, Peter J. Bygrave, Rita Bylsma, Josh E.

Campbell, Roberto Car, David H. Case, Renu Chadha, Jason C. Cole, Katherine Cosburn, Herma M. Cuppen, Farren Curtis, Graeme M. Day, Robert A. DiStasio Jr, Alexander Dzyabchenko, Bouke P. van Eijck, Dennis M. Elking, Joost A. van den Ende, Julio C. Facelli, Marta B. Ferraro, Laszlo Fusti-Molnar, Christina-Anna Gatsiou, Thomas S. Gee, René de Gelder, Luca M. Ghiringhelli, Hitoshi Goto, Stefan Grimme, Rui Guo, Detlef W. M. Hofmann, Johannes Hoja, Rebecca K. Hylton, Luca Iuzzolino, Wojciech Jankiewicz, Daniël T. de Jong, John Kendrick, Niek J. J. de Klerk, Hsin-Yu Ko, Liudmila N. Kuleshova, Xiayue Li, Sanjaya Lohani, Frank J. J. Leusen, Albert M. Lund, Jian Lv, Yanming Ma, Noa Marom, Artëm E. Masunov, Patrick McCabe, David P. McMahon, Hugo Meekes, Michael P. Metz, Alston J. Misquitta, Sharmarke Mohamed, Bartomeu Monserrat, Richard J. Needs, Marcus A. Neumann, Jonas Nyman, Shigeaki Obata, Harald Oberhofer, Artem R. Oganov, Anita M. Orendt, Gabriel I. Pagola, Constantinos C. Pantelides, Chris J. Pickard, Rafal Podeszwa, Louise S. Price, Sarah L. Price, Angeles Pulido, Murray G. Read, Karsten Reuter, Elia Schneider, Christoph Schober, Gregory P. Shields, Pawanpreet Singh, Isaac J. Sugden, Krzysztof Szalewicz, Christopher R. Taylor, Alexandre Tkatchenko, Mark E. Tuckerman, Francesca Vacarro, Manolis Vasileiadis, Alvaro Vazquez-Mayagoitia, Leslie Vogt, Yanchao Wang, Rona E. Watson, Gilles A. de Wijs, Jack Yang, Qiang Zhu, and Colin R. Groom. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Cryst. B*, 72 (4):439–459, aug 2016. ISSN 2052-5206. doi: 10.1107/S2052520616007447. URL `http://scripts.iucr.org/cgi-bin/paper?S2052520616007447`.

[88] Artem R Oganov, J Christian Schön, Martin Jansen, Scott M Woodley, William W Tipton, and Richard G Hennig. First Blind Test of Inorganic Crystal Structure Prediction Methods. In Artem R. Oganov, editor, *Modern Methods of Crystal Structure Prediction*, pages 223–231. Wiley, Weinheim, Germany, 2011. ISBN 9783527409396.

[89] S. Wen and G.J.O. Beran. Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-fly Ab initio force field parametrization. *J. Chem. Theory Comput.*, 7(11):3733–3742, 2011. ISSN 15499618 15499626. doi: 10.1021/ct200541h.

[90] Jay W Ponder, Chuanjie Wu, Pengyu Ren, Vijay S Pande, John D Chodera, Michael J Schnieders, Imran Haque, David L Mobley, Daniel S Lambrecht, Robert A DiStasio, Martin Head-Gordon, Gary N I Clark, Margaret E Johnson, and Teresa Head-Gordon. Current status of the AMOEBA polarizable force field. *J. Phys. Chem. B*, 114(8): 2549–64, mar 2010. ISSN 1520-5207. doi: 10.1021/jp910674d.

[91] Johnny C Wu, Gaurav Chattree, and Pengyu Ren. Automation of AMOEBA polarizable force field parameterization for small molecules. *Theor. Chem. Acc.*, 131(3):1138, feb 2012. ISSN 1432-881X. doi: 10.1007/s00214-012-1138-6. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3322661&tool=pmcentrez&rendertype=abstract`.

[92] MOLPRO, version 2012.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, see http://www.molpro.net.

[93] J. W. Ponder, TINKER v6.3, 2014, `http://dasher.wustl.edu/tinker/`. Accessed August 10, 2015.

[94] Thom H Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, jan 1989. ISSN 0021-9606. doi: 10.1063/1.456153. URL `https://doi.org/10.1063/1.456153`.

[95] Y. Heit and G.J.O. Beran. Exploiting space-group symmetry in fragment-based molecular crystal calculations. *J. Comp. Chem.*, 35(30):2205–2213, 2014. ISSN 1096987X 01928651. doi: 10.1002/jcc.23737.

[96] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision A.1. Gaussian Inc. Wallingford CT 2009.

[97] Qiang Zhu, Artem R. Oganov, Colin W Glass, and Harold T Stokes. Constrained evolutionary algorithm for structure prediction of molecular crystals: methodology and applications. *Acta Cryst. B*, 68:215–226, jun 2012. ISSN 1600-5740. doi: 10.1107/S0108768112017466. URL `http://www.ncbi.nlm.nih.gov/pubmed/22610672`.

[98] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, 118 (45):11225–11236, nov 1996. ISSN 0002-7863. doi: 10.1021/ja9621760. URL

`https://doi.org/10.1021/ja9621760`.

[99] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougoussis, Anton Kokalj, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P Seitsonen, Alexander Smogunov, Paolo Umari, and Renata M Wentzcovitch. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter*, 21(39):395502, 2009. ISSN 0953-8984. doi: 10.1088/0953-8984/21/39/395502. URL `http://dx.doi.org/10.1088/0953-8984/21/39/395502`.

[100] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical Review Letters*, 77(18):3865–3868, oct 1996. doi: 10.1103/PhysRevLett.77.3865. URL `https://link.aps.org/doi/10.1103/PhysRevLett.77.3865`.

[101] Jens Antony and Stefan Grimme. Density functional theory including dispersion corrections for intermolecular interactions in a large benchmark set of biologically relevant molecules. *Phys. Chem. Chem. Phys.*, 8(45):5287, 2006. ISSN 1463-9076. doi: 10.1039/b612585a. URL `http://xlink.rsc.org/?DOI=b612585a`.

[102] Clare F Macrae, Ian J Bruno, James A Chisholm, Paul R Edgington, Patrick McCabe, Elna Pidcock, Lucia Rodriguez-Monge, Robin Taylor, Jacco van de Streek, and Peter A Wood. *Journal of Applied Crystallography*, 41(2):466–470, apr 2008. doi: 10.1107/S0021889807067908. URL `https://doi.org/10.1107/S0021889807067908`.

[103] R. T. Downs and M. S. Somayazulu. Carbon Dioxide at 1.0 GPa. *Acta Cryst. C*, 54(7):897–898, jul 1998. ISSN 0108-2701. doi: 10.1107/S0108270198001140. URL `http://scripts.iucr.org/cgi-bin/paper?S0108270198001140`.

[104] Jinjin Li, Olaseni Sode, and So Hirata. Second-Order Many-Body Perturbation Study on Thermal Expansion of Solid Carbon Dioxide. *J. Chem. Theory Comput.*, 11(1): 224–229, 2015. ISSN 1549-9618. doi: 10.1021/ct500983k. URL `http://pubs.acs.org/doi/abs/10.1021/ct500983k`.

[105] François Gygi. First-principles simulations of organic compounds: Solid CO2 under pressure. *Comp. Mat. Sci.*, 10(1-4):63–66, 1998. ISSN 09270256. doi: 10.1016/S0927-0256(97)00086-4.

[106] Jinjin Li, Olaseni Sode, Gregory A Voth, and So Hirata. Corrigendum: A solid–solid phase transition in carbon dioxide at high pressures and intermediate temperatures. *Nature Commun.*, 6(May):8907, nov 2015. ISSN 2041-1723. doi: 10.1038/

ncomms9907. URL `http://www.nature.com/doifinder/10.1038/ncomms9907`.

[107] Olaseni Sode, Murat Keceli, Kiyoshi Yagi, and So Hirata. Fermi resonance in solid CO_2 under pressure. *J. Chem. Phys.*, 138(7):074501, 2013. ISSN 00219606. doi: 10.1063/1.4790537.

[108] F. Bolduan, H. D. Hochheimer, and H. J. Jodl. High pressure Raman study of solid CS_2. *J. Chem. Phys.*, 84(12):6997–7004, jun 1986. ISSN 0021-9606. doi: 10.1063/1.450817. URL `http://aip.scitation.org/doi/10.1063/1.450817`.

[109] K Aoki, Y Kakudate, M Yoshida, S Usuba, K Tanaka, and S Fujiwara. Raman Scattering Observations of Phase Transitions and Polymerizations in Acetylene at High Pressure. *Sol. State Commun.*, 64(10):1329–1331, 1987.

[110] A. F. Goncharov, E. Gregoryanz, H.-K. Mao, and R. J. Hemley. Vibrational dynamics of solid molecular nitrogen to megabar pressures. *Low Temp. Phys.*, 27(9):866–869, sep 2001. ISSN 1063-777X. doi: 10.1063/1.1414578. URL `http://aip.scitation.org/doi/10.1063/1.1414578`.

[111] M. I. Eremets, A. G. Gavriliuk, N. R. Serebryanaya, I. A. Trojan, D. A. Dzivenko, R. Boehler, H. K. Mao, and R. J. Hemley. Structural transformation of molecular nitrogen to a single-bonded atomic state at high pressures. *J. Chem. Phys.*, 121(22): 11296, 2004. ISSN 00219606. doi: 10.1063/1.1814074. URL `http://scitation.aip.org/content/aip/journal/jcp/121/22/10.1063/1.1814074`.

[112] Eugene Gregoryanz, Chrystèle Sanloup, Roberto Bini, Jörg Kreutz, Hans J Jodl, Maddury Somayazulu, Ho-Kwang Mao, and Russell J Hemley. On the ϵ-ζ transition of nitrogen. *J. Chem. Phys.*, 124(11):116102, mar 2006. ISSN 0021-9606. doi: 10.1063/1.2177247. URL `http://www.ncbi.nlm.nih.gov/pubmed/16555920http://aip.scitation.org/doi/10.1063/1.2177247`.

[113] Eugene Gregoryanz, Alexander F Goncharov, Russell J Hemley, Ho-kwang Mao, Maddury Somayazulu, and Guoyin Shen. Raman, infrared, and x-ray evidence for new phases of nitrogen at high pressures and temperatures. *Phys. Rev. B*, 66 (22):224108, dec 2002. ISSN 0163-1829. doi: 10.1103/PhysRevB.66.224108. URL `https://link.aps.org/doi/10.1103/PhysRevB.66.224108`.

[114] Watit Sontising, Yonaton N. Heit, Jessica L. McKinley, and Gregory J. O. Beran. Theoretical predictions suggest carbon dioxide phases III and VII are identical. *Chem. Sci.*, pages 1–13, 2017. ISSN 2041-6520. doi: 10.1039/C7SC03267F. URL `http://xlink.rsc.org/?DOI=C7SC03267F`.

[115] A. Erba, L. Maschio, S. Salustro, and S. Casassa. A post-Hartree–Fock study of pressure-induced phase transitions in solid nitrogen: The case of the α, γ, and ϵ low-pressure phases. *J. Chem. Phys.*, 134(7):074502, 2011. ISSN 00219606. doi:

10.1063/1.3553206. URL `http://link.aip.org/link/JCPSA6/v134/i7/p074502/s1&Agg=doi`.

[116] A. Erba, L. Maschio, C. Pisani, and S. Casassa. Pressure-induced transitions in solid nitrogen: Role of dispersive interactions. *Phys. Rev. B*, 84(1):012101, jul 2011. ISSN 1098-0121. doi: 10.1103/PhysRevB.84.012101. URL `http://link.aps.org/doi/10.1103/PhysRevB.84.012101`.

[117] Mark S Gordon, Dmitri G Fedorov, Spencer R Pruitt, and L Slipchenko. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.*, 112:632–672, aug 2012. ISSN 1520-6890. doi: 10.1021/cr200093j. URL `http://www.ncbi.nlm.nih.gov/pubmed/21866983`.

[118] Michael A. Collins and Ryan P. A. Bettens. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.*, 115:5067–5642, 2015. ISSN 0009-2665. doi: 10.1021/cr500455b. URL `http://pubs.acs.org/doi/abs/10.1021/cr500455b`.

[119] Krishnan Raghavachari and Arjun Saha. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.*, 115:5643–5677, 2015. ISSN 0009-2665. doi: 10.1021/cr500606e. URL `http://pubs.acs.org/doi/abs/10.1021/cr500606e`.

[120] So Hirata. Fast electron-correlation methods for molecular crystals: an application to the α, β_1, and β_2 modifications of solid formic acid. *J. Chem. Phys.*, 129(20):204104, nov 2008. ISSN 1089-7690. doi: 10.1063/1.3021077. URL `http://www.ncbi.nlm.nih.gov/pubmed/19045849`.

[121] Olaseni Sode, Murat Keçeli, So Hirata, and Kiyoshi Yagi. Coupled-cluster and many-body perturbation study of energies, structures, and phonon dispersions of solid hydrogen fluoride. *Int. J. Quant. Chem.*, 109(9):1928–1939, 2009. ISSN 00207608. doi: 10.1002/qua.22022. URL `http://doi.wiley.com/10.1002/qua.22022`.

[122] Olaseni Sode and So Hirata. Second-order many-body perturbation study of solid hydrogen fluoride under pressure. *Phys. Chem. Chem. Phys.*, 14:7765–7779, 2012.

[123] Yanqiang Han, Jinfeng Liu, Lei Huang, Xiao He, and Jinjin Li. Predicting the phase diagram of solid carbon dioxide at high pressure from first principles. *npj Quantum Materials*, 4(1):10, dec 2019. ISSN 2397-4648. doi: 10.1038/s41535-019-0149-0. URL `http://www.nature.com/articles/s41535-019-0149-0`.

[124] Michael A Salim, Soohaeng Yoo Willow, and So Hirata. Ice Ih anomalies: Thermal contraction, anomalous volume isotope effect, and pressure-induced amorphization. *J. Chem. Phys.*, 144(20):204503, may 2016. ISSN 0021-9606. doi: 10.1063/1.4951687. URL `http://dx.doi.org/10.1063/1.4951687http://scitation.aip.org/content/aip/journal/jcp/144/20/10.1063/1.4951687`.

[125] A D Becke. On the large-gradient behavior of the density functional exchange energy. *The Journal of Chemical Physics*, 85(12):7184–7187, dec 1986. ISSN 0021-9606. doi: 10.1063/1.451353. URL `https://doi.org/10.1063/1.451353`.

[126] A Otero-de-la Roza and Erin R. Johnson. Van der Waals interactions in solids using the exchange-hole dipole moment model. *J. Chem. Phys.*, 136(17):174109, 2012. ISSN 1089-7690. doi: 10.1063/1.4705760. URL `http://www.ncbi.nlm.nih.gov/pubmed/22583212`.

[127] P Giannozzi, O Andreussi, T Brumme, O Bunau, M Buongiorno Nardelli, M Calandra, R Car, C Cavazzoni, D Ceresoli, M Cococcioni, N Colonna, I Carnimeo, A Dal Corso, S de Gironcoli, P Delugas, R A DiStasio Jr, A Ferretti, A Floris, G Fratesi, G Fugallo, R Gebauer, U Gerstmann, F Giustino, T Gorni, J Jia, M Kawamura, H-Y Ko, A Kokalj, E Küçükbenli, M Lazzeri, M Marsili, N Marzari, F Mauri, N L Nguyen, H-V Nguyen, A Otero de-la Roza, L Paulatto, S Poncé, D Rocca, R Sabatini, B Santra, M Schlipf, A P Seitsonen, A Smogunov, I Timrov, T Thonhauser, P Umari, N Vast, X Wu, and S Baroni. Advanced capabilities for materials modelling with quantum espresso. *Journal of Physics: Condensed Matter*, 29(46):465901, 2017. URL `http://stacks.iop.org/0953-8984/29/i=46/a=465901`.

[128] Robert M Parrish, Lori A Burns, Daniel G A Smith, Andrew C Simmonett, A Eugene DePrince, Edward G Hohenstein, Uğur Bozkaya, Alexander Yu. Sokolov, Roberto Di Remigio, Ryan M Richard, Jérôme F Gonthier, Andrew M James, Harley R McAlexander, Ashutosh Kumar, Masaaki Saitow, Xiao Wang, Benjamin P Pritchard, Prakash Verma, Henry F Schaefer, Konrad Patkowski, Rollin A King, Edward F Valeev, Francesco A Evangelista, Justin M Turney, T Daniel Crawford, and C David Sherrill. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *Journal of Chemical Theory and Computation*, 13(7):3185–3197, jul 2017. ISSN 1549-9618. doi: 10.1021/acs.jctc.7b00174. URL `https://doi.org/10.1021/acs.jctc.7b00174`.

[129] Roberto Dovesi, Alessandro Erba, Roberto Orlando, Claudio M Zicovich-Wilson, Bartolomeo Civalleri, Lorenzo Maschio, Michel Rérat, Silvia Casassa, Jacopo Baima, Simone Salustro, and Bernard Kirtman. Quantum-mechanical condensed matter simulations with CRYSTAL. *WIREs Computational Molecular Science*, 8(4):e1360, jul 2018. ISSN 1759-0876. doi: 10.1002/wcms.1360. URL `https://doi.org/10.1002/wcms.1360`.

[130] Trygve Helgaker, Wim Klopper, Henrik Koch, and Jozef Noga. Basis-set convergence of correlated calculations on water. *The Journal of Chemical Physics*, 106(23):9639–9646, jun 1997. ISSN 0021-9606. doi: 10.1063/1.473863. URL `https://doi.org/10.1063/1.473863`.

[131] Michael F Peintinger, Daniel Vilela Oliveira, and Thomas Bredow. Consistent Gaussian basis sets of triple-zeta valence with polarization quality for solid-state calcula-

tions. *Journal of Computational Chemistry*, 34(6):451–459, mar 2013. ISSN 0192-8651. doi: 10.1002/jcc.23153. URL `https://doi.org/10.1002/jcc.23153`.

[132] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005. ISSN 1463-9076. doi: 10.1039/B508541A. URL `http://dx.doi.org/10.1039/B508541A`.

[133] Klaus-Peter Huber and Gerhard. Herzberg. *Molecular spectra and molecular structure, volume 4. Constants of diatomic molecules / by K.P. Huber and G. Herzberg. volume 4. Constants of diatomic molecules / by K.P. Huber and G. Herzberg.* Van Nostrand Reinhold, New York, NY, 1979. ISBN 0442233949 9780442233945.

[134] Chris J Pickard and R J Needs. Ab initio random structure searching. *J. Phys. Condens. Mat.*, 23(5):053201, 2011. ISSN 0953-8984. doi: 10.1088/0953-8984/23/5/053201. URL `http://stacks.iop.org/0953-8984/23/i=5/a=053201?key=crossref.44087776832bc9a0edd0dc492a67d376`.

[135] Aurora J. Cruz-Cabeza, Susan M. Reutzel-Edens, and Joel Bernstein. Facts and Fictions About Polymorphism. *Chem. Soc. Rev.*, 44:8619–8635, 2015. ISSN 0306-0012. doi: 10.1039/C5CS00227C. URL `http://xlink.rsc.org/?DOI=C5CS00227C`.

[136] Jonas Nyman and Graeme M Day. Static and lattice vibrational energy differences between polymorphs. *CrystEngComm*, 17(28):5154–5165, 2015. doi: 10.1039/C5CE00045A. URL `http://dx.doi.org/10.1039/C5CE00045A`.

[137] A. K. McMahan and R. LeSar. Pressure Dissociation of Solid Nitrogen under 1 Mbar. *Phys. Rev. Lett.*, 54(17):1929–1932, apr 1985. ISSN 0031-9007. doi: 10.1103/PhysRevLett.54.1929. URL `https://link.aps.org/doi/10.1103/PhysRevLett.54.1929`.

[138] C. Mailhiot, L. H. Yang, and A. K. McMahan. Polymeric nitrogen. *Phys. Rev. B*, 46 (22):14419–14435, 1992. ISSN 01631829. doi: 10.1103/PhysRevB.46.14419.

[139] C. A. Swenson. New Modification of Solid Nitrogen. *J. Chem. Phys.*, 23(10): 1963–1964, oct 1955. ISSN 0021-9606. doi: 10.1063/1.1740623. URL `http://aip.scitation.org/doi/10.1063/1.1740623`.

[140] Steven Buchsbaum, Robert L. Mills, and David Schiferl. Phase diagram of nitrogen determined by Raman spectroscopy from 15 to 300 K at pressures to 52 GPa. *J. Phys. Chem.*, 88(12):2522–2525, jun 1984. ISSN 0022-3654. doi: 10.1021/j150656a018. URL `http://pubs.acs.org/doi/abs/10.1021/j150656a018`.

[141] David Schiferl, Steven Buchsbaum, and Robert L Mills. Phase transitions in nitrogen observed by Raman spectroscopy from 0.4 to 27.4 GPa at 15 K. *J. Phys. Chem.*, 89 (11):2324–2330, may 1985. ISSN 0022-3654. doi: 10.1021/j100257a036. URL `http:`

//pubs.acs.org/doi/abs/10.1021/j100257a036.

[142] Carsten Müller and Denis Usvyat. Incrementally corrected periodic local MP2 calculations: I. The cohesive energy of molecular crystals. *J. Chem. Theory Comput.*, 9 (12):5590–5598, 2013. ISSN 15499618. doi: 10.1021/ct400797w.

[143] Johannes Hoja, Anthony M. Reilly, and Alexandre Tkatchenko. First-principles modeling of molecular crystals: structures and stabilities, temperature and pressure. *WIREs Comput. Mol. Sci.*, 7(1):e1294, jan 2017. ISSN 17590876. doi: 10.1002/wcms.1294. URL `http://doi.wiley.com/10.1002/wcms.1294`.

[144] Y. Huang and G.J.O. Beran. Reliable prediction of three-body intermolecular interactions using dispersion-corrected second-order Møller-Plesset perturbation theory. *J. Chem. Phys.*, 143(4), 2015. ISSN 00219606. doi: 10.1063/1.4927304.

[145] R. W. G. Wyckoff. *Crystal Structures*, volume 1. Interscience Publishers, New York, New York, 2nd ed edition, 1963.

[146] R. L. Mills and A. F. Schuch. Crystal Structure of Gamma Nitrogen. *Phys. Rev. Lett.*, 23(20):1154–1156, nov 1969. ISSN 0031-9007. doi: 10.1103/PhysRevLett.23.1154. URL `https://link.aps.org/doi/10.1103/PhysRevLett.23.1154`.

[147] H. Olijnyk. High pressure x-ray diffraction studies on solid N2 up to 43.9 GPa. *J. Chem. Phys.*, 93(12):8968–8972, 1990. ISSN 00219606. doi: 10.1063/1.459236.

[148] Mikhail I. Eremets, Russell J. Hemley, Ho-kwang Mao, and Eugene Gregoryanz. Semiconducting non-molecular nitrogen up to 240 GPa and its low-pressure stability. *Nature*, 411(6834):170–174, may 2001. ISSN 0028-0836. doi: 10.1038/35075531. URL `http://www.nature.com/articles/35075531`.

[149] Eugene Gregoryanz, Alexander F. Goncharov, Russell J. Hemley, and Ho-kwang Mao. High-pressure amorphous nitrogen. *Phys. Rev. B*, 64(5):052103, jul 2001. ISSN 0163-1829. doi: 10.1103/PhysRevB.64.052103. URL `https://link.aps.org/doi/10.1103/PhysRevB.64.052103`.

[150] Mikhail I Eremets, Alexander G Gavriliuk, Ivan A Trojan, Dymitro A Dzivenko, and Reinhard Boehler. Single-bonded cubic form of nitrogen. *Nature Mater.*, 3(8):558–563, aug 2004. ISSN 1476-1122. doi: 10.1038/nmat1146. URL `http://www.nature.com/articles/nmat1146`.

[151] G. W. Stinton, I. Loa, L. F. Lundegaard, and M. I. McMahon. The crystal structures of δ and δ^* nitrogen. *J. Chem. Phys.*, 131(10):104511, 2009. ISSN 00219606. doi: 10.1063/1.3204074. URL `http://scitation.aip.org/content/aip/journal/jcp/131/10/10.1063/1.3204074`.

[152] Dane Tomasino, Minseob Kim, Jesse Smith, and Choong-Shik Yoo. Pressure-Induced

Symmetry-Lowering Transition in Dense Nitrogen to Layered Polymeric Nitrogen (LP-N) with Colossal Raman Intensity. *Phys. Rev. Lett.*, 113(20):205502, nov 2014. ISSN 0031-9007. doi: 10.1103/PhysRevLett.113.205502. URL `https://link.aps.org/doi/10.1103/PhysRevLett.113.205502`.

[153] D. Laniel, G. Geneste, G. Weck, M. Mezouar, and P. Loubeyre. Hexagonal Layered Polymeric Nitrogen Phase Synthesized near 250 GPa. *Phys. Rev. Lett.*, 122(6):066001, feb 2019. ISSN 0031-9007. doi: 10.1103/PhysRevLett.122.066001. URL `https://doi.org/10.1103/PhysRevLett.122.066001https://link.aps.org/doi/10.1103/PhysRevLett.122.066001`.

[154] Dominique Laniel, Bjoern Winkler, Timofey Fedotenko, Anna Pakhomova, Stella Chariton, Victor Milman, Vitali Prakapenka, Leonid Dubrovinsky, and Natalia Dubrovinskaia. High-pressure polymeric nitrogen allotrope with the black phosphorus structure. *Phys Rev Lett*, page accepted, 2020. URL `http://arxiv.org/abs/2003.02758`.

[155] Watit Sontising and Gregory J O Beran. PHYSICAL REVIEW MATERIALS 3 , 095002 (2019) Theoretical assessment of the structure and stability of the λ phase of nitrogen. *Physical Review Materials*, 3(9):95002, 2019. ISSN 0031-899X. doi: 10.1103/PhysRevMaterials.3.095002. URL `https://doi.org/10.1103/PhysRevMaterials.3.095002`.

[156] Robin Reichlin, David Schiferl, Sue Martin, Craig Vanderborgh, and Robert L. Mills. Optical Studies of Nitrogen to 130 GPa. *Phys. Rev. Lett.*, 55(14):1464–1467, sep 1985. ISSN 0031-9007. doi: 10.1103/PhysRevLett.55.1464. URL `https://link.aps.org/doi/10.1103/PhysRevLett.55.1464`.

[157] J Hooper, A G Hu, F Zhang, and T K Woo. Genetic algorithm and first-principles DFT study of the high-pressure molecular zeta phase of nitrogen. *Phys. Rev. B*, 80(10):104117, 2009. ISSN 1098-0121. doi: Artn104117\rDoi10.1103/Physrevb.80.104117.

[158] Sharon Elizabeth Ashbrook and David McKay. Combining Solid-State NMR Spectroscopy with First-Principles Calculations – A Guide to NMR Crystallography. *Chem. Commun.*, 52:7186–7204, 2016. ISSN 1359-7345. doi: 10.1039/C6CC02542K. URL `http://pubs.rsc.org/en/content/articlehtml/2016/cc/c6cc02542k`.

[159] Beatriz H. Cogollo-Olivo, Sananda Biswas, Sandro Scandolo, and Javier A. Montoya. Ab initio Determination of the Phase Diagram of CO 2 at High Pressures and Temperatures. *Phys. Rev. Lett.*, 124(9):095701, mar 2020. ISSN 0031-9007. doi: 10.1103/PhysRevLett.124.095701. URL `http://arxiv.org/abs/1908.11352https://link.aps.org/doi/10.1103/PhysRevLett.124.095701`.

[160] Harold T. Stokes and Dorian M. Hatch. FINDSYM : program for identifying the space-group symmetry of a crystal. *J. Appl. Crystall.*, 38(1):237–238, jan 2005. ISSN 0021-

8898. doi: 10.1107/S0021889804031528. URL `http://scripts.iucr.org/cgi-bin/paper?S0021889804031528`.

[161] H. T. Stokes, D. M. Hatch, and B. J. Campbell, FINDSYM, ISOTROPY Software Suite, iso.byu.edu.

[162] Joel Bernstein. *Polymorphism in molecular crystals.* Clarendon Press, Oxford, 2002.

[163] Claudio Cazorla and Jordi Boronat. Simulation and understanding of atomic and molecular quantum crystals. *Reviews of Modern Physics*, 2017. ISSN 15390756. doi: 10.1103/RevModPhys.89.035003.

[164] K.D. Nanda and G.J.O. Beran. Prediction of organic molecular crystal geometries from MP2-level fragment quantum mechanical/molecular mechanical calculations. *J. Chem. Phys.*, 137(17), 2012. ISSN 00219606. doi: 10.1063/1.4764063.